

PALACKÝ UNIVERSITY OLMOUC

Faculty of Science

Department of Physical Chemistry



Docking of Growth Regulators to Plant Receptors

Doctoral Thesis

Author:	Mgr. Ing. Václav Bazgier
Supervisor:	prof. Ing. Miroslav Strnad, DSc.
Consultant:	doc. RNDr. Karel Berka, Ph.D.
Study programme:	Chemistry
Study field:	Physical chemistry
Study form:	Daily
Date of submission:	29. 3. 2018

Olomouc 2018

Bibliografická identifikace:

Jméno a příjmení autora	Václav Bazgier
Název práce	Dokování růstových regulátorů do rostlinných receptorů
Typ práce	Disertační
Pracoviště	Katedra fyzikální chemie
Vedoucí práce	prof. Ing. Miroslav Strnad, DSc.
Konzultant	doc. RNDr. Karel Berka, Ph.D.
Rok obhajoby práce	2018
Klíčová slova	Molekulární dokování, kinázy, proteiny, inhibitory, skórovací funkce
Počet stran	50
Počet příloh	12 publikace
Jazyk	Anglický
Abstrakt	<p>Cílem disertační práce bylo nalezení vhodných ligandů pro vybrané nejen rostlinné receptory pomocí různých technik molekulového dokování a vylepšení techniky molekulového dokování návrhem nové skórovací funkce pro vybrané systémy a otestování její kvality.</p>

Bibliographical identification:

Author's first name and surname	Václav Bazgier
Title	Docking of Growth Regulators to Plant Receptors
Type of thesis	Doctoral
Department	Department of Physical Chemistry
Supervisor	prof. Ing. Miroslav Strnad, DSc.
Consultant	doc. RNDr. Karel Berka, Ph.D.
The year of presentation	2018
Keywords	Molecular docking, scoring function, kinase, proteins, inhibitors
Number of pages	50
Number of appendices	12 publications
Language	English
Abstract	<p>The main goals of this work are identification of new ligands of plant-receptors and other proteins or enzymes using several molecular docking approaches, and to design and validate new scoring function on multiple systems to enhance molecular docking techniques.</p>

To my wife Jana and sons Viktor and Vincent and those two underway.

Declaration of the author

I declare that I have worked out this thesis by myself using cited references. Neither the thesis nor any of its part was previously used for obtaining academic degree.

Olomouc, March 2018

Václav Bazgier

I cesta může být cíl...

Acknowledgement

Firstly, I would like to thank my consultant Karel Berka for introducing me to molecular modelling and his patience. Secondly, I would like to thank Pavel Banas for his peaceful patience. Thirdly, I must thank to prof. Miroslav Strnad and prof. Michal Otyepka for support during my studies. Most of all I would like to thank my family for the continuous support and trust in me during my studies.

This work was supported by Internal Grant Agency of Palacký University grants PrF_2013_028, IGA_PrF_2014023, IGA_PrF_2015_027, IGA_PrF_2016_028, IGA_PrF_2017_028, and IGA_PrF_2018_032.

Content

Motivation.....	1
Introduction	3
Molecular Docking Theory	3
Ligand Pose Generation	3
Stochastic Search.....	3
Systematic Search	4
Scoring Functions	4
Force-Field Scoring Function.....	4
Empirical Scoring Function	5
Knowledge-based Scoring Function	6
Quantum Mechanical based Scoring Function	6
Molecular Docking Workflow	7
Preparation of Receptor.....	7
Preparation of the Ligand Collection	8
Grid Definition	9
Interpretation and Visualization of Docking Results.....	9
Validation of Docking Results	9
Methods	11
Molecular Docking Workflow Specifics	11
Molecular Docking Approaches	11
Crystal-like Docking	11
Conformation-wise Docking.....	12
Ensemble Docking.....	12
Blind Docking.....	12
Results.....	13
Projects based on Crystal-like Approach.....	13
Brassinosteroid receptor	13
Kinases.....	15
CDK2 and Abl kinase pyrazolo[4,3- <i>e</i>][1,2,4]triazine inhibitors.....	16
CRK3 kinase 2,6,9-trisubstituted purine inhibitors.....	17
CDK2, CDK5 and AurA kinase pyrazolo[4, 3- <i>d</i>]pyrimidine inhibitors.....	18

CDK2 kinase 7-azaindole based tri-heterocyclic inhibitors	20
Projects based on Conformation-wise Approach.....	21
CDK2 kinase 2,6,9-trisubstituted purine inhibitors.....	21
Cytokinin Oxidase/Dehydrogenase.....	22
Project based on Ensemble Docking Approach	25
Histamine H1 receptor	25
Projects based on Blind Docking Approach	27
Sodium/Potassium Pump	27
Improvement of DOCK Scoring Function	29
Summary	31
Shrnutí.....	33
List of Publications	35
References	37
List of Symbols and Shortcuts.....	45
List of Appendices	47

Motivation

Molecular docking plays a significant role in *in silico* drug design, in which new compounds of biological function are devised using computers. It is a useful tool for prediction of interaction between receptors and potential new chemical compounds. Molecular docking can be used for study of different biological targets and ligands and thus it can help to shorten time for ligand selection and screening in *in vitro* or *in vivo* systems and to construct structure activity relationship. This thesis is focused on usage of molecular docking techniques on receptors of plant growth regulators and other proteins.

Molecular docking is based on description of interactions between the ligand and its receptor. Theoretical part of this thesis is focused on molecular docking background and workflow description. Methodology chapter is focused on usage of different molecular docking strategies and approaches, which we have used thorough individual docking projects as described in the Results chapter:

- First approach, crystal-like docking technique, which uses the position of the similar ligand in crystal structure, is shown on brassinosteroid receptors (BRs, Appendices A ,B and C), kinases (CDK2+ABL2, Appendices D; CRK3, Appendix E; CDK2+Aurora Kinase, Appendix F; CDK2/Cyclin E, Appendix G).
- Second approach, conformation-wise docking technique, which stems from the ambiguity of some ligand crystal poses where the ligand position is not completely certain, and the docking needs to consider multiple possibilities of ligand pose and configuration. We will show this approach on cyclin-dependent kinase 2 (CDK2, Appendix H) and cytokinin oxidase/dehydrogenase receptor (CKX, Appendix I).
- Third approach, ensemble docking technique, which is focused on to catch flexibility of active site by usage of ensemble of structures. We will demonstrate this approach on histamine H1 receptor (H1R, Appendix J).
- Fourth approach, blind docking technique, when active sites of receptor are unknown - sodium potassium pump (NKA, Appendix K).
- Lastly, based on our experience with docking, we have developed new scoring function with exponential repulsion – ExpDock (Appendix L), which can be successfully used for docking in small pockets, where conventional docking on grid does not work properly.

Introduction

Molecular Docking Theory

In this chapter, a short theoretical background about the molecular modelling and molecular docking will be presented. Molecular docking is a method of the molecular modelling discipline.^{1,2,3} The interaction between the active site of receptor and ligand is quantum mechanical in nature. However, the complexity of a biological system cannot be described by quantum mechanical approaches because of the enormous computational cost. Therefore, simplifications of the calculation of receptor-ligand interactions should be used. Major simplification is usually the use of rigid receptor represented by the grid of pre-calculated interaction energies for the individual atoms of the ligand. Also, the calculation of energy is simplified so that an arbitrary score value is typically used instead of proper physically sound energy.

Molecular docking procedure consists of two crucial steps. First step is preparation of the collection of possible orientations and conformations of ligands within the receptor. The second step is the evaluation of collection of poses by so called scoring function.

Ligand Pose Generation

In the search for conformation of the ligands within the active site, the structural parameters of ligands are modified by rotation of dihedral angles, translation and rotation of the ligand. Bonds and angles are usually taken as constant. Two approaches can be recognized – either stochastic/random or systematic.

Stochastic Search

Stochastic search uses random modifications of structural parameters of the ligands. General algorithm generate set of ligand conformation and it is trying to cover the whole range of the energy landscape. Using this strategy, it is possible to increase the probability of finding a global minimum. Well-known algorithms for stochastic search is family of genetic algorithms (GA). The basic principle of GA is in the use of theory of evolution and natural selection for selection of best poses. Firstly, all structural parameters are initialized to the data structures called the gene. Based on this first gene the random search algorithm generates a next set of genes covering a wide area of the energy landscape. The genes representing the structures with the lowest energy values are selected for next generation. The GA algorithm run recursively and when the condition is met (i.e. number of runs, or energy difference between two consecutive runs), calculation is stopped. Other known and used algorithms are Monte Carlo sampling⁴ and simulated annealing⁵. Examples of programs using stochastic search are AutoDock⁶, AutoDock Vina⁷, GOLD⁸, or LigandFit⁹.

Systematic Search

Systematic search generally use degrees of freedom to prepare all possible conformation of ligand depending on the binding site of protein by systematic variation of the starting structure of the ligand.¹⁰ Generally, brute systematic search approach is limited by the number of rotatable bonds that are numerable. This is usually overcome by systematic building from individual rigid building blocks that are docked first and followed by their connection by flexible linkers. Examples of programs using systematic search algorithms are DOCK¹¹, GLIDE¹², eHiTS¹³ and FlexX¹⁴.

Scoring Functions

Scoring function is function to calculate the binding score/energy between the proposed ligand pose and receptor. The final binding energy is the sum of the most important physical-chemical parameters. There are four general classes of scoring function:

- Force-field based
- Empirical methods
- Knowledge based
- Quantum mechanical-based

This work is mainly focused on Force-Field, but all classes of scoring functions are briefly described in the following sub-chapters. It is important to emphasize that there are many different scoring functions today.¹⁵

Force-Field Scoring Function

Force field-based methods use terms derived from molecular mechanics calculations to capture the physicochemical forces governing protein-ligand interactions. Basic principle of this class of scoring function is to computation of non-bonded interaction such as van der Waals and Coulombic potentials between receptor and ligand. The total energy of the studied system can be written as:

$$E_{Total} = E_{bonded} + E_{nonbonded} \quad (1)$$

Where E_{bonded} is term with summation of all bond terms:

$$E_{bonded} = \sum_{interactions} + \sum_{angles} + \sum_{torsion\ angles} \quad (2)$$

Each term means:

$\sum_{interactions}$ – Energy of all bond interactions between atoms within the molecule. Bonds are usually kept constant within docking.

\sum_{angles} – Energy of all angles within the molecule. Angles are usually kept constant within docking.

$\sum_{torsion\ angles}$ - Energy of all torsion angles within the molecule. It is usually neglected during docking

$E_{nonbonded}$ is term accounting for nonbonding interactions:

$$E_{nonbonded} = \sum_{vdw} + \sum_{els} \quad (3)$$

Σ_{vdw} - Energy of all van der Waals interaction pairs, where

$$E_{vdw} = \sum \varepsilon_{ij} \left[\left(\frac{R_{ij}}{r_{ij}} \right)^{12} - 2 * \left(\frac{R_{ij}}{r_{ij}} \right)^6 \right], \quad (4)$$

where ε_{ij} and R_{ij} stand for the potential well-depth and radius of van der Waals pair parameters, respectively, r_{ij} is the distance between atoms and Σ_{els} is coulombic energy of all pairs of electrostatic interaction atoms

Especially nonbonded interactions are important during molecular docking as these interactions are governing the ligand-receptor interactions.

Detailed description of individual force fields can be found elsewhere^{16,17,3} (Table 1) and depend on the core type AMBER¹⁸, OPLS¹⁹, UFF²⁰ and MMFF94²¹.

Table 1. Force-field-based programs for molecular docking

Program name	License	Link
DOCK ²²	Free for academic use	http://dock.compbio.ucsf.edu
AutoDock ⁶	Free for academic use	http://autodock.scripps.edu
GoldScore ⁸	Licenses required	https://www.ccdc.cam.ac.uk
ICM ²³	Licenses required	https://www.molsoft.com
LigandFit ⁹	Free for academic use	https://www.phenix-online.org
Molegro Virtual docker ²⁴	Licenses required	http://www.molegro.com/
SYBYL ²⁵	Licenses required	https://www.certara.com/
MedusaScore ²⁶	Unknown license	Unknown address

Empirical Scoring Function

The classical empirical scoring function is using experimental measured data of binding affinity of the receptor – ligand complex.²⁷ Main idea of this function is reproduction and approximation of measured data. Each term of the empirical scoring function describes one type of molecular properties, which can be tuned by synthetic chemist to emulate physicochemical events involved in the formation of the ligand-receptor complex. Therefore empirical scoring function includes number of hydrogen bonds, metal bonds, lipophilic interactions, and others including desolvation and entropic effects. The example of the implementation of empirical scoring function is ChemScore in the GOLD software²⁸

$$\begin{aligned} \text{ChemScore} = & S_{H\text{-bond}} + S_{\text{metal}} + S_{\text{lipophilic}} + \\ & + P_{\text{rotor}} + P_{\text{strain}} + P_{\text{clash}} + P_{\text{covalent}} + P_{\text{constraint}} \end{aligned} \quad (5)$$

where, S_x is linearized terms for description of specific interactions such as hydrogen bond, interaction with metal, lipophilic interaction and others. P_x terms represent the several types of physical contributions to binding. The list of programs that use the empirical scoring function is listed below (Table 2).

Table 2. Empirical scoring-based programs for molecular docking

Program name	License	Link
Glide ¹²	Licenses required	https://www.schrodinger.com/glide
ChemScore ²⁹	Licenses required	https://www.eyesopen.com/
SCORE ³⁰	Unknown license	Unknown address
LUDI ³¹	Unknown license	Unknown address
LigScore ³²	GNU Lesser GPL	https://integrativemodeling.org/
FlexX ¹⁴	Licenses required	https://www.biosolveit.de/FlexX/
eHiTS ¹³	Free for academic use	http://www.simbiosys.ca/

Knowledge-based Scoring Function

The core of this kind of scoring functions is the knowledge of individual relationships based on the analysis of structural data obtained from databases of protein-ligand complexes. The main idea of knowledge-based binding energy is an assumption that close interactions between a certain type of atoms or functional groups that are attracted to each other occur more frequently (more likely) than would be in the case of random distribution.³³ Such binding energy is further calculated from the Boltzman's equation:

$$\Delta W_{i,j}(r) = W_{i,j}(r) - W(r) = -\ln \frac{g_{i,j}(r)}{g(r)}, \quad (6)$$

where $\Delta W_{i,j}(r)$ is the interacting free energy between atoms i and j at the point of interaction in position r , $W(r)$ is the free energy between these atoms in the position r , $W(r)$ is the average free energy at position r , $g_{ij}(r)$ is the average occurrence of the pair of atoms, $g(r)$ is the occurrence of atoms in position r . DrugScore is typical implementation of this type of scoring function³⁴ and other programs are listed below (Table 3).

Table 3. Knowledge-based programs for molecular docking

Program name	License	Link
AutoDock Vina ⁷	Free for academic use	http://vina.scripps.edu
SMoG ³⁵	Free for academic use	http://smog-server.org/
DrugScore ³⁶	Free for use	http://pc1664.pharmazie.uni-marburg.de/drugscore/
MotifScore ³⁷	Free for academic use	https://sourceforge.net/projects/msdock/
PoseScore ³⁸	GNU Lesser GPL	https://integrativemodeling.org/
RF_Score ³⁹	Free for academic use	https://github.com/oddt/rfscorevs

Quantum Mechanical based Scoring Function

Nowadays exist more than three main classes of scoring function mentioned above. Brand new and interesting classes are Quantum mechanism based scoring functions (QM).⁴⁰ The best known application of this class is using PM6 semiempirical method from the laboratory of prof. Hobza.⁴¹

Molecular Docking Workflow

Similarly like in laboratory, molecular docking use work protocol for the project realization. Each docking project consists of several steps. The most important steps are following (Figure 1):

- Preparation of collection of ligands and receptors
- Localization of active site or generally docking sites for docking grid definition
- Molecular docking (search algorithm, scoring function)
- Analysis of resulting ligand poses and interaction energies

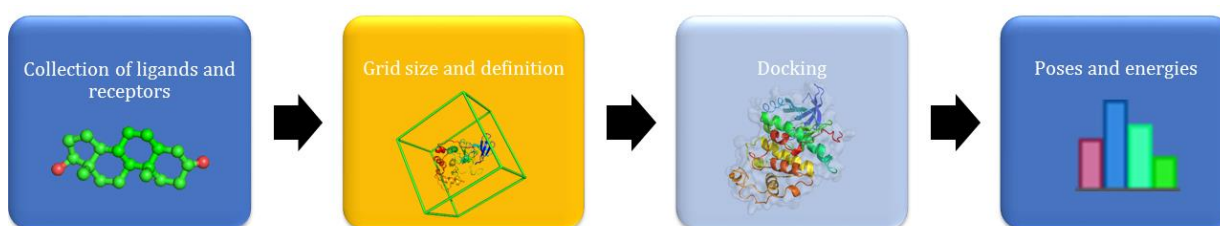


Figure 1. Schematic overview of drug discovery workflow.

Preparation of Receptor

Firstly, a very important step is identification of active site of receptor and important amino acids around. The starting point for obtaining a 3D structure of receptor is public database of protein structures like wwPDB⁴². For identification of active site it is possible use specific software tools^{43,44,45,46} or use software for visualization of protein to manual identification⁴⁷. The active site of the receptor is most commonly identified by the presence of natural ligand or inhibitor within the protein structure. Also identification of amino acids in contact with ligand and analysis of their flexibility may play key role later in molecular docking. It is also possible to prepare a mutant structures based on modification of amino acids around active site using designed software tools⁴⁷.

After identification of the active site, the protein structure must be cleaned from crystallization artefacts and shortcomings. The presence of water molecules can also play a role, therefore the structural water molecules position should be analyzed and such water molecules should be retained within the structure.

Next, hydrogen atoms and partial atom charges are assigned depending on the docking method (Table 4).

If the required 3D structure of receptor is not available, than homology model needs to be prepared based on a related structure using known reliable approaches and applications^{48,49,50}.

Table 4. List of programs for preparation of receptors and different atomic models

Programs for receptor preparation	Atom charge models
AutoDock Tools ⁶	Restrained ElectroStatic Potential (RESP) ⁵¹
UCSF Chimera ⁵²	AM1-BCC ⁵³
PyMol ⁴⁷	Gasteiger ⁵⁴
Maestro suite ⁵⁵	

Preparation of the Ligand Collection

Some care is also needed for the selection and preparation of ligand collection that will be docked into the prepared receptor. The selection of the ligands can be driven by either set of experimentally prepared or would be prepared ligands in case of focused studies or generally from subset of all available compounds from large datasets in case of exploratory studies in a quest of new scaffolds. Since molecular docking requires ligands' correct 3D structure, all possible tautomers have to be generated at the beginning of the process, followed by the preparation of minimized 3D structure of each ligand from 1D (SMILES, InChI) or 2D source (SDF, CDX).

In the case of manual preparation of ligands, it is possible to draw each structure individually or modify starting compound. Multiple programs are capable of such task, e.g. ChemSketch⁵⁶, Marvin Suite⁵⁷, Maestro⁵⁵, etc. In the case that 1D or 2D source is available, there are multiple programs for preparation of 3D structure, e.g. Open Babel⁵⁸, LigPrep⁵⁹, RDKit⁶⁰ and others. In some cases it is necessary to optimize the ligand using molecular dynamics⁶¹. It is worthwhile to check subset of prepared 3D structures prior docking as the wrongly prepared ligand structure will prepare only wrong docking results.

In the case of exploratory studies, when we do not have focused ligand library at hand, it is possible to use freely available databases of compounds. Nowadays several different specialized compound databases exist in different variants of representation such as 1D, 2D or even 3D. A common functionality of these databases is the ability to select ligands based on various parameters. This means that it is possible download specifically selected compounds based on molecular weight, count of all atoms or heavy atoms, similarity with the pattern, Lipinski's rule of five⁶², or another parameter and their various combinations. Available databases with user-friendly web interface are e.g. ZINC⁶³, PubChem⁶⁴, ChEMBL⁶⁵, ChemSpider⁶⁶, DrugBank⁶⁷, Binding DB⁶⁸, eMolecules⁶⁹, or ChemBank⁷⁰. After selection of the ligands, it is possible to download, process and prepare them for molecular docking.

Preparation of ligands for docking is very similar to the preparation of receptors. That means addition of hydrogens and partial charges to each atom. Sometimes it is necessary to modify or fix the dihedral angles in the ligands using programs like AutoDock Tools.

Grid Definition

Molecular docking programs like AutoDock⁶ or DOCK²² use for calculation of final pose and binding energy pre-calculated N-dimensional matrix called grid. This grid contains pre-calculated protein interaction values for any atom type that would fit to the grid point. The pre-calculation reduces the time of final scoring as it transfers more complex calculations of nonbonded interaction in any spot of the active site to simple addition of interaction energy values for each atom in ligand. Grid size and position often depend on localization of active site or/and maximum size of ligand in the collection.

Interpretation and Visualization of Docking Results

Generally, molecular docking outputs have two main parts. First part, it is file with final poses localized inside grid and sorted by the energy. Second part is the binding energy of each pose usually expressed in any arbitrary score or in kilocalories per mole (kcal/mol) when these energies are fitted to Gibbs free energy of binding (ΔG)⁷¹. Some program like DOCK or AutoDock provide additional output information. Very useful is usually the value of root-mean-square deviation (RMSD) which indicate the shift of the output pose from either the best score poses or if used externally for instance for the shift from pre-defined motif, e.g. ligand in crystal structure.

Poses can be viewed in visualization software like PyMOL, Maestro or Chimera, where manual evaluation and interpretation can be performed, for example measurement of the hydrogen bond distance between ligand and amino acid of receptor. The molecular docking results can be further analyzed and used for correct interpretation of the binding between receptor-ligand. Results from molecular docking can be also used for comparison to another drug design methods like pharmacophore⁷² or QSAR⁷³.

Validation of Docking Results

Validation of molecular docking results with experimental values is important. The most typical values from the experimental measurements are half maximal inhibitory concentration (IC₅₀) or the inhibitory constant (K_i) in the case of development of inhibitors. These experimental values can be transferred to each other and then these values can be compared with molecular docking results energies using statistical methods, e.g. regression⁷⁴.

In the case of development or optimization of molecular docking program or scoring function itself, it is possible to use prepared datasets of receptor-ligand complexes. Directory of useful decoys⁷⁵ (DUD) with 40 protein-ligand complexes, SB2012 docking validation dataset containing 1043 protein-ligand complexes⁷⁶ or Wang dataset with 100 protein-ligand it is freely available for download on internet¹⁵. This dataset usually contains separate receptor, ligand and its experimental binding energy and decoys – compounds that look alike ligands but are not binding to a given receptor.

A separate point of validation docking results is the control of pose from docking to a pose in a crystal structure. This procedure is often called re-docking. This step is meant to visualize, whether docking will be able to repeat the experimentally detected pose.

Methods

Molecular Docking Workflow Specifics

In this dissertation thesis, the molecular docking workflow of all projects is typically composed of several parts. As soon as the biological target is set as a goal of new project, preparatory works begin.

Firstly, selection of relevant 3D structures of biological target is manually explored. If the active site in structure is known from the presence of any ligand, then the list of the amino acids probably interacting with the ligand is checked for completeness and flexibility.

Secondly, preparation of receptor using AutoDock Tools or UCSF Chimera (add hydrogens, add charges) starts. A considerable step is the preparation and localization of the active site. Later, the collection of ligands' 3D structures is collected and prepared together with nature ligand of receptor for internal validation.

Molecular docking programs AutoDock Vina and/or DOCK with the optimized version of our new scoring function expDOCK are used in the majority of the projects realized in this thesis. The biggest advantage of AutoDock Vina is its speed, when the calculation time is in the order of tens of seconds per ligand. On the other hand, the docking with DOCK has extremely slowly only the grid preparation, the calculation of binding energy of set of ligands using scoring function is fast.

Final step after computation is the interpretation of results, validation and reporting.

Molecular Docking Approaches

Molecular docking method is typically used for generation and favorable scoring of flexible ligand within defined rigid binding site. However, this idealized situation seldom leads to the success on the first trial. To obtain successful docking and scoring of poses, it is usually beneficial to use as much of additional external information as possible.

Here, we list a set of approaches we have pursued during individual molecular docking projects. We have sorted them according to the methodology and describe them briefly in the following subchapters together with their advantages and disadvantages (Table 4).

Crystal-like Docking

This class of molecular docking is using the standard docking protocol and interpretation of results. The only difference is that not the pose with the best energy, but the set of "crystal-like" poses from collection of results are selected for comparison. Term "crystal-like" means that the preferred docked ligand pose has the smallest RMSD value against the natural ligand present in a crystal or with certain typical binding motif. An example of such motif is for instance of well-known binding motif within CDK2 kinases. The goal of these modifications is preserved by the crystal-like position.

Conformation-wise Docking

Different approach for crystal-like docking is using conformation-wise protocol. In some crystal structures, it is possible that the natural ligand exists in two or more positions or configurations. It means that all possible positions in the active site must be generated and compared. It is important to decide at the beginning of this approach which pose will be preferred. This approach is applicable in the case of almost symmetrical planar ligands or in the case of multiple options of hydrogenation of the ligand and/or active site. The next steps are similar to previous approach.

Ensemble Docking

Classical view of molecular docking is focused on the docking of several ligands into a single receptor. The collection of ligands is typically prepared as a flexible part of the molecular docking and receptor is rigid. Main goal of the ensemble docking is study of interactions of ligand (still flexible) and different snapshot of receptor over time of movement in order to cover the flexibility of the receptor. Molecular docking protocol of this approach firstly utilize some means to generate ensemble of structures and this is followed by docking into individual snapshots in the ensemble of structures.

Blind Docking

Basic approach of blind docking is based on ignorance of the position of the active site. Main goal of blind docking is identification of potential active site of receptor. Methodology consists of several steps to identify a few potential active sites. Typically, first round is provided with a grid that can encompass whole protein, while following rounds investigate hot spots with multiple ligands present within the receptor. Data results are interpreted and based on these results is selected the best tip on active site of receptor.

Table 4. Advantages and Disadvantages of Molecular Docking Approaches

Molecular docking protocol	Advantages	Disadvantages
Crystal-like	<ul style="list-style-type: none">• Similarity with natural ligand	<ul style="list-style-type: none">• Only nature ligand-like binding can be obtained
Conformation-wise	<ul style="list-style-type: none">• Extension of crystal-like docking for multiple pose situations	<ul style="list-style-type: none">• More complicated evaluation
Ensemble	<ul style="list-style-type: none">• Allow flexibility in receptor	<ul style="list-style-type: none">• A complex problem• Processing results
Blind	<ul style="list-style-type: none">• Identification of possible binding motifs	<ul style="list-style-type: none">• Correct interpretation requires experimental validation• Multiple grid setup

Results

This chapter is divided into several parts describing our individual approaches to the molecular docking. Firstly, the crystal-like approach is discussed on molecular docking of brassinosteroid receptor and several kinases. Secondly, conformation wise approach is shown on study of cytokinin oxidase/dehydrogenase receptor. Thirdly, ensemble docking approach is reflected on the study of possible antihistaminics interacting with histamine H1 receptor. Next, blind docking approach is demonstrated on localization of multiple binding sites on sodium potassium pump. The last part is dedicated to improvement and refactorization of the molecular docking program DOCK and development and testing of new scoring function.

Projects based on Crystal-like Approach

The core of this approach is in search of crystal-like pose of each of newly designed ligands. Ligand must be present in the crystal structure with a typical bonding motif. The molecular docking then proceeds. Only those poses that share binding motif with the crystal one are scored and compared with experimental data. Herein, we have used this approach for the design of brassinolide analogs for the brassinosteroid receptor, and for the design of several kinase inhibitors.

Brassinosteroid receptor

Brassinosteroids are stress-response plant hormones⁷⁷, but they are acting in low dosages and they are known, as steroids, to be hardly synthesizable. Therefore, it is hard to obtain them in a quantity that would allow their use as new intelligent fertilizers. Brassinosteroid-Insensitive 1 (BRI1) receptor structure with brassinolide bound (PDBID: 3RGZ with 2.281 Å resolution)⁷⁸ has shown how brassinosteroids bind to this receptor which then transfers the signal to the plant cell. BRI1 structure has shown possible new way for the development of new and more easily synthesizable brassinosteroid analogs with use of molecular docking approach. The structure also showed the positions of the major binding spots on the brassinolide molecule and that the binding pocket on the surface of BRI1 is almost completely complementary to brassinolide (Figure 2). This explains why this brassinosteroid is unparalleled in effect.

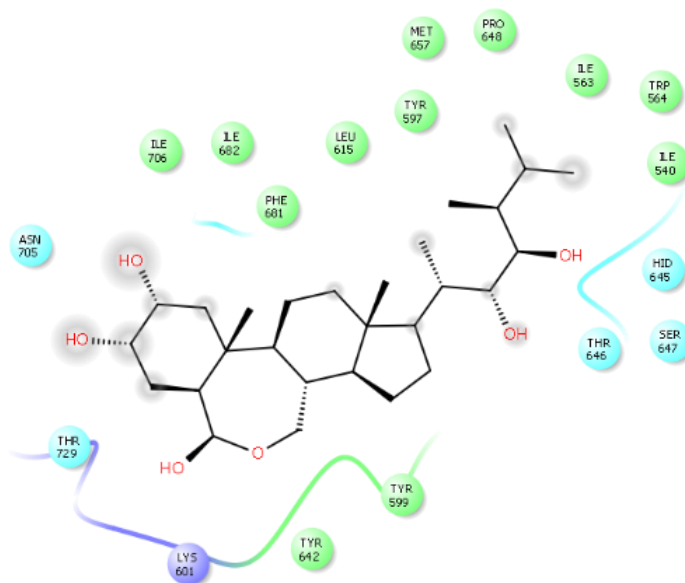


Figure 2. Binding pattern of brassinolide within BRI1 receptor in 2D visualization. Image was produced in Maestro program.

We have gathered the collection of brassinolide analogues with modification of side-chain based on pattern we analyzed – first with shortened side chains and second with elongated side chains.

The first group of designed new analogs with shortened side-chains were docked and then analyzed. Molecular docking showed that this group of shortened side-chain ligands bind in two distinguishable poses – the highest energy poses, where the ligand bound more deeply into the binding pocket, and weaker ones in crystal-like pose (Figure 4). Only the latter results were confirmed experimentally.⁷⁹ (Appendix A)

On the other hand, second group of prepared analogs contained an extension of side-chain with bulkier functional groups. Molecular docking results showed that this group of analogs bind in to active site in a crystal-like manner with stronger interaction energy than previously unbeatable natural ligand brassinolide (Figure 4, 5). Again, the best crystal-like poses were compared with experimental values and confirmed molecular docking results. The disadvantage of this collection of analogs is their worse preparation in laboratory, but it has shown the proof-of-principle of future brassinosteroid analog design.^{80,81} (Appendices B and C)

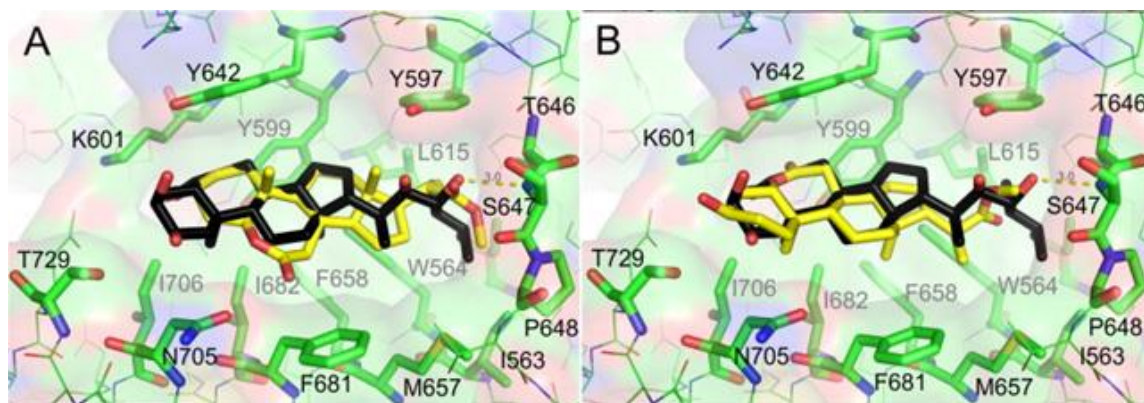


Figure 4. High energy pose (left) and crystal-like pose (right) of the shortened side-chain brassinolide analog in comparison with brassinolide (black). Figure reprinted from Ref. 79 (Appendix A).

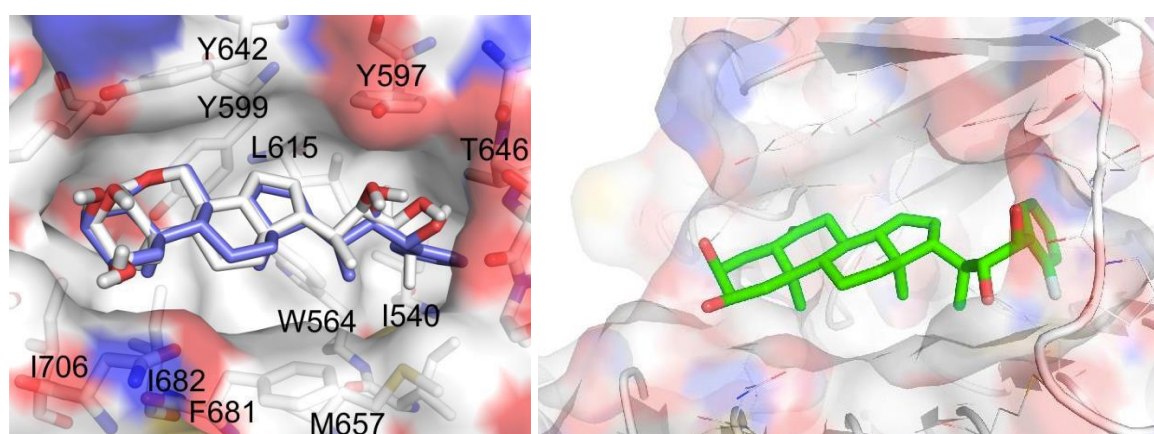


Figure 5. Almost perfect crystal-like pose of brassinolide analog with phenyl functional group (violet) in comparison with brassinolide (white). Figure reprinted from Ref. 80 (Appendix B) and Ref. 81 (Appendix C).

Kinases

Kinases are enzymes that play crucial roles during cell cycle.⁸² They catalyze the transfer of phosphate groups from phosphate-donating molecules to specific proteins and this process is used for regulation of many signaling pathways. Kinase inhibitors have been successfully used to inhibit proliferation and angiogenesis for cancer therapy. However, kinase deregulation has been demonstrated to play a key role in almost all major disease areas. Kinase inhibitor drug discovery programs have therefore broadened their focus to expand the range of kinase targets. The realization of many of these projects showed that the results of molecular docking did not correlate well with the experimental IC₅₀ data, but when only crystal-like poses were used, the agreement with the experiment was substantial. Here, we will show, how can be used crystal-like docking for different kinase targets.

CDK2 and Abl kinase pyrazolo[4,3-e][1,2,4]triazine inhibitors

The first kinase project was focused on a preparation of new cyclin-dependent kinase 2 (CDK2) and Abelson tyrosine-protein kinase (Abl) inhibitors. Both kinases are important for cell cycle regulation and play role in chronic myelogenous leukaemia^{83,84}.

Our collection of ligands were sulfonamide derivatives of pyrazolo[4,3-e][1,2,4]triazine (Figure 7, Appendix D). The crystal structures of CDK2 with ligand NU6102 (PDBID: 2C60)⁸⁵ and Abl with ligand PD180970 (PDBID: 2HZI)⁸⁶ were used as the receptors for docking. The aim of this study was to confirm binding pose of new derivatives to active sites of receptors and compare the experimental results with molecular docking values and to select the best inhibitors for further research.

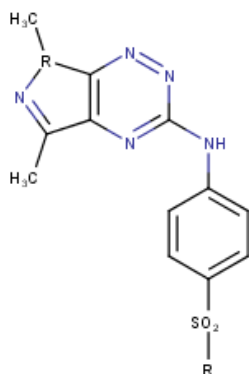


Figure 7. Scaffold of proposed kinase inhibitors based on sulfonamide derivatives of pyrazolo[4,3-e][1,2,4]triazine. Figure reprinted from Ref.⁸⁷ (Appendix D).

Molecular docking confirmed and explained negative results obtained in the biochemical testing of some of the compounds into CDK2. Results showed that designed new compounds do not keep the binding motif well-known in CDK2 active site. Well-known purine CDK inhibitors such as roscovitine⁸⁸ or following structures⁸⁹ (Figure 8 left) typically form donor-acceptor-donor H-bonding motif formed by the N7 group and the secondary amino group at C6 that interacts with the hinge region of CDK2 with backbone of Glu81 to Leu83 at residues, gives the known CDK2 inhibitors a well-defined binding orientation within the CDK2 active site. However, pyrazolo[4,3-e][1,2,4]triazines cannot be substituted at the corresponding positions, it is likely that the methyl substitution of the pyrazole ring would both abolish one of these H-bonds and also form an unfavorable steric interaction that would discourage the binding of pyrazolo[4,3-e][1,2,4]triazines compounds in the active site (Figure 8 right).

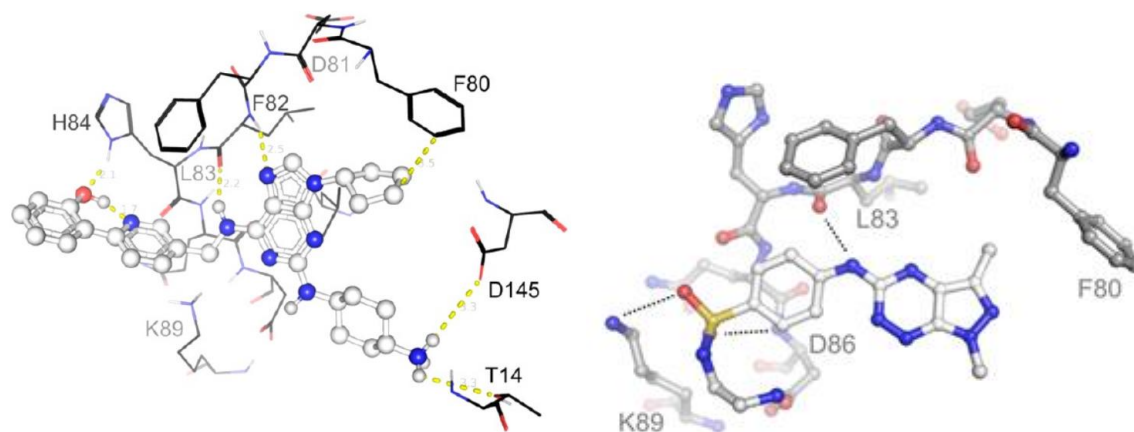


Figure 8. (left) Binding pattern of purine CDK2 inhibitor from our study Gucky et al [89] (Appendix H). Interaction to backbone of Leu83 via H-bonds is shown, H-bond between aromatic hydrogen at C8 with Asp81 backbone is weaker. (right) Binding pose of pyrazolo[4,3-e][1,2,4]triazine compound showing lack of interactions with CDK2 hinge pattern.[87]

On the other hand, molecular docking described the binding motif and important interaction of pyrazolo[4,3-e][1,2,4]triazine compounds with ABL2 kinase and confirmed favorable biochemical testing. The compounds bound to ATP-binding site of Abl kinase interacting with the protein via H-bonds with backbone of Met318 in crystal-like pose. In addition, for the best inhibiting compounds, we observed new possible H-bond between Ile313 and the secondary amino groups on the sulfonamide appendages. The output of this project identified two compounds as potential selective inhibitors of the ABL2 protein kinase (Figure 9).

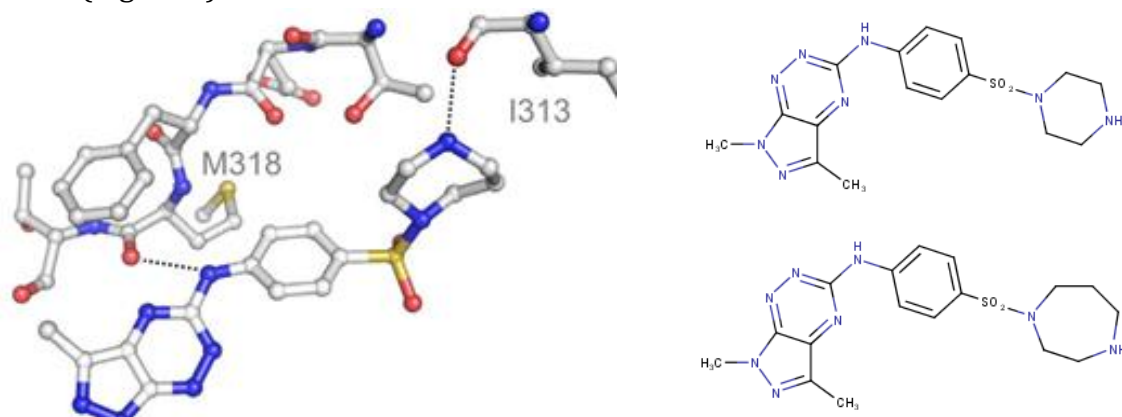


Figure 9. (left) binding pose of best compound within Abl kinase ATP-binding site. (right) 2D structures of two new potentially inhibitors of the ABL2 protein kinase. Figure reprinted from Ref. ⁸⁷ (Appendix D).

CRK3 kinase 2,6,9-trisubstituted purine inhibitors

Next project based on crystal-like docking was focused on Cdc2-related protein kinase 3 (CRK3) and preparing of new analogs of 2,6,9-trisubstituted purines. CRKs belong to the group of serine/threonine protein kinases that play key roles in regulating the cell cycle and proliferation. CRK3 kinases affect progression of Leishmania disease and their

inhibitors can be important for preparing of new treatment⁹⁰. Collection of new compounds is based on substitution of purine skeleton (Figure 10).

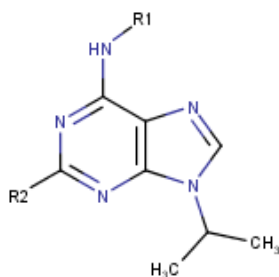


Figure 10. Scaffold of proposed kinase inhibitors based on sulfonamide derivatives of pyrazolo[4,3-*e*][1,2,4]triazine. Figure reprinted from Ref. ⁹¹ (Appendix E).

In this project, it was necessary to prepare *Leishmania major* CRK3 kinase (LmCRK3; UNIPROT O96526) using a homology modeling based on its homology to the human CDK2 kinase (PDBID: 4KD1)⁹². It should be noted that active site of CDK2 and new model of LmCRK3 is very similar but even though it features several amino acid changes, we could use CDK2 crystal poses as a template. Importantly, the H-bonding hinge motif observed in CDK complex is reproduced in CRK3 via main chains of Glu100 and Val102.

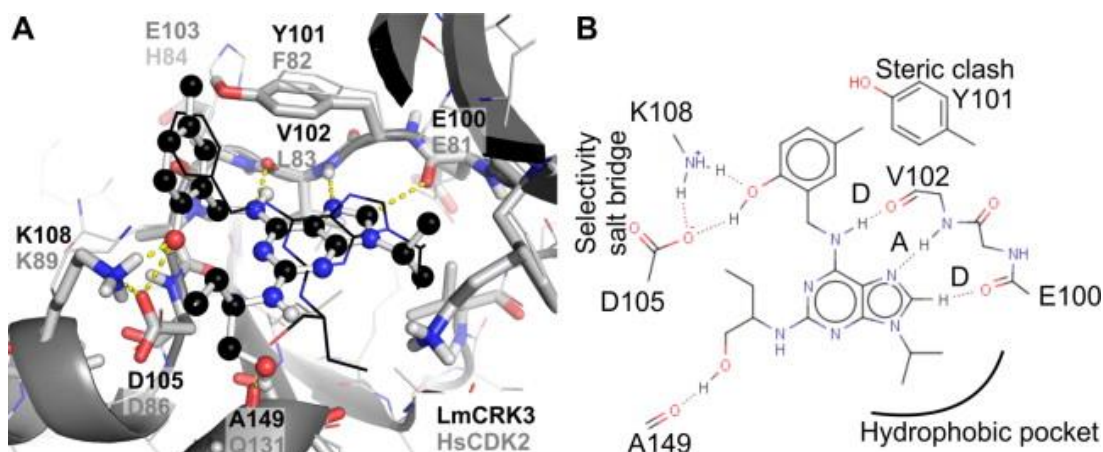


Figure 11. (A) Docked binding pose of new compound 10i (ball-and-stick representation) in LmCRK3 compared with that of roscovitine in CDK2 (lines representation). (B) The selected compound interactions in 2D visualization. Figure reprinted from Ref. ⁹¹ (Appendix E).

Molecular docking showed that all new analogues hold the binding motif of crystal-ligand roscovitine in active site of CDK2 kinase (Figure 11). In this case, docking results explained why *o*-hydroxylated N6 benzyl group and 1-(hydroxymethyl)propylamino group at C2 helps the selectivity and inhibition activity of designed analogs.

CDK2, CDK5 and AurA kinase pyrazolo[4, 3-*d*]pyrimidine inhibitors

Preparation of new selectively inhibiting compounds of CDK2 and CDK5 and Aurora A (AurA) kinase is another project where crystal-like docking was used. Finding of the new inhibitor CDK2 and CDK5 (PDBID: 3DDQ and 1W98)^{93,94} and AurA (PDBID: 3O50)⁹⁵ kinases could potentially become powerful tools for the treatment of cancer^{96,97}. All new

compounds are based on pyrazolo[4, 3-*d*]pyrimidine (Figure 12). Based on laboratory testing, the most effective compound was selected as potentially new inhibitor. On new analogs it was surprising that they are bind to the active site of AurA kinase receptor and inhibit it, whereas its predecessor roscovitine didn't. Molecular docking offers an explanation based on the description of individual interactions between analogues and amino acids around active site.

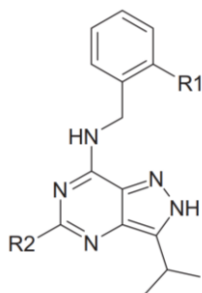


Figure 12. General scaffold of prepared analogs based on pyrazolo [4, 3-*d*] pyrimidine. Figure reprinted from Ref. ⁹⁸ (Appendix F).

Results of molecular docking shows that best compound binds to ATP-binding site of CDK2/cyclin A. Binding is very similar to orientation of crystal ligand roscovitine forming two H-bonds with the main chains of Leu83 and Glu81. Its binding is further stabilized by additional H-bonds involving 2-aminobenzylamino moiety, one with Asp86 as a donor and the other with Lys89 as an acceptor. This could be explanation why best compound binds more powerfully than roscovitine to CDK2.

The interactions between best compound and AurA kinase shows four important H-bonds between active site and compound: one between the Ala213 carbonyl group and N7, another between the Ala213 amino group and N1, a third between the Glu211 carbonyl group and N2, and a fourth between the Pro214 carbonyl group and the 2-aminobenzylamino substituent (Figure 13).

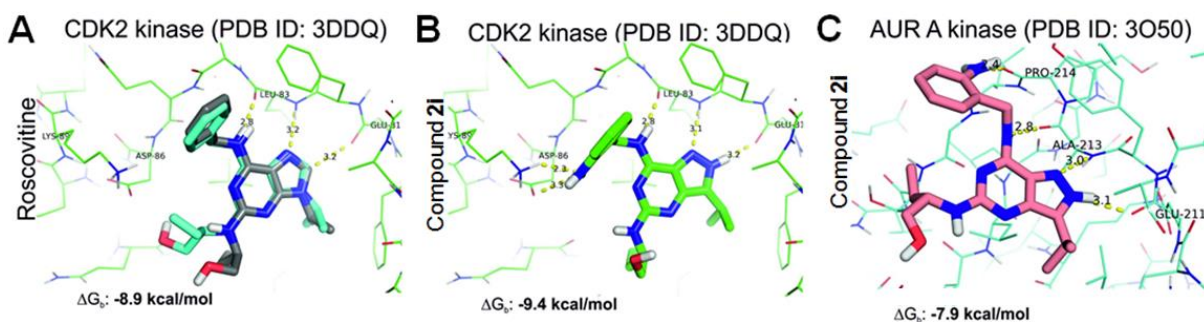


Figure 13. (A) Result of re-docking of roscovitine to active site of CDK2 (gray is crystal like compound and cyan is actual docking pose). (B) Final pose of best pyrazolo[4, 3-*d*]pyrimidine compound to active site of CDK2. (C) Pose of the same best compound in active site of AurA kinase. Figure reprinted from Ref. ⁹⁸ (Appendix F).

The docking thus explained why best pyrazolo[4, 3-*d*]pyrimidine compound binds more tightly than roscovitine to the ATP-binding site of CDK2 and suggest that it may bind to the AurA kinase in a similar fashion.

CDK2 kinase 7-azaindole based tri-heterocyclic inhibitors

This CDK2 project using crystal-like docking was focused on study of new 7-azaindole based tri-heterocyclic compound series. As was already stated earlier, CDKs have oncogenic potential and are responsive to pharmacological inhibition, their inhibitors have therefore attracted great interest as potential anticancer drugs⁹⁹. Main target in this study was CDK2 kinase with Cyclin E (PDBID: 2A4L)¹⁰⁰. Active site of CDK2 kinase is formed by a planar hydrophobic pocket which contains H-bond donor-acceptor-donor pattern of main chains of Glu81 and Leu83.

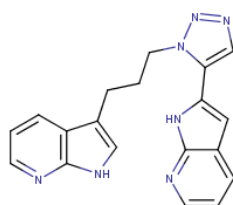


Figure 14. The best compound from the set of 7-azaindole based tri-heterocyclic compounds. Figure reprinted from Ref. ¹⁰¹ (Appendix G).

Molecular docking showed that all prepared compound bind to active site with crystal-like pose interacting interaction of 7-azaindole ring with CDK2 hinge region of main chains of Glu81 and Leu83. The ligand forms a U-shape with two branches formed by 7-azaindoles and triazole part in center of ligand. Triazole shows two interactions – directly with Thr14 and more distantly with Asp145. Second azaindole branch is more flexible with one interaction with main chain of Glu12 and it is more exposed to water (Figure 15).

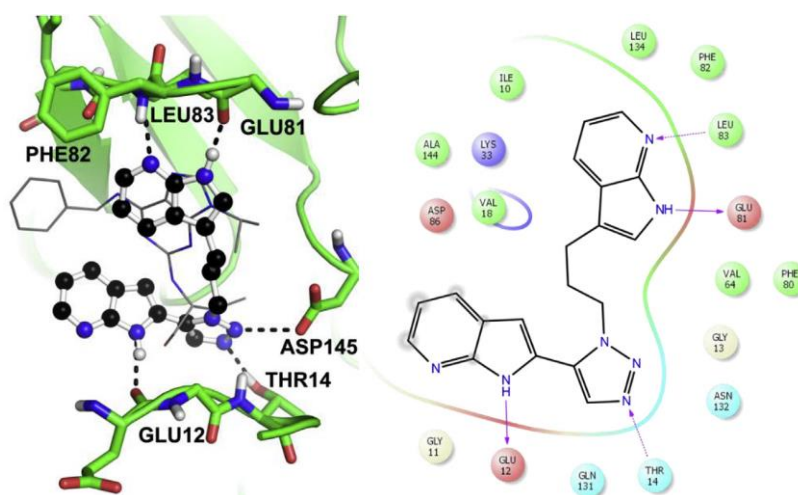


Figure 15. Binding pose of best 7-azaindole tri-heterocyclic compound (ball and sticks) in the active site of CDK2 in comparison with roscovitine (lines). Figure reprinted from Ref. ¹⁰¹ (Appendix G).

Projects based on Conformation-wise Approach

The conformation-wise method extends crystal-like approach. In some case, crystal do not contain entirely unique conformation. In such case, all possible conformations of ligand and protein must be considered. The preferred conformation is later selected on behalf of sensitive evaluation of the individual interaction patterns. This approach has been used to study of cyklin-dependent kinase 2 and cytokinin oxidase/dehydrogenase inhibitors.

CDK2 kinase 2,6,9-trisubstituted purine inhibitors

Last CDK2 project was focused on study of identification and validation of new analogs of 2,6,9-trisubstituted purine inhibitors (Figure 21). Human CDK2 in complex with roscovitine was used as docking receptor for molecular docking (PDBID: 2A4L)⁸⁸.

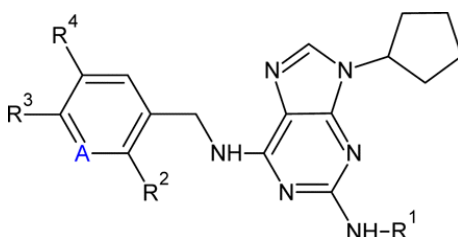


Figure 21. Scaffold of proposed kinase inhibitors based on 2,6,9-Trisubstituted purine inhibitors. Figure reprinted from Ref. ⁸⁹ (Appendix H).

The early realization of this project showed that the results of molecular docking did not correlate well with the experimental IC₅₀ data ($r^2 < 0.2$, data not shown). We analyzed binding motifs of newly designed analogs with crystal poses of roscovitine. In all cases, the binding motif was present interacting with conserved H-bonds of the Leu83 and Glu81 backbones like roscovitine crystal structure, but the binding motif was not the best pose selected by docking.

Therefore, we used a three-step docking procedure, considering the conserved binding motif. Firstly, after the first round of docking, the pose of the best crystal-like position was selected. Secondly, the ligand was constrained by the atoms forming conserved hydrogen bonds and it was treated like a flexible side chain and only a dummy atom of hydrogen was docked to allow adjustment of the flexible part of ligand. Thirdly, the final pose of the ligand stemming from the previous step was then rescored to obtain binding free energies while retaining conserved binding motif.

But even this approach did not work for all compounds. We have observed that biphenyl moiety was usually docked with 90° between the rings, which is energetically unfavorable. For this reason, we have run torsional angle screening using quantum chemical method (DF-PBE/6-311++g(3df,3pf) basis set) to observe torsion angle energy profile and to select minimal energy structures (Figure 22).

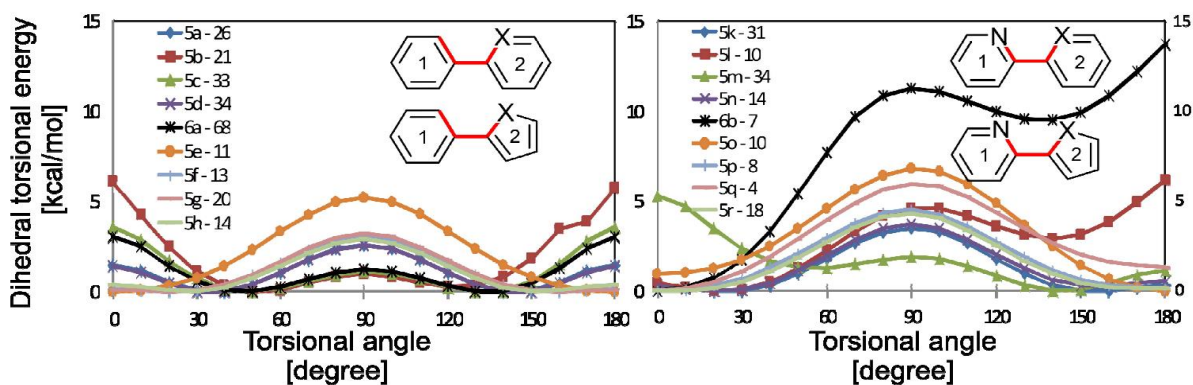


Figure 22. Dihedral scans between the two rings for compounds with two aromatic rings. Inclusion of nitrogen in pyridine in ring 1 enlarges dihedral barrier thus rigidifying the structure.

We have then constrained minimal energy conformations and rerun the molecular docking. Finally, we have received agreement with the experiment and identified prospective binding pose (Figure 8 left).

Conformation wise approach therefore is not straightforward application of molecular docking, but since individual interactions are important for binding, the investment into the analysis of the crystal structure and into the ligand conformations can be the only possibility, how to reach consensus with experimental finding and to show structure activity relationships.

Cytokinin Oxidase/Dehydrogenase

Cytokinins play a very important role in plants such as influencing cell division, root and shoot development and many others.¹⁰² Thidiazuron was used as template for preparing of new collection of potentially new inhibitors of *Arabidopsis thaliana* cytokinin oxidase/dehydrogenase (Figure 19). *Zea mays* cytokinin oxidase/dehydrogenases (ZmCKX1; PDB ID: 2QKN and ZmCKX4a; 4O95 and 4OAL)^{103,104,104} were used to construct homology model of *Arabidopsis thaliana* cytokinin oxidase/dehydrogenase 2 (AtCKX2).

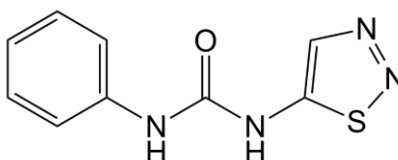


Figure 19. 2D structure of thidiazuron (TDZ) scaffold

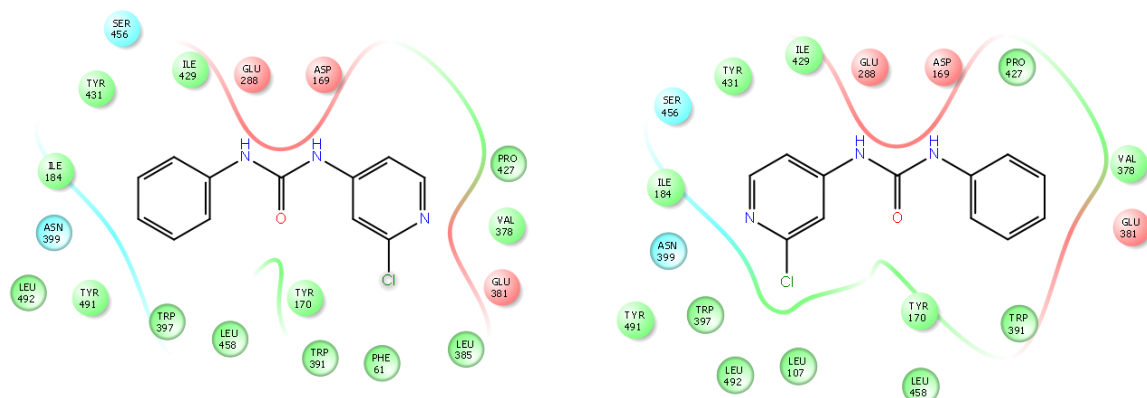
Active site of ZmCKX1 (PDB ID: 2QKN) is made from two pockets connected with narrow planar tunnel controlled by Asp169 interacting with nitrogen atoms with urea-derivative ligand within the active site. Smaller pocket is occupied by flavin adenine dinucleotide

(FAD) cofactor with flavin moiety in direct π -stacking contact with ligand. Larger pocket is more exposed, and it is guarded by Glu 381. Crystal structure contain two possible binding mode of 1-phenyl-3-(2-chloropyridin-4-yl)urea (ClPhU) ligand – both with contact Asp169, but either phenyl or 2-chloropyridin-4-yl moiety interact with FAD (Figure 20).

In order to analyze, which binding pose is the major one, we have redocked ClPhU ligand back into the clean structure. Binding pose with 2-chloropyridin-4-yl rather than phenyl interacting with Glu381 amino acid showed far better binding energy, therefore we can establish that smaller ring will preferentially bind to the smaller pocket with FAD cofactor.

A homology model of AtCKX2, prepared based on the ZmCKX1 structure, revealed that it shares a similar shape and amino acids at almost the same positions but with shifted numbering (AtCKX2/ZmCKX1 amino acids: Asp150/Asp169, Val345/Val378, Trp358/Trp391, Asn366/Asn399, Tyr460/Tyr491, Leu461/Leu492 and especially Glu348/Glu381). The binding pose of ClPhU was also retained.

ClPhU in ZmCKX1 in preferred orientation ClPhU in ZmCKX1 in poor orientation



ClPhU in AtCKX2 in preferred orientation ClPhU in AtCKX2 in poor orientation

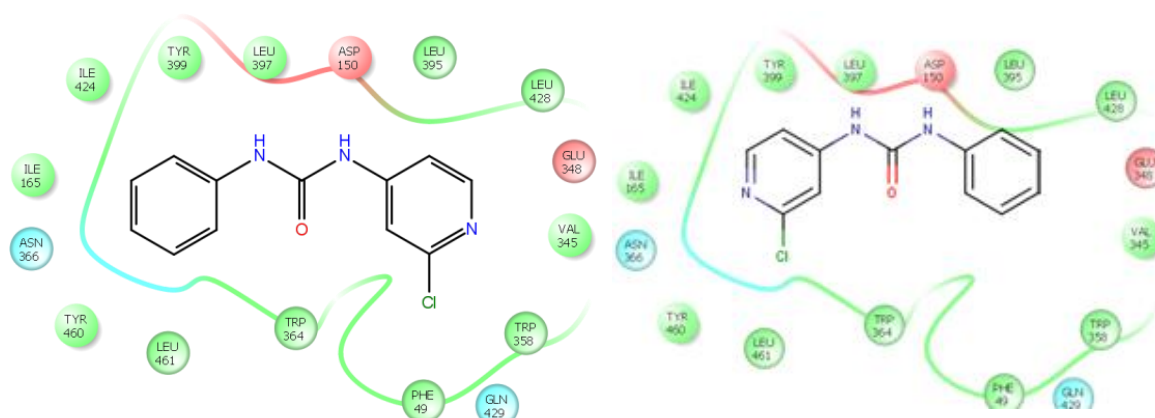


Figure 2. Orientations of ClPhU ligand in an active site of ZmCKX1/AtCKX2.

TDZ derivatives all docked into the active site of the AtCKX2 model with apparent preferred orientation of the smaller 1,2,3-thiadiazolyl ring towards the isoalloxazine ring of FAD. Both urea nitrogen atoms bound to Asp150. The best compounds (noted as HETDZ and 3FMTDZ) from in vitro test (IC_{50} value is $3.9\mu\text{M}$ and $5.5\mu\text{M}$, respectively) had also the best binding energy from molecular docking (ΔG -8.5 kcal/mol and -9.3 kcal/mol, respectively). Interestingly, these two best compounds worked similarly well on ZmCKX1, whereas their inhibition of ZmCKX4a was reversed by the order of magnitude. Molecular docking results and later experimental validation showed that in the case of ZmCKX4a, the poses of those compounds are actually reversed (Figure 21). The orientation of the thidiazuron derivatives therefore plays an important role in the inhibition of CKX enzymes.

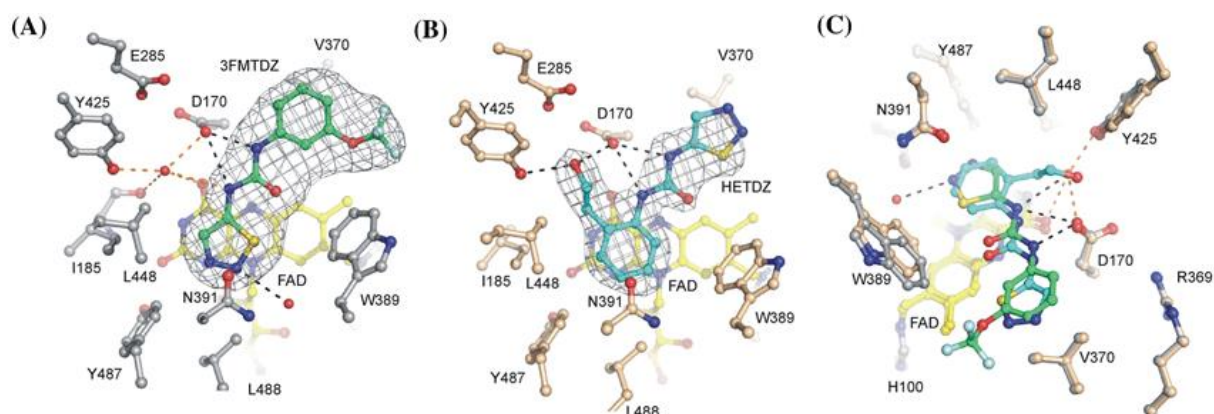


Figure 21. Docked inhibitors 3FMTDZ and HETDZ in the active site of ZmCKX4. Both ligands form a hydrogen bond to the oxygen atom of Asp170 via two nitrogen atoms of the urea backbone. The FAD is colored yellow and neighboring residues are labeled. Figure reprinted from Ref. ¹⁰⁵ (Appendix I).

Project based on Ensemble Docking Approach

Ensemble docking approach enables to dynamically capture the flexibility of the receptor, especially with focus on the active site. Ensemble docking approach use series of receptor structures generated for example by molecular dynamics simulation. Molecular docking is later provided into an ensemble of receptor structures and the best binding poses are reported. This approach thus builds on conformation selection hypothesis, where the active site of receptor is fluctuating, and ligand can be bound only to a subset of receptor conformations. Ensemble docking approach was used to study of potential new inhibitors of histamine H1 receptor.

Histamine H1 receptor

Histamine H1 receptor (H1R) is an important chemical mediator in allergic reactions and inflammation. This project was focused on the design and study of new histamine H1 receptor inhibitors based on pigments from sea urchins shells spinochrome (Figure 23) and spinochrome dimers (Figure 24). Histamine H1-receptor (H1R) structure (PDBID: 3RZE)¹⁰⁶ with doxepin ligand was used for generation of membrane model of this receptor. It was necessary to repair missing loops on top of the ligand-binding pocket of the transmembrane receptor domain and to delete specific viral domain that was used for crystallization using molecular modelling prior molecular docking. As the loop in the close vicinity of the active site was modelled, molecular dynamics simulation was used to sample its position as a lid on the access to the active site.

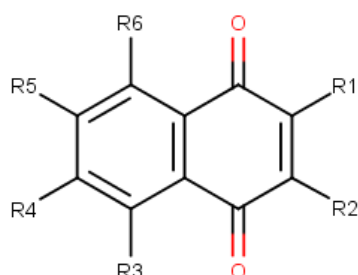


Figure 23. Spinochrome core of monomer structure. Figure reprinted from Ref.¹⁰⁷ (Appendix J).

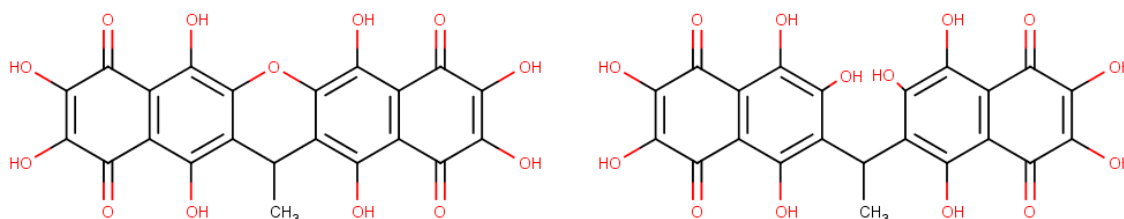


Figure 24. Spinochrome dimers anhydroethylidene 6,6'-bis(2,3,7-trihydroxynaphthazarin) and ethylidene-6,6'-bis(2,3,7-trihydroxynaphthazarin). Figure reprinted from Ref.¹⁰⁷ (Appendix J).

Molecular docking results show that all prepared compounds as monomers and dimers bind to receptor. The best binding potential had spinochrome dimers (Figure 25). On the other hand, it is important emphasize that only monomers could fully bind to active site of receptors, whereas dimers bind rather to the flexible loops regions. As monomers bind better to the active site, they may compete with histamines more effectively than dimers and therefore may be more effective inhibitors than the dimers.

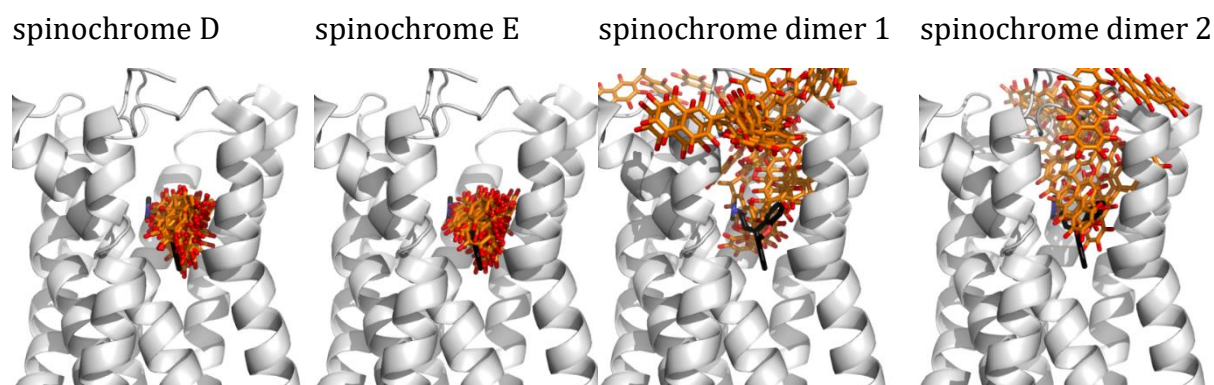


Figure 25. Docking poses of individual spinochrome in H1R crystal structure. While monomeric spinochrome ligands bind preferentially to the similar position as doxepin (black sticks with balls), dimeric spinochrome ligands bound mostly to the loops in the access of the active site. Figure reprinted from Ref.¹⁰⁷ (Appendix J).

Projects based on Blind Docking Approach

Blind docking approach is focused on finding and identifying a potentially new active site. The approach is composed of several steps, where each step iteratively identifies individual candidates for active site. The final set of candidates is then evaluated. The results of the approach have to be validated and compare with experimental techniques. The interaction sites of fluorone dyes with sodium/potassium pump were identified using this approach.

Sodium/Potassium Pump

Sodium/potassium pump or Na^+/K^+ -ATPase (shortly NKA) is one of the most important enzymes involved in the animal cell metabolism. It is a member of the P-type ATPase superfamily of the active cation transport proteins. In this project, we have studied how this protein interacts with fluorone dyes, which can be used as reversible inhibitors of NKA (fluorescein, eosin Y, erythrosine B and rose bengal; Figure 26). NKA is composed of several biological units, where some have known structures and others have to be modeled. Specifically, nucleotide-binding domain (N-domain) and phosphorylation domain (P-domain) together form so called C45 loop, that can be used as a soluble model of ATP-binding active site. Molecular modelling was used to prepare C45 loop. Human variant of whole NKA and of C45 loop were prepared in two conformation - in open and closed form using homology model based on spiny dogfish NKA (PDBID: 2ZZE)¹⁰⁸.

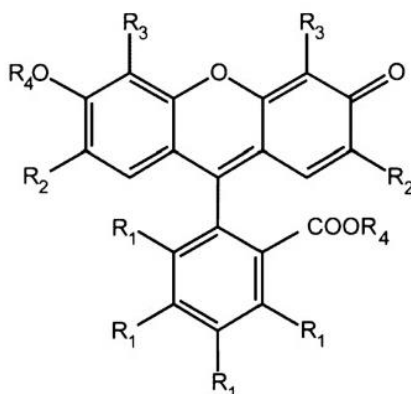


Figure 26. Basic skeleton of chemical structure of fluorone dyes. Figure reprinted from Ref.¹⁰⁹ (Appendix K).

In blind docking approach, the grid size plays a key role in molecular docking studies. By default, the grid size is set to the nearest active site. Here, it was necessary to set the grid to complete enzyme and then highlight individual active sites identified on the whole structure and later to validate them. Molecular docking results show that there were at least two binding sites for eosin Y, erythrosine B, rose bengal, and fluorescein. The first binding site was located on the C45 loop close of ATP binding site and the other active site was located on the extracellular part of the enzyme. The results were experimentally verified using fluorescence techniques.

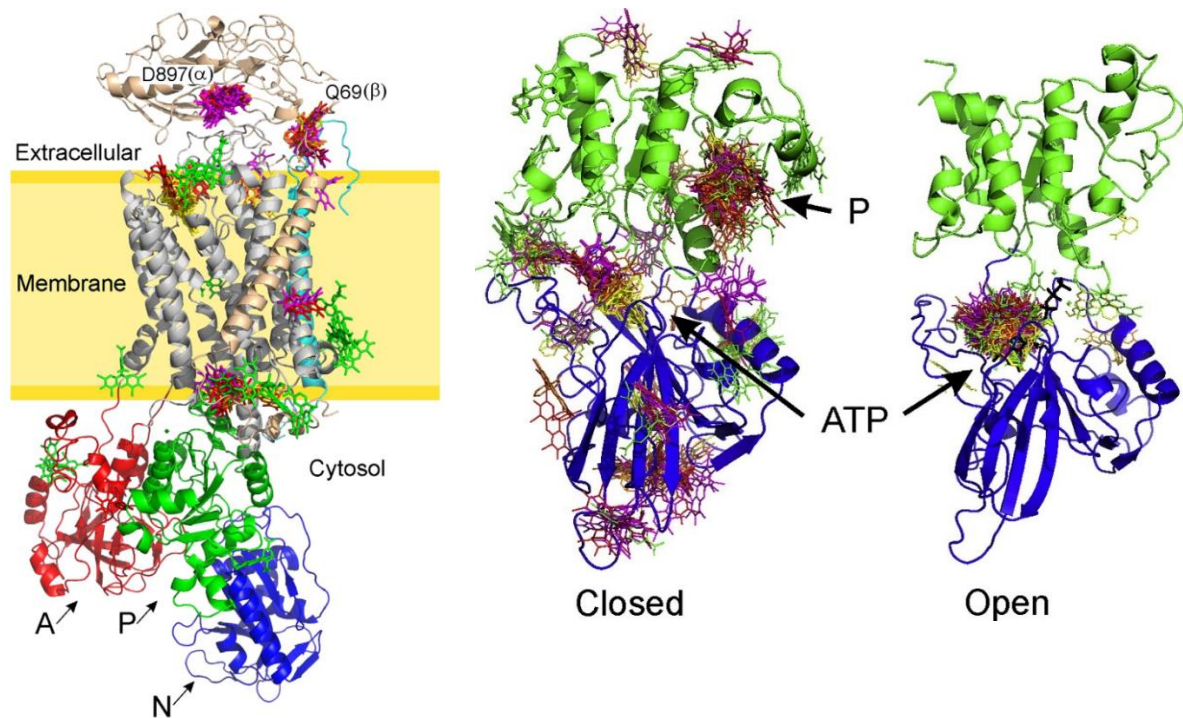


Figure 27. Molecular docking to NKA structure (left) and C45 loop in closed and open conformations (center and right). P domain is shown in green, and N domain in blue. Fluorescein in yellow, eosin Y in orange, erythrosine B in red, rose bengal in magenta, and 5(6)-carboxyeosin in green are colored ligands. Most strongly bound poses for all ligands are shown in thicker sticks. ATP is shown in black. Two possible binding poses have been detected in closed structure — near ATP binding site and on P domain, while only one major can be found in the case of open structure in the ATP-binding pocket. Figure reprinted from Ref.¹⁰⁹ (Appendix K).

It is not without interest, that fluorone dyes selectively bind to ATP binding place, which can explain competitive inhibition of NKA by fluorone dyes. Blind docking approach thus can help to generate hypothesis of binding that can be validated by experimental techniques such as fluorescence measurement, as was further used in following study of docking of RH421 dye, that was confirmed by fluorescence resonance energy transfer¹¹⁰.

Improvement of DOCK Scoring Function

During the implementation of various molecular docking projects, it has been observed, that there is a possibility to improving of molecular docking scoring function. Especially in projects with kinases, where unclear results were observed in the former docking studies where crystal like poses were not reproduced unless crystal-like docking approach was used. Such situation is typically caused by a narrow active site where molecular docking programs such as AutoDock 4⁶, Autodock Vina⁷ or DOCK²² have a problem with the repulsion part of Lennard-Jones potential¹¹¹ that is usually calculated with r^{12} or scaled r^{12} terms. These repulsion functions are however too steep in comparison to realistic exponential function. (Figure 28)

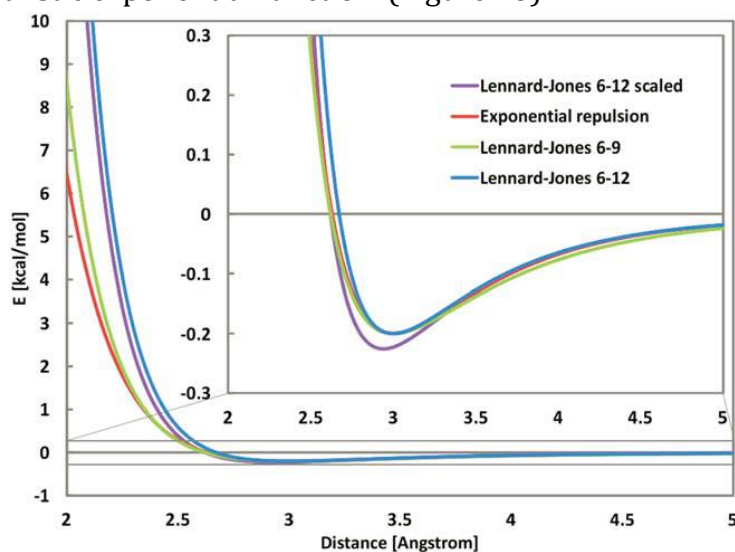


Figure 28. Different representations of van der Waals potentials. Figure reprinted from Ref. ¹¹² (Appendix L).

Based on incorrect calculation of repulsion, other final conformations are preferred over the more expected and corrected ones (Figure 29).



Figure 29. Scheme showing the difference of docking of ligand into wide and narrow binding site. Whereas in former case, the ligand can relax, in later case the cavity is too narrow, and atoms cannot relax. Figure reprinted from Ref. ¹¹² (Appendix L).

Based on this experience, we proposed a new scoring function with the implemented exponential repulsion part of calculation in non-bonded term. This implementation made the physical course of repulsion interaction more physical. It means, that the repulsive interaction was described by an exponential function.

$$E_{vdw} = \sqrt{\varepsilon_i \varepsilon_j} \left[e^{12 \left(1 - \frac{r_{ij}}{R_{ij}} \right)} - 2 \left(\frac{(R_i + R_j)^6}{r_{ij}^6 + ((R_i + R_j)/3)^6} \right) \right] \quad (7)$$

Generally, the exponential function is a very time-consuming process for computer calculations, so it was necessary to optimize and re-design algorithms used for computation, especially loops and data storages. Software package DOCK 6.6 was selected for this implementation and refactoring of the calculation process. The final modification of the program DOCK 6.6 has been tested by several datasets focused on tuning of scoring functions and molecular docking programs such as DUD, SB2012 and Wang dataset.

Especially, Directory of Useful Decoys (DUD) testing of enrichment of active compounds within first 20% of compounds (EF_{20%}) clearly shown that exponential repulsion scoring function significantly improved prediction of the binding modes of proteins bearing narrow active sites such as kinases, serine proteases, HIV proteases and others. On the other hand, this modification did not improve the better prediction of receptors which have larger binding cavities such as nuclear receptors or those that include metal active sites such as metalloenzymes. This is because the modification did not concern to metals and their interaction.

We can conclude that usage of the exponential repulsion represents modest improvement of the force field-based scoring functions, but further modifications especially in the electrostatic interactions would require further development.

Summary

This dissertation thesis is focused on molecular docking. This *in silico* method allows the use of computer technology to study interactions of small molecular ligands with receptors - biological polymers - proteins, nucleic acids and others.

The first part of this work is focused on to the motivation to study interaction of small molecules with biologically interesting targets. The other parts of the theoretical introduction deal with the background of the calculation starting with generation of conformations of the ligands and subsequently continuing with specification and execution of the scoring function. The last part describes the basic molecular docking workflow and the description of preparation of receptor and ligand collection.

The methodological part describes individual approaches of molecular docking. This is, on one hand, the use of the known position of the ligand in the crystal structure of the active site of the receptor. Then we extend this approach with several possible conformations of the ligand within the active site. None of the previous approaches take into account the influence of receptor flexibility. This can be overcome by ensemble docking approach which docks into a set of structures obtained from, for example, molecular dynamics. Lastly, the "blind" docking approach shows how molecular docking can help to identify suitable receptor binding sites if their position is not known.

In the result section, the outcomes of individual projects are discussed based on previously discussed approaches of molecular docking. These projects have been addressed for individual biological systems such as brassinosteroid receptor, kinases, cytokinin oxidase/dehydrogenase, histamine receptor and sodium/potassium pump. The final part of the result part is devoted to the design and subsequent implementation of a new scoring function within the DOCK program which considers the physically justified exponential repulsion - expDOCK. After its implementation, this new scoring feature has been tested on available datasets and achieves particularly good results when used on tight closed cavity systems.

We demonstrate on each example that molecular docking methods can successfully complement experimental approaches and can help to establish and solve the hypotheses of the effect of proposed substances. At the same time, however, their results must always be critically controlled and often initial docking has to be adjusted to achieve a desirable description of the real situation. Molecular docking cannot be used as a "black box" method. But its ability to visualize and evaluate predicted ligand binding to the receptor is an invaluable aid in rational design of new substances.

Shrnutí

Předkládaná disertační práce se věnuje problematice molekulárního dokování. Tato *in silico* metoda umožňuje pomocí výpočetní techniky studovat interakce malých molekul s receptory, což jsou molekulárně větší celky například biologické polymery – proteiny, nukleové kyseliny a další.

První část této práce je věnována motivaci studia interakce malých molekul s biologicky zajímavými cíli. Další části teoretického úvodu se zabývá pozadím výpočtu, tedy generováním prostorových konformací ligandů a následně specifikaci skórovacích funkcí a jejich rozdělení do jednotlivých tříd. Poslední část popisuje základního pracovního postupu molekulárního dokování a popisu přípravy kolekce receptorů a ligandů.

V metodické části jsou popsány jednotlivé přístupy molekulárního dokování, které se následně využívají k samotnému dokování. Jedná se jednak o využití známé pozice ligandu v krystalové struktuře aktivního místa receptoru. Následně tento přístup rozšiřujeme pomocí několika možných konformací ligandu v rámci aktivního místa. Ani jeden z předchozích přístupů nezohledňuje vliv flexibility receptoru. Tu zohledňuje dokování do souboru struktur, tedy dokovat ligand do jednotlivých snímků receptoru získaných například z molekulové dynamiky. Poslední přístup dokování „naslepo“ pak ukazuje, jak se pomocí molekulárního dokování dají hledat vhodná vazebná místa na receptoru, jestliže není známa jejich poloha.

Ve výsledkové části jsou následně na základě uvedených přístupů molekulárního dokování diskutovány výstupy z jednotlivých realizovaných projektů. Tyto projekty byly řešeny pro jednotlivé biologické systémy, jako například brassinosteroidní receptor, kinázy, cytokinin oxidázy/dehydrogenázy, histaminový receptor a sodno/draselná pumpa. Finální část výsledkové části byl věnován návrhu a následné implementaci nové skórovací funkce v rámci programu DOCK zohledňující fyzikální průběh repulzní interakce pomocí fyzikálně oprávněnější exponenciální repulze – tzv expDOCK. Po implementaci byla tato nová skórovací funkce testována na dostupných datasetech a dosahuje zvláště dobrých výsledků, pokud je použita na systémy s těsnými uzavřenými kavitami.

Na jednotlivých příkladech tak ukazujeme, že techniky molekulárního dokování mohou úspěšně doplňovat experimentální přístupy a mohou pomáhat ustanovovat a řešit hypotézy účinku jednotlivých navrhovaných látek. Současně se ale musejí jejich výsledky neustále kriticky kontrolovat a mnohdy se musí prvotní dokování upravovat, aby se dosáhlo kýženého popisu skutečné situace. Molekulární dokování tedy nelze používat jako metodu „černé skříňky“. Ale jeho možnost vizualizace a ohodnocení predikované vazby ligandu do receptoru je neocenitelným pomocníkem při racionálním návrhu nových látek.

List of Publications

1. Pozharitskaya O, Shikov A, Makarova M, Ivanova S, Kosman V, Makarov V, **Bazgier V**, Berka K, Otyepka M, Ulrichová J: Antiallergic Effects of Pigments Isolated from Green Sea Urchin (*Strongylocentrotus droebachiensis*) Shells. *Planta Med.*, 79(18), 1698-1704, 2013.
2. Havlíková M, Huličiak M, **Bazgier V**, Berka K, Kubala M: Fluorone dyes have binding sites on both cytoplasmic and extracellular domains of Na,K-ATPase. *Biochim. Biophys. Acta, Biomembr.*, 1828(2), 568-576, 2013.
3. Gucký T, Jorda R, Zatloukal M, **Bazgier V**, Berka K, Řezníčková E, Béres T, Strnad M, Kryštof V: A Novel Series of Highly Potent 2,6,9-Trisubstituted Purine Cyclin-Dependent Kinase Inhibitors. *J. Med. Chem.*, 56(15), 6234-6247, 2013.
4. Mojzych M, Šubertová V, Bielawska A, Bielawski K, **Bazgier V**, Berka K, Gucký T, Fornal E, Kryštof V: Synthesis and kinase inhibitory activity of new sulfonamide derivatives of pyrazolo[4,3-e][1,2,4]triazines. *Eur. J. Med. Chem.*, 78, 217-224, 2014.
5. Kvasnica M, Oklešťková J, **Bazgier V**, Rárová L, Berka K, Strnad M: Biological activities of new monohydroxylated brassinosteroid analogues with a carboxylic group in the side chain. *Steroids*, 85, 58-64, 2014.
6. Řezníčková E, Weitensteiner S, Havlíček L, Jorda R, Gucký T, Berka K, **Bazgier V**, Zahler S, Kryštof V, Strnad M: Characterization of a Pyrazolo[4,3- d]pyrimidine Inhibitor of Cyclin-Dependent Kinases 2 and 5 and Aurora A With Pro-Apoptotic and Anti-Angiogenic Activity In Vitro. *Chem. Biol. Drug. Des.*, 86(6), 1528–1540, 2015.
7. Řezníčková E, Popa A, Gucký T, Zatloukal M, Havlíček L, **Bazgier V**, Berka K, Jorda R, Popa I, Nasereddin A, Jaffe CL, Kryštof V, Strnad M: 2,6,9-Trisubstituted purines as CRK3 kinase inhibitors with antileishmanial activity in vitro. *Bioorg. Med. Chem. Lett.*, 25(11), 2298-2301, 2015.
8. Nisler J, Kopečný D, Končítíková R, Zatloukal M, **Bazgier V**, Berka K, Zalabák D, Briozzo P, Strnad M, Spíchal L: Novel thidiazuron-derived inhibitors of cytokinin oxidase/dehydrogenase. *Plant. Mol. Biol.*, 92(1), 235–248, 2016.
9. Lamie PF, Ali WAM, **Bazgier V**, Rárová L: Novel N-substituted indole Schiff bases as dual inhibitors of cyclooxygenase-2 and 5-lipoxygenase enzymes: Synthesis, biological activities in vitro and docking study. *Eur. J. Med. Chem.*, 123, 803–813, 2016.
10. Kvasnica M, Oklešťková J, **Bazgier V**, Rárová L, Kořínková P, Mikulík J, Buděšínský M, Béres T, Berka K, Lu Q, Russinova E, Strnad M: Design, synthesis and biological activities of new brassinosteroid analogues with phenyl group in the side chain. *Org. Biomol. Chem.*, 14(37), 8691-8701, 2016.
11. Baltus CB, Jorda R, Marot C, Berka K, **Bazgier V**, Kryštof V, Prié G, Viaud-Massuard M: Synthesis, Biological Evaluation and Molecular Modeling of a Novel

- Series of 7-Azaindole Based Tri-Heterocyclic Compounds as Potent CDK2/Cyclin E Inhibitors. *Eur. J. Med. Chem.*, 108(27), 701–719, 2016.
12. Carazo Fernández A, Hyrsova L, Dušek J, Chodounská H, Horvátová A, Berka K, **Bazgier V**, Gan-Schreier H, Chamulitrat W, Kudová E, Pávek P: Acetylated deoxycholic (DCA) and cholic (CA) acids are potent ligands of pregnane X (PXR) receptor. *Toxicol. Lett.*, 265, 86–96, 2017.
 13. **Bazgier V**, Berka K, Otyepka M, Banáš P: Exponential repulsion improves structural predictability of molecular docking. *J. Comput. Chem.*, 37(28), 2485–2494, 2016.
 14. Korinkova P, **Bazgier V**, Oklestkova J, Rarova L, Strnad M, Kvasnica M: Synthesis of novel aryl brassinosteroids through alkene crossmetathesis and preliminary biological study. *Steroids*, 127, 46-55, 2017.
 15. Huličiak M, **Bazgier V**, Berka K, Kubala M: RH421 binds into the ATP-binding site on the Na⁺/K⁺-ATPase. *Biochim. Biophys. Acta, Biomembranes*, 1859(10), 2113–2122, 2017.
 16. Šrejber M, Navrátilová V, Paloncýová M, **Bazgier V**, Berka K, Anzenbacher P, Otyepka M: Membrane-attached mammalian cytochromes P450: An overview of the membrane's effects on structure, drug binding, and interactions with redox partners. *J Inorg Biochem, ASAP*, 2018.
 17. Pravda L, Sehnal D, Toušek D, Navrátilová V, **Bazgier V**, Berka K, Svobodová-Vařeková R, Koča J, Otyepka M: MOLEonline: A Web-based Tool for Analyzing Channels, Tunnels and Pores (2018 update). *Nucleic Acids Res*, submitted, 2018.

References

- (1) Ferreira, L.; dos Santos, R.; Oliva, G.; Andricopulo, A. Molecular Docking and Structure-Based Drug Design Strategies. *Molecules* **2015**, *20* (12), 13384–13421.
- (2) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A Review of Protein-Small Molecule Docking Methods. *J. Comput. Aided. Mol. Des.* **2002**, *16* (3), 151–166.
- (3) Yuriev, E.; Holien, J.; Ramsland, P. A. Improvements, Trends, and New Ideas in Molecular Docking: 2012-2013 in Review. *J. Mol. Recognit.* **2015**, *28* (10), 581–604.
- (4) Abagyan, R.; Totrov, M. Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins. *J. Mol. Biol.* **1994**, *235* (3), 983–1002.
- (5) Goodsell, D. S.; Olson, A. J. Automated Docking of Substrates to Proteins by Simulated Annealing. *Proteins Struct. Funct. Genet.* **1990**, *8* (3), 195–202.
- (6) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785–2791.
- (7) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2009**, NA-NA.
- (8) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and Validation of a Genetic Algorithm for Flexible Docking 1 Edited by F. E. Cohen. *J. Mol. Biol.* **1997**, *267* (3), 727–748.
- (9) Venkatachalam, C. M.; Jiang, X.; Oldfield, T.; Waldman, M. LigandFit: A Novel Method for the Shape-Directed Rapid Docking of Ligands to Protein Active Sites. *J. Mol. Graph. Model.* **2003**, *21* (4), 289–307.
- (10) Yuriev, E.; Agostino, M.; Ramsland, P. A. Challenges and Advances in Computational Docking: 2009 in Review. *J. Mol. Recognit.* **2011**, *24* (2), 149–164.
- (11) Ewing, T. J. A.; Makino, S.; Skillman, A. G.; Kuntz, I. D. DOCK 4.0: Search Strategies for Automated Molecular Docking of Flexible Molecule Databases. *J. Comput. Aided. Mol. Des.* **2001**, *15* (5), 411–428.
- (12) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749.
- (13) Zsoldos, Z.; Reid, D.; Simon, A.; Sadjad, S. B.; Johnson, A. P. eHiTS: A New Fast, Exhaustive Flexible Ligand Docking System. *J. Mol. Graph. Model.* **2007**, *26* (1), 198–212.
- (14) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *J. Mol. Biol.* **1996**, *261* (3), 470–489.
- (15) Wang, R.; Lu, Y.; Wang, S. Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *J. Med. Chem.* **2003**, *46* (12), 2287–2303.

- (16) Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W. Computational Methods in Drug Discovery. *Pharmacol. Rev.* **2013**, *66* (1), 334–395.
- (17) Liu, J.; Wang, R. Classification of Current Scoring Functions. *J. Chem. Inf. Model.* **2015**, *55* (3), 475–482.
- (18) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules J . Am . Chem . Soc . 1995 , 117 , 5179–5197. *J. Am. Chem. Soc.* **1996**, *118* (9), 2309–2309.
- (19) Jorgensen, W.; Tirado-Rives, J. The OPLS Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem.* **1988**, *110* (6), 1657–1666.
- (20) Rappi, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard III, W. A.; Skigg, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. SOC* **1992**, *114*, 10024–10039.
- (21) Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17* (5–6), 490–519.
- (22) Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. DOCK 6: Combining Techniques to Model RNA-Small Molecule Complexes. *RNA* **2009**, *15* (6), 1219–1230.
- (23) Abagyan, R.; Totrov, M.; Kuznetsov, D. ICM?A New Method for Protein Modeling and Design: Applications to Docking and Structure Prediction from the Distorted Native Conformation. *J. Comput. Chem.* **1994**, *15* (5), 488–506.
- (24) Sochacka, J. Docking of Thiopurine Derivatives to Human Serum Albumin and Binding Site Analysis with Molegro Virtual Docker. *Acta Pol Pharm* **2014**, *71* (2), 343–349.
- (25) Wang, R.; Lu, Y.; Wang, S. Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *J. Med. Chem.* **2003**, *46* (12), 2287–2303.
- (26) Yin, S.; Biedermannova, L.; Vondrasek, J.; Dokholyan, N. V. MedusaScore: An Accurate Force Field-Based Scoring Function for Virtual Drug Screening. *J. Chem. Inf. Model.* **2008**, *48* (8), 1656–1662.
- (27) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V; Mee, R. P. Empirical Scoring Functions .1. The Development of a Fast Empirical Scoring Function to Estimate the Binding Affinity of Ligands in Receptor Complexes. *J. Comput. Aided. Mol. Des.* **1997**, *11* (5), 425–445.
- (28) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved Protein-Ligand Docking Using GOLD. *Proteins Struct. Funct. Bioinforma.* **2003**, *52* (4), 609–623.
- (29) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V; Mee, R. P. Empirical Scoring Functions .1. The Development of a Fast Empirical Scoring Function to Estimate the Binding Affinity of Ligands in Receptor Complexes. *J. Comput. Aided. Mol. Des.* **1997**, *11* (5), 425–445.
- (30) Wang, R.; Liu, L.; Lai, L.; Tang, Y. SCORE: A New Empirical Method for Estimating the Binding Affinity of a Protein-Ligand Complex. *J. Mol. Model.* **1998**, *4* (12), 379–394.

- (31) Böhm, H. J. The Development of a Simple Empirical Scoring Function to Estimate the Binding Constant for a Protein-Ligand Complex of Known Three-Dimensional Structure. *J. Comput. Aided. Mol. Des.* **1994**, *8* (3), 243–256.
- (32) Krammer, A.; Kirchhoff, P. D.; Jiang, X.; Venkatachalam, C. M.; Waldman, M. LigScore: A Novel Scoring Function for Predicting Binding Affinities. *J. Mol. Graph. Model.* **2005**, *23* (5), 395–407.
- (33) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-Based Scoring Function to Predict Protein-Ligand Interactions. *J. Mol. Biol.* **2000**, *295* (2), 337–356.
- (34) Velec, H. F. G.; Gohlke, H.; Klebe, G. DrugScore CSD Knowledge-Based Scoring Function Derived from Small Molecule Crystal Data with Superior Recognition Rate of Near-Native Ligand Poses and Better Affinity Prediction. *J. Med. Chem.* **2005**, *48* (20), 6296–6303.
- (35) DeWitte, R. S.; Shakhnovich, E. I. SMOG: De Novo Design Method Based on Simple, Fast, and Accurate Free Energy Estimates. 1. Methodology and Supporting Evidence. *J. Am. Chem. Soc.* **1996**, *118* (47), 11733–11744.
- (36) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-Based Scoring Function to Predict Protein-Ligand Interactions. *J. Mol. Biol.* **2000**, *295* (2), 337–356.
- (37) Xie, Z.-R.; Hwang, M.-J. An Interaction-Motif-Based Scoring Function for Protein-Ligand Docking. *BMC Bioinformatics* **2010**, *11* (1), 298.
- (38) Fan, H.; Schneidman-Duhovny, D.; Irwin, J. J.; Dong, G.; Shoichet, B. K.; Sali, A. Statistical Potential for Modeling and Ranking of Protein-Ligand Interactions. *J. Chem. Inf. Model.* **2011**, *51* (12), 3078–3092.
- (39) Ballester, P. J.; Mitchell, J. B. O. A Machine Learning Approach to Predicting Protein-ligand Binding Affinity with Applications to Molecular Docking. *Bioinformatics* **2010**, *26* (9), 1169–1175.
- (40) Lepšík, M.; Řezáč, J.; Kolář, M.; Pecina, A.; Hobza, P.; Fanfrlík, J. The Semiempirical Quantum Mechanical Scoring Function for In Silico Drug Design. *Chempluschem* **2013**, *78* (9), 921–931.
- (41) S. Brahmkshatriya, P.; Dobeš, P.; Fanfrlík, J.; Řezáč, J.; Paruch, K.; Bronowska, A.; Lepšík, M.; Hobza, P. Quantum Mechanical Scoring: Structural and Energetic Insights into Cyclin-Dependent Kinase 2 Inhibition by Pyrazolo[1,5-A]pyrimidines. *Curr. Comput. Aided-Drug Des.* **2013**, *9* (1), 118–129.
- (42) Berman, H.; Henrick, K.; Nakamura, H. Announcing the Worldwide Protein Data Bank. *Nat. Struct. Mol. Biol.* **2003**, *10* (12), 980–980.
- (43) Laurie, A. T. R.; Jackson, R. M. Q-SiteFinder: An Energy-Based Method for the Prediction of Protein-Ligand Binding Sites. *Bioinformatics* **2005**, *21* (9), 1908–1916.
- (44) Wass, M. N.; Kelley, L. A.; Sternberg, M. J. E. 3DLigandSite: Predicting Ligand-Binding Sites Using Similar Structures. *Nucleic Acids Res.* **2010**, *38* (suppl_2), W469–W473.
- (45) Weisel, M.; Proschak, E.; Schneider, G. PocketPicker: Analysis of Ligand Binding-Sites with Shape Descriptors. *Chem. Cent. J.* **2007**, *1* (1), 7.
- (46) Bentham Science Publisher, B. S. P. Methods for the Prediction of Protein-Ligand Binding Sites for Structure-Based Drug Design and Virtual Ligand Screening. *Curr.*

- Protein Pept. Sci.* **2006**, 7 (5), 395–406.
- (47) The PyMOL Molecular Graphics System. Schrödinger, LLC. 2017.
- (48) Webb, B.; Sali, A. Protein Structure Modeling with MODELLER; 2014; pp 1–15.
- (49) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL Workspace: A Web-Based Environment for Protein Structure Homology Modelling. *Bioinformatics* **2006**, 22 (2), 195–201.
- (50) Schwede, T. SWISS-MODEL: An Automated Protein Homology-Modeling Server. *Nucleic Acids Res.* **2003**, 31 (13), 3381–3385.
- (51) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, 97 (40), 10269–10280.
- (52) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera?A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, 25 (13), 1605–1612.
- (53) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. *J. Comput. Chem.* **2000**, 21 (2), 132–146.
- (54) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Electronegativity—a Rapid Access to Atomic Charges. *Tetrahedron* **1980**, 36 (22), 3219–3228.
- (55) Schrödinger Release 2017-4: Maestro. Maestro, Schrödinger, LLC, New York, NY 2017.
- (56) ACD/Structure Elucidator. Advanced Chemistry Development, Inc., Toronto, ON, Canada 2015.
- (57) Marvin. ChemAxon 2015.
- (58) O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminform.* **2011**, 3 (1), 33.
- (59) Schrödinger Release 2017-4: LigPrep. Schrödinger, LLC, New York, NY 2017.
- (60) Landrum, G. RDKit: Open-Source Cheminformatics; 2017.
- (61) Gaussian Software Package. Gaussian, Inc.: Wallingford 2017.
- (62) Lipinski, C. A. Lead- and Drug-like Compounds: The Rule-of-Five Revolution. *Drug Discov. Today Technol.* **2004**, 1 (4), 337–341.
- (63) Sterling, T.; Irwin, J. J. ZINC 15 – Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, 55 (11), 2324–2337.
- (64) Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; et al. PubChem Substance and Compound Databases. *Nucleic Acids Res.* **2016**, 44 (D1), D1202–D1213.
- (65) Bento, A. P.; Gaulton, A.; Hersey, A.; Bellis, L. J.; Chambers, J.; Davies, M.; Krüger, F. A.; Light, Y.; Mak, L.; McGlinchey, S.; et al. The ChEMBL Bioactivity Database: An Update. *Nucleic Acids Res.* **2014**, 42 (D1), D1083–D1090.
- (66) Pence, H. E.; Williams, A. ChemSpider: An Online Chemical Information Resource. *J. Chem. Educ.* **2010**, 87 (11), 1123–1124.

- (67) Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; et al. DrugBank 5.0: A Major Update to the DrugBank Database for 2018. *Nucleic Acids Res.* **2018**, *46* (D1), D1074–D1082.
- (68) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: A Web-Accessible Database of Experimentally Determined Protein-Ligand Binding Affinities. *Nucleic Acids Res.* **2007**, *35* (Database), D198–D201.
- (69) eMolecules <https://www.emolecules.com/>.
- (70) Seiler, K. P.; George, G. A.; Happ, M. P.; Bodycombe, N. E.; Carrinski, H. A.; Norton, S.; Brudz, S.; Sullivan, J. P.; Muhlich, J.; Serrano, M.; et al. ChemBank: A Small-Molecule Screening and Cheminformatics Resource Database. *Nucleic Acids Res.* **2007**, *36* (Database), D351–D359.
- (71) Mohan, V.; Gibbs, A.; Cummings, M.; Jaeger, E.; DesJarlais, R. Docking: Successes and Challenges. *Curr. Pharm. Des.* **2005**, *11* (3), 323–333.
- (72) Yang, S.-Y. Pharmacophore Modeling and Applications in Drug Discovery: Challenges and Recent Advances. *Drug Discov. Today* **2010**, *15* (11–12), 444–450.
- (73) Leo, A.; Hoekman, D. H. Exploring QSAR: Fundamentals and Applications in Chemistry and Biology. *An Am. Chem. Soc. Publ.* **1995**, *1*.
- (74) Tye, H. Application of Statistical “design of Experiments” Methods in Drug Discovery. *Drug Discov. Today* **2004**, *9* (11), 485–491.
- (75) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006**, *49* (23), 6789–6801.
- (76) Mukherjee, S.; Balius, T. E.; Rizzo, R. C. Docking Validation Resources: Protein Family and Ligand Flexibility Experiments. *J. Chem. Inf. Model.* **2010**, *50* (11), 1986–2000.
- (77) Bajguz, A. Brassinosteroids – Occurrence and Chemical Structures in Plants. In *Brassinosteroids: A Class of Plant Hormone*; Springer Netherlands: Dordrecht, 2011; pp 1–27.
- (78) She, J.; Han, Z.; Kim, T.-W.; Wang, J.; Cheng, W.; Chang, J.; Shi, S.; Wang, J.; Yang, M.; Wang, Z.-Y.; et al. Structural Insight into Brassinosteroid Perception by BRI1. *Nature* **2011**, *474* (7352), 472–476.
- (79) Kvasnica, M.; Oklešťková, J.; Bazgier, V.; Rárová, L.; Berka, K.; Strnad, M. Biological Activities of New Monohydroxylated Brassinosteroid Analogues with a Carboxylic Group in the Side Chain. *Steroids* **2014**, *85*, 58–64.
- (80) Kvasnica, M.; Oklešťkova, J.; Bazgier, V.; Rárová, L.; Korinkova, P.; Mikulík, J.; Budesinsky, M.; Béres, T.; Berka, K.; Lu, Q.; et al. Design, Synthesis and Biological Activities of New Brassinosteroid Analogues with a Phenyl Group in the Side Chain. *Org. Biomol. Chem.* **2016**, *14* (37), 8691–8701.
- (81) Korinkova, P.; Bazgier, V.; Oklešťkova, J.; Rarova, L.; Strnad, M.; Kvasnica, M. Synthesis of Novel Aryl Brassinosteroids through Alkene Cross-Metathesis and Preliminary Biological Study. *Steroids* **2017**, *127*, 46–55.
- (82) Ferguson, F. M.; Gray, N. S. Kinase Inhibitors: The Road Ahead. *Nat. Rev. Drug Discov.* **2018**.

- (83) Zhang, J.; Yang, P. L.; Gray, N. S. Targeting Cancer with Small Molecule Kinase Inhibitors. *Nat. Rev. Cancer* **2009**, *9* (1), 28–39.
- (84) Capdeville, R.; Buchdunger, E.; Zimmermann, J.; Matter, A. Glivec (STI571, Imatinib), a Rationally Developed, Targeted Anticancer Drug. *Nat. Rev. Drug Discov.* **2002**, *1* (7), 493–502.
- (85) Debreczeni, J.E., Salah, E., Papagrigoriou, E., Burgess, N., von Delft, F., Gileadi, O., Sundstrom, M., Edwards, A., Arrowsmith, C., Weigelt, J., Knapp, S. Crystal Structure of Human Mitogen-Activated Protein Kinase Kinase Kinase 3 Isoform 2 Fox Domain at 1.25 Å Resolution. 2005.
- (86) Cowan-Jacob, S. W.; Fendrich, G.; Floersheimer, A.; Furet, P.; Liebetanz, J.; Rummel, G.; Rheinberger, P.; Centeleghe, M.; Fabbro, D.; Manley, P. W. Structural Biology Contributions to the Discovery of Drugs to Treat Chronic Myelogenous Leukaemia. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2007**, *63* (1), 80–93.
- (87) Mojzych, M.; Šubertová, V.; Bielawska, A.; Bielawski, K.; Bazgier, V.; Berka, K.; Gucký, T.; Fornal, E.; Kryštof, V. Synthesis and Kinase Inhibitory Activity of New Sulfonamide Derivatives of pyrazolo[4,3-e][1,2,4]triazines. *Eur. J. Med. Chem.* **2014**, *78*, 217–224.
- (88) Azevedo, W. F.; Leclerc, S.; Meijer, L.; Havlicek, L.; Strnad, M.; Kim, S.-H. Inhibition of Cyclin-Dependent Kinases by Purine Analogues. Crystal Structure of Human cdk2 Complexed with Roscovitine. *Eur. J. Biochem.* **1997**, *243* (1–2), 518–526.
- (89) Gucký, T.; Jorda, R.; Zatloukal, M.; Bazgier, V.; Berka, K.; Řezníčková, E.; Béres, T.; Strnad, M.; Kryštof, V. A Novel Series of Highly Potent 2,6,9-Trisubstituted Purine Cyclin-Dependent Kinase Inhibitors. *J. Med. Chem.* **2013**, *56* (15), 6234–6247.
- (90) den Boer, M.; Argaw, D.; Jannin, J.; Alvar, J. Leishmaniasis Impact and Treatment Access. *Clin. Microbiol. Infect.* **2011**, *17* (10), 1471–1477.
- (91) Řezníčková, E.; Popa, A.; Gucký, T.; Zatloukal, M.; Havlíček, L.; Bazgier, V.; Berka, K.; Jorda, R.; Popa, I.; Nasereddin, A.; et al. 2,6,9-Trisubstituted Purines as CRK3 Kinase Inhibitors with Antileishmanial Activity in Vitro. *Bioorg. Med. Chem. Lett.* **2015**, *25* (11), 2298–2301.
- (92) Martin, M. P.; Olesen, S. H.; Georg, G. I.; Schönbrunn, E. Cyclin-Dependent Kinase Inhibitor Dinaciclib Interacts with the Acetyl-Lysine Recognition Site of Bromodomains. *ACS Chem. Biol.* **2013**, *8* (11), 2360–2365.
- (93) Bettayeb, K.; Oumata, N.; Echalié, A.; Ferandin, Y.; Endicott, J. A.; Galons, H.; Meijer, L. CR8, a Potent and Selective, Roscovitine-Derived Inhibitor of Cyclin-Dependent Kinases. *Oncogene* **2008**, *27* (44), 5797–5807.
- (94) Honda, R.; Lowe, E. D.; Dubinina, E.; Skamnaki, V.; Cook, A.; Brown, N. R.; Johnson, L. N. The Structure of Cyclin E1/CDK2: Implications for CDK2 Activation and CDK2-Independent Roles. *EMBO J.* **2005**, *24* (3), 452–463.
- (95) Cee, V. J.; Schenkel, L. B.; Hodous, B. L.; Deak, H. L.; Nguyen, H. N.; Olivieri, P. R.; Romero, K.; Bak, A.; Be, X.; Bellon, S.; et al. Discovery of a Potent, Selective, and Orally Bioavailable Pyridinyl-Pyrimidine Phthalazine Aurora Kinase Inhibitor. *J. Med. Chem.* **2010**, *53* (17), 6368–6377.
- (96) Malumbres, M.; Barbacid, M. Cell Cycle, CDKs and Cancer: A Changing Paradigm. *Nat. Rev. Cancer* **2009**, *9* (3), 153–166.

- (97) Dar, A. A.; Goff, L. W.; Majid, S.; Berlin, J.; El-Rifai, W. Aurora Kinase Inhibitors - Rising Stars in Cancer Therapeutics? *Mol. Cancer Ther.* **2010**, *9* (2), 268–278.
- (98) Řezníčková, E.; Weitensteiner, S.; Havlíček, L.; Jorda, R.; Gucký, T.; Berka, K.; Bazgier, V.; Zahler, S.; Kryštof, V.; Strnad, M. Characterization of a Pyrazolo[4,3- D]Pyrimidine Inhibitor of Cyclin-Dependent Kinases 2 and 5 and Aurora A With Pro-Apoptotic and Anti-Angiogenic Activity In Vitro. *Chem. Biol. Drug Des.* **2015**, *86*(6), 1528–1540.
- (99) Senderowicz, A. M. Small Molecule Modulators of Cyclin-Dependent Kinases for Cancer Therapy. *Oncogene* **2000**, *19* (56), 6600–6606.
- (100) Azevedo, W. F.; Leclerc, S.; Meijer, L.; Havlicek, L.; Strnad, M.; Kim, S.-H. Inhibition of Cyclin-Dependent Kinases by Purine Analogues. Crystal Structure of Human cdk2 Complexed with Roscovitine. *Eur. J. Biochem.* **1997**, *243* (1–2), 518–526.
- (101) Baltus, C. B.; Jorda, R.; Marot, C.; Berka, K.; Bazgier, V.; Kryštof, V.; Prié, G.; Viaud-Massuard, M.-C. Synthesis, Biological Evaluation and Molecular Modeling of a Novel Series of 7-Azaindole Based Tri-Heterocyclic Compounds as Potent CDK2/Cyclin E Inhibitors. *Eur. J. Med. Chem.* **2016**, *108*, 701–719.
- (102) Mok, D. W.; Mok, M. C. CYTOKININ METABOLISM AND ACTION. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2001**, *52* (1), 89–118.
- (103) Kopečný, D.; Briozzo, P.; Popelková, H.; Šebela, M.; Končítíková, R.; Spíchal, L.; Nisler, J.; Madzak, C.; Frébort, I.; Laloue, M.; et al. Phenyl- and Benzylurea Cytokinins as Competitive Inhibitors of Cytokinin Oxidase/dehydrogenase: A Structural Study. *Biochimie* **2010**, *92* (8), 1052–1062.
- (104) Kopečný, D.; Končítíková, R.; Popelka, H.; Briozzo, P.; Vigouroux, A.; Kopečná, M.; Zalabák, D.; Šebela, M.; Skopalová, J.; Frébort, I.; et al. Kinetic and Structural Investigation of the Cytokinin Oxidase/dehydrogenase Active Site. *FEBS J.* **2016**, *283* (2), 361–377.
- (105) Nisler, J.; Kopečný, D.; Končítíková, R.; Zatloukal, M.; Bazgier, V.; Berka, K.; Zalabák, D.; Briozzo, P.; Strnad, M.; Spíchal, L. Novel Thidiazuron-Derived Inhibitors of Cytokinin Oxidase/dehydrogenase. *Plant Mol. Biol.* **2016**, *92* (1–2), 235–248.
- (106) Shimamura, T.; Shiroishi, M.; Weyand, S.; Tsujimoto, H.; Winter, G.; Katritch, V.; Abagyan, R.; Cherezov, V.; Liu, W.; Han, G. W.; et al. Structure of the Human Histamine H1 Receptor Complex with Doxepin. *Nature* **2011**, *475* (7354), 65–70.
- (107) Pozharitskaya, O.; Shikov, A.; Makarova, M.; Ivanova, S.; Kosman, V.; Makarov, V.; Bazgier, V.; Berka, K.; Otyepka, M.; Ulrichová, J. Antiallergic Effects of Pigments Isolated from Green Sea Urchin (*Strongylocentrotus Droebachiensis*) Shells. *Planta Med.* **2013**, *79* (18), 1698–1704.
- (108) Shinoda, T.; Ogawa, H.; Cornelius, F.; Toyoshima, C. Crystal Structure of the Sodium-potassium Pump at 2.4 Å Resolution. *Nature* **2009**, *459* (7245), 446–450.
- (109) Havlíková, M.; Huličiak, M.; Bazgier, V.; Berka, K.; Kubala, M. Fluorone Dyes Have Binding Sites on Both Cytoplasmic and Extracellular Domains of Na,K-ATPase. *Biochim. Biophys. Acta - Biomembr.* **2013**, *1828* (2), 568–576.
- (110) Huličiak, M.; Bazgier, V.; Berka, K.; Kubala, M. RH421 Binds into the ATP-Binding Site on the Na + /K + -ATPase. *Biochim. Biophys. Acta - Biomembr.* **2017**, *1859* (10), 2113–2122.

- (111) Jones, J. E. On the Determination of Molecular Fields. II. From the Equation of State of a Gas. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1924**, *106* (738), 463–477.
- (112) Bazgier, V.; Berka, K.; Otyepka, M.; Banáš, P. Exponential Repulsion Improves Structural Predictability of Molecular Docking. *J. Comput. Chem.* **2016**, *37* (28), 2485–2494.

List of Symbols and Shortcuts

ABL	Abl Protein Tyrosine Kinase
ALK5	Anaplastic lymphoma kinase
BRI	Brassinosteroid-insensitive1
CDK2	Cycline-Dependent Kinase 2
CRK3	Cdc2-related protein kinase 3
DUD	A Directory of Useful Decoys
HIV	Human immunodeficiency virus
InChI	IUPAC International Chemical Identifier
MOL	Molecule file format
PDB	Protein Data Bank
PDBQT	Protein Data Bank (Q) Charge (T) Atom type
RMSD	Root-mean-square deviation
SDF	Structure Data Format
SMILES	Simplified molecular input line entry specification
TDZ	Thidiazuron

List of Appendices

- A. Kvasnica M, Oklešťková J, **Bazgier V**, Rárová L, Berka K, Strnad M. *Steroids*, 85, 58-64, 2014.
- B. Kvasnica M, Oklešťková J, **Bazgier V**, Rárová L, Kořínková P, Mikulík J, Buděšínský M, Béres T, Berka K, Lu Q, Russinova E, Strnad M. *Org. Biomol. Chem.*, 14(37), 8691-8701, 2016.
- C. Korinkova P†, **Bazgier V**†, Oklestkova J, Rarova L, Strnad M, Kvasnica M. *Steroids*, 127, 46-55, 2017.
- D. Mojzych M, Šubertová V, Bielawska A, Bielawski K, **Bazgier V**, Berka K, Gucký T, Fornal E, Kryštof V. *Eur. J. Med. Chem.*, 78, 217-224, 2014.
- E. Řezníčková E, Popa A, Gucký T, Zatloukal M, Havlíček L, **Bazgier V**, Berka K, Jorda R, Popa I, Nasereddin A, Jaffe CL, Kryštof V, Strnad M. *Bioorg. Med. Chem. Lett.*, 25(11), 2298-2301, 2015.
- F. Řezníčková E, Weitensteiner S, Havlíček L, Jorda R, Gucký T, Berka K, **Bazgier V**, Zahler S, Kryštof V, Strnad M. *Chem. Biol. Drug. Des.*, 86(6), 1528-1540, 2015.
- G. Baltus CB, Jorda R, Marot C, Berka K, **Bazgier V**, Kryštof V, Prié G, Viaud-Massuard M. *Eur. J. Med. Chem.*, 108(27), 701-719, 2016.
- H. Nisler J, Kopečný D, Končítíková R, Zatloukal M, **Bazgier V**, Berka K, Zalabák D, Briozzo P, Strnad M, Spíchal L. *Plant. Mol. Biol.*, 92(1), 235-248, 2016.
- I. Gucký T, Jorda R, Zatloukal M, **Bazgier V**, Berka K, Řezníčková E, Béres T, Strnad M, Kryštof V. *J. Med. Chem.*, 56(15), 6234-6247, 2013.
- J. Pozharitskaya O, Shikov A, Makarova M, Ivanova S, Kosman V, Makarov V, **Bazgier V**, Berka K, Otyepka M, Ulrichová J. *Planta Med.*, 79(18), 1698-1704, 2013.
- K. Havlíková M, Huličiak M, **Bazgier V**, Berka K, Kubala M. *Biochim. Biophys. Acta, Biomembr.*, 1828(2), 568-576, 2013.
- L. **Bazgier V**, Berka K, Otyepka M, Banáš P. *J. Comput. Chem.*, 37(28), 2485-2494, 2016.

† - shared first authorship

Appendices

Appendix A

Kvasnica M, Oklešťková J, **Bazgier V**, Rárová L, Berka K, Strnad M:

Biological activities of new monohydroxylated brassinosteroid analogues with a
carboxylic group in the side chain.

Steroids, 85, 58-64, 2014

DOI: 10.1016/j.steroids.2014.04.007

IF = 2.282



Biological activities of new monohydroxylated brassinosteroid analogues with a carboxylic group in the side chain



Miroslav Kvasnica^a, Jana Oklestkova^{a,*}, Vaclav Bazgier^{a,b}, Lucie Rarova^a, Karel Berka^c, Miroslav Strnad^a

^aLaboratory of Growth Regulators, Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany ASCR & Palacký University, Šlechtitelů 11, 78371 Olomouc, Czech Republic

^bDepartment of Physical Chemistry, Faculty of Science, Palacký University, tř. 17. Listopadu 12, 77146 Olomouc, Czech Republic

^cRegional Centre of Advanced Technologies and Material, Department of Physical Chemistry, Faculty of Science, Palacký University, 17. Listopadu 12, 77146 Olomouc, Czech Republic

ARTICLE INFO

Article history:

Received 16 December 2013

Received in revised form 3 April 2014

Accepted 10 April 2014

Available online 21 April 2014

Keywords:

Anticancer activity

Brassinosteroids

Organic synthesis

Molecular docking

Receptor kinase BRI1

Plant bioassays

ABSTRACT

Thirteen monohydroxylated brassinosteroids analogues were synthesized and tested for their biological activity in plant and animal systems. The cytotoxic activity of the products was studied using human normal and cancer cell lines with 28-homocastasterone as positive control, their brassinolide type activity was established using the bean second-internode test with 24-epibrassinolide as standard.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Brassinosteroids (BRs) represent a large group of plant steroids which include more than 70 structurally and functionally related compounds [1]. BRs have been found in a wide range of plant species, including higher and lower plants, and have been detected in various plant parts such as pollen, seeds, leaves, stems, roots, flowers and insect galls. They demonstrate various kinds of regulatory action on the growth and development of plants, such as the stimulation of cell enlargement and cell division, improvement of the biomass formation, yield and quality of seeds, and plant adaptability [2]. At the molecular level, BRs change the gene expression and the metabolism of nucleic acids and proteins. BRs have structures similar to those of animal steroid hormones. Plants perceive steroids at cell membrane, using the membrane-integral receptor kinase brassinosteroid insensitive 1 (BRI1) [3–5]. The encoded protein, BRI1, belongs to a large family of plant LRR (leucine-rich repeat) receptor-like kinases, characterized by an extracellular LRR domain, a single-pass transmembrane segment and a cytoplasmic kinase domain. BRI1 has been established as an authentic brassinosteroid receptor by genetic and biochemical investigations [6]. Crystal structures of BRI1 in both free (PDB ID: 3RIZ, 3RGX,) and

brassinolide-bound (PDB ID: 3RJ0, 3BRZ,) forms are available, following independent X-ray diffraction structural determinations by two groups [6,7]. The structure of the ligand-binding domain resembles a superhelix of 25 twisted LRRs. A 70-amino acid island domain between LRRs 21 and 22 folds back into the interior of the superhelix, creating a surface pocket where the brassinosteroids bind. These recently published structures of *Arabidopsis thaliana* BRI1 enable the rational design of brassinosteroid-like antagonists and agonists. Recent studies have indicated that BRs have antiviral, antiproliferative and antibacterial activity [8]. BRs analogues have been reported to have antiviral activity against herpes simplex virus type 1, arenaviruses as well as against replication of vesicular stomatitis virus in Vero cells [9–11]. Natural types of BRs and their analogues affected the viability, proliferation, apoptosis and expression of some cell cycle related proteins in cancer and normal human cell lines [12–14]. It was shown that natural BRs, 24-epibrassinolide and 28-homocastasterone, inhibited *in vitro* angiogenesis of human endothelial cells [15].

The aim of our study relates to the synthesis of new brassinosteroid monohydroxylated derivatives and to study of their biological and anticancer properties. The plant growth promoting activity of synthetic analogs was assayed using the bean second internode bioassay. The antiproliferative activity of the new BRs analogs was evaluated *in vitro* using cancer cell lines of different histopathological origins and normal human fibroblasts. All derivative

* Corresponding author. Tel.: +420 58563483; fax: +420 585634870.

E-mail address: jana.oklestkova@upol.cz (J. Oklestkova).

structures were further subjected to docking studies using AutoDock Vina [16] in order to find structural patterns responsible for the results.

2. Experimental

2.1. General methods

The melting points were determined on a Hund H 600 apparatus (Helmut Hund, Germany). The elemental analyses (C, H, N) were carried out on a Perkin-Elmer 2400 II elemental analyzer. Optical rotations were measured on an Autopol IV polarimeter (Rudolf Research Analytical, Flanders, USA) at 25 °C in chloroform and $[\alpha]_D$ values are given in 10^{-1} deg cm^2 g. The infrared spectra were recorded on a Bruker IFS 55 spectrometer in chloroform. The wave numbers are given in cm^{-1} . The ^1H NMR spectra were taken in CDCl_3 on a Bruker AVANCE-400 (at 400 MHz) instrument with tetramethylsilane as an internal reference. Chemical shifts are given in ppm (δ -scale), coupling constants (J) in Hz. All of the values were obtained by first-order analysis. The mass spectra (ESI) were obtained with a LTQ Orbitrap XL (Thermo Fisher Scientific). For column chromatography, neutral silica gel (60 μm , Fluka) was used. The HPLC system consisted of a Gilson semi-preparative HPLC system including a quaternary pump, liquid handler, UV-VIS, RI, and ELSD detectors. The semi preparative column was based on silica gel (Labio Biosphere PSI 200).

Reagents and solvents were purchased from Sigma-Aldrich and were not purified.

2.2. Hydroboration of **1**

Solution of borane in THF (1 M, 20 mL) was added dropwise to the solution of methyl ester **1** (1 g, 2.49 mmol) in dry THF (30 mL). The reaction mixture was stirred at room temperature for 6 h. Then, the reaction mixture was quenched by water (10 mL) and cooled to 10 °C. Aqueous solution of NaOH (10%, 20 mL) was added followed by dropwise addition of 30% hydrogen peroxide (30 mL). This reaction mixture was stirred overnight at room temperature. It was then poured into a mixture of sodium sulfite (2 g) in water (50 mL) and acetic acid (1 mL) and stirred for additional 30 min. After then, the reaction mixture was diluted with ethyl acetate, extracted twice with water, the extract was dried over anhydrous sodium sulfate, and solvents evaporated under reduced pressure. A crude mixture of products was dissolved in a mixture of THF and acetone (5:1, 10 mL) and 5% aqueous HCl was added. The reaction mixture was heated under reflux for 3 h. Then it was diluted with ethyl acetate, extracted twice with water, the extract was dried over anhydrous sodium sulfate, and the solvents were evaporated under reduced pressure. Two main products were separated on HPLC in ethyl acetate–hexane (1:1). Two hydroxy compounds **2a** and **2b** were obtained.

2.2.1. Methyl (20S)-3 α -hydroxy-6-oxo-5 α -pregnane-20-carboxylate (**2a**)

Yield: 380 mg (40%), m.p. 198–200 °C (MeOH), $[\alpha]_D = -6^\circ$ (c 0.27). IR ν (cm^{-1}) 3615, 1727, 1706, 1165. ^1H NMR (CDCl_3) δ 0.68, 0.73 (both s, 3H, CH_3), 1.19 (d, 3H, $J = 6.9$ Hz, CH_3), 1.97 (m, 1H), 2.02 (dd, 1H, $J = 12.8$, $J' = 1.0$ Hz, H-7 β), 2.29 (dd, 1H, $J = 12.8$, $J' = 4.5$ Hz, H-7 α), 2.43 (dq, 1H, $J = 10.3$, $J' = 6.8$ Hz, H-20), 2.73 (dd, 1H, $J = 8.9$, $J' = 6.9$ Hz, H-5 α), 3.65 (s, 3H, OCH_3), 4.16 (t, 1H, $J = 2.1$ Hz, H-3 β). ^{13}C NMR δ 12.17 (CH_3), 12.30 (CH_3), 17.06 (CH_3), 21.01 (CH_2), 23.96 (CH_2), 26.95 (CH_2), 27.69 (CH_2), 28.16 (CH_2), 31.67 (CH_2), 37.90 (CH), 39.27 (CH_2), 41.49 (C), 42.35 (CH), 43.00 (C), 46.75 (CH_2), 51.34 (CH), 51.66 (CH_3), 52.79 (CH), 53.78 (CH), 56.37 (CH), 65.35 (CH), 177.13 (C), 212.37 (C). HRMS: (ESI+)

calculated for $\text{C}_{23}\text{H}_{36}\text{O}_4\text{Na}$ (M^+Na) 399.25058. Found 399.25059. Anal. Calcd for $\text{C}_{23}\text{H}_{36}\text{O}_4$: C, 73.37; H, 9.64. Found: C, 73.45; H, 9.60%.

2.2.2. Methyl (20S)-2 α -hydroxy-6-oxo-5 α -pregnane-20-carboxylate (**2b**)

Yield: 400 mg (46%), m.p. 153–155 °C (MeOH), $[\alpha]_D = -8^\circ$ (c 0.62). IR ν (cm^{-1}) 3608, 1727, 1708, 1168. ^1H NMR (CDCl_3) δ 0.68, 0.75 (both s, 3H, CH_3), 1.20 (d, 3H, $J = 6.9$ Hz, CH_3), 1.95–2.10 (m, 4H), 2.15 (dd, 1H, $J = 11.9$, $J' = 2.6$ Hz, H-5 α), 2.29 (dd, 1H, $J = 13.1$, $J' = 4.6$ Hz, H-7 α), 2.43 (dq, 1H, $J = 10.3$, $J' = 6.8$ Hz, H-20), 3.65 (s, 3H, OCH_3), 3.77 (m, 1H, $W_{1/2} = 31.5$ Hz, H-2 β). ^{13}C NMR δ 12.16 (CH_3), 14.14 (CH_3), 17.06 (CH_3), 19.49 (CH_2), 21.20 (CH_2), 23.96 (CH_2), 26.95 (CH_2), 34.57 (C), 37.57 (CH), 39.18 (CH_2), 42.31 (CH), 42.98 (CH_2), 46.56 (CH_2), 47.07 (C), 51.35 (CH_3), 52.76 (CH), 54.12 (CH), 56.26 (CH), 57.95 (CH), 66.93 (CH), 177.13 (C), 211.58 (C). HRMS: (ESI+) calculated for $\text{C}_{23}\text{H}_{36}\text{O}_4\text{Na}$ (M^+Na) 399.25058. Found 399.25057. Anal. Calcd for $\text{C}_{23}\text{H}_{36}\text{O}_4$: C, 73.37; H, 9.64. Found: C, 73.42; H, 9.57%.

2.3. General procedure for Baeyer–Villiger oxidation

A solution of trifluoroperoxyacetic acid, freshly prepared from trifluoroacetic anhydride (0.45 mL, 3.2 mmol) and hydrogen peroxide (30%; 0.1 mL) in dichloromethane (3 mL), was added to a solution of 6-oxocompound (0.40 mmol) in dichloromethane (8 mL). After standing at room temperature for 5 h, the reaction mixture was poured into water and taken up in chloroform (50 mL). The chloroform extract was washed with water, a saturated solution of potassium hydrogen carbonate, water, and dried over anhydrous sodium sulfate. Solvents were evaporated under reduced pressure and products were separated and purified on HPLC in ethyl acetate–hexane 7:3.

2.4. Baeyer–Villiger oxidation of **2a**

Ketone **2a** (150 mg; 0.40 mmol) was used for Baeyer–Villiger oxidation as described in general procedure. Reaction afforded two lactones **3a** and **3b**.

2.4.1. Methyl (20S)-3 α -hydroxy-7-oxa-7a-homo-6-oxo-5 α -pregnane-20-carboxylate (**3a**)

Yield: 69 mg (44%), m.p. 185–187 °C (MeOH), $[\alpha]_D = +10^\circ$ (c 0.27). IR ν (cm^{-1}) 3615, 1723, 1182. ^1H NMR (CDCl_3) δ 0.72, 0.90 (both s, 3H, CH_3), 1.19 (d, 3H, $J = 6.9$ Hz, CH_3), 1.93 (dt, 1H, $J = 6.6$, $J' = 3.0$ Hz), 2.14 (ddd, 1H, $J = 15.0$, $J' = 12.3$, $J'' = 2.6$ Hz, H-4a), 2.43 (dq, 1H, $J = 10.3$, $J' = 6.9$ Hz, H-20), 3.19 (dd, 1H, $J = 12.2$, $J' = 4.4$ Hz, H-5 α), 3.65 (s, 3H, OCH_3), 4.07–4.11 (m, 2H, H-7 α , H-7 β), 4.17 (m, 1H, $W_{1/2} = 10.9$ Hz, H-3 β). ^{13}C NMR δ 11.97 (CH_3), 14.54 (CH_3), 17.02 (CH_3), 22.11 (CH_2), 24.88 (CH_2), 26.92 (CH_2), 28.27 (CH_2), 32.53 (CH_2), 32.92 (CH_2), 36.34 (C), 39.47 (CH_2), 39.50 (CH), 41.79 (CH), 42.35 (CH), 42.68 (C), 51.15 (CH), 51.39 (CH_3), 52.80 (CH), 58.36 (CH), 64.87 (CH), 70.32 (CH_2), 176.60 (C), 177.02 (C). HRMS: (ESI+) calculated for $\text{C}_{23}\text{H}_{36}\text{O}_5\text{Na}$ (M^+Na) 415.24550. Found 415.24540. Anal. Calcd for $\text{C}_{23}\text{H}_{36}\text{O}_5$: C, 70.38; H, 9.24. Found: C, 70.29; H, 9.30%.

2.4.2. Methyl (20S)-3 α -hydroxy-6-oxa-7a-homo-7-oxo-5 α -pregnane-20-carboxylate (**3b**)

Yield: 50 mg (32%), m.p. 225–227 °C (MeOH), $[\alpha]_D = +10^\circ$ (c 0.22). IR ν (cm^{-1}) 3614, 1723, 1168. ^1H NMR (CDCl_3) δ 0.71, 0.90 (both s, 3H, CH_3), 1.18 (d, 3H, $J = 6.9$ Hz, CH_3), 1.83 (ddd, 1H, $J = 14.2$, $J' = 11.4$, $J'' = 2.7$ Hz, H-4a), 1.92 (d, 1H, $J = 12.3$, $J' = 3.0$ Hz), 2.03 (ddd, 1H, $J = 14.2$, $J' = 5.1$, $J'' = 3.1$ Hz, H-4b), 2.43 (dq, 1H, $J = 10.3$, $J' = 6.9$ Hz, H-20), 2.46–2.53 (m, 2H, H-7 α , H-7 β), 3.65 (s, 3H, OCH_3), 4.22 (m, 1H, H-3 β), 4.62 (dd, 1H, $J = 11.4$, $J' = 5.1$ Hz,

H-5 α). ^{13}C NMR δ 11.56 (CH₃), 11.97 (CH₃), 16.99 (CH₃), 22.17 (CH₂), 25.34 (CH₂), 26.57 (CH₂), 27.99 (CH₂), 31.22 (CH₂), 34.87 (CH), 35.69 (CH₂), 38.10 (CH₂), 39.48 (CH₂), 39.94 (C), 42.33 (CH), 42.73 (C), 51.33 (CH₃), 53.13 (CH), 55.18 (CH), 58.02 (CH), 66.29 (CH), 79.62 (CH), 175.07 (C), 177.05 (C). HRMS: (ESI⁺) calculated for C₂₃H₃₆O₅Na (M⁺+Na) 415.24550. Found 415.24542. Anal. Calcd for C₂₃H₃₆O₅: C, 70.38; H, 9.24. Found: C, 70.31; H, 9.35%.

2.5. Baeyer–Villiger oxidation of **2b**

Ketone **2b** (150 mg; 0.40 mmol) was used for Baeyer–Villiger oxidation as described in general procedure. The reaction afforded two lactones **5a** and **5b**.

2.5.1. Methyl (20S)-2 α -hydroxy-7-oxa-7a-homo-6-oxo-5 α -pregnane-20-carboxylate (**5a**)

Yield: 62 mg (40%), m.p. 173–175 °C (MeOH), $[\alpha]_{\text{D}}^{20} = +20^{\circ}$ (c 0.31). IR ν (cm⁻¹) 3610, 1726, 1170. ^1H NMR (CDCl₃) δ 0.71, 0.92 (both s, 3H, CH₃), 1.19 (d, 3H, $J = 6.9$ Hz, CH₃), 1.58–1.82 (m, 8H), 1.94 (dt, 1H, $J = 6.6$, $J' = 4.4$ Hz), 1.99–2.07 (m, 2H), 2.23 (m, 1H), 2.43 (dq, 1H, $J = 10.3$, $J' = 6.9$ Hz, H-20), 2.67 (dd, 1H, $J = 11.9$, $J' = 4.6$ Hz, H-5 α), 3.65 (s, 3H, OCH₃), 3.67 (tt, 1H, $J = 11.4$, $J' = 4.0$ Hz, H-2 β), 4.03 (dd, 1H, $J = 12.7$, $J' = 8.9$ Hz, H-7 α), 4.09 (dd, 1H, $J = 12.7$, $J' = 2.1$ Hz, H-7 β). ^{13}C NMR δ 11.95 (CH₃), 16.04 (CH₃), 17.02 (CH₃), 22.19 (CH₂), 24.86 (CH₂), 24.91 (CH₂), 26.91 (CH₂), 34.09 (CH₂), 38.94 (C), 39.20 (CH), 39.34 (CH₂), 42.31 (CH), 42.64 (C), 46.79 (CH), 49.53 (CH₂), 51.07 (CH), 51.39 (CH₃), 52.77 (CH), 58.89 (CH), 67.02 (CH), 70.33 (CH₂), 176.03 (C), 176.98 (C). HRMS: (ESI⁺) calculated for C₂₃H₃₆O₅Na (M⁺+Na) 415.24550. Found 415.24534. Anal. Calcd for C₂₃H₃₆O₅: C, 70.38; H, 9.24. Found: C, 70.45; H, 9.31%.

2.5.2. Methyl (20S)-2 α -hydroxy-6-oxa-7a-homo-7-oxo-5 α -pregnane-20-carboxylate (**5b**)

Yield: 77 mg (49%), m.p. 201–203 °C (MeOH), $[\alpha]_{\text{D}}^{20} = +6^{\circ}$ (c 0.13). IR ν (cm⁻¹) 3611, 1725, 1168, 1041. ^1H NMR (CDCl₃) δ 0.71, 0.95 (both s, 3H, CH₃), 1.19 (d, 3H, $J = 6.9$ Hz, CH₃), 1.90–2.04 (m, 3H), 2.20 (m, 1H, H-1 α), 2.40–2.54 (m, 3H, H-7 α , H-7 β , H-20), 3.65 (s, 3H, OCH₃), 3.74 (tt, 1H, $J = 11.3$, $J' = 4.1$ Hz, H-2 β), 4.23 (dd, 1H, $J = 11.1$, $J' = 5.2$ Hz, H-5 α). ^{13}C NMR δ 11.96 (CH₃), 13.26 (CH₃), 17.00 (CH₃), 22.26 (CH₂), 25.29 (CH₂), 26.58 (CH₂), 27.84 (CH₂), 33.27 (CH₂), 34.44 (CH), 38.10 (CH₂), 39.37 (CH₂), 41.20 (C), 42.32 (CH), 42.71 (C), 47.14 (CH₂), 51.39 (CH₃), 53.11 (CH), 55.10 (CH), 58.38 (CH), 66.60 (CH), 82.58 (CH), 174.47 (C), 177.02 (C). HRMS: (ESI⁺) calculated for C₂₃H₃₆O₅Na (M⁺+Na) 415.24550. Found 415.24536. Anal. Calcd for C₂₃H₃₆O₅: C, 70.38; H, 9.24. Found: C, 70.37; H, 9.36%.

2.6. General procedure for ester hydrolysis

Methyl ester (0.15 mmol) in methanol (5 mL) with potassium hydroxide (80 mg, 1.5 mmol) was heated at reflux under nitrogen for 10 h. The solution was neutralized with 5% aqueous HCl (10 mL) and the reaction mixture was left to stand at room temperature for 30 min. The mixture was evaporated to dryness, water (20 mL) was added, and the product was extracted with chloroform (2 \times 50 mL). The chloroform solution was dried with anhydrous sodium sulfate and evaporated under reduced pressure. Product crystallized from ethanol (unless otherwise stated).

2.6.1. (20S)-3 α -Hydroxy-7-oxa-7a-homo-6-oxo-5 α -pregnane-20-carboxylic acid (**4a**)

Compound **4a** was prepared from methyl ester **3a** (50 mg; 0.13 mmol) according to the above general procedure for ester hydrolysis. Yield: 35 mg (73%), m.p. 129–131 °C (lyophilized from *t*-BuOH), $[\alpha]_{\text{D}}^{25} = +25^{\circ}$ (c 0.14; CHCl₃:MeOH–1:1). IR (KBr) ν (cm⁻¹)

3430, 1726, 1711, 1182. ^1H NMR (CDCl₃) δ 0.74, 0.88 (both s, 3H, CH₃), 1.21 (d, 3H, $J = 6.8$ Hz, CH₃), 1.90–2.13 (m, 7H), 2.43 (dq, 1H, $J = 10.3$, $J' = 6.9$ Hz, H-20), 3.18 (dd, 1H, $J = 12.2$, $J' = 4.3$ Hz, H-5 α), 3.95 (m, 1H, H-3 β), 4.06–4.10 (m, 2H, H-7 α , H-7 β). ^{13}C NMR δ 11.97 (CH₃), 14.56 (CH₃), 17.00 (CH₃), 22.10 (CH₂), 24.92 (CH₂), 27.08 (CH₂), 28.20 (CH₂), 32.49 (CH₂), 32.91 (CH₂), 36.32 (C), 39.48 (CH₂), 39.53 (CH), 41.76 (CH), 42.23 (CH), 42.76 (C), 51.20 (CH), 52.37 (CH), 58.32 (CH), 64.94 (CH), 70.29 (CH₂), 176.62 (C), 181.14 (C). HRMS: (ESI⁺) calculated for C₂₂H₃₄O₅Na (M⁺+Na) 401.22985. Found 401.22976. Anal. Calcd for C₂₂H₃₄O₅: C, 69.81; H, 9.05. Found: C, 69.77; H, 9.12%.

2.6.2. (20S)-3 α -Hydroxy-6-oxa-7a-homo-7-oxo-5 α -pregnane-20-carboxylic acid (**4b**)

Compound **4b** was prepared from methyl ester **3b** (70 mg; 0.18 mmol) according to the above general procedure for ester hydrolysis. Yield: 50 mg (74%), m.p. 272–275 °C (EtOH), $[\alpha]_{\text{D}}^{25} = +5^{\circ}$ (c 0.21; CHCl₃:MeOH–1:1). IR (KBr) ν (cm⁻¹) 3428, 1736, 1699, 1151, 1033. ^1H NMR (CDCl₃ + MeOD) δ 0.71, 0.90 (both s, 3H, CH₃), 1.18 (d, 3H, $J = 6.9$ Hz, CH₃), 1.83 (ddd, 1H, $J = 14.2$, $J' = 11.4$, $J'' = 2.7$ Hz, H-4 α), 1.92 (d, 1H, $J = 12.3$, $J' = 3.0$ Hz), 2.03 (ddd, 1H, $J = 14.2$, $J' = 5.1$, $J'' = 3.1$ Hz, H-4 β), 2.43 (dq, 1H, $J = 10.3$, $J' = 6.9$ Hz, H-20), 2.46–2.53 (m, 2H, H-7 α , H-7 β), 3.65 (s, 3H, OCH₃), 4.22 (m, 1H, H-3 β), 4.62 (dd, 1H, $J = 11.4$, $J' = 5.1$ Hz, H-5 α). ^{13}C NMR δ 11.56 (CH₃), 11.97 (CH₃), 16.99 (CH₃), 22.17 (CH₂), 25.34 (CH₂), 26.57 (CH₂), 27.99 (CH₂), 31.22 (CH₂), 34.87 (CH), 35.69 (CH₂), 38.10 (CH₂), 39.48 (CH₂), 39.94 (C), 42.33 (CH), 42.73 (C), 51.33 (CH₃), 53.13 (CH), 55.18 (CH), 58.02 (CH), 66.29 (CH), 79.62 (CH), 175.07 (C), 177.05 (C). HRMS: (ESI⁺) calculated for C₂₂H₃₄O₅Na (M⁺+Na) 401.22985. Found 401.22971. Anal. Calcd for C₂₂H₃₄O₅: C, 69.81; H, 9.05. Found: C, 69.77; H, 9.09%.

2.6.3. (20S)-2 α -Hydroxy-7-oxa-7a-homo-6-oxo-5 α -pregnane-20-carboxylic acid (**6a**)

Compound **6a** was prepared from methyl ester **5a** (60 mg; 0.15 mmol) according to the above general procedure for ester hydrolysis. Yield: 42 mg (73%), m.p. 276–278 °C (EtOH), $[\alpha]_{\text{D}}^{25} = +19^{\circ}$ (c 0.34; CHCl₃:MeOH–1:1). IR (KBr) ν (cm⁻¹) 3435, 1728, 1684, 1201, 1177. ^1H NMR (CDCl₃ + MeOD) δ 0.73, 0.91 (both s, 3H, CH₃), 1.21 (d, 3H, $J = 6.8$ Hz, CH₃), 1.91–2.04 (m, 3H), 2.21 (m, 1H), 2.38 (dq, 1H, $J = 13.5$, $J' = 6.8$ Hz, H-20), 2.71 (dd, 1H, $J = 11.7$, $J' = 4.4$ Hz, H-5 α), 3.67 (tt, 1H, $J = 11.3$, $J' = 3.9$ Hz, H-2 β), 4.05–4.12 (m, 2H, H-7 α , H-7 β). ^{13}C NMR δ 11.63 (CH₃), 15.64 (CH₃), 16.77 (CH₃), 21.96 (CH₂), 24.64 (2 \times CH₂), 26.81 (CH₂), 33.32 (CH₂), 38.62 (C), 38.92 (CH), 39.19 (CH₂), 42.26 (CH), 42.41 (C), 46.60 (CH), 48.90 (CH₂), 50.83 (CH), 52.31 (CH), 58.59 (CH), 66.21 (CH), 70.34 (CH₂), 177.03 (C), 179.05 (C). HRMS: (ESI⁻) calculated for C₂₂H₃₃O₅ (M⁺-H) 377.23335. Found 377.23300. Anal. Calcd for C₂₂H₃₄O₅: C, 69.81; H, 9.05. Found: C, 69.72; H, 9.16%.

2.6.4. (20S)-2 α -Hydroxy-6-oxa-7a-homo-7-oxo-5 α -pregnane-20-carboxylic acid (**6b**)

Compound **6b** was prepared from methyl ester **5b** (40 mg; 0.10 mmol) according to the above general procedure for ester hydrolysis. Yield: 33 mg (85%), m.p. 284–285 °C (EtOH), $[\alpha]_{\text{D}}^{25} = +12^{\circ}$ (c 0.15; CHCl₃:MeOH–1:1). IR (KBr) ν (cm⁻¹) 3425, 1720, 1697, 1188, 1038. ^1H NMR (CDCl₃ + MeOD) δ 0.73, 0.94 (both s, 3H, CH₃), 1.21 (d, 3H, $J = 6.9$ Hz, CH₃), 1.89–2.02 (m, 3H), 2.17 (m, 1H, H-1 α), 2.39 (dq, 1H, $J = 10.3$, $J' = 6.9$ Hz, H-20), 2.47–2.52 (m, 2H, H-7 α , H-7 β), 3.68 (dd, 1H, $J = 14.9$, $J' = 7.6$, $J'' = 3.7$ Hz, H-2 β), 4.28 (dd, 1H, $J = 10.7$, $J' = 4.7$ Hz, H-5 α). ^{13}C NMR δ 11.64 (CH₃), 12.85 (CH₃), 16.78 (CH₃), 22.01 (CH₂), 25.09 (CH₂), 26.47 (CH₂), 27.60 (CH₂), 32.55 (CH₂), 34.26 (CH), 37.79 (CH₂), 39.20 (CH₂), 40.82 (C), 42.26 (CH), 42.47 (C), 46.49 (CH₂), 52.65 (CH), 54.85 (CH), 58.04 (CH), 65.73 (CH), 82.87 (CH), 175.44 (C), 179.08 (C). HRMS: (ESI⁺) calculated for C₂₂H₃₄O₅Na (M⁺+Na) 401.22985.

Found 401.22968. Anal. Calcd for $C_{22}H_{34}O_5$: C, 69.81; H, 9.05. Found: C, 69.71; H, 9.20%.

2.6.5. Methyl (20S)-3 α -(4-nitrobenzoyloxy)-6-oxo-5 α -pregnane-20-carboxylate (**8**)

Diethyl azodicarboxylate (40% in toluene; 0.6 mL; 1.53 mmol) was added to a solution of hydroxyketone **7** (200 mg; 0.53 mmol), triphenylphosphine (400 mg; 1.53 mmol), and 4-nitrobenzoic acid (260 mg; 1.56 mmol) in dry THF. Reaction was stirred at 80 °C for 4 h. THF was evaporated under reduced pressure and the mixture was subjected to column chromatography in ethyl acetate-hexane 1:3 to afford 4-nitrobenzoate **8** (245 mg; 88%), m.p. 193–195 °C (MeOH), $[\alpha]_D = -23^\circ$ (c 0.26). IR ν (cm^{-1}) 1720, 1609, 1600, 1530, 1277. 1H NMR ($CDCl_3$) δ 0.63, 0.74 (both s, 3H, CH_3), 1.13 (d, 3H, $J = 6.9$ Hz, CH_3), 1.85–2.01 (m, 4H), 2.27 (dd, 1H, $J = 13.2$, $J' = 4.5$ Hz, H-7 α), 2.37 (dq, 1H, $J = 10.3$, $J' = 6.8$ Hz, H-20), 2.56 (dd, 1H, $J = 12.4$, $J' = 3.1$ Hz, H-5 α), 3.58 (s, 3H, OCH_3), 5.35 (m, 1H, $W_{1/2} = 5.3$ Hz, H-3 β), 8.12 (m, 2H, $2 \times H_{Ar}$), 8.24 (m, 2H, $2 \times H_{Ar}$). ^{13}C NMR δ 12.20 (CH_3), 12.43 (CH_3), 17.09 (CH_3), 21.09 (CH_2), 23.97 (CH_2), 25.10 (CH_2), 25.40 (CH_2), 26.98 (CH_2), 32.69 (CH_2), 37.88 (CH), 39.19 (CH_2), 41.30 (C), 42.32 (CH), 43.02 (C), 46.61 (CH_2), 51.39 (CH), 52.83 (CH_3), 52.87 (CH), 53.94 (CH), 56.32 (CH), 70.93 (CH), 123.58 ($2 \times CH$), 130.59 ($2 \times CH$), 136.14 (C), 150.51 (C), 163.75 (C), 177.12 (C), 210.99 (C). HRMS: (ESI+) calculated for $C_{30}H_{39}NO_7Na$ ($M^+ + Na$) 548.26187. Found 548.26169. Anal. Calcd for $C_{30}H_{39}NO_7$: C, 68.55; H, 7.48; N, 2.66. Found: C, 68.50; H, 7.56; N, 2.54%.

2.7. Baeyer–Villiger oxidation of **8**

Ketone **8** (200 mg; 0.38 mmol) was used for Baeyer–Villiger oxidation as described in general procedure. The reaction afforded two lactones **9a** and **9b**.

2.7.1. Methyl (20S)-3 α -(4-nitrobenzoyloxy)-7-oxa-7a-homo-6-oxo-5 α -pregnane-20-carboxylate (**9a**)

Yield: 109 mg (53%), m.p. 227–230 °C (MeOH), $[\alpha]_D = +25^\circ$ (c 1.27). IR ν (cm^{-1}) 1724, 1609, 1600, 1531, 1278. 1H NMR ($CDCl_3$) δ 0.73, 0.97 (both s, 3H, CH_3), 1.19 (d, 3H, $J = 6.9$ Hz, CH_3), 2.33 (ddd, 1H, $J = 15.1$, $J' = 12.3$, $J'' = 2.6$ Hz, H-4a), 2.44 (dq, 1H, $J = 10.2$, $J' = 6.8$ Hz, H-20), 3.10 (dd, 1H, $J = 12.2$, $J' = 4.4$ Hz, H-5 α), 3.65 (s, 3H, OCH_3), 4.06–4.17 (m, 2H, H-7 α , H-7 β), 5.39 (m, 1H, $W_{1/2} = 5.2$ Hz, H-3 β), 8.20 (m, 2H, $2 \times H_{Ar}$), 8.32 (m, 2H, $2 \times H_{Ar}$). ^{13}C NMR δ 11.95 (CH_3), 14.57 (CH_3), 17.01 (CH_3), 22.16 (CH_2), 24.85 (CH_2), 25.30 (CH_2), 26.88 (CH_2), 29.77 (CH_2), 33.99 (CH_2), 36.33 (C), 39.32 (CH_2), 39.46 (CH), 42.27 (CH), 42.63 (C), 42.89 (CH), 51.07 (CH), 51.38 (CH_3), 52.79 (CH), 58.78 (CH), 70.46 (CH_2), 70.52 (CH), 123.61 ($2 \times CH$), 130.54 ($2 \times CH$), 135.99 (C), 150.52 (C), 163.78 (C), 175.55 (C), 176.91 (C). HRMS: (ESI+) calculated for $C_{30}H_{39}NO_8Na$ ($M^+ + Na$) 564.25679. Found 564.25603. Anal. Calcd for $C_{30}H_{39}NO_8$: C, 66.52; H, 7.26; N, 2.59. Found: C, 66.46; H, 7.36; N, 2.55%.

2.7.2. Methyl (20S)-3 α -(4-nitrobenzoyloxy)-7-oxo-7a-homo-6-oxo-5 α -pregnane-20-carboxylate (**9b**)

Yield: 70 mg (34%), m.p. 205–208 °C (MeOH), $[\alpha]_D = -10^\circ$ (c 1.64). IR ν (cm^{-1}) 1725, 1609, 1600, 1531, 1272. 1H NMR ($CDCl_3$) δ 0.73, 0.98 (both s, 3H, CH_3), 1.18 (d, 3H, $J = 6.9$ Hz, CH_3), 2.19–2.31 (m, 2H), 2.43 (dq, 1H, $J = 10.3$, $J' = 6.9$ Hz, H-20), 2.46–2.58 (m, 2H, H-7 α , H-7 β), 3.65 (s, 3H, OCH_3), 4.55 (dd, 1H, $J = 11.3$, $J' = 5.3$ Hz, H-5 α), 5.44 (m, 1H, $W_{1/2} = 5.1$ Hz, H-3 β), 8.19 (m, 2H, $2 \times H_{Ar}$), 8.33 (m, 2H, $2 \times H_{Ar}$). ^{13}C NMR δ 11.60 (CH_3), 11.92 (CH_3), 16.96 (CH_3), 22.20 (CH_2), 24.95 (CH_2), 25.29 (CH_2), 26.53 (CH_2), 32.07 (CH_2), 32.93 (CH_2), 34.76 (CH), 38.15 (CH_2), 39.31 (CH_2), 39.86 (C), 42.24 (CH), 42.67 (C), 51.35 (CH_3), 53.11 (CH), 55.04 (CH), 58.30 (CH), 71.61 (CH), 79.60 (CH), 123.62 ($2 \times CH$),

130.58 ($2 \times CH$), 135.72 (C), 150.56 (C), 163.71 (C), 174.31 (C), 176.95 (C). HRMS: (ESI+) calculated for $C_{30}H_{39}NO_8Na$ ($M^+ + Na$) 564.25679. Found 564.25612. Anal. Calcd for $C_{30}H_{39}NO_8$: C, 66.52; H, 7.26; N, 2.59. Found: C, 66.43; H, 7.35; N, 2.50%.

2.8. Hydrolysis of 4-nitrobenzoates **4a** and **4b**

Compound **4a** and **4b** were prepared from methyl esters **9a** and **9b** (each 50 mg; 0.09 mmol) according to general procedure for hydrolysis of ester. Spectral characteristics for both acids were in agreement with the data reported in experiments above.

2.9. Molecular docking

Docking was performed to obtain prediction of conformation and energy ranking between BRI1 receptor (PDB ID: 3RGZ) and the steroid molecule. The docking studies were carried out using AutoDock Vina 1.05 [16]. All 3D structures of BRI1 ligands were obtained with Marvin 5.10.3 [7], software which can be used for drawing, displaying and characterization of chemical structure, substructures and reactions. Ligands were prepared as derivatives of natural ligand brassinolide (BLD). Polar hydrogens were added to all ligands and proteins with the AutoDock Tools (ADT) [17] program prior to docking with Autodock Vina program. Grid box with size of 40 Å were centered on active site of protein. Exhaustiveness parameter was set to 20 (default 8). After docking we compared the docked ligand with brassinolide crystal-like poses and the best crystal-like poses of each ligand were analyzed.

2.10. Cell cultures

The screening cell lines: T-lymphoblastic leukemia CEM; breast carcinoma MCF7 (estrogen-sensitive); cervical carcinoma cell line HeLa; and human foreskin fibroblasts BJ were obtained from the American Type Culture Collection (Manassas, VA, USA). Cells were cultured in DMEM (Dulbecco's Modified Eagle Medium, Sigma, MO, USA). Media used were supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and 1% penicillin–streptomycin. The cell lines were maintained under standard cell culture conditions at 37 °C and 5% CO_2 in a humid environment. Cells were subcultured twice or three times a week using the standard trypsinization procedure.

2.11. Calcein AM cytotoxicity assay

Suspensions with approximately 1.0×10^5 cells/mL were distributed in 96-well microtiter plates and after 24 h of stabilization the BRs analogues tested were added at the desired concentrations in DMSO. Control cultures were treated with DMSO alone, and the final concentration of DMSO in the reaction mixture never exceeded 0.6%. In most cases six serial 4-fold dilutions of the test substances were added at time zero in 20 μ L aliquots to the microtiter plate wells and the highest final concentration in the wells was 50 μ M. After incubation for 72 h, Calcein AM solution (2 μ M, Molecular Probes) was added and the cells were incubated for a further hour. The fluorescence of viable cells was then quantified using a Fluoroskan Ascent instrument (Labsystems, Finland). The percentage of surviving cells in each well was calculated by dividing the intensity of the fluorescence signals from the exposed wells by the intensity of signals from control wells and multiplying by 100. These ratios were then used to construct dose–response curves from which IC_{50} values, the concentrations of the respective compounds that were lethal to 50% of the tumor cells, were calculated.

Table 1
Activity in the bean second-internode bioassay.

Compound	Maximal prolongation of the second internode (mm)	SD (mm)	Amount applied for maximal prolongation (mol)
24-Epibrassinolide	54.8	±10.8	10 ⁻¹⁰
2a	13.1	±6.6	10 ⁻⁹
2b	17.1	±7.7	10 ⁻¹⁰
3a	14.5	±4.6	10 ⁻¹¹
3b	9.7	±3.6	10 ⁻¹¹
4a	10.1	±6.1	10 ⁻¹⁰
4b	3.9	±1.1	10 ⁻⁹
5a	8.1	±5.3	10 ⁻⁸
5b	6.6	±3.2	10 ⁻⁸
6a	11.6	±2.2	10 ⁻⁹
6b	8.3	±4.8	10 ⁻⁹
8	3.4	±0.9	10 ⁻¹²
9a	7.3	±6.3	10 ⁻¹⁰
9b	4.4	±0.8	10 ⁻⁹

2.12. The bean second-internode bioassay

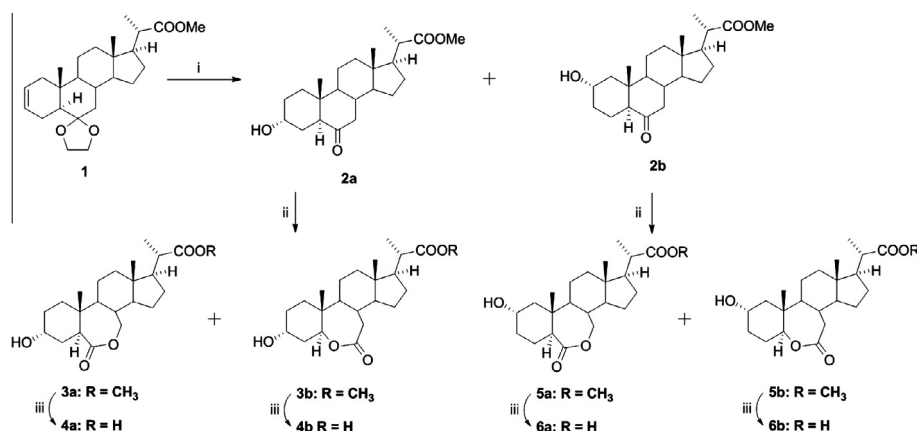
Brassinolide-type activity was measured by the bean second-internode bioassay modified by us [18]. Seeds of bean (*Phaseolus vulgaris* L., cv. Pinto) germinated for 2 days were selected for

uniformity from a large population of seedlings and then transferred into pots containing perlite and 1/10 diluted Hoagland solution (half concentration, pH 5.7). Seedlings were grown in a light-controlled cultivation room (25–27 °C, light 48 W m⁻², light/dark period 16/8 h). Seven days old plants with second internodes 2 mm long were treated with different amounts of tested compounds in 5 μL fractionated lanolin. The substances were applied in microdrops to the scar left after the removal of bract from the base of the second internode. The control plants were treated with lanolin alone. At least seven plants were used for each experiment and the assays were repeated at least three times. The length of the second internodes was measured after 5 days and the difference in length between treated and control plants provided a measure of activity (Table 1).

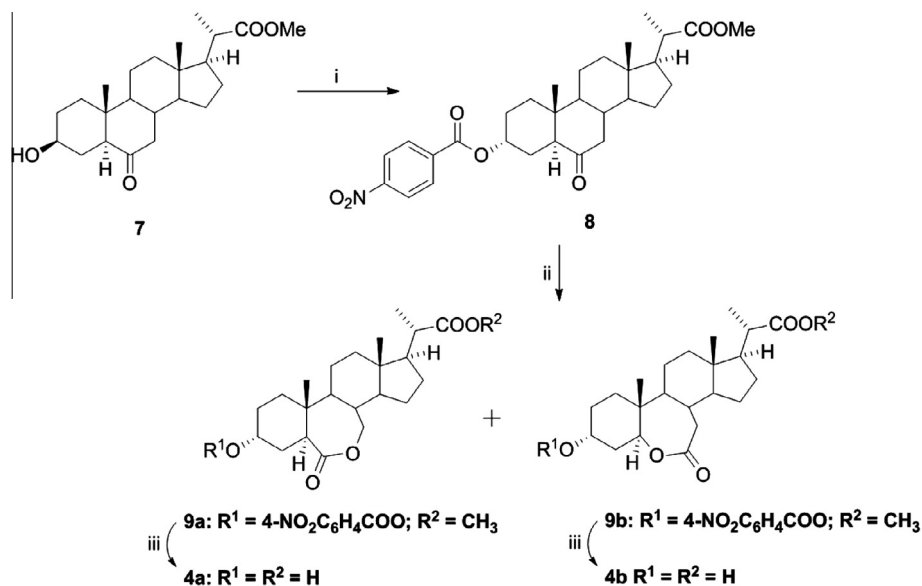
3. Results and discussion

3.1. Chemistry

For the preparation of above mentioned compounds we set out from the known [19] methyl (20S)-6,6-ethylenedioxy-5α-pregn-2-en-20-carboxylate (**1**). This olefin was used for hydroboration reaction [20] was followed by acidic hydrolysis of the ketal group. Two



Scheme 1. (i) (a) B₂H₆, THF, r.t.; (b) 30% H₂O₂, NaOH, H₂O, r.t.; Na₂SO₃, AcOH, H₂O, r.t.; (ii) TFAA, 30% H₂O₂, CH₂Cl₂, r.t.; (iii) (a) KOH, MeOH, refl.; (b) 5% HCl, H₂O, r.t.



Scheme 2. (i) DEAD, Ph₃P, 4-nitrobenzoic acid, THF, 80 °C; (ii) TFAA, 30% H₂O₂, CH₂Cl₂, r.t.; (iii) (a) KOH, MeOH, refl.; (b) 5% HCl, H₂O, r.t.

main products were isolated and characterized as 3α -hydroxyester **2a** and 2α -hydroxyester **2b**. Their structure was determined by 1D and 2D NMR spectra. Both hydroxyesters were then used for Baeyer–Villiger oxidation. In both cases, two isomers of lactone were obtained, **3a** and **3b** from hydroxyester **2a**, and **5a** and **5b** from hydroxyester **2b**. Unlike $2\alpha,3\alpha$ -dihydroxy derivatives when the mostly desired 7-oxa-6-oxo isomer is obtained in very high yield, monohydroxy derivatives afforded both isomers in ratio almost 1:1, even a higher yield of 6-oxa-7-oxo isomer **5b** in case of 2α -hydroxyester **2b**. This fact proved that presence of hydroxy group on A ring is essential for oxidation to a less sterically hindered position. [21] All four lactones were then hydrolyzed with potassium hydroxide to prepare four acids **4a**, **4b**, **6a**, and **6b**.

Acids **4a** and **4b** were prepared also by alternative way using Mitsunobu acylation. As a starting material we used known [22] methyl (20S)- 3β -hydroxy-6-oxo- 5α -pregnan-20-carboxylate (**7**). Standard Mitsunobu procedure [23] used also for steroidal alcohols afforded 4-nitrobenzoate **8** in a very high yield. The following reactions were the same as in previous experiments. Baeyer–Villiger oxidation was followed with alkaline hydrolysis of lactones **9a** and **9b** prepared. Both products were identical with acids **4a** and **4b**. All chemical reactions and conditions can be seen in Scheme 1 and 2. All compounds were characterized by NMR, IR and MS techniques, together with elemental analysis and optical rotation.

3.2. Biological activity and docking

Antiproliferative activity was determined by comparing human normal fibroblasts (BJ) and cancer cell lines (T-lymphoblastic leukemia CEM, breast carcinoma MCF7, cervical carcinoma cell line HeLa). Cells of all these lines were exposed to six 3-fold dilutions of each drug for 72 h prior to determination of cell survival. The IC_{50} (concentration leading to 50% inhibition of viability) values obtained from Calcein AM cytotoxicity assay were calculated. 28-homocastasterone was used as a positive control, which is most potent natural brassinosteroid towards CEM cells (IC_{50} 13 μ M, [12]). All tested monohydroxylated BRs had no detectable activity,

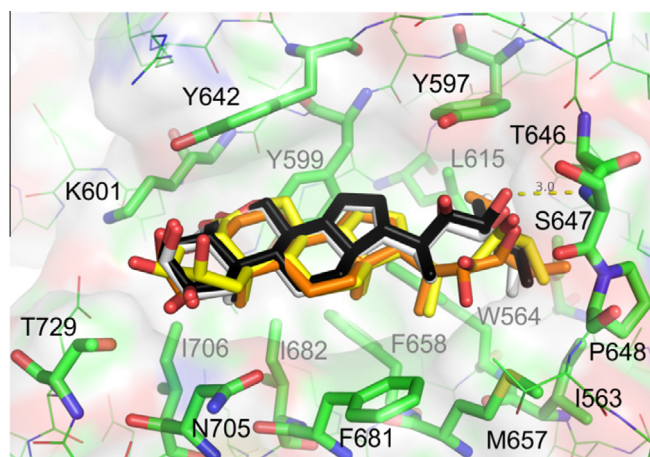


Fig. 1. Comparison of brassinolide analogues with brassinolide crystal pose (black; PDB ID: 3RGZ). Brassinolide interacts with mostly nonpolar groove in a receptor. Only one hydrogen bond is present between the 23-hydroxyl group and main chain nitrogen in S647 residue. Brassinolide docked into the same pose as in crystal (cf. white and black). Both 24-epibrassinolide (yellow) and 28-homocastasterone (orange) showed similar pose with slightly less binding affinity accountable to the small variation in the tail region, while still retaining the hydrogen bond found in crystal. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

even when tested in concentrations of up to 50 μ M. No BRs derivative mediated loss of viability was observed in the BJ fibroblasts.

Brassinolide-type activity was measured by the second internode bean bioassay [24]. 24-Epibrassinolide showed very good biological activity and therefore was used as a positive control. This compound elicited nearly 200% increase in elongation of the

Table 2

Resulting binding free energies for best poses and for crystal-like pose and its ranking between all poses. Star (*) denotes structures with common first pose binding more deeply in the receptor groove similarly as shown on Fig. 2.

Compound	$\Delta G_{\text{bind best}}$ (kcal/mol)	$\Delta G_{\text{bind X-tal}}$ (kcal/mol)
Brassinolide	-10.6	-10.6 (1.)
24-Epibrassinolide	-9.8	-9.8 (1.)
28-Homocastasterone	-9.6	-9.3 (2.)
2a	-10.1*	-9.1 (2.)
2b	-10.1*	-9.0 (2.)
3a	-9.7*	-8.9 (4.)
3b	-10.5*	-9.0 (4.)
4a	-9.8*	-9.0 (4.)
4b	-10.4*	-8.2 (11.)
5a	-9.8*	-8.9 (2.)
5b	-10.2*	-8.0 (10.)
6a	-9.9*	-9.0 (3.)
6b	-10.1*	-9.0 (3.)
8	-8.9	-8.6 (2.)
9a	-8.9	-8.3 (3.)
9b	-8.9	-7.7 (14.)

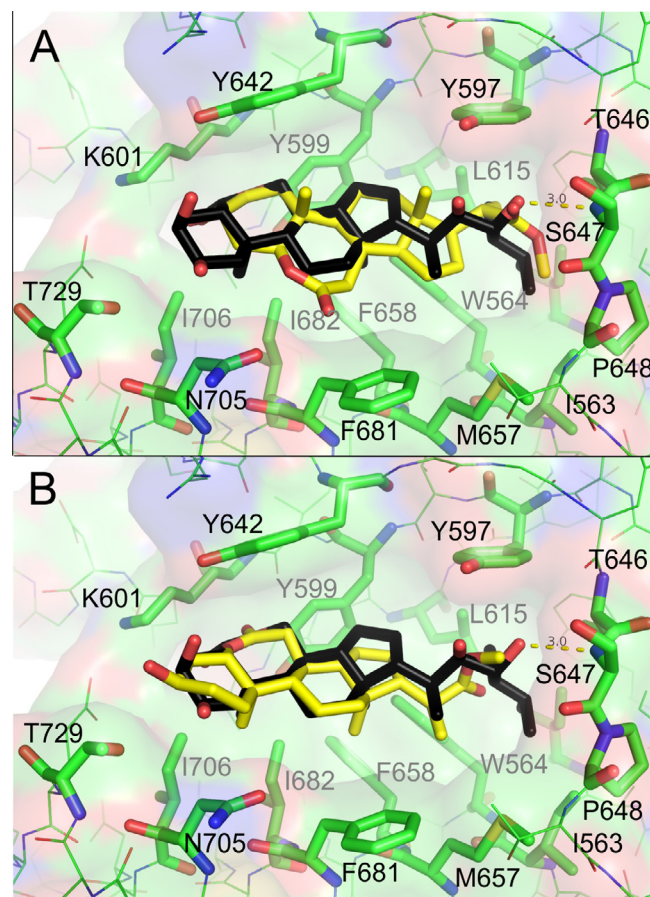


Fig. 2. Comparison of monosubstituted brassinolide analogue **3b** (yellow) with brassinolide crystal pose (black; PDB ID: 3RGZ). Two major poses were detected – (A) more deeply immersed pose with best binding free energy and with hydrogen bonding pattern to S647, (B) brassinolide-like binding pose with lower affinity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

second internode, causing its elongation, swelling, twisting and splitting [25]. The newly prepared monohydroxylated BRs analogs were very weakly active inducing elongation of the second bean internode only (Table 1); the best activity was found in **2b** (+17.1 mm) and in **3a** (+14.5 mm). In the case of **2b**, applied amount was 10^{-10} mol and in the case of **3a** 10^{-11} mol per plant.

The weak activity of derivatives was further analyzed with use of molecular docking of their structures into the receptor domain of kinase brassinosteroid insensitive 1 (BRI1). Firstly, we docked into the structure of brassinolide, 24-epibrassinolide and 28-homocastasterone in order to analyze their best fit. All these compounds docked into the receptor groove in a similar fashion to brassinolide crystal pose (Fig. 1) with only a slightly less affinity to the receptor than brassinolide itself (Table 2). Binding pose is mainly dictated by the number of nonpolar residues in a receptor groove forming a hydrophobic pocket interacting with nonpolar parts of rings and tail of brassinolide and its derivatives. Hydroxyl groups are mostly pointing outward from the receptor groove to the solvent, whereas 23-hydroxyl group interacts with the main chain nitrogen in S647 residue. Brassinolide, 24-epibrassinolide and 28-homocastasterone all have the nonpolar tail pointing deeper into the hydrophobic pocket close to W564 and L615 residues. In such a pose, 2,3-dihydroxy group, thought to be important for biological response [7], positioned in the same area pointing outward from the binding pocket. It was thus shown, that Autodock Vina is able to dock ligands into the binding pocket in crystal-like pose within energetically best poses (Table 2).

However in the case of the monohydroxylated derivatives synthesized in this work, we have found that these compounds showed two distinct binding poses: (i) one with the best binding free energy characterized with a more deeply immersed pose and retained hydrogen bonding pattern to S647, and (ii) a brassinolide-like binding pose with a lower affinity lacking interactions with hydrophobic pocket due to the shorter and more polar tail (Fig. 2). Moreover, the introduction of 4-nitrobenzoic acid on 3-hydroxyl group further lowered the affinity towards the BRI receptor. As most of the interactions between the receptor binding groove and the ligand are mediated through nonpolar interactions, the larger the area of contact between ligand and pocket would govern the affinity. In such case, the shorter aliphatic tail is moving prepared the monohydroxylated brassinosteroid derivatives out of the productive pose more deeply into the receptor groove, thus effectively blocking 2- or 3-hydroxyl groups, which are thought to be important for the biological response of the receptor [7] and diminishing their proposed function in brassinolide-type activity. In future, we will focus on the aliphatic tail elongation in the brassinosteroid derivatives synthesis, as this seems to be the key feature of the brassinolide-like substrates interacting with BRI receptor.

Acknowledgments

The program “Návrát” for Research, Development, and Innovations” (LK21306), which is funded by the Ministry of Education, Youth and Sports of the Czech Republic, is highly appreciated.

This work was supported by project of the Ministry of Education, Youth and Sports CR NPUI LO1204 and by the Czech Science Foundation (GACR 14-27669P).

V.B. acknowledges support by the Student Project PrF_2013_028 of Palacky University.

K.B. acknowledges support by Operational Program Research and Development for Innovations – European Regional Development Fund (CZ.1.05/2.1.00/03.0058).

References

- [1] (a) Bajguz A. Metabolism of brassinosteroids in plants. *Plant Physiol Biochem* 2007;45:95–107; (b) Khripach VA, Zabinskij VN, de Groot AE. Brassinosteroids – A New Class of Plant Hormones. Academic Press; 1999; (c) Back TG, Pharis RP. Structure–activity studies of brassinosteroids and the search for novel analogues and mimetics with improved bioactivity. *J Plant Growth Regul* 2003;22:350–61; (d) Sakurai A, Yokota T, Clouse SD. Brassinosteroids – Steroidal Plant Hormones. Tokyo: Springer-Verlag; 1999.
- [2] Clouse SD, Sasse JM. Brassinosteroids: essential regulators of plant growth and development. *Ann Rev Plant Physiol Plant Mol Biol* 1998;49:427–51.
- [3] Li JM, Chory J. A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. *Cell* 1997;90:929–38.
- [4] Kinoshita T, Cano-Delgado A, Seto H, Hiranuma S, Fujioka S, Yoshida S, Chory J. Binding of brassinosteroids to the extracellular domain of plant receptor kinase BRI1. *Nature* 2005;433:167–71.
- [5] She J, Han ZF, Zhou B, Chai JJ. Structural basis for differential recognition of brassinolide by its receptors. *Protein Cell* 2013;4:475–82.
- [6] She J, Han ZF, Kim TW, Wang JJ, Cheng W, Chang JB, et al. Structural insight into brassinosteroid perception by BRI1. *Nature* 2011;474:472–6.
- [7] Hothorn M, Belkhadir Y, Dreux M, Dabi T, Noel JP, Wilson IA, Chory J. Structural basis of steroid hormone perception by the receptor kinase BRI1. *Nature* 2011;474:467–71.
- [8] Bhardwaj R, Arora N, Uppal P, Sharma I, Kanwar MK. Prospects of brassinosteroids in medicinal applications. In: Hayat S, Ahmad A, editors. *Brassinosteroids: A class of Plant Hormone*. Germany: Springer; 2011. p. 439–59.
- [9] Wachsman MB, Lopez EM, Ramirez JA, Galagovsky LR, Coto CE. Antiviral effect of brassinosteroids against herpes virus and arenaviruses. *Antivir Chem Chemother* 2000;11:71–7.
- [10] Wachsman MB, Lopez EM, Ramirez JA, Galagovsky LR, Coto CE. Antiviral activity of brassinosteroids derivatives against measles virus in cell cultures. *Antivir Chem Chemother* 2002;13:61–6.
- [11] Wachsman MB, Castilla V, Talarico LB, Ramirez JA, Galagovsky LR, Coto CE. Antiherpetic mode of action of (22S,23S)-3 β -bromo-5 α ,22,23-trihydroxystigmastan-6-one in vitro. *Int J Antimicrob Agents* 2004;5:524–6.
- [12] Malikova J, Swaczynova J, Kolar Z, Strnad M. Anticancer and antiproliferative activity of natural brassinosteroids. *Phytochemistry* 2008;69:418–26.
- [13] Steigerova J, Oklestkova J, Levkova M, Rarova L, Kolar Z, Strnad M. Brassinosteroids cause cell cycle arrest and apoptosis of human breast cancer cells. *Chem Biol Interact* 2010;188:487–96.
- [14] Steigerova J, Rarova L, Oklestkova J, Krizova K, Levkova M, Svachova M, Kolar Z, Strnad M. Mechanisms of natural brassinosteroid-induced apoptosis of prostate cancer cells. *Food Chem Toxicol* 2012;50:4068–76.
- [15] Rarova L, Zahler S, Liebl J, Kryštof V, Sedláč D, Bartůněk P, Kohout L, Strnad M. Brassinosteroids inhibit in vitro angiogenesis in human endothelial cells. *Steroids* 2012;77:1502–9.
- [16] Trott O, Olson AJ. Software News and Update AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010;31:455–61.
- [17] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 2009;16:2785–91.
- [18] Sisa M, Hnilickova J, Swaczynova J, Kohout L. Syntheses of new androstane brassinosteroids with 17 β -ester group – butyrates, heptafluorobutyrate, and laurates. *Steroids* 2005;70:755–62.
- [19] Eignerova B, Slavikova B, Budesinsky M, Dracinsky M, Klepetarova B, Stastna E, Kotora M. Synthesis of fluorinated brassinosteroids based on alkene cross-metathesis and preliminary biological assessment. *J Med Chem* 2009;52:5753–7.
- [20] Hanson JR, Liman MD, Nagaratnam S. Regio- and stereochemical effects in the hydroboration of Δ^2 -steroidal allylic and homoallylic alcohols. *J Chem Res (S)* 1997;22:282–3.
- [21] Takatsuto S, Ikekawa N. Remote substituent effect on the regioselectivity in the Baeyer–Villiger oxidation of 5 α -cholestan-6-one derivatives. *Tetrahedron Lett* 1983;24:917–20.
- [22] Cerny V, Strnad M, Kaminek M. Preparation of 2 α , 3 α -dihydroxy-7-oxa-6-oxo-23,24-dinor-B-homo-5 α -cholanolic acid, its esters and amides as brassinolide analogs. *Collect Czech Chem Commun* 1986;51:687–97.
- [23] Ismaili J, Boisvert M, Longpre F, Carange J, Gall CL, Martinoli MG, Daoust B. Brassinosteroids and analogs as neuroprotectors: synthesis and structure–activity relationships. *Steroids* 2012;77:91–9.
- [24] Mitchell JW, Mandava N, Worley JF, Plimmer JR, Smith MV. Brassins a new family of plant hormones from rape pollen. *Nature* 1970;225:1065–6.
- [25] Swaczynová J, Novák O, Hauserová E, Fuksová K, Šiša M, Strnad M. New techniques for the estimation of naturally occurring brassinosteroids. *J Plant Growth Regul* 2006;26:1–14.

Appendix B

Kvasnica M, Oklešťková J, **Bazgier V**, Rárová L, Kořínková P, Mikulík J, Buděšínský M,
Béres T, Berka K, Lu Q, Russinova E, Strnad M:

Design, synthesis and biological activities of new brassinosteroid analogues with phenyl
group in the side chain.

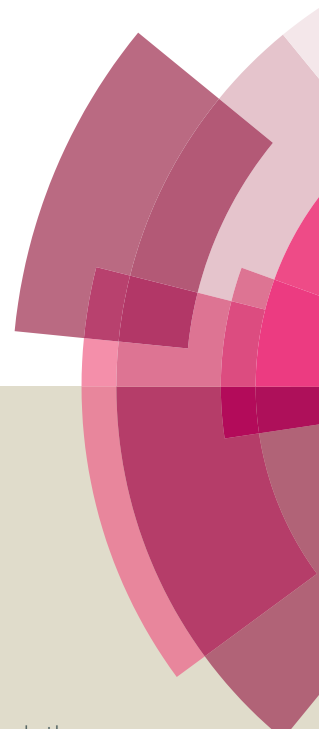
Org. Biomol. Chem. , 14(37), 8691-8701, 2016.

DOI: 10.1039/C6OB01479H

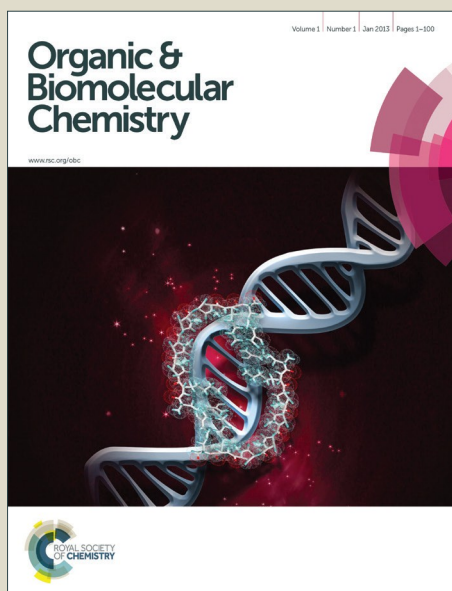
IF = 3.564

Organic & Biomolecular Chemistry

Accepted Manuscript



This article can be cited before page numbers have been issued, to do this please use: M. Kvasnica, J. Oklestkova, V. Bazgier, L. Rárová, P. Korinkova, J. Mikulik, M. Budesinsky, T. Beres, K. Berka, Q. Lu, E. Russinova and M. Strnad, *Org. Biomol. Chem.*, 2016, DOI: 10.1039/C6OB01479H.



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Design, synthesis and biological activities of new brassinosteroid analogues with phenyl group in the side chain

M. Kvasnica,^{a,†} J. Oklestkova,^a V. Bazgier,^{a,b} L. Rárová,^c P. Kořínková,^a J. Mikulík,^a M. Budesinsky,^d T. Béres,^c K. Berka,^{b,e} Q. Lu,^{f,g} E. Russinova,^{f,g} M. Strnad^a

00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x
www.rsc.org/

We have prepared and studied a series of new brassinosteroid derivatives with *p*-substituted phenyl group in the side chain. To have the best comparison between molecular docking and biological activities both types of brassinosteroid were synthesized; 6-ketones, 10 examples, and B-lactones, 8 examples. The phenyl group was introduced into steroid skeleton by Horner–Wadsworth–Emmons. The docking studies were carried out using AutoDock Vina 1.05. Plant biological activities were established using different brassinosteroid bioassays in comparison with natural brassinosteroids. Differences in the production of plant hormone ethylene were also observed in etiolated pea seedlings after treatment with new brassinosteroids. The most active compounds were lactone 8f and 6-oxoderivatives 8c and 9c, their biological activities were comparable or even better than naturally occurring brassinolide. Finally cytotoxicity of new derivatives was studied using human normal and cancer cell lines.

Introduction

Brassinosteroids (BRs, Fig. 1) are a class of plant steroid hormone, that are now known to be essential for many aspects of plant growth and development, such as cell division, elongation and differentiation, pollen tube growth, seed germination, regulation of gene expression, enzyme activation and photosynthesis.^{1–3} They are also involved in defense against a wide range of biotic and abiotic stresses, such as water, temperature, oxidative stresses and high salinity.^{4,5} Moreover, recent studies have shown that natural BRs have potential application to medicine due to antiviral,^{6,7} immunomodulatory and neuroprotective activity,^{8,9} and anti-proliferative effects in animal cells in vitro.^{10–15} In contrast to the

mammalian nuclear steroid receptor, BRs are perceived at the cell surface by the transmembrane receptor complex formed by the receptor kinase BRI1 and its co-receptor BAK1.^{16–19} The BRI1 receptor has a binding site for BRs located in the extracellular ectodomain. There, the nonpolar side of BRs fit into a highly nonpolar cavity of the receptor cleft and hydroxyl groups of BRs are exposed to the solvent or towards interaction with BAK1 or SERK1.^{18,20} This structural knowledge, formed the basis of the idea of the brassinosteroid side chain modifications using a nonpolar group such as the phenyl group. Moreover, some analogues containing cycloalkyl substituents at C-24 (replacing the isopropyl group of brassinolide **1**) exhibited significant activity in rice lamina inclination biotest, where the compound with the cyclohexyl group, showed the lowest activity, whereas the cyclopentyl analogue was comparably active as brassinolide in this bioassay.^{21,22}

The aim of this study was to synthesize new brassinosteroid derivatives with a *p*-substituted phenyl group in the side chain and study their biological properties. The phenyl group was chosen owing to its successful molecular docking into the active site of BRI1 using AutoDock Vina. Some compounds showed marked interaction with the BRI1 receptor. The biological activities of newly prepared derivatives were confirmed using by plant bioassay (pea inhibition biotest, Arabidopsis root and hypocotyl sensitivity bioassay and BES-1 dephosphorylation assay). Their cytotoxic activities were studied using human normal and cancer cell lines.

^a Laboratory of Growth Regulators, Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany ASCR & Palacký University, Šlechtitelů 27, 78371 Olomouc, Czech Republic.

^b Department of Physical Chemistry, Faculty of Science, Palacký University, tř. 17. Listopadu 12, 77146 Olomouc, Czech Republic.

^c Department of Chemical Biology and Genetics, Centre of the Region Haná for Biotechnological and Agricultural Research, Palacký University, Šlechtitelů 27, 78371 Olomouc, Czech Republic.

^d Institute of Organic Chemistry and Biochemistry, ASCR, Flemingovo n. 2, 16610 Prague 6, Czech Republic.

^e Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry Palacký University in Olomouc, 17. listopadu 1131, Olomouc CZ779 00, Czech Republic.

^f Department of Plant Systems Biology, VIB, 9052 Gent, Belgium.

^g Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Gent, Belgium.

† Corresponding author: kvasnica@ueb.cas.cz

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

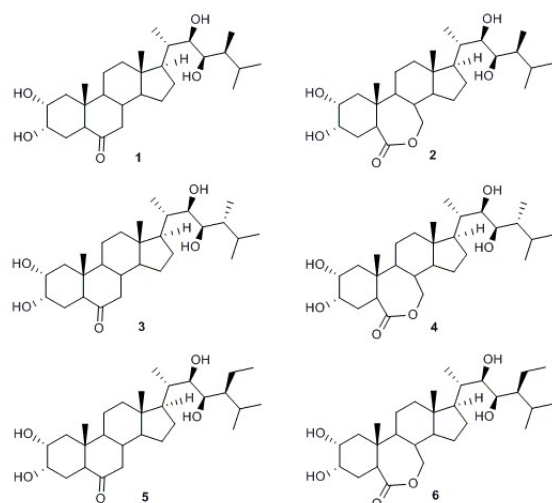


Fig. 1: Structures of most common natural brassinosteroids; castasterone (**1**), brassinolide (**2**), 24-epicastasterone (**3**), 24-epibrassinolide (**4**), 28-homocastasterone (**5**), 28-homobrassinolide (**6**).

Results and discussion

Chemistry

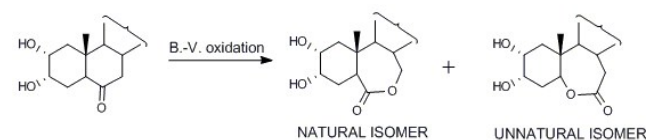
For the preparation of desired phenyl analogues of brassinosteroids, we set out from the known aldehyde **7** ((2*OS*)-6,6-ethylenedioxy-5 α -pregn-2-en-20-carbaldehyde) prepared according to a published procedure.²³ With this aldehyde Horner–Wadsworth–Emmons (HWE) reaction was then carried out with different commercially available *p*-substituted benzylphosphonates. Based on the character of the BRI1 non-polar part of the cavity for side chain, we used these substituents in *para* position: fluorine, chlorine, bromine, iodine, nitro, methyl, methoxy, nitrile, and isopropyl. Benzyl triphenylphosphonium chloride was used for preparation of non-substituted aryl analogues.

Despite standard use of sodium hydride as a base in the HWE reaction, we observed isomerization of the methyl group on C-20 (see supporting information for detailed reaction conditions and analysis). Using this base along with steric hindrance of the reaction site for bulky aryl group, the reaction time increased and this led to enolization of aldehyde. The aryl group of phosphonates used, is crucial for this epimerization as it was not observed when smaller stabilizing groups were used (e.g. COOR, CN).²⁴ Using *n*-butyllithium instead, solved the problem and only desired aryl diens **8a–17a** with 22*E* configuration were obtained. In almost all cases, the reaction gave products in good yields (80–90 %). Only compound **12a** was prepared in lower yield (65 %) due to the presence of reactive iodine.

Further hydrolysis of ketal group gave aryl dienons **8b–17b** which were subjected to dihydroxylation in almost quantitative yield. Simultaneous Sharpless dihydroxylation of both double bonds was used to minimize formation of unnatural configuration of 22 and 23 hydroxy groups. As a chiral ligand, we used hydroquinidine 4-chlorobenzoate. The reaction rate was increased by addition of methansulfonamide.²⁵ Without methansulfonamide, the reaction took more than 48 hours. Such reaction conditions allowed us to isolate only the desired 22*R*,23*R*-isomers **8c–17c** in good yields (75–83 %). The correct configuration was determined according to a published alternative preparation of **8c**.²⁶ The unnatural isomers (22*S*,23*S*) were formed only in trace amounts and therefore not

isolated. Dihydroxylation of the double bond on the A-ring is stereo controlled by A-ring conformation and the presence of a methyl group on C10. The attack of reagent is always from the bottom side of molecule and thus only 2 α ,3 α -diol was detected.

As brassinosteroids with lactone in the B-ring are known to show higher biological activity than corresponding 6-ketones, we decided to prepare them as well. The direct Baeyer-Villiger oxidation of tetrahydroxy-ketones with freshly prepared trifluoroperoxyacetic acid led to both isomers in a ratio approx. 10–15:1 (with natural isomer favoured). However, these mixtures were inseparable even by HPLC (Scheme 1).



Scheme 1: Baeyer-Villiger oxidation of B-ring showing formation of natural and unnatural isomeric lactones.

For this reason, we had to prepare tetraacetates **8d–17d** first and carry out the oxidation reaction on these. This “detour” allowed us to prepare and easily separate the required lactones **8e–11e**, **13e**, **14e**, **16e**, and **17e** in good yields (79–88 %). Unfortunately, two aryl-ketones **12d** and **15d** were unstable during the reaction and only a mixture of products was obtained in both cases. The instability of iodophenyl derivative **12d**, was predictable owing to the presence of easily oxidizable iodine. Due to the presence of the activating group, the methoxyphenyl derivative **15d** is more nucleophilic on the benzene ring than other derivatives and thus also undergoes electrophilic aromatic substitution reaction (substitution with trifluoroacetate). These two reactions also failed using 3-chloroperoxybenzoic acid.

The last reaction step was basic hydrolysis of tetraacetoxy-lactones to corresponding tetrahydroxy-lactones **8f–11f**, **13f**, **14f**, **16f**, and **17f** in almost quantitative yields. The reaction steps can be seen in Scheme 2. All compounds were characterized by NMR, IR and MS techniques. Compounds used for biological testing were also characterized for purity by elemental analysis and melting point.

<Scheme 2>

Scheme 2: Synthesis of brassinosteroid phenyl analogues: a) benzyltriphenylphosphonium bromide or diethyl arylphosphonates, *n*-BuLi/THF; b) 5% HCl/THF; c) OsO₄, CH₃SO₂NH₂, K₃[Fe(CN)₆], K₂CO₃, hydroquinidine 4-chlorobenzoate/*t*-BuOH, H₂O; d) Ac₂O/pyridine; e) trifluoroperoxyacetic acid/ CH₂Cl₂; f) NaOH/THF, H₂O.

Molecular docking

Molecular docking is a useful tool for understanding of the pose and energetics of a protein-ligand complex. The binding site of BRI1 is located on the surface of the receptor ectodomain as a nonpolar cleft lined by nonpolar aromatic and aliphatic residues (I540, I563, W564, Y599, Y642, M657, F681, I682, I706), whereas hydroxyl groups form the cleft ridge (Y597, Y599, Y642, S647). Brassinolide fits into the cleft via its nonpolar side and displays its hydroxyl groups towards the solvent and protein partners (Fig. 2).

Since there is a space left in the cleft around the brassinolide chain, we were running molecular docking of BRs derivatives with a phenyl ring on the tail replacing 1',2'-dimethylpropyl moiety of brassinolide. Molecular docking predicted similar or better binding

energies than for brassinolide for compounds **8c**, **8f**, **9c**, **9f**, **10c**, **10f**, and **14f** (Fig. 3). This implies that derivatives with a phenyl ring on the tail or phenyl ring with small groups such as fluorine, chlorine or methyl should be accommodated within the BRI1 cleft at least as easily as brassinolide itself and hence they were the best candidates for showing similar binding experimentally. On the other hand, derivatives with larger groups did not fit well into the cleft and the docked pose often revealed the tail out of the cleft completely (for the pose of all compounds see sup. inf.). In all docking cases, lactones showed better binding energies than 6-ketones due to better fit to cavity, close to Y599. The best compound to emerge from molecular docking was compound **8f**.

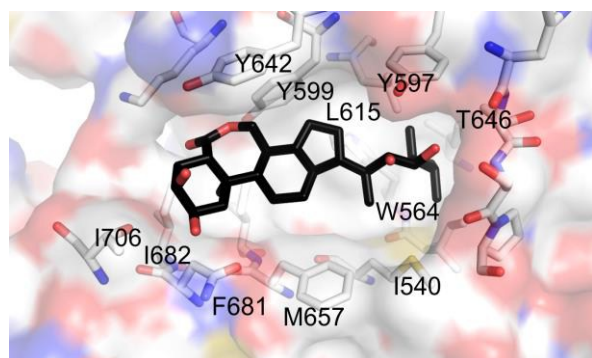


Fig. 2: 3D view of brassinolide with amino acid around active site.

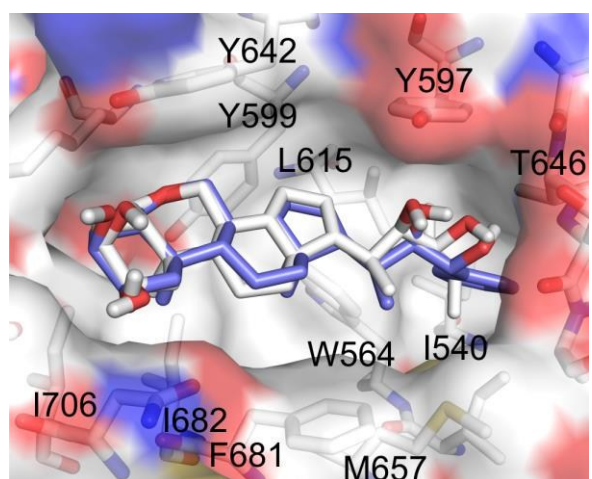


Fig. 3: Pose of **8f** within BRI1 binding site. Structure of brassinolide and BRI1 binding site are shown in white, structure of **8f** is shown in blue.

Biology

The response of plant tissues to applied BRs varies with BR concentration. In most cases, low concentration induces elongation and curvature as a result of cell division and cell elongation. Most BR bioassays are based on this effect.²⁷ However, there are other ways of regulating growth by BRs. For example, BRs inhibit the growth of etiolated pea seedlings at high concentration and this is probably caused by increased ethylene production. Ethylene effects alteration of the normal planes of cell growth. Radial swelling or abnormal radial expansion of stem, such as that seen in the response of etiolated pea seedlings to ethylene application, results from inhibited elongation, increased radial expansion and probably also it accounts for leaf epinasty.²⁸ These effects are known as the "triple response" of etiolated seedlings to ethylene.²⁹

Tab. 1: IC₅₀ (mol/L) values obtained from the pea inhibition biotest.

Compound	IC ₅₀ (mol.L ⁻¹)
Brassinolide	2.2.10 ⁻⁵ ± 2.10 ⁻⁶
8c	2.5.10 ⁻⁶ ± 3.10 ⁻⁷
8f	1.8.10 ⁻⁶ ± 5.10 ⁻⁸
9c	2.0.10 ⁻⁶ ± 3.10 ⁻⁷
9f	2.6.10 ⁻⁶ ± 8.10 ⁻⁸
10c	1.8.10 ⁻⁵ ± 3.10 ⁻⁶
10f	2.3.10 ⁻⁵ ± 2.10 ⁻⁶
11c	2.7.10 ⁻⁴ ± 2.10 ⁻⁵
11f	1.7.10 ⁻⁶ ± 3.10 ⁻⁷
12c	2.1.10 ⁻² ± 4.10 ⁻³
13c	4.0.10 ⁻² ± 5.10 ⁻⁴
13f	no inhibition
14c	1.8.10 ⁻⁴ ± 7.10 ⁻⁶
14f	1.4.10 ⁻⁶ ± 2.10 ⁻⁷
15c	no inhibition
16c	2.1.10 ⁻⁴ ± 5.10 ⁻⁶
16f	no inhibition
17c	2.0.10 ⁻³ ± 6.10 ⁻⁵
17f	2.1.10 ⁻⁴ ± 5.10 ⁻⁶

The IC₅₀ values obtained from the pea inhibition biotest are summarized in Table 1. The most active lactones were **8f**, **9f**, **11f** and **14f** (IC₅₀ 2.56.10⁻⁶ - 1.4.10⁻⁶ mol.L⁻¹) whereas brassinolide (BL), the most active natural BRs used as a positive control, was about ten times less active (IC₅₀ 2.2.10⁻⁵ mol.L⁻¹). 6-oxoderivatives **8c** and **9c** were also active in this bioassay. Their IC₅₀ values were comparable with active lactones (Table 1). Compounds **15c**, **13f** and **16f** showed no inhibition of etiolated pea plants. Dose response curves for the most active BRs derivatives are shown in Fig. 4.

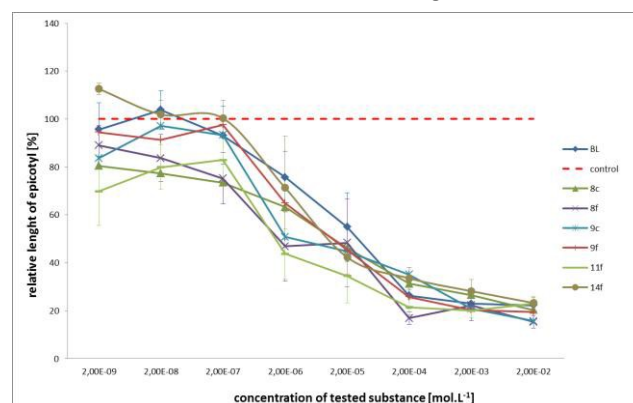


Fig. 4: Effect of selected brassinosteroid derivatives on the inhibition of etiolated pea seedlings. Error bars represent s.d.

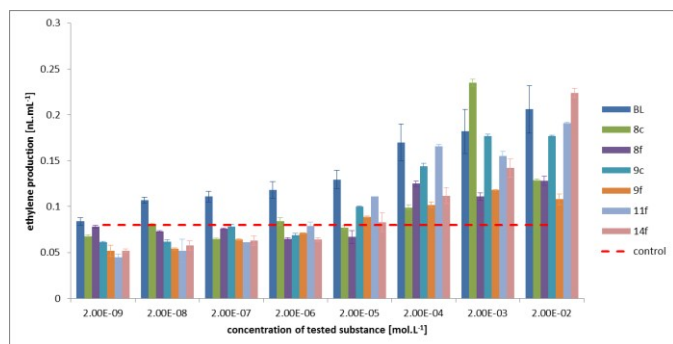


Fig. 5: Effect of selected brassinosteroid derivatives on ethylene production (nL.mL^{-1}) in etiolated pea seedlings determined by GC-FID 24 h after ventilation. Error bars represent s.d.

Levels of ethylene were measured in cultivation vessels during the incubation of etiolated pea plants after treatment with different BR derivatives (Fig. 5). The highest concentration of ethylene (235 and 224 nL.L^{-1}) was determined after 24 h treatment by **8c** and **14f** compared to 206 nL.L^{-1} for BL treatment. While level of this gaseous plant hormone produced by untreated control pea plants, was found to be significantly lower (about 80 nL.L^{-1} , Fig.5). Artica et al.³⁰ also observed stimulated ethylene production in etiolated mung bean segments 4 hours after treatment with $1 \mu\text{M}$ BL and this increased production became greater over the following 20 h.

The most potent compounds (**8c**, **8f** and **9c**) were tested in Arabidopsis brassinosteroid sensitivity and BES-1 dephosphorylation bioassays. The potency of these compounds was as follows: **BL** \geq **8f** $>$ **9c** $>$ **8c** in Arabidopsis sensitivity bioassays. The effects of tested compounds on the Arabidopsis roots and hypocotyls, are shown in supplementary information (Fig. S3 and Fig. S4). Compound **8f** significantly increased dephosphorylation of BES1 (Fig. 6), which is an important transcription factor in the BR signalling pathway. Taken together, these results confirm that the biological activity of compound **8f** is comparable with natural BR-brassinolide.

The antiproliferative activity of prepared brassinosteroid derivatives was tested using several models of normal and cancer cell lines. We compared the *in vitro* cytotoxic activity of selected analogues against human foreskin fibroblasts (BJ) and cancer cell lines of various histopathological origins, including T-lymphoblastic leukemia CEM, breast carcinoma (MCF7) and cervical carcinoma (HeLa). Cells were exposed to six 3-fold dilutions of each drug for 72 h prior to determination of cell survival. The IC_{50} (concentration leading to 50% inhibition of viability) values obtained from Calcein AM cytotoxicity assay are presented in supplementary information (Table S2). Most tested BRs analogues had no detectable cytotoxic activity, even when tested in concentrations of up to $50 \mu\text{M}$. Only compounds **10f**, **11f** and **13f** showed moderate cytotoxic activity against CEM and HeLa cell lines (IC_{50} around $35 \mu\text{M}$). No BRs derivative mediated loss of viability was observed in the BJ fibroblasts. 24-Epibrassinolide was used as a control. It is a natural brassinosteroid with modest cytotoxicity against CEM cells (IC_{50} $44 \mu\text{M}$).¹⁵

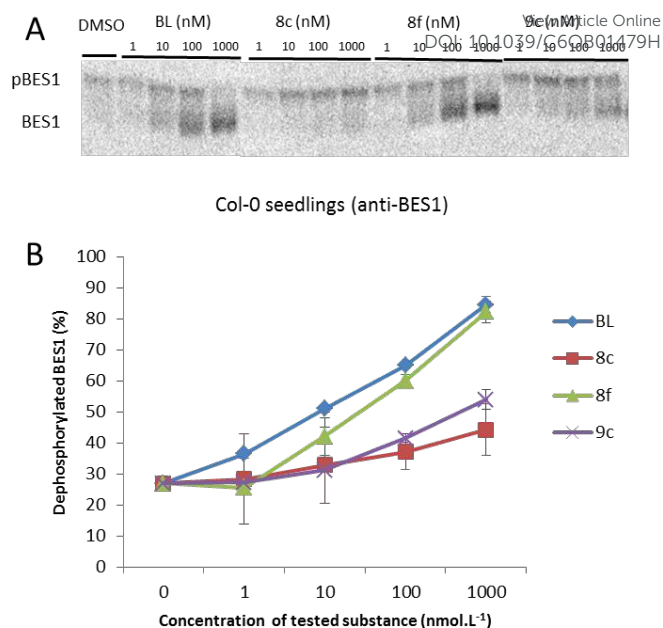


Fig. 6: **A.** Immunoblot analysis of BES1 in *Arabidopsis thaliana* (Col-0) seedlings showing dephosphorylation of BES1 after BRs treatment. **B.** Graph shows percentage of dephosphorylated BES1 relative to the total BES1 detected in Arabidopsis. The data are the average of two biological repeats. Error bars indicate s.e.m. pBES1, phosphorylated BES1.

Experimental

General methods

The melting points were determined on a Stuart SMP30 instrument (Bibby Scientific Ltd., UK). Elemental analyses were performed using an EA 1108 elemental analyzer (Fison Instruments); the values (C, H, N) agreed with the calculated values within acceptable limits. The infrared spectra were recorded on a Thermo Scientific Nicolet spectrometer iZ10 using the ATR technique. The wave numbers are given in cm^{-1} . The NMR spectra were taken on a JEOL JNM-ECA 500 (JEOL, Tokyo, Japan; ^1H , 500 MHz; ^{13}C , 125 MHz) spectrometer equipped with a 5 mm JEOL Royal probe. ^1H NMR and ^{13}C NMR chemical shifts (δ) were calibrated using tetramethylsilane (TMS, ^1H $\delta = 0$ ppm) or solvents: CDCl_3 (^1H $\delta = 7.26$ ppm, ^{13}C $\delta = 77.00$ ppm) or DMSO-d_6 (^1H $\delta = 2.46$ ppm, ^{13}C $\delta = 40.00$ ppm). Chemical shifts are given in ppm (δ -scale), coupling constants (J) in Hz. All values were obtained by first-order analysis. For API HRMS analysis, the samples were dissolved in chloroform (or chloroform : methanol; 1:1; v/v, in case of hydroxylated compounds) to a concentration $10 \mu\text{g.mL}^{-1}$. The ASAP (Atmospheric Solids Analysis Probe) was dipped into the sample solution, placed into the ion source and analysed in fullscan mode. The source of the Synapt G2-Si Mass Spectrometer (Waters, Manchester, UK) was operated in positive ionisation mode (ASAP+), if not stated otherwise, at source temperature of 120°C . The Corona needle current was kept at $5 \mu\text{A}$ and the collision energy at value 4. The probe temperature was ramped up from 50°C to 600°C in 3 minutes. Data were acquired from 50 to 1000 Da with 1.0 s scan time in High Resolution Mode. The data were processed using the Masslynx 4.1 software (Waters). Mass accuracy of 1 ppm or less was achieved with the described instrumentation for all compounds. Merck silica gel Kieselgel 60 (230–400 mesh) was

used for column chromatography. The HPLC system consisted of a Waters semi-preparative HPLC system including quaternary pump, liquid handler, UV-VIS and ELSD detectors. The semi preparative column was filled with silica gel. Reagents and solvents were purchased from Sigma–Aldrich and were not purified. For experimental procedures and data for compounds of series b, d, e, and f see supplementary information.

(22E)-6,6-Ethylenedioxy-23-phenyl-24-nor-5 α -chola-2,22-diene (8a)

To a suspension of benzyl triphenylphosphonium chloride (0.81 mmol) in dried THF (10 mL) was added *n*-BuLi (1.6 M solution in *n*-hexane, 0.81 mmol) at 0 °C and stirred for 1 h. To the resultant red solution was added a solution of aldehyde **7** (200 mg, 0.54 mmol) in THF (10 mL) and stirred for 4 h at 25 °C. The reaction mixture was quenched by water and extracted with Et₂O (2 × 10 mL). The combined organic fractions were washed with brine and dried over anhydrous magnesium sulfate. Evaporation of the volatiles under reduced pressure followed by column chromatography on silica gel (Et₂O/cyclohexane - 1/19) gave 211 mg (88 %) of the title compound **8a** as a colorless oil: IR ν (cm⁻¹) 2933, 1655, 1598. ¹H NMR (CDCl₃) δ 0.74, 0.89 (both s, 3H, CH₃), 1.11 (d, 3H, J = 6.7 Hz, CH₃), 1.68-1.74 (m, 2H), 1.78 (m, 2H), 1.93-2.03 (m, 3H), 2.09 (m, 1H), 2.26 (m, 1H, Σ J = 37.6 Hz), 3.78 (m, 1H, Σ J = 24.1 Hz, OCH), 3.88-3.99 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 6.06 (dd, 1H, J = 15.7, J' = 8.7 Hz, H-22), 6.29 (d, 1H, J = 15.7 Hz, H-23), 7.17 (m, 1H, Ar-H), 7.25-7.34 (m, 4H, 4×Ar-H). ¹³C NMR δ 12.25 (CH₃), 13.60 (CH₃), 20.41 (CH₃), 20.87 (CH₂), 21.43 (CH₂), 24.15 (CH₂), 28.36 (CH₂), 33.35 (CH), 35.91 (C), 39.68 (CH₂), 40.49 (CH), 41.20 (CH₂), 41.24 (CH₂), 42.65 (C), 48.08 (CH), 53.50 (CH), 55.93 (CH), 56.00 (CH), 64.08 (CH₂), 65.56 (CH₂), 110.02 (C), 124.80 (CH), 125.70 (CH), 125.90 (2×CH), 126.64 (CH), 127.19 (CH), 128.43 (2×CH), 137.30 (CH), 138.08 (C). HRMS: (API+) calculated for C₃₁H₄₃O₂ ([M+H]⁺) 447.3263, Found 447.3266.

General procedure for Wadsworth-Horner-Emmons (WHE) reaction

To a suspension of substituted diethyl phenylphosphonate (1.08 mmol) in dried THF (10 mL) was added *n*-BuLi (500 μ L, 1.6 M solution in *n*-hexane, 0.81 mmol) at 0 °C and stirred for 1 h. To the resultant yellow solution was added a solution of aldehyde **7** (200 mg, 0.54 mmol) in THF (10 mL) and stirred for 4 h at room temperature. The reaction mixture was quenched by water and extracted with Et₂O (2 × 10 mL). The combined organic fractions were washed with brine and dried over anhydrous magnesium sulfate. Evaporation of the volatiles under reduced pressure followed by column chromatography on silica gel gave the desired product.

(22E)-6,6-Ethylenedioxy-23-(4-fluorophenyl)-24-nor-5 α -chola-2,22-diene (9a)

The general procedure for WHE reaction with diethyl 4-fluorophenylphosphonate and chromatography on silica (Et₂O/cyclohexane - 1/19) yielded 221 mg (89 %) of the title compound **9a** as a colorless oil: IR ν (cm⁻¹) 2933, 1654, 1599. ¹H NMR (CDCl₃) δ 0.74, 0.89 (both s, 3H, CH₃), 1.11 (d, 3H, J = 6.7 Hz, CH₃), 1.71 (m, 2H), 1.78 (m, 2H), 1.92-2.03 (m, 3H), 2.09 (m, 1H), 2.23 (m, 1H, Σ J = 37.6 Hz), 3.77 (q, 1H, J = 6.8 Hz, OCH), 3.87-3.99 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 5.97 (dd, 1H, J = 15.6, J' = 8.9 Hz, H-22), 6.25 (d, 1H, J = 15.6 Hz, H-23), 6.94-6.98 (m, 2H, 2×Ar-H), 7.25-7.28 (m, 2H, 2×Ar-H). ¹³C NMR δ 12.21 (CH₃), 13.57 (CH₃), 20.36 (CH₃), 20.86 (CH₂), 21.42 (CH₂), 24.12 (CH₂), 28.37

(CH₂), 33.34 (CH), 35.90 (C), 39.68 (CH₂), 40.45 (CH), 41.20 (CH₂), 41.24 (CH₂), 42.63 (C), 48.08 (CH), 53.50 (CH), 55.92 (CH), 56.00 (CH), 64.07 (CH₂), 65.55 (CH₂), 109.99 (C), 115.22 (d, J = 21.6 Hz, 2×CH), 124.77 (CH), 125.69 (CH), 126.04 (CH), 127.25 (d, J = 7.2 Hz, 2×CH), 134.18 (d, J = 3.6 Hz, C), 137.01 (d, J = 2.4 Hz, CH), 161.77 (d, J = 244.7 Hz, C). ¹⁹F NMR {¹H} δ -116.36 (s, 1F). HRMS: (API+) calculated for C₃₁H₄₂FO₂ ([M+H]⁺) 465.3169, Found 465.3170.

(22E)-6,6-Ethylenedioxy-23-(4-chlorophenyl)-24-nor-5 α -chola-2,22-diene (10a)

The general procedure for WHE reaction with diethyl 4-chlorophenylphosphonate and chromatography on silica (Et₂O/cyclohexane - 1/19) yielded 225 mg (87 %) of the title compound **10a** as a colorless oil: IR ν (cm⁻¹) 2933, 1656, 1593. ¹H NMR (CDCl₃) δ 0.74, 0.89 (both s, 3H, CH₃), 1.11 (d, 3H, J = 6.7 Hz, CH₃), 1.71 (m, 2H), 1.78 (m, 2H), 1.92-2.03 (m, 3H), 2.09 (m, 1H), 2.25 (m, 1H, Σ J = 37.6 Hz), 3.78 (q, 1H, J = 6.8 Hz, OCH), 3.88-4.00 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 6.04 (dd, 1H, J = 15.7, J' = 8.9 Hz, H-22), 6.25 (d, 1H, J = 15.7 Hz, H-23), 7.24 (m, 2H, 2×Ar-H), 7.26 (m, 2H, 2×Ar-H). ¹³C NMR δ 12.24 (CH₃), 13.60 (CH₃), 20.30 (CH₃), 20.86 (CH₂), 21.43 (CH₂), 24.14 (CH₂), 28.36 (CH₂), 33.35 (CH), 35.91 (C), 39.68 (CH₂), 40.50 (CH), 41.21 (CH₂), 41.24 (CH₂), 42.67 (C), 48.08 (CH), 53.48 (CH), 55.85 (CH), 55.98 (CH), 64.09 (CH₂), 65.58 (CH₂), 110.01 (C), 124.79 (CH), 125.70 (CH), 126.06 (CH), 127.11 (2×CH), 128.53 (2×CH), 132.13 (C), 136.56 (C), 138.02 (CH). HRMS: (API+) calculated for C₃₁H₄₂ClO₂ ([M+H]⁺) 481.2873, Found 481.2878.

(22E)-6,6-Ethylenedioxy-23-(4-bromophenyl)-24-nor-5 α -chola-2,22-diene (11a)

General procedure for WHE reaction with diethyl 4-bromophenylphosphonate and chromatography on silica (Et₂O/cyclohexane - 1/19) afforded 231 mg (82 %) of the title compound **11a** as a colorless oil: IR ν (cm⁻¹) 2933, 1656, 1595. ¹H NMR (CDCl₃) δ 0.74, 0.89 (both s, 3H, CH₃), 1.11 (d, 3H, J = 6.7 Hz, CH₃), 1.71 (m, 2H), 1.78 (m, 2H), 1.92-2.03 (m, 3H), 2.09 (m, 1H), 2.25 (m, 1H, Σ J = 37.6 Hz), 3.78 (q, 1H, J = 6.8 Hz, OCH), 3.88-4.00 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 6.05 (dd, 1H, J = 15.7, J' = 8.7 Hz, H-22), 6.23 (d, 1H, J = 15.7 Hz, H-23), 7.18 (m, 2H, 2×Ar-H), 7.39 (m, 2H, 2×Ar-H). ¹³C NMR δ 12.23 (CH₃), 13.59 (CH₃), 20.25 (CH₃), 20.85 (CH₂), 21.42 (CH₂), 24.13 (CH₂), 28.35 (CH₂), 33.34 (CH), 35.90 (C), 39.68 (CH₂), 40.51 (CH), 41.20 (CH₂), 41.23 (CH₂), 42.67 (C), 48.07 (CH), 53.47 (CH), 55.82 (CH), 55.97 (CH), 64.08 (CH₂), 65.57 (CH₂), 110.00 (C), 120.21 (C), 124.78 (CH), 125.69 (CH), 126.10 (CH), 127.46 (2×CH), 131.46 (2×CH), 137.01 (C), 138.15 (CH). HRMS: (API+) calculated for C₃₁H₄₂⁷⁹BrO₂ ([M+H]⁺) 525.2368, Found 525.2370.

(22E)-6,6-Ethylenedioxy-23-(4-iodophenyl)-24-nor-5 α -chola-2,22-diene (12a)

General procedure for WHE reaction with diethyl 4-iodophenylphosphonate and chromatography on silica (Et₂O/cyclohexane - 1/19) afforded 200 mg (65 %) of the title compound **12a** as a colorless oil: IR ν (cm⁻¹) 2933, 1656, 1595. ¹H NMR (CDCl₃) δ 0.74, 0.89 (both s, 3H, CH₃), 1.11 (d, 3H, J = 6.7 Hz, CH₃), 1.71 (m, 2H), 1.78 (m, 2H), 1.92-2.03 (m, 3H), 2.09 (m, 1H), 2.24 (m, 1H, Σ J = 37.1 Hz), 3.78 (q, 1H, J = 6.8 Hz, OCH), 3.88-4.00 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 6.06 (dd, 1H, J = 15.7, J' = 8.7 Hz, H-22), 6.21 (d, 1H, J = 15.7 Hz, H-23), 7.06 (m, 2H, 2×Ar-H), 7.59 (m, 2H, 2×Ar-H). ¹³C NMR δ 12.23 (CH₃), 13.59 (CH₃), 20.24 (CH₃), 20.85 (CH₂), 21.42 (CH₂), 24.13 (CH₂), 28.33 (CH₂), 33.34 (CH), 35.90 (C), 39.67 (CH₂), 40.50 (CH), 41.20 (CH₂), 41.23 (CH₂),

42.67 (C), 48.07 (CH), 53.47 (CH), 55.81 (CH), 55.96 (CH), 64.08 (CH₂), 65.57 (CH₂), 91.52 (C), 110.00 (C), 124.78 (CH), 125.70 (CH), 126.20 (CH), 127.75 (2×CH), 137.43 (2×CH), 137.60 (C), 138.29 (CH). HRMS: (API+) calculated for C₃₁H₄₂O₂ ([M+H]⁺) 573.2229, Found 573.2230.

(22E)-6,6-Ethylenedioxy-23-(4-nitrophenyl)-24-nor-5α-chola-2,22-diene (13a)

General procedure for WHE reaction with diethyl 4-nitrophenylphosphonate and chromatography on silica (Et₂O/cyclohexane - 1/19) afforded 230 mg (87 %) of the title compound **13a** as a colorless oil: IR ν (cm⁻¹) 2930, 1646, 1594, 1510, 1346. ¹H NMR (CDCl₃) δ 0.76, 0.89 (both s, 3H, CH₃), 1.15 (d, 3H, J = 6.5 Hz, CH₃), 1.72 (m, 2H), 1.78 (m, 2H), 1.92-2.04 (m, 3H), 2.09 (m, 1H), 2.32 (m, 1H, ΣJ = 37.0 Hz), 3.78 (q, 1H, J = 6.8 Hz, OCH), 3.88-4.00 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 6.28 (dd, 1H, J = 15.7, J' = 8.6 Hz, H-22), 6.37 (d, 1H, J = 15.7 Hz, H-23), 7.43 (m, 2H, 2×Ar-H), 8.15 (m, 2H, 2×Ar-H). ¹³C NMR δ 12.26 (CH₃), 13.59 (CH₃), 20.02 (CH₃), 20.85 (CH₂), 21.42 (CH₂), 24.14 (CH₂), 28.32 (CH₂), 33.34 (CH), 35.90 (C), 39.68 (CH₂), 40.75 (CH), 41.21 (CH₂), 41.23 (CH₂), 42.79 (C), 48.08 (CH), 53.46 (CH), 55.63 (CH), 55.92 (CH), 64.09 (CH₂), 65.58 (CH₂), 109.97 (C), 123.94 (2×CH), 124.74 (CH), 125.70 (2×CH), 126.31 (2×CH), 142.50 (CH), 144.67 (C), 146.31 (C). HRMS: (API+) calculated for C₃₁H₄₂NO₄ ([M+H]⁺) 492.3114, Found 492.3118.

(22E)-6,6-Ethylenedioxy-23-(4-methylphenyl)-24-nor-5α-chola-2,22-diene (14a)

General procedure for WHE reaction with diethyl 4-methylphenylphosphonate and chromatography on silica (Et₂O/cyclohexane - 1/19) afforded 200 mg (81 %) of the title compound **14a** as a colorless oil: IR ν (cm⁻¹) 2934, 1656, 1595. ¹H NMR (CDCl₃) δ 0.74, 0.89 (both s, 3H, CH₃), 1.11 (d, 3H, J = 6.7 Hz, CH₃), 1.71 (m, 2H), 1.77 (m, 2H), 1.92-2.03 (m, 3H), 2.09 (m, 1H), 2.23 (m, 1H, ΣJ = 37.2 Hz), 2.31 (s, 3H, CH₃), 3.77 (q, 1H, J = 6.5 Hz, OCH), 3.87-3.98 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 6.00 (dd, 1H, J = 15.6, J' = 8.9 Hz, H-22), 6.26 (d, 1H, J = 15.6 Hz, H-23), 7.08 (d, 2H, J = 7.9 Hz, 2×Ar-H), 7.21 (d, 2H, J = 7.9 Hz, 2×Ar-H). ¹³C NMR δ 12.20 (CH₃), 13.56 (CH₃), 20.42 (CH₃), 20.84 (CH₂), 21.07 (CH₃), 21.40 (CH₂), 24.11 (CH₂), 28.32 (CH₂), 33.32 (CH), 35.87 (C), 39.66 (CH₂), 40.42 (CH), 41.16 (CH₂), 41.21 (CH₂), 42.58 (C), 48.04 (CH), 53.47 (CH), 55.96 (2×CH), 64.03 (CH₂), 65.53 (CH₂), 109.97 (C), 124.77 (CH), 125.67 (CH), 125.76 (2×CH), 126.96 (CH), 129.08 (2×CH), 135.24 (C), 136.25 (CH), one aromatic C not detected. HRMS: (API+) calculated for C₃₂H₄₅O₂ ([M+H]⁺) 461.3420, Found 461.3422.

(22E)-6,6-Ethylenedioxy-23-(4-methoxyphenyl)-24-nor-5α-chola-2,22-diene (15a)

General procedure for WHE reaction with diethyl 4-methoxyphenylphosphonate and chromatography on silica (Et₂O/cyclohexane - 1/19) afforded 203 mg (80 %) of the title compound **15a** as a colorless oil: IR ν (cm⁻¹) 2934, 1656, 1594. ¹H NMR (CDCl₃) δ 0.74, 0.89 (both s, 3H, CH₃), 1.11 (d, 3H, J = 6.7 Hz, CH₃), 1.71 (m, 2H), 1.78 (m, 2H), 1.92-2.03 (m, 3H), 2.09 (m, 1H), 2.22 (m, 1H, ΣJ = 37.4 Hz), 3.77 (q, 1H, J = 6.7 Hz, OCH), 3.79 (s, 3H, OCH₃), 3.87-3.99 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 5.91 (dd, 1H, J = 15.7, J' = 8.7 Hz, H-22), 6.24 (d, 1H, J = 15.7 Hz, H-23), 6.82 (m, 2H, 2×Ar-H), 7.25 (m, 2H, 2×Ar-H). ¹³C NMR δ 12.21 (CH₃), 13.58 (CH₃), 20.50 (CH₃), 20.85 (CH₂), 21.41 (CH₂), 24.12 (CH₂), 28.38 (CH₂), 33.33 (CH), 35.88 (C), 39.67 (CH₂), 40.43 (CH), 41.18 (CH₂), 41.22 (CH₂), 42.58 (C), 48.06 (CH), 53.48 (CH), 55.24

(CH₃), 55.99 (CH), 56.02 (CH), 64.05 (CH₂), 65.54 (CH₂), 109.99 (C), 113.83 (2×CH), 124.78 (CH), 125.68 (CH), 126.48 (CH), 126.92 (2×CH), 130.87 (C), 135.23 (CH), 158.48 (C). HRMS: (API+) calculated for C₃₂H₄₅O₃ ([M+H]⁺) 477.3369, Found 477.3371.

(22E)-6,6-Ethylenedioxy-23-(4-isopropylphenyl)-24-nor-5α-chola-2,22-diene (16a)

The general procedure for WHE reaction with diethyl 4-isopropylphenylphosphonate and chromatography on silica (Et₂O/cyclohexane - 1/19) afforded 231 mg (88 %) of the title compound **16a** as a colorless oil: IR ν (cm⁻¹) 2935, 1656, 1593. ¹H NMR (CDCl₃) δ 0.74, 0.89 (both s, 3H, CH₃), 1.11 (d, 3H, J = 6.7 Hz, CH₃), 1.229, 1.231 (both d, 3H, J = 7.0 Hz, CH₃), 1.71 (m, 2H), 1.78 (m, 2H), 1.93-2.03 (m, 3H), 2.09 (m, 1H), 2.24 (m, 1H, ΣJ = 37.6 Hz), 2.87 (septet, 1H, J = 7.0 Hz, CH(CH₃)₂), 3.78 (q, 1H, J = 6.8 Hz, OCH), 3.88-4.00 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 6.01 (dd, 1H, J = 15.8, J' = 8.6 Hz, H-22), 6.27 (d, 1H, J = 15.8 Hz, H-23), 7.15 (d, 2H, J = 8.3 Hz, 2×Ar-H), 7.26 (d, 2H, J = 8.3 Hz, 2×Ar-H). ¹³C NMR δ 12.24 (CH₃), 13.60 (CH₃), 20.49 (CH₃), 20.86 (CH₂), 21.43 (CH₂), 23.97 (2×CH₃), 24.14 (CH₂), 28.35 (CH₂), 33.35 (CH), 33.79 (CH), 35.91 (C), 39.68 (CH₂), 40.50 (CH), 41.19 (CH₂), 41.24 (CH₂), 42.61 (C), 48.07 (CH), 53.49 (CH), 55.97 (CH), 56.00 (CH), 64.07 (CH₂), 65.56 (CH₂), 110.02 (C), 124.81 (CH), 125.69 (CH), 125.86 (2×CH), 126.50 (2×CH), 126.97 (CH), 135.69 (C), 136.44 (CH), 147.43 (C). HRMS: (API+) calculated for C₃₄H₄₉O₂ ([M+H]⁺) 489.3733, Found 489.3736.

(22E)-6,6-Ethylenedioxy-23-(4-cyanophenyl)-24-nor-5α-chola-2,22-diene (17a)

The general procedure for WHE reaction with diethyl 4-cyanophenylphosphonate and chromatography on silica (Et₂O/cyclohexane - 1/19) afforded 220 mg (87 %) of the title compound **17a** as an amorphous solid: IR ν (cm⁻¹) 2933, 2224, 1646, 1604. ¹H NMR (CDCl₃) δ 0.75, 0.89 (both s, 3H, CH₃), 1.13 (d, 3H, J = 6.7 Hz, CH₃), 1.71 (m, 2H), 1.78 (m, 2H), 1.93-2.03 (m, 3H), 2.09 (m, 1H), 2.30 (m, 1H, ΣJ = 37.2 Hz), 3.78 (q, 1H, J = 6.8 Hz, OCH), 3.88-4.00 (m, 3H, 3×OCH), 5.54 (m, 1H, H-3), 5.66 (m, 1H, H-2), 6.21 (dd, 1H, J = 15.7, J' = 8.6 Hz, H-22), 6.31 (d, 1H, J = 15.7 Hz, H-23), 7.39 (m, 2H, 2×Ar-H), 7.56 (m, 2H, 2×Ar-H). ¹³C NMR δ 12.25 (CH₃), 13.59 (CH₃), 20.07 (CH₃), 20.84 (CH₂), 21.42 (CH₂), 24.13 (CH₂), 28.32 (CH₂), 33.34 (CH), 35.90 (C), 39.67 (CH₂), 40.66 (CH), 41.20 (CH₂), 41.22 (CH₂), 42.76 (C), 48.08 (CH), 53.45 (CH), 55.65 (CH), 55.92 (CH), 64.09 (CH₂), 65.58 (CH₂), 109.77 (C), 109.97 (C), 119.21 (C), 124.75 (CH), 125.70 (CH), 125.99 (CH), 126.36 (2×CH), 132.30 (2×CH), 141.46 (CH), 142.62 (C). HRMS: (API+) calculated for C₃₂H₄₂NO₂ ([M+H]⁺) 472.3216, Found 472.3218.

General procedure for dihydroxylation of dienes

To a solution of diene (160 mg), hydroquinidine 4-chlorobenzoate (45 mg; 0.097 mmol), methansulfonamide (65 mg; 0.68 mmol), potassium carbonate (280 mg; 2.03 mmol), and potassium ferricyanide (700 mg; 2.13 mmol) in the mixture of *t*-butanol and water (15 mL; 1:1 v/v) was added 0.2 mL of osmium tetroxide in *t*-butanol (1 g/20 mL; 0.039 mmol). Reaction mixture was stirred at room temperature for 24 h. A saturated solution of sodium sulfite (3 mL) was then added. After an additional 30 minutes of stirring, the reaction mixture was diluted with ethyl acetate (30 mL) and extracted with water (2 × 20 mL). The combined organic fractions were dried over anhydrous magnesium sulfate and evaporated under reduced pressure. Column chromatography on silica gel gave the desired product.

(22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-phenyl-24-nor-5 α -cholan-6-one (8c)

The general procedure for dihydroxylation of **8b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 144 mg (77 %) of the title compound **8c** as a white solid: m. p. 268-270 °C (EtOH), IR ν (cm⁻¹) 3344, 2940, 1708, 1496. ¹H NMR (DMSO-d₆) δ 0.30, 0.59 (both s, 3H, CH₃), 0.86 (d, 3H, J = 6.4 Hz, CH₃), 1.80 (m, 1H), 1.86 (m, 1H), 1.97 (dd, 1H, J = 13.1, J' = 4.6 Hz), 2.07 (t, 1H, J = 12.6 Hz), 2.58 (dd, 1H, J = 12.1, J' = 3.2 Hz), 3.44 (m, 1H), 3.48 (dd, 1H, J = 8.4, J' = 4.4 Hz), 3.74 (m, 1H), 4.19 (d, 1H, J = 2.8 Hz, OH), 4.32 (d, 1H, J = 6.1 Hz, OH), 4.35 (dd, 1H, J = 8.6, J' = 3.9 Hz), 4.51 (d, 1H, J = 4.3 Hz, OH), 5.14 (d, 1H, J = 3.9 Hz, OH), 7.22-7.27 (m, 3H), 7.31 (m, 2H). ¹³C NMR δ 11.39 (CH₃), 12.47 (CH₃), 13.36 (CH₃), 20.79 (CH₂), 23.28 (CH₂), 26.82 (CH₂), 27.22 (CH₂), 36.29 (CH), 37.02 (CH), 39.14 (C), 41.84 (C), 42.03 (CH₂), 45.95 (CH₂), 50.28 (CH), 51.88 (CH), 52.85 (CH), 55.95 (CH), 67.10 (CH), 67.49 (CH), 75.16 (CH), 76.30 (CH), 127.01 (2 \times CH), 127.21 (CH), 128.07 (2 \times CH), 143.28 (C), 211.57 (C). One CH₂ covered by DMSO multiplet. HRMS: (API+) calculated for C₂₉H₄₃O₅ ([M+H]⁺) 471.3110, Found 471.3108. Anal. Calcd for C₂₉H₄₂O₅: C, 74.01; H, 8.99. Found: C, 73.95; H, 9.06 %.

(22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-(4-fluorophenyl)-24-nor-5 α -cholan-6-one (9c)

The general procedure for dihydroxylation of **9b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 149 mg (80 %) of the title compound **9c** as a white solid: m. p. 277-279 °C (EtOH), IR ν (cm⁻¹) 3251, 2937, 1709, 1607, 1513. ¹H NMR (DMSO-d₆) δ 0.33, 0.60 (both s, 3H, CH₃), 0.84 (d, 3H, J = 6.7 Hz, CH₃), 1.80 (m, 1H), 1.85 (m, 1H), 1.98 (dd, 1H, J = 13.0, J' = 4.8 Hz), 2.07 (t, 1H, J = 12.6 Hz), 2.58 (dd, 1H, J = 12.2, J' = 3.1 Hz), 3.42-3.47 (m, 2H), 3.74 (m, 1H), 4.19 (d, 1H, J = 2.4 Hz, OH), 4.32 (d, 1H, J = 6.1 Hz, OH), 4.37 (dd, 1H, J = 8.7, J' = 3.5 Hz), 4.54 (d, 1H, J = 4.3 Hz, OH), 5.14 (d, 1H, J = 4.0 Hz, OH), 7.14 (m, 2H), 7.30 (m, 2H). ¹³C NMR δ 11.42 (CH₃), 12.44 (CH₃), 13.38 (CH₃), 20.80 (CH₂), 23.30 (CH₂), 26.84 (CH₂), 27.28 (CH₂), 36.35 (CH), 37.04 (CH), 39.15 (C), 41.86 (C), 42.06 (CH₂), 45.97 (CH₂), 50.31 (CH), 51.89 (CH), 52.88 (CH), 55.97 (CH), 67.11 (CH), 67.51 (CH), 74.40 (CH), 76.34 (CH), 114.84 (d, J = 21.6 Hz, 2 \times CH), 128.86 (d, J = 8.4 Hz, 2 \times CH), 139.58 (d, J = 2.4 Hz, C), 161.26 (d, J = 242.3 Hz, C), 211.62 (C). One CH₂ covered by DMSO multiplet. ¹⁹F NMR {¹H} δ -115.37 (s, 1F). HRMS: (API+) calculated for C₂₉H₄₂FO₅ ([M+H]⁺) 489.3016, Found 489.3017. Anal. Calcd for C₂₉H₄₁FO₅: C, 71.28; H, 8.46. Found: C, 71.88; H, 8.55 %.

(22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-(4-chlorophenyl)-24-nor-5 α -cholan-6-one (10c)

The general procedure for dihydroxylation of **10b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 152 mg (82 %) of the title compound **10c** as a white solid: m. p. 251-253 °C (EtOH), IR ν (cm⁻¹) 3220, 2940, 1712, 1598, 1494. ¹H NMR (DMSO-d₆) δ 0.34, 0.60 (both s, 3H, CH₃), 0.85 (d, 3H, J = 6.7 Hz, CH₃), 1.80 (m, 1H), 1.85 (m, 1H), 1.98 (dd, 1H, J = 13.0, J' = 4.8 Hz), 2.07 (t, 1H, J = 12.6 Hz), 2.58 (dd, 1H, J = 12.2, J' = 3.4 Hz), 3.42-3.48 (m, 2H), 3.74 (m, 1H), 4.19 (d, 1H, J = 2.4 Hz, OH), 4.33 (d, 1H, J = 6.1 Hz, OH), 4.37 (dd, 1H, J = 8.4, J' = 3.8 Hz), 4.58 (d, 1H, J = 4.6 Hz, OH), 5.20 (d, 1H, J = 3.8 Hz, OH), 7.29 (m, 2H), 7.37 (m, 2H). ¹³C NMR δ 11.46 (CH₃), 12.49 (CH₃), 13.38 (CH₃), 20.80 (CH₂), 23.30 (CH₂), 26.84 (CH₂), 27.29 (CH₂), 36.47 (CH), 37.05 (CH), 39.15 (C), 41.87 (C), 42.07 (CH₂), 45.98 (CH₂), 50.31 (CH), 51.89 (CH), 52.89 (CH), 55.97 (CH), 67.12 (CH), 67.51 (CH), 74.44 (CH), 76.21 (CH), 128.10 (2 \times CH), 128.87 (2 \times CH), 131.52 (C), 142.41 (C), 211.62 (C). One CH₂ covered by DMSO multiplet. HRMS: (API+) calculated for C₂₉H₄₂ClO₅ ([M+H]⁺)

505.2721, Found 505.2723. Anal. Calcd for C₂₉H₄₁ClO₅: C, 68.96; H, 8.18. Found: C, 68.90; H, 8.28 %. DOI: 10.1039/C6OB01479H

(22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-(4-bromophenyl)-24-nor-5 α -cholan-6-one (11c)

The general procedure for dihydroxylation of **11b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 146 mg (80 %) of the title compound **11c** as a white solid: m. p. 247-249 °C (*i*-PrOH), IR ν (cm⁻¹) 3237, 2941, 1709, 1600, 1490. ¹H NMR (DMSO-d₆) δ 0.34, 0.60 (both s, 3H, CH₃), 0.84 (d, 3H, J = 6.7 Hz, CH₃), 1.79 (m, 1H), 1.87 (m, 1H), 1.98 (dd, 1H, J = 13.0, J' = 4.6 Hz), 2.07 (t, 1H, J = 12.5 Hz), 2.58 (dd, 1H, J = 12.2, J' = 3.1 Hz), 3.42-3.47 (m, 2H), 3.74 (m, 1H), 4.19 (d, 1H, J = 2.8 Hz, OH), 4.33 (d, 1H, J = 6.1 Hz, OH), 4.36 (dd, 1H, J = 8.6, J' = 3.8 Hz), 4.58 (d, 1H, J = 4.3 Hz, OH), 5.20 (d, 1H, J = 3.8 Hz, OH), 7.23 (d, 2H, J = 8.6 Hz, 2 \times Ar-H), 7.51 (d, 2H, J = 8.6 Hz, 2 \times Ar-H). ¹³C NMR δ 11.49 (CH₃), 12.50 (CH₃), 13.38 (CH₃), 20.81 (CH₂), 23.31 (CH₂), 26.84 (CH₂), 27.30 (CH₂), 36.50 (CH), 37.06 (CH), 39.15 (C), 41.87 (C), 42.08 (CH₂), 45.99 (CH₂), 50.32 (CH), 51.89 (CH), 52.89 (CH), 55.98 (CH), 67.13 (CH), 67.52 (CH), 74.51 (CH), 76.16 (CH), 120.08 (C), 129.26 (2 \times CH), 131.01 (2 \times CH), 142.83 (C), 211.62 (C). One CH₂ covered by DMSO multiplet. HRMS: (API+) calculated for C₂₉H₄₂⁷⁹BrO₅ ([M+H]⁺) 549.2216, Found 549.2216. Anal. Calcd for C₂₉H₄₁BrO₅: C, 63.38; H, 7.52. Found: C, 63.29; H, 7.55 %.

(22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-(4-iodophenyl)-24-nor-5 α -cholan-6-one (12c)

The general procedure for dihydroxylation of **12b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 150 mg (83 %) of the title compound **12c** as a white solid: m. p. 253-255 °C (EtOH), IR ν (cm⁻¹) 3193, 2943, 1710, 1590. ¹H NMR (DMSO-d₆) δ 0.35, 0.60 (both s, 3H, CH₃), 0.84 (d, 3H, J = 6.7 Hz, CH₃), 1.79 (m, 1H), 1.86 (m, 1H), 1.98 (dd, 1H, J = 13.2, J' = 4.9 Hz), 2.07 (t, 1H, J = 12.5 Hz), 2.58 (dd, 1H, J = 12.2, J' = 3.4 Hz), 3.41-3.48 (m, 2H), 3.74 (m, 1H), 4.19 (d, 1H, J = 2.8 Hz, OH), 4.32 (d, 1H, J = 6.1 Hz, OH), 4.34 (dd, 1H, J = 8.6, J' = 3.7 Hz), 4.56 (d, 1H, J = 4.3 Hz, OH), 5.18 (d, 1H, J = 3.8 Hz, OH), 7.09 (m, 2H, 2 \times Ar-H), 7.67 (m, 2H, 2 \times Ar-H). ¹³C NMR δ 11.49 (CH₃), 12.48 (CH₃), 13.36 (CH₃), 20.80 (CH₂), 23.29 (CH₂), 26.83 (CH₂), 27.28 (CH₂), 36.51 (CH), 37.03 (CH), 39.14 (C), 41.85 (C), 42.05 (CH₂), 45.96 (CH₂), 50.30 (CH), 51.87 (CH), 52.86 (CH), 55.95 (CH), 67.11 (CH), 67.49 (CH), 74.60 (CH), 76.07 (CH), 92.95 (C), 129.40 (2 \times CH), 136.83 (2 \times CH), 143.19 (C), 211.57 (C). One CH₂ covered by DMSO multiplet. HRMS: (API+) calculated for C₂₉H₄₂I O₅ ([M+H]⁺) 597.2077, Found 597.2076. Anal. Calcd for C₂₉H₄₁I O₅: C, 58.39; H, 6.93. Found: C, 58.32; H, 7.01 %.

(22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-(4-nitrophenyl)-24-nor-5 α -cholan-6-one (13c)

The general procedure for dihydroxylation of **13b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 144 mg (78 %) of the title compound **13c** as a white solid: m. p. 244-246 °C (EtOH), IR ν (cm⁻¹) 3180, 2940, 1711, 1605, 1523, 1349. ¹H NMR (DMSO-d₆) δ 0.34, 0.59 (both s, 3H, CH₃), 0.87 (d, 3H, J = 6.7 Hz, CH₃), 1.80 (m, 1H), 1.86 (m, 1H), 1.97 (dd, 1H, J = 13.2, J' = 4.6 Hz), 2.06 (t, 1H, J = 12.5 Hz), 2.58 (dd, 1H, J = 12.2, J' = 3.4 Hz), 3.45 (m, 1H), 3.50 (m, 1H), 3.74 (m, 1H), 4.19 (d, 1H, J = 2.8 Hz, OH), 4.33 (d, 1H, J = 6.1 Hz, OH), 4.53 (dd, 1H, J = 7.9, J' = 3.5 Hz), 4.74 (d, 1H, J = 4.6 Hz, OH), 5.47 (d, 1H, J = 3.5 Hz, OH), 7.57 (m, 2H), 8.19 (m, 2H). ¹³C NMR δ 11.46 (CH₃), 12.64 (CH₃), 13.37 (CH₃), 20.80 (CH₂), 23.30 (CH₂), 26.83 (CH₂), 27.32 (CH₂), 36.77 (CH), 37.03 (CH), 39.14 (C), 41.86 (C), 42.09 (CH₂), 45.97 (CH₂), 50.31 (CH), 51.92 (CH), 52.87 (CH), 55.95 (CH), 67.11 (CH), 67.51 (CH), 74.46 (CH), 76.09 (CH), 123.31 (2 \times CH), 128.25 (2 \times CH), 146.62 (C), 151.53 (C), 211.61 (C).

One CH₂ covered by DMSO multiplet. HRMS: (ESI-) calculated for C₂₉H₄₁NO₇ ([M⁻]) 515.2883, Found 515.2888. Anal. Calcd for C₂₉H₄₁NO₇: C, 67.55; H, 8.01. Found: C, 67.45; H, 8.09 %.

(22R, 23R)-2α,3α,22,23-tetrahydroxy-23-(4-methylphenyl)-24-nor-5α-cholan-6-one (14c)

The general procedure for dihydroxylation of **14b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 145 mg (78 %) of the title compound **14c** as a white solid: m. p. 271-272 °C (EtOH), IR ν (cm⁻¹) 3215, 2937, 1710, 1610, 1516. ¹H NMR (DMSO-d₆) δ 0.32, 0.59 (both s, 3H, CH₃), 0.84 (d, 3H, J = 6.7 Hz, CH₃), 1.80 (m, 1H), 1.86 (m, 1H), 1.98 (dd, 1H, J = 13.2, J' = 4.6 Hz), 2.07 (t, 1H, J = 12.5 Hz), 2.27 (s, 3H, CH₃), 2.58 (dd, 1H, J = 12.2, J' = 3.4 Hz), 3.42-3.48 (m, 2H), 3.74 (m, 1H), 4.19 (d, 1H, J = 2.8 Hz, OH), 4.31 (dd, 1H, J = 8.6, J' = 3.8 Hz), 4.34 (d, 1H, J = 6.1 Hz, OH), 4.46 (d, 1H, J = 4.4 Hz, OH), 5.00 (d, 1H, J = 3.8 Hz, OH), 7.10-7.15 (m, 4H). ¹³C NMR δ 11.50 (CH₃), 12.44 (CH₃), 13.40 (CH₃), 20.83 (CH₂, CH₃), 23.32 (CH₂), 26.85 (CH₂), 27.28 (CH₂), 36.38 (CH), 37.07 (CH), 39.15 (C), 41.89 (C), 42.07 (CH₂), 46.00 (CH₂), 50.33 (CH), 51.90 (CH), 52.91 (CH), 56.00 (CH), 67.14 (CH), 67.52 (CH), 74.96 (CH), 76.36 (CH), 126.99 (2×CH), 128.71 (2×CH), 136.15 (C), 140.20 (C), 211.68 (C). One CH₂ covered by DMSO multiplet. HRMS: (API+) calculated for C₃₀H₄₅O₅ ([M+H]⁺) 485.3267, Found 485.3270. Anal. Calcd for C₃₀H₄₅O₅: C, 74.34; H, 9.15. Found: C, 74.30; H, 9.20 %.

(22R, 23R)-2α,3α,22,23-tetrahydroxy-23-(4-methoxyphenyl)-24-nor-5α-cholan-6-one (15c)

The general procedure for dihydroxylation of **15b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 150 mg (81 %) of the title compound **15c** as a white solid: m. p. 242-244 °C (EtOH), IR ν (cm⁻¹) 3197, 2934, 1710, 1616, 1513. ¹H NMR (DMSO-d₆) δ 0.33, 0.60 (both s, 3H, CH₃), 0.84 (d, 3H, J = 6.7 Hz, CH₃), 1.80 (m, 1H), 1.86 (m, 1H), 1.98 (dd, 1H, J = 13.2, J' = 4.6 Hz), 2.07 (t, 1H, J = 12.7 Hz), 2.58 (dd, 1H, J = 12.2, J' = 3.4 Hz), 3.42-3.49 (m, 2H), 3.72 (s, 3H, CH₃), 3.74 (m, 1H), 4.18 (d, 1H, J = 2.5 Hz, OH), 4.30 (dd, 1H, J = 8.6, J' = 3.8 Hz), 4.32 (d, 1H, J = 6.4 Hz, OH), 4.44 (d, 1H, J = 4.1 Hz, OH), 4.96 (d, 1H, J = 3.8 Hz, OH), 6.87 (m, 2H), 7.17 (m, 2H). ¹³C NMR δ 11.48 (CH₃), 12.38 (CH₃), 13.37 (CH₃), 20.80 (CH₂), 23.30 (CH₂), 26.83 (CH₂), 27.26 (CH₂), 36.31 (CH), 37.04 (CH), 39.14 (C), 41.86 (C), 42.04 (CH₂), 45.97 (CH₂), 50.30 (CH), 51.89 (CH), 52.88 (CH), 54.97 (CH₃), 55.98 (CH), 67.11 (CH), 67.50 (CH), 74.59 (CH), 76.29 (CH), 113.43 (2×CH), 128.12 (2×CH), 135.17 (C), 158.29 (C), 211.59 (C). One CH₂ covered by DMSO multiplet. HRMS: (API+) calculated for C₃₀H₄₅O₆ ([M+H]⁺) 501.3218, Found 501.3216. Anal. Calcd for C₃₀H₄₅O₆: C, 71.97; H, 8.86. Found: C, 71.91; H, 8.97 %.

(22R, 23R)-2α,3α,22,23-tetrahydroxy-23-(4-isopropylphenyl)-24-nor-5α-cholan-6-one (16c)

The general procedure for dihydroxylation of **16b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 146 mg (79 %) of the title compound **16c** as a white solid: m. p. 247-249 °C (EtOH), IR ν (cm⁻¹) 3215, 2938, 1710, 1617, 1513. ¹H NMR (DMSO-d₆) δ 0.31, 0.59 (both s, 3H, CH₃), 0.84 (d, 3H, J = 6.7 Hz, CH₃), 1.17 (d, 6H, J = 6.9 Hz, 2×CH₃), 1.77-1.86 (m, 2H), 1.98 (dd, 1H, J = 13.2, J' = 4.6 Hz), 2.07 (t, 1H, J = 12.5 Hz), 2.58 (dd, 1H, J = 12.1, J' = 3.2 Hz), 2.86 (septet, 1H, J = 6.9 Hz), 3.44 (m, 1H), 3.50 (dt, 1H, J = 8.3, J' = 4.1 Hz), 3.74 (m, 1H), 4.19 (d, 1H, J = 2.8 Hz, OH), 4.31 (dd, 1H, J = 8.6, J' = 3.8 Hz), 4.33 (d, 1H, J = 6.1 Hz, OH), 4.44 (d, 1H, J = 4.0 Hz, OH), 5.00 (d, 1H, J = 3.8 Hz, OH), 7.17 (s, 4H). ¹³C NMR δ 11.46 (CH₃), 12.50 (CH₃), 13.37 (CH₃), 20.80 (CH₂), 23.32 (CH₂), 23.92 (CH₃), 24.02 (CH₃), 26.85 (CH₂), 27.20 (CH₂), 33.11 (CH), 36.33 (CH), 37.05 (CH), 39.14 (C), 41.87 (C), 42.07 (CH₂), 45.98 (CH₂), 50.32 (CH), 51.97

(CH), 52.89 (CH), 55.96 (CH), 67.12 (CH), 67.52 (CH), 75.05 (CH), 76.18 (CH), 125.97 (2×CH), 127.04 (2×CH), 140.66 (C), 147.21 (C), 211.62 (C). One CH₂ covered by DMSO multiplet. HRMS: (API+) calculated for C₃₂H₄₉O₅ ([M+H]⁺) 513.3580, Found 513.3585. Anal. Calcd for C₃₂H₄₉O₅: C, 74.96; H, 9.44. Found: C, 74.89; H, 9.51 %.

(22R, 23R)-2α,3α,22,23-tetrahydroxy-23-(4-cyanophenyl)-24-nor-5α-cholan-6-one (17c)

The general procedure for dihydroxylation of **17b** and chromatography on silica (MeOH/CHCl₃ - 1/16) afforded 139 mg (75 %) of the title compound **17c** as a white solid: m. p. 260-262 °C (EtOH), IR ν (cm⁻¹) 3198, 2936, 2230, 1706, 1611, 1510. ¹H NMR (DMSO-d₆) δ 0.33, 0.59 (both s, 3H, CH₃), 0.85 (d, 3H, J = 6.7 Hz, CH₃), 1.74-1.88 (m, 2H), 1.98 (dd, 1H, J = 13.2, J' = 4.8 Hz), 2.06 (t, 1H, J = 12.5 Hz), 2.57 (dd, 1H, J = 12.2, J' = 3.4 Hz), 3.46-3.50 (m, 2H), 3.74 (m, 1H), 4.20 (d, 1H, J = 2.4 Hz, OH), 4.34 (d, 1H, J = 6.1 Hz, OH), 4.46 (dd, 1H, J = 8.1, J' = 4.3 Hz), 4.69 (d, 1H, J = 4.3 Hz, OH), 5.04 (d, 1H, J = 4.0 Hz, OH), 7.48 (d, 2H, J = 8.3 Hz, 2×Ar-H), 7.79 (d, 2H, J = 8.3 Hz, 2×Ar-H). ¹³C NMR δ 11.47 (CH₃), 12.64 (CH₃), 13.41 (CH₃), 20.83 (CH₂), 23.33 (CH₂), 26.87 (CH₂), 27.33 (CH₂), 36.71 (CH), 37.08 (CH), 39.14 (C), 41.89 (C), 42.11 (CH₂), 46.00 (CH₂), 50.35 (CH), 51.94 (CH), 52.91 (CH), 55.98 (CH), 67.15 (CH), 67.54 (CH), 74.74 (CH), 76.12 (CH), 109.86 (C), 119.03 (C), 128.05 (2×CH), 132.14 (2×CH), 149.34 (C), 211.68 (C). One CH₂ covered by DMSO multiplet. HRMS: (API+) calculated for C₃₀H₄₂NO₅ ([M+H]⁺) 496.3063, Found 496.3064. Anal. Calcd for C₃₀H₄₁NO₅: C, 72.70; H, 8.34. Found: C, 72.69; H, 8.39 %.

Molecular docking

Docking was performed to obtain prediction of conformation and energy ranking between BRI1 receptor (PDB ID: 3RGZ) and the steroid molecule. The docking studies were carried out using AutoDock Vina 1.05.³¹ All 3D structures of BRI1 ligands were obtained with Marvin 5.10.3³², software which can be used for drawing, displaying and characterization of chemical structure, substructures and reactions. Ligands were prepared as derivatives of natural ligand brassinolide (BLD). Polar hydrogens were added to all ligands and proteins with the AutoDock Tools (ADT)¹⁶ program prior to docking with Autodock Vina program. Grid box with size of 40 Å were centered on active site of protein. The exhaustiveness parameter was set to 20 (default 8). After docking, we compared the docked ligand with brassinolide crystal-like poses and the best crystal-like poses of each ligand were analyzed.

The pea inhibition biotest

Pea seedlings (*Pisum arvense* L. sort Arvica) germinating for 2 days were selected for uniformity from a large population and then transferred into pots containing perlite and 1/10 diluted Hoagland solution (half concentration, pH 5.7) After 24 h in a dark cultivation room (24 °C, humidity 75%) the seedlings were treated with different amounts of tested compounds in 5 μl fractionated lanolin. The substances were applied as microdrops to the scar left after the removal of bract. The control plants were treated with lanolin alone. At least seven plants were used for each experiment and the assays were repeated at least three times. The inhibition of etiolated pea stems were measured after 4 days and the difference in length between treated and control plants provided a measure of activity. For each treatment, 8 seedlings were analyzed in two biological replicates. The mean values were subjected to the statistical analysis using the Student's t test.

Determination of ethylene production

For measurement of ethylene production, pea seedlings (8 plants/tested amount of substance) were placed in a 0.5 L glass container for 24 h in the dark. One milliliter of headspace gas from the chamber was withdrawn for each measurement and injected into a gas chromatograph (Agilent Technologies, GC System) equipped with a flame ionic detector (FID) and a capillary column (HP-AL/S stationary phase, 15 μm , i.d. = 0.535). The chromatographic analytical parameters were as follows: column temperature: 150 $^{\circ}\text{C}$; detector temperature: 220 $^{\circ}\text{C}$; and helium was used as carrier gas. The area under the resultant peak (y -axis) versus sensitivity (x -axis; $\text{nL}\cdot\text{mL}^{-1}$) was representing a quantitative measure of ethylene concentration. The measurements were done in triplicates and data were statistically analyzed using the Student's t test.

Arabidopsis brassinosteroid sensitivity assays

Arabidopsis thaliana L. (Heyhn.) (Columbia ecotype, Col-0; referred to Arabidopsis) seedlings were stratified for 2 d at 4 $^{\circ}\text{C}$ and germinated on vertical half-strength Murashige and Skoog (1% w/v sucrose) agar plates with different concentrations of BL (Fuji Chemical Industries) and BR derivatives at 22 $^{\circ}\text{C}$ in a 16 h/8 h light-dark cycle for 5 d. For the hypocotyl assay, after stratification, plants were exposed to light for 6 h and grown in dark for 5 days. Roots and hypocotyls were then straightened on solid media plates, scanned with an Epson high-resolution scanner and the entire root and hypocotyl length measured with ImageJ (<http://rsbweb.nih.gov/ij/>). For each treatment, more than 25 seedlings were analyzed in two biological repeats. P values were calculated with a two-tailed Student t -test using Excel software.

BES-1 dephosphorylation assay

For BES1 dephosphorylation studies, thirty to sixty 5-day-old Arabidopsis seedlings grown on BL and new BR derivatives in continuous light were used. DMSO was used as the control solvent. The protein extraction and Western blot analysis were carried out as previously described.³³ Endogenous BES1 was detected using rabbit polyclonal anti-BES1 antibodies (1:1000)³⁴ and HRP-conjugated anti-rabbit antibodies (1:10000; NA934, GE Healthcare). Signals were detected using ECL (ECL plus, GE healthcare).

Cell Cultures

The screening cell lines: T-lymphoblastic leukemia CEM; breast carcinoma MCF7 (estrogen-sensitive); cervical carcinoma cell line HeLa; and human foreskin fibroblasts BJ were obtained from the American Type Culture Collection (Manassas, VA, USA). Cells were cultured in DMEM (Dulbecco's Modified Eagle Medium, Sigma, MO, USA). Media used were supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and 1% penicillin-streptomycin. The cell lines were maintained under standard cell culture conditions at 37 $^{\circ}\text{C}$ and 5% CO_2 in a humid environment. Cells were subcultured twice or three times a week using the standard trypsinization procedure.

Calcein AM cytotoxicity assay

Suspensions with approximately 1.0×10^5 cells/mL were distributed in 96-well microtiter plates and after 24 h of stabilization the BRs analogues tested were added at the desired concentrations in DMSO. Control cultures were treated with DMSO alone, and the final concentration of DMSO in the reaction mixture never exceeded 0.6 %. In most cases, six serial 3-fold dilutions of the test

substances were added at time zero in 20 μL aliquots to the microtiter plate wells and the highest final concentration in the wells was 50 μM . After incubation for 72 h, Calcein AM solution (2 μM , Molecular Probes) was added and the cells were incubated for a further hour. The fluorescence of viable cells was then quantified using a Fluoroskan Ascent instrument (Labsystems, Finland). The percentage of surviving cells in each well was calculated by dividing the intensity of the fluorescence signals from the exposed wells by the intensity of signals from control wells and multiplying by 100. These ratios were then used to construct dose-response curves from which IC_{50} values, the concentrations of the respective compounds that were lethal to 50 % of the tumor cells, were calculated.

Conclusions

Several novel brassinosteroid 23-phenyl analogues were synthesized based on molecular docking into BRI1 receptor. The introduction of a phenyl group with no or small non-polar substituents (fluorine, chlorine, methyl) resulted in new compounds with plant growth promoting activities comparable with natural brassinosteroids. Results of biological screenings showed that molecular docking into BRI1 is a powerful tool for prediction and design of new compounds with strong brassinosteroid activities. New active compounds might be good candidates for potential application in agriculture to improve growth and yield or to increase resistance of plants against various biotic and abiotic stresses. Because recent progress in the chemical synthesis also leads to overcome economically restriction, which are currently the major constraints to use BRs at large scale in the fields.

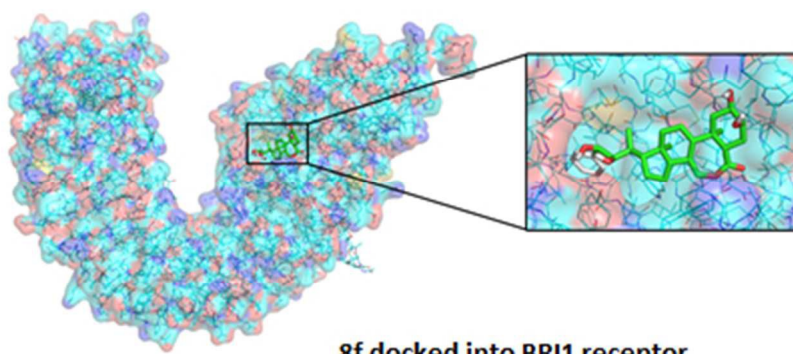
Acknowledgements

We thank Yanhai Yin for sharing BES1 antibodies. This work was supported by projects of the Ministry of Education, Youth and Sports CR NPUI LO1204, by grant GJ15-08202Y and 14-27669P of the Grant Agency of the Czech Republic. The work was also supported by student projects IGA_PrF_2016_018 and IGA_PrF_2016_028 of Palacky University in Olomouc. This work was supported by ELIXIR CZ research infrastructure project (MEYS Grant No: LM2015047) including access to computing and storage facilities. Qing Lu is indebted to the Chinese Scholarship Council (CSC) for his predoctoral fellowship.

References

- 1 Y. Coll, F. Coll, A. Amoros, M. Pujol, *Biologia*, 2015, **70**, 726.
- 2 J. Oklestkova, L. Rarova, M. Kvasnica, M. Strnad, *Phytochem. Rev.*, 2015, **14**, 1053.
- 3 S.D. Clouse, *J. Plant Growth Reg.*, 2015, **34**, 828.
- 4 Q. Fariduddin, M. Yusuf, I. Ahmad, A. Ahmad, *Biologia Plantarum*, 2014, **58**, 9.
- 5 A. Bajguz, S. Hayat, *Plant Physiol. Biochem.*, 2009, **47**, 1.
- 6 M.B. Wachsman, E.M.F. Lopez, J.A. Ramirez, L.R. Galagovsky, C.E. Coto, *Antivir. Chem. Chemother.*, 2000, **11**, 71;
- 7 M.B. Wachsman, V. Castilla, *Antiviral properties of brassinosteroids*. In *Brassinosteroids: practical applications in agriculture and human health*. A.B. Pereira-Netto (ed.), Bentham Science Publishers, Sharjah, 2012, pp 57–71.
- 8 F.M. Michelini, A. Berra, L.E. Alché, *J. Steroid Biochem. Mol. Biol.*, 2008, **108**, 164.

- 9 F.M. Michelini, P. Zorrilla, C. Robello, L.E. Alché, *Bioorg. Med. Chem.*, 2013, **21**, 560.
- 10 J. Malikova, J. Swaczynova, Z. Kolar, M. Strnad, *Phytochemistry*, 2008, **69**, 418.
- 11 L. Rarova, S. Zahler, J. Liebl, V. Krystof, D. Sedlak, P. Bartunek, M. Strnad, *Steroids*, 2012, **77**, 1502.
- 12 J. Steigerova, J. Oklestkova, M. Levkova, L. Rarova, Z. Kolar, M. Strnad, *Chem. Biol. Interact.*, 2010, **5**, 487.
- 13 P. Obakan, E.D. Arisan, A. Calcabrini, E. Agostinelli, S. Bolkent, N. Palavan-Unsal, *Amino Acids*, 2014, **46**, 553.
- 14 P. Obakan, E.D. Arisan, A. Coker-Gurkan, N. Palavan-Unsal, *Prostate*, 2014, **74**, 1622.
- 15 L. Rarova, J. Steigerova, M. Kvasnica, P. Bartunek, K. Krizova, H. Chodounska, Z. Kolar, D. Sedlak, J. Oklestkova, M. Strnad, *J. Steroid Biochem. Mol. Biol.*, 2016, **159**, 154.
- 16 S. Martins, E.M.N. Dohmann, A. Cayrel, A. Johnson, W. Fisher, J. Chory, G. Vert, *Nature Communications*, 2015, **6**, 6151.
- 17 J. Santiago, C. Henzler, M. Hothorn, *Science*, 2013, **341**, 889.
- 18 M. Hothorn, Y. Belkhadir, M. Dreux, T. Dabi, I.A. Wilson, J. Chory, *Nature*, 2011, **474**, 467.
- 19 J. She, Z. Han, T.-W. Kim, J. Wang, W. Cheng, J. Chang, S. Shi, J. Wang, M. Yang, Z.-Y. Wang, J. Chai, *Nature*, 2011, **474**, 472.
- 20 <ftp://ftp.wwpdb.org/pub/pdb/data/structures/divided/pdb/m7/pdb4m7e.ent.gz>
- 21 T.G. Back, L. Janzen, S.K. Nakajima, R.P. Pharis, *J. Org. Chem.*, 2000, **65**, 3047.
- 22 T.G. Back, R.P. Pharis, *J. Plant Growth Regul.*, 2003, **22**, 350.
- 23 B. Eignerova, B. Slavikova, M. Budesinsky, M. Dracinsky, B. Klepetarova, E. Stastna, M. Kotora, *J. Med. Chem.*, 2009, **52**, 5753.
- 24 S. Yamamoto, B. Watanabe, J. Otsuki, Y. Nakagawa, M. Akamatsu, H. Miyagawa, *Bioorg. Med. Chem.*, 2006, **14**, 1761.
- 25 K.B. Sharpless, W. Amberg, Y.L. Bennani, G.A. Criapino, J. Hartung, K.-S. Jeong, H.-L. Kwong, K. Morikawa, Z.-M. Wang, D. Xu, X.-L. Zhang, *J. Org. Chem.*, 1992, **57**, 2768.
- 26 L.-F. Huang, W.-S. Zhou, *J. Chem. Soc. Perkin Trans. 1*, 1994, 3579.
- 27 V.A. Khripach, V.N. Zhabinskii, A.E. de Groot, Brassinosteroids, a new class of plant hormones. Academic, San Diego, 1999, pp 301-309.
- 28 L. Kohout, M. Strand, M. Kamínek, Types of Brassinosteroid and their bioassays. In: *Brassinosteroids, chemistry, bioactivity and applications*. H.G. Cutler, T. Yokota, G. Adam (ed.), American Chemical Society, Washington, 1990, pp 56-74.
- 29 P. Guzmán, J.R. Ecker, *Plant Cell*, 1990, **2**, 513-523.
- 30 R.N. Arteca, D.-S. Tai, C. Schlagnhauser, N.B. Mandava, *Physiol. Plant.*, 1983, **59**, 539-544.
- 31 O. Trott, A.J. Olson, *J. Comput. Chem.*, 2010, **31**, 455;
- 32 G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, *J. Comput. Chem.*, 2009, **16**, 2785;
- 33 N.G. Irani, S. Di Rubbo, E. Mylle, J. Van den Begin, J. Schneider-Pizoń, J. Hnilickova, Miroslav Sisa, D. Buyst, J. Villarrasa-Blasi, A.-M. Szatmari, D. Van Damme, K. Mishev, M.-C. Codreanu, L. Kohout, M. Strnad, A. I Caño-Delgado, J. Friml, A. Madder, E. Russinova, *Nature Chem. Biol.*, 2012, **8**, 583.
- 34 Y. Yin, Z.Y. Wang, S. Mora-Garcia, J. Li, S. Yoshida, T. Asami, J. Chory, *Cell*, 2002, **109**, 181.



110x53mm (96 x 96 DPI)

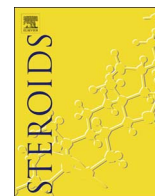
Appendix C

Korinkova P, **Bazgier V**, Oklestkova J, Rarova L, Strnad M, Kvasnica M:
Synthesis of novel aryl brassinosteroids through alkene crossmetathesis and
preliminary biological study.

Steroids, 127, 46-55, 2017.

DOI: 10.1016/j.steroids.2017.08.010

IF = 2.282



Synthesis of novel aryl brassinosteroids through alkene cross-metathesis and preliminary biological study



Petra Korinkova^{a,1}, Vaclav Bazgier^{b,1}, Jana Oklestkova^a, Lucie Rarova^c, Miroslav Strnad^a, Miroslav Kvasnica^{a,*}

^a Laboratory of Growth Regulators, Centre of the Region Haná for Biotechnological and Agricultural Research, Palacký University & Institute of Experimental Botany ASCR, Šlechtitelů 27, 783 71 Olomouc, Czech Republic

^b Regional Centre of Advanced Technologies and Material, Department of Physical Chemistry, Faculty of Science, Palacký University, 17. listopadu 12, 771 46 Olomouc, Czech Republic

^c Department of Chemical Biology and Genetics, Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacký University, Šlechtitelů 27, 783 71 Olomouc, Czech Republic

ARTICLE INFO

Keywords:

Brassinosteroids
Cross-metathesis
Organic synthesis
Molecular docking
BR11 receptor kinase
Plant bioassays

ABSTRACT

A series of phenyl analogues of brassinosteroids was prepared via alkene cross-metathesis using commercially available styrenes and 24-nor-5 α -chola-2,22-dien-6-one. All derivatives were successfully docked into the active site of BRI1 using AutoDock Vina. Plant growth promoting activity was measured using the pea inhibition biotest and Arabidopsis root sensitivity assay and then was compared with naturally occurring brassinosteroids. Differences in the production of plant hormone ethylene were also observed in etiolated pea seedlings after treatment with the new and also five known brassinosteroid phenyl analogues. Antiproliferative activity was also studied using normal human fibroblast and human cancer cell lines.

1. Introduction

Brassinosteroids (BRs, Fig. 1) represent a large group of plant steroids with more than 70 structurally and functionally related compounds [1]. BRs have been found at low concentrations throughout the plant kingdom, widely distributed in higher and lower plants, and have been detected in various plant parts such as pollen, seeds, leaves, stems, roots, and flowers. They are essential for many aspects of plant growth and development, such as cell division, elongation and differentiation, pollen tube growth, seed germination, regulation of gene expression, enzyme activation and photosynthesis [2–5]. At the molecular level, BRs change the gene expression and the metabolism of nucleic acids and proteins. BRs have structures similar to those of animal steroid hormone. Unlike animals, plants perceive steroids at cell membrane, using the membrane-integral receptor kinase brassinosteroid insensitive 1 (BRI1) [6–8]. The encoded protein, BRI1, belongs to a large family of plant LRR (leucine-rich repeat) receptor-like kinases, characterized by an extracellular LRR domain, a single-pass transmembrane segment and a cytoplasmic kinase domain. BRI1 has been established as an authentic brassinosteroid receptor by genetic and biochemical investigations [9]. Crystal structures of BRI1 in both free (PDB ID: 3RIZ, 3RGX), and brassinolide-bound (PDB ID: 3RJ0, 3BRZ), forms are available, following

independent X-ray diffraction structural determinations by two groups [9,10]. The structure of the ligand-binding domain resembles a superhelix of 25 twisted LRRs. A 70-amino acid island domain between LRRs 21 and 22 folds back into the interior of the superhelix, creating a surface pocket where the brassinosteroids bind. These recently published structures of *Arabidopsis thaliana* BRI1 enable the rational design of brassinosteroid-like antagonists and agonists. Recent studies [11,12] have indicated that molecular docking is a powerful tool to predict how effective incorporation of different functional groups into brassinosteroid skeleton is and to design new types of BRs with biological activities comparable to natural BRs [11].

The aim of our study is related to the synthesis of new aryl analogues of BRs by alkene cross-metathesis and to study of their biological properties. Alkene cross-metathesis was chosen for preparation of all aryl analogues as an efficient method for construction of the new side chains using different commercially available substituted styrenes. The biological activities of newly prepared derivatives were evaluated using plant bioassays (pea inhibition biotest and Arabidopsis root sensitivity bioassay) and Calcein AM cytotoxicity assay. All derivative structures were subjected to docking studies using AutoDock Vina [13] in order to analyze the results with theoretical studies.

* Corresponding author at: Laboratory of Growth Regulators, Institute of Experimental Botany ASCR & Palacký University, Šlechtitelů 27, 78371 Olomouc, Czech Republic.
E-mail address: kvasnica@ueb.cas.cz (M. Kvasnica).

¹ These authors contributed equally to this work.

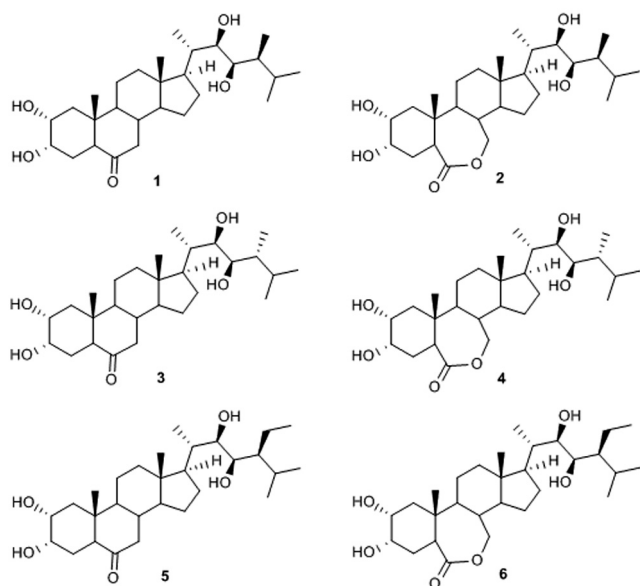


Fig. 1. Structures of most common natural brassinosteroids; castasterone (1), brassinolide (2), 24-epicastasterone (3), 24-epibrassinolide (4), 28-homocastasterone (5), 28-homobrassinolide (6).

2. Experimental

2.1. General methods

The melting points were determined on a Stuart SMP30 instrument (Bibby Scientific Ltd., UK). Elemental analyses were performed using an EA 1108 elemental analyzer (Fison Instruments); the values (C, H, N) agreed with the calculated values within acceptable limits. The infrared spectra were recorded on a Thermo Scientific Nicolet spectrometer iZ10 using the ATR technique. The wave numbers are given in cm^{-1} . The NMR spectra were taken on a JEOL JNM-ECA 500 (JEOL, Tokyo, Japan; ^1H , 500 MHz; ^{13}C , 125 MHz) spectrometer equipped with a 5 mm JEOL Royal probe. ^1H NMR and ^{13}C NMR chemical shifts (δ) were calibrated using tetramethylsilane (TMS, ^1H δ = 0 ppm) or solvents: CDCl_3 (^1H δ = 7.26 ppm, ^{13}C δ = 77.00 ppm) or $\text{DMSO}-d_6$ (^1H δ = 2.46 ppm, ^{13}C δ = 40.00 ppm). Chemical shifts are given in ppm (δ -scale), coupling constants (J) in Hz. All values were obtained by first-order analysis. For API HRMS analysis, the samples were dissolved in chloroform (or chloroform: methanol; 1:1; v/v, in case of hydroxylated compounds) to a concentration $10 \mu\text{g}\cdot\text{mL}^{-1}$. The ASAP (Atmospheric Solids Analysis Probe) was dipped into the sample solution, placed into the ion source and analysed in fullscan mode. The source of the Synapt G2-Si Mass Spectrometer (Waters, Manchester, UK) was operated in positive ionisation mode (ASAP+), if not stated otherwise, at source temperature of 120°C . The Corona needle current was kept at 5 μA and the collision energy at value 4. The probe temperature was ramped up from 50°C to 600°C in 3 min. Data were acquired from 50 to 1000 Da with 1.0 s scan time in High Resolution Mode. The data were processed using the Masslynx 4.1 software (Waters). Mass accuracy of 1 ppm or less was achieved with the described instrumentation for all compounds. Merck silica gel Kieselgel 60 (230–400 mesh) was used for column chromatography. The HPLC system consisted of a Waters semi-preparative HPLC system including quaternary pump, liquid handler, UV-VIS and ELSD detectors. The semi preparative column was filled with silica gel. Reagents and solvents were purchased from Sigma-Aldrich and were not purified.

2.1.1. General procedure for cross metathesis

Hoveyda-Grubbs 2nd generation catalyst (19 mg; 0.03 mmol) was added to a solution of dien **7** (100 mg; 0.31 mmol) and styrene

(2.48 mmol) in dichloroethane (5 mL). The reaction mixture was heated at 80°C for 5 h. Then, another portion of H-G catalyst (19 mg; 0.03 mmol) was added and the reaction mixture was heated at 80°C for additional 5 h. Then, the solvent was evaporated and crude solid was purified by column chromatography on silica gel (mobile phase – 3% ethyl acetate in cyclohexane, R_f of products 0.18–0.25). In some cases, stated in each experiment, HPLC had to be used due to very close retention time of product and starting material (mobile phase – 0.5% ethyl acetate in cyclohexane).

2.1.2. (22E)-23-phenyl-24-nor-5 α -chola-2,22-dien-6-one (**8a**)

The general procedure with styrene afforded 120 mg (81%) of the title compound **8a** as a colorless oil. ^1H NMR (CDCl_3) δ 0.72, 0.74 (both s, 3H, CH_3), 1.14 (d, 3H, J = 6.4 Hz, CH_3), 1.69–1.80 (m, 2H), 1.96–2.04 (m, 4H), 2.07 (dt, 1H, J = 12.6, J' = 3.3 Hz), 2.23–2.31 (m, 2H), 2.34–2.37, (m, 2H), 5.57 (m, 1H, H-3), 5.69 (m, 1H, H-2), 6.06 (dd, 1H, J = 15.9, J' = 8.9 Hz, H-22), 6.30 (d, 1H, J = 15.9 Hz, H-23), 7.19 (m, 1H, Ar-H), 7.27–7.35 (m, 4H, 4 \times Ar-H). ^{13}C NMR δ 12.19 (C-18), 13.50 (C-19), 20.38 (C-21), 21.10, 21.70, 23.92, 28.21, 37.68, 39.35, 39.37, 40.04, 40.39, 42.92, 46.96, 53.40, 53.83, 55.83, 56.75, 124.49 (C-3), 124.95 (C-2), 125.92 (2 \times C), 126.72, 127.42 (C-23), 128.45 (2 \times C), 136.94 (C-22), 137.98, 211.98 (C-6). Spectral data in agreement with literature [11].

2.1.3. (22E)-23-(2-fluorophenyl)-24-nor-5 α -chola-2,22-dien-6-one (**8b**)

The general procedure with *o*-fluorostyrene afforded 84 mg (65%) of the title compound **9a** as a colorless oil: IR ν (cm^{-1}) 2930, 1702, 1655, 1593, 1560, 965. ^1H NMR (CDCl_3) δ 0.73, 0.75 (both s, 3H, CH_3); 1.15 (d, 3H, J = 6.7 Hz, CH_3); 1.72–1.81 (m, 2H); 1.97–2.06 (m, 4H); 2.08 (dt, 1H, J = 12.6, J' = 3.5 Hz); 2.24–2.33 (m, 2H); 2.34–2.38 (m, 2H); 5.58 (m, 1H, H-3); 5.70 (m, 1H, H-2); 6.14 (dd, 1H, J = 15.9, J' = 8.6 Hz, H-22); 6.47 (d, 1H, J = 15.9 Hz, H-23); 7.01 (ddd, J = 10.9, J'' = 7.8, J''' = 0.9 Hz, Ar-H); 7.07 (td, 1H, J = 7.8, J' = 1.2 Hz, Ar-H); 7.16 (m, 1H, Ar-H); 7.41 (td, 1H, J = 7.8, J' = 1.8 Hz, Ar-H). ^{13}C NMR δ 12.19 (C-18), 13.48 (C-19), 20.27 (C-21), 21.08, 21.70, 23.91, 28.17, 37.65, 39.32, 39.35, 40.01, 40.76, 42.92, 46.93, 53.37, 53.80, 55.68, 56.71, 115.56 (d, J = 22.8 Hz), 119.79 (d, J = 3.6 Hz), 123.91 (d, J = 3.6 Hz), 124.48 (C-3), 124.94 (C-2), 125.62 (d, J = 13.2 Hz), 126.92 (d, J = 3.6 Hz), 127.86 (d, J = 8.4 Hz), 139.50 (d, J = 3.6 Hz), 159.94 (d, J = 248.3 Hz, C-F), 211.90 (C-6). ^{19}F NMR (^1H) δ -118.83 (s, 1F). HRMS: (API+) calculated for $\text{C}_{29}\text{H}_{38}\text{FO}$ ($[\text{M}+\text{H}]^+$) 421.2907, Found 421.2910.

2.1.4. (22E)-23-(3-fluorophenyl)-24-nor-5 α -chola-2,22-dien-6-one (**8c**)

The general procedure with *m*-fluorostyrene afforded 94 mg (73%) of the title compound **10a** as a colorless oil: IR ν (cm^{-1}) 2933, 1705, 1656, 1593, 1560, 966. ^1H NMR (CDCl_3) δ 0.73, 0.74 (both s, 3H, CH_3); 1.14 (d, 3H, J = 6.7 Hz, CH_3); 1.71–1.80 (m, 2H); 1.97–2.04 (m, 4H); 2.07 (dt, 1H, J = 12.5, J' = 3.2 Hz); 2.23–2.31 (m, 2H); 2.34–2.38 (m, 2H); 5.58 (m, 1H, H-3); 5.70 (m, 1H, H-2); 6.08 (dd, 1H, J = 15.7, J' = 8.7 Hz, H-22); 6.28 (d, 1H, J = 15.7 Hz, H-23); 6.88 (td, 1H, J = 8.3, J' = 2.6 Hz, Ar-H); 7.03 (m, 1H, Ar-H); 7.08 (b d, 1H, J = 8.3 Hz, Ar-H); 7.24 (td, 1H, J = 7.8, J' = 6.1 Hz, Ar-H). ^{13}C NMR δ 12.16 (C-18), 13.46 (C-19), 20.22 (C-21), 21.06, 21.68, 23.87, 28.16, 37.62, 39.30, 39.32, 39.98, 40.30, 42.90, 46.90, 53.33, 53.77, 55.67, 56.66, 112.25 (d, J = 21.6 Hz), 113.44 (d, J = 21.6 Hz), 121.81 (d, J = 2.4 Hz), 124.45 (C-3), 124.92 (C-2), 126.46 (d, J = 2.4 Hz), 129.79 (d, J = 8.4 Hz), 138.28, 140.35 (d, J = 7.2 Hz), 163.08 (d, J = 244.7 Hz, C-F), 211.82 (C-6). ^{19}F NMR (^1H) δ -113.78 (s, 1F). HRMS: (API+) calculated for $\text{C}_{29}\text{H}_{38}\text{FO}$ ($[\text{M}+\text{H}]^+$) 421.2907, Found 421.2910.

2.1.5. (22E)-23-(4-fluorophenyl)-24-nor-5 α -chola-2,22-dien-6-one (**8d**)

The general procedure with *p*-fluorostyrene afforded 96 mg (75%) of the title compound **8** as a colorless oil. ^1H NMR (CDCl_3) δ 0.72, 0.73 (both s, 3H, CH_3), 1.13 (d, 3H, J = 6.7 Hz, CH_3), 1.71–1.80 (m, 2H),

1.96–2.04 (m, 4H), 2.08 (dt, 1H, $J = 12.5$, $J' = 3.2$ Hz), 2.22–2.31 (m, 2H), 2.33–2.37 (m, 2H), 5.57 (m, 1H, H-3), 5.69 (m, 1H, H-2), 5.97 (dd, 1H, $J = 15.7$, $J' = 8.7$ Hz, H-22), 6.26 (d, 1H, $J = 15.7$ Hz, H-23), 6.97 (m, 2H, 2 × Ar-H), 7.27 (m, 2H, 2 × Ar-H). ^{13}C NMR δ 12.17 (C-18), 13.48 (C-19), 20.35 (C-21), 21.08, 21.69, 23.88, 28.21, 37.65, 39.32, 39.34, 40.01, 40.33, 42.89, 46.92, 53.36, 53.80, 55.79, 56.71, 115.25 (d, $J = 20.4$ Hz, 2 × C), 124.47 (C-3), 124.93 (C-2), 126.25 (C-23), 127.28 (d, $J = 8.4$ Hz, 2 × C), 134.07 (d, $J = 3.6$ Hz), 136.65 (d, $J = 2.4$ Hz, C-22), 161.79 (d, $J = 244.7$ Hz, C-F), 211.93 (C-6). Spectral data in agreement with literature [11].

2.1.6. (22E)-23-(2-chlorophenyl)-24-nor-5 α -chola-2,22-dien-6-one (8e)

The general procedure with *o*-chlorostyrene afforded 86 mg (64%) of the title compound **11a** as a colorless oil: IR ν (cm^{-1}) 2941, 1701, 1655, 1593, 1563, 964. ^1H NMR (CDCl_3) δ 0.73, 0.75 (both s, 3H, CH_3); 1.16 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.81 (m, 2H); 1.96–2.04 (m, 4H); 2.08 (dt, 1H, $J = 12.5$, $J' = 3.2$ Hz); 2.22–2.33 (m, 2H); 2.34–2.38 (m, 2H); 5.58 (m, 1H, H-3); 5.70 (m, 1H, H-2); 6.04 (dd, 1H, $J = 15.6$, $J' = 8.9$ Hz, H-22); 6.69 (d, 1H, $J = 15.9$ Hz, H-23); 7.13 (m, 1H, $\Sigma J = 16.8$ Hz, Ar-H); 7.19 (m, 1H, $\Sigma J = 15.9$ Hz, Ar-H); 7.33 (dd, 1H, $J = 7.9$, $J' = 1.2$ Hz, Ar-H); 7.47 (dd, 1H, $J = 7.9$, $J' = 1.5$ Hz, Ar-H). ^{13}C NMR δ 12.19 (C-18), 13.46 (C-19), 20.26 (C-21), 21.07, 21.68, 23.91, 28.17, 37.62, 39.30, 39.32, 39.98, 40.51, 42.93, 46.91, 53.34, 53.77, 55.63, 56.68, 123.80, 124.46 (C-3), 124.91 (C-2), 126.53, 126.64, 127.73, 129.52, 132.58, 136.00, 139.72, 211.87 (C-6). HRMS: (API+) calculated for $\text{C}_{29}\text{H}_{38}\text{ClO}$ ($[\text{M}+\text{H}]^+$) 437.2611, Found 437.2614.

2.1.7. (22E)-23-(3-chlorophenyl)-24-nor-5 α -chola-2,22-dien-6-one (8f)

The general procedure with *m*-chlorostyrene afforded 94 mg (70%) of the title compound **12a** as a colorless oil: IR ν (cm^{-1}) 2944, 1701, 1654, 1593, 1562, 964. ^1H NMR (CDCl_3) δ 0.72, 0.73 (both s, 3H, CH_3); 1.13 (d, 3H, $J = 6.7$ Hz, CH_3); 1.70–1.80 (m, 2H); 1.96–2.05 (m, 4H); 2.07 (dt, 1H, $J = 12.5$, $J' = 3.2$ Hz); 2.22–2.31 (m, 2H); 2.33–2.37 (m, 2H); 5.58 (m, 1H, H-3); 5.69 (m, 1H, H-2); 6.07 (dd, 1H, $J = 15.9$, $J' = 8.9$ Hz, H-22); 6.25 (d, 1H, $J = 15.9$ Hz, H-23); 7.14–7.22 (m, 3H, 3 × Ar-H); 7.31 (t, 1H, $J = 1.7$ Hz, Ar-H). ^{13}C NMR δ 12.18 (C-18), 13.48 (C-19), 20.24 (C-21), 21.08, 21.69, 23.89, 28.19, 37.64, 39.31, 39.33, 40.01, 40.39, 42.92, 46.92, 53.34, 53.79, 55.66, 56.68, 124.23, 124.47 (C-3), 124.94 (C-2), 126.22, 126.63, 127.74, 129.63, 134.36, 138.47, 139.85, 211.92 (C-6). HRMS: (API+) calculated for $\text{C}_{29}\text{H}_{38}\text{ClO}$ ($[\text{M}+\text{H}]^+$) 437.2611, Found 437.2615.

2.1.8. (22E)-23-(4-chlorophenyl)-24-nor-5 α -chola-2,22-dien-6-one (8g)

The general procedure with *p*-chlorostyrene afforded 100 mg (75%) of the title compound **8** as a colorless oil. ^1H NMR (CDCl_3) δ 0.72, 0.73 (both s, 3H, CH_3), 1.13 (d, 3H, $J = 6.7$ Hz, CH_3), 1.70–1.80 (m, 2H), 1.96–2.04 (m, 4H), 2.07 (dt, 1H, $J = 12.5$, $J' = 3.4$ Hz), 2.22–2.30 (m, 2H), 2.34–2.37 (m, 2H), 5.57 (m, 1H, H-3), 5.69 (m, 1H, H-2), 6.03 (dd, 1H, $J = 15.9$, $J' = 8.9$ Hz, H-22), 6.27 (d, 1H, $J = 15.9$ Hz, H-23), 7.24 (s, 4H, 4 × Ar-H). ^{13}C NMR δ 12.19 (C-18), 13.50 (C-19), 20.27 (C-21), 21.09, 21.70, 23.91, 28.21, 37.66, 39.33, 39.36, 40.04, 40.39, 42.93, 46.94, 53.37, 53.82, 55.75, 56.72, 124.49 (C-3), 124.96 (C-2), 126.30 (C-23), 127.12 (2 × C), 128.55 (2 × C), 132.22, 136.45, 137.64 (C-22), 211.96 (C-6). Spectral data in agreement with literature [11].

2.1.9. (22E)-23-(2-bromophenyl)-24-nor-5 α -chola-2,22-dien-6-one (8h)

The general procedure with *o*-bromostyrene afforded 87 mg (59%) of the title compound **13a** as a colorless oil: IR ν (cm^{-1}) 2944, 1699, 1654, 1593, 1564, 964. ^1H NMR (CDCl_3) δ 0.72, 0.75 (both s, 3H, CH_3); 1.15 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.81 (m, 2H); 1.96–2.04 (m, 4H); 2.08 (dt, 1H, $J = 12.5$, $J' = 3.1$ Hz); 2.22–2.38 (m, 4H); 5.57 (m, 1H, H-3); 5.69 (m, 1H, H-2); 5.99 (dd, 1H, $J = 15.6$, $J' = 8.9$ Hz, H-22); 6.63 (d, 1H, $J = 15.6$ Hz, H-23); 7.05 (m, 1H, $\Sigma J = 16.8$ Hz, Ar-H); 7.23 (m, 1H, $\Sigma J = 16.2$ Hz, Ar-H); 7.45 (dd, 1H, $J = 7.6$, $J' = 1.5$ Hz, Ar-H); 7.52 (m, 1H, $\Sigma J = 9.2$ Hz, Ar-H). ^{13}C NMR δ 12.24 (C-18), 13.51

(C-19), 20.28 (C-21), 21.10, 21.70, 23.96, 28.21, 37.66, 39.34, 39.36, 40.03, 40.47, 42.98, 46.95, 53.39, 53.83, 55.66, 56.72, 123.26, 124.50 (C-3), 124.96 (C-2), 126.50, 126.81, 127.33, 128.06, 132.77, 137.80, 139.93, 211.97 (C-6). HRMS: (API+) calculated for $\text{C}_{29}\text{H}_{38}^{79}\text{BrO}$ ($[\text{M}+\text{H}]^+$) 481.2106, Found 481.2107.

2.1.10. (22E)-23-(3-bromophenyl)-24-nor-5 α -chola-2,22-dien-6-one (8i)

The general procedure with *m*-bromostyrene afforded 97 mg (66%) of the title compound **14a** as a colorless oil: IR ν (cm^{-1}) 2944, 1700, 1655, 1588, 1560, 964. ^1H NMR (CDCl_3) δ 0.72, 0.73 (both s, 3H, CH_3); 1.13 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.80 (m, 2H); 1.96–2.04 (m, 4H); 2.06 (dt, 1H, $J = 12.5$, $J' = 3.4$ Hz); 2.22–2.31 (m, 2H); 2.34–2.37 (m, 2H); 5.58 (m, 1H, H-3); 5.69 (m, 1H, H-2); 6.06 (dd, 1H, $J = 15.7$, $J' = 8.7$ Hz, H-22); 6.23 (d, 1H, $J = 15.7$ Hz, H-23); 7.14 (t, 1H, $J = 7.8$ Hz, Ar-H); 7.22 (m, 1H, Ar-H); 7.30 (m, 1H, Ar-H); 7.47 (t, 1H, $J = 1.7$ Hz, Ar-H). ^{13}C NMR δ 12.19 (C-18), 13.50 (C-19), 20.25 (C-21), 21.09, 21.70, 23.91, 28.20, 37.66, 39.34, 39.36, 40.03, 40.41, 42.95, 46.94, 53.37, 53.82, 55.69, 56.71, 122.69, 124.48 (C-3), 124.69, 124.96 (C-2), 126.15, 128.70, 129.56, 129.95, 138.55, 140.17, 211.92 (C-6). HRMS: (API+) calculated for $\text{C}_{29}\text{H}_{38}^{79}\text{BrO}$ ($[\text{M}+\text{H}]^+$) 481.2106, Found 481.2108.

2.1.11. (22E)-23-(4-bromophenyl)-24-nor-5 α -chola-2,22-dien-6-one (8j)

The general procedure with *p*-bromostyrene afforded 107 mg (73%) of the title compound **8** as a colorless oil. ^1H NMR (CDCl_3) δ 0.72, 0.73 (both s, 3H, CH_3), 1.13 (d, 3H, $J = 6.7$ Hz, CH_3), 1.70–1.79 (m, 2H), 1.96–2.04 (m, 4H), 2.07 (dt, 1H, $J = 12.5$, $J' = 3.3$ Hz), 2.22–2.30 (m, 2H), 2.33–2.37 (m, 2H), 5.57 (m, 1H, H-3), 5.69 (m, 1H, H-2), 6.05 (dd, 1H, $J = 15.7$, $J' = 8.7$ Hz, H-22), 6.24 (d, 1H, $J = 15.7$ Hz, H-23), 7.18 (m, 2H, 2 × Ar-H), 7.39 (m, 2H, 2 × Ar-H). ^{13}C NMR δ 12.19 (C-18), 13.49 (C-19), 20.24 (C-21), 21.08, 21.70, 23.90, 28.19, 37.65, 39.32, 39.35, 40.02, 40.40, 42.92, 46.93, 53.36, 53.81, 55.71, 56.70, 120.30, 124.47 (C-3), 124.95 (C-2), 126.33 (C-23), 127.47 (2 × C), 131.48 (2 × C), 136.89, 137.77 (C-22), 211.92 (C-6). Spectral data in agreement with literature [11].

2.1.12. (22E)-23-(2-methylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8k)

The general procedure with *o*-methylstyrene afforded 83 mg (65%) of the title compound **15a** as a colorless oil: IR ν (cm^{-1}) 2932, 1701, 1653, 1593, 1563, 964. ^1H NMR (CDCl_3) δ 0.72, 0.74 (both s, 3H, CH_3); 1.14 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.80 (m, 2H); 1.96–2.04 (m, 4H); 2.07 (dt, 1H, $J = 12.6$, $J' = 3.6$ Hz); 2.22–2.31 (m, 2H); 2.32 (s, 3H, Ar- CH_3); 2.33–2.37 (m, 2H); 5.57 (m, 1H, H-3); 5.69 (m, 1H, $J = 15.6$ Hz, H-2); 5.89 (dd, 1H, $J = 15.6$, $J' = 8.9$ Hz, H-22); 6.49 (d, 1H, $J = 15.6$ Hz, H-23); 7.10–7.15 (m, 3H, 3 × Ar-H); 7.37 (d, 1H, $J = 7.3$ Hz, Ar-H). ^{13}C NMR δ 12.20 (C-18), 13.48 (C-19), 19.58 (Ar- CH_3), 20.53 (C-21), 21.08, 21.69, 23.92, 28.30, 37.65, 39.31, 39.34, 40.01, 40.72, 42.87, 46.93, 53.35, 53.78, 55.71, 56.73, 124.48 (C-3), 124.93 (C-2), 125.28, 125.45, 125.92, 126.67, 130.06, 134.94, 137.15, 138.43, 211.98 (C-6). HRMS: (API+) calculated for $\text{C}_{30}\text{H}_{41}\text{O}$ ($[\text{M}+\text{H}]^+$) 417.3157, Found 417.3160.

2.1.13. (22E)-23-(3-methylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8l)

The general procedure with *m*-methylstyrene afforded 88 mg (69%) of the title compound **16a** as a colorless oil: IR ν (cm^{-1}) 2933, 1700, 1653, 1593, 1562, 964. ^1H NMR (CDCl_3) δ 0.71, 0.73 (both s, 3H, CH_3); 1.12 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.79 (m, 2H); 1.95–2.04 (m, 4H); 2.06 (dt, 1H, $J = 12.6$, $J' = 3.5$ Hz); 2.20–2.30 (m, 2H); 2.32 (s, 3H, Ar- CH_3); 2.33–2.37 (m, 2H); 5.57 (m, 1H, H-3); 5.68 (m, 1H, H-2); 6.04 (dd, 1H, $J = 15.9$, $J' = 8.9$ Hz, H-22); 6.27 (d, 1H, $J = 15.9$ Hz, H-23); 7.00 (d, v, $J = 7.3$ Hz, Ar-H); 7.11–7.19 (m, 3H, 3 × Ar-H). ^{13}C NMR δ 12.19 (C-18), 13.49 (C-19), 20.40 (C-21), 21.09, 21.40 (Ar- CH_3), 21.70, 23.90, 28.20, 37.66, 39.32, 39.35, 40.03, 40.40, 42.88, 46.94, 53.37, 53.80, 55.82, 56.73, 123.09, 124.49 (C-3), 124.94 (C-2), 126.60, 127.45, 127.51, 128.36, 136.73, 137.89, 137.95, 211.99 (C-6). HRMS: (API+) calculated for $\text{C}_{30}\text{H}_{41}\text{O}$ ($[\text{M}+\text{H}]^+$) 417.3157, Found 417.3161.

2.1.14. (22E)-23-(4-methylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8m)

The general procedure with *p*-methylstyrene afforded 91 mg (71%) of the title compound **8** as a colorless oil. ^1H NMR (CDCl_3) δ ^1H NMR (CDCl_3) δ 0.72, 0.73 (both s, 3H, CH_3), 1.12 (d, 3H, $J = 6.7$ Hz, CH_3), 1.70–1.79 (m, 2H), 1.96–2.04 (m, 4H), 2.07 (dt, 1H, $J = 12.5$, $J' = 3.4$ Hz), 2.22–2.30 (m, 2H), 2.32 (s, 3H, CH_3), 2.33–2.37 (m, 2H), 5.57 (m, 1H, H-3), 5.69 (m, 1H, H-2), 6.00 (dd, 1H, $J = 15.9$, $J' = 8.9$ Hz, H-22), 6.27 (d, 1H, $J = 15.9$ Hz, H-23), 7.09 (d, 2H, $J = 7.9$ Hz, $2 \times \text{Ar-H}$), 7.22 (d, 2H, $J = 7.9$ Hz, $2 \times \text{Ar-H}$). ^{13}C NMR δ 12.19 (C-18), 13.49 (C-19), 20.42 (C-21), 21.09 (CH_2 , Ar- CH_3), 21.70, 23.90, 28.19, 37.68, 39.34, 39.36, 40.04, 40.34, 42.88, 46.95, 53.39, 53.81, 55.89, 56.75, 124.50 (C-3), 124.95 (C-2), 125.80 ($2 \times \text{C}$), 127.21 (C-23), 129.13 ($2 \times \text{C}$), 135.18, 135.94 (C-22), 136.40, 212.00 (C-6). Spectral data in agreement with literature [11].

2.1.15. (22E)-23-(2-trifluoromethylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8n)

The general procedure with *o*-trifluoromethylstyrene afforded 82 mg (57%) of the title compound **17a** as a colorless oil: IR ν (cm^{-1}) 2942, 1701, 1653, 1593, 1560, 1128, 964. ^1H NMR (CDCl_3) δ 0.73, 0.76 (both s, 3H, CH_3); 1.15 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.81 (m, 2H); 1.96–2.04 (m, 4H); 2.07 (dt, 1H, $J = 12.6$, $J' = 3.6$ Hz); 2.23–2.31 (m, 2H); 2.31–2.38 (m, 2H); 5.58 (m, 1H, H-3); 5.69 (m, 1H, H-2); 6.00 (dd, 1H, $J = 15.6$, $J' = 8.9$ Hz, H-22); 6.68 (qd, 1H, $J = 15.6$, $J' = 2.1$ Hz, H-23); 7.29 (t, 1H, $J = 7.7$ Hz, Ar-H); 7.46 (t, 1H, $J = 7.7$ Hz, Ar-H); 7.55 (d, 1H, $J = 7.7$ Hz, Ar-H); 7.60 (d, 1H, $J = 7.7$ Hz, Ar-H). ^{13}C NMR δ 12.19 (C-18), 13.48 (C-19), 20.17 (C-21), 21.08, 21.70, 23.94, 28.02, 37.63, 39.32 ($2 \times \text{C}$), 40.00, 40.59, 42.95, 46.92, 53.37, 53.80, 55.68, 56.69, 123.72, 124.38 (q, $J = 273.5$ Hz), 124.47 (C-3), 124.93 (C-2), 125.58 (q, $J = 6.0$ Hz), 126.44, 126.92 (q, $J = 30.0$ Hz), 127.28, 131.65, 137.26, 141.19, 211.90 (C-6). ^{19}F NMR (δ) -59.56 (s, 1F). HRMS: (API+) calculated for $\text{C}_{30}\text{H}_{38}\text{F}_3\text{O}$ ($[\text{M} + \text{H}]^+$) 471.2875, Found 471.2876.

2.1.16. (22E)-23-(3-trifluoromethylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8o)

The general procedure with *m*-trifluoromethylstyrene afforded 95 mg (66%) of the title compound **18** as a colorless oil: IR ν (cm^{-1}) 2940, 1701, 1655, 1593, 1561, 1124, 964. ^1H NMR (CDCl_3) δ 0.72, 0.74 (both s, 3H, CH_3); 1.14 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.80 (m, 2H); 1.97–2.05 (m, 4H); 2.07 (dt, 1H, $J = 12.6$, $J' = 3.3$ Hz); 2.22–2.32 (m, 2H); 2.34–2.37 (m, 2H); 5.58 (m, 1H, H-3); 5.69 (m, 1H, H-2); 6.14 (dd, 1H, $J = 15.7$, $J' = 8.7$ Hz, H-22); 6.34 (d, 1H, $J = 15.7$ Hz, H-23); 7.37–7.44 (m, 2H, $2 \times \text{Ar-H}$); 7.48 (d, 1H, $J = 7.3$ Hz, Ar-H); 7.56 (b s, 1H, Ar-H). ^{13}C NMR δ 12.19 (C-18), 13.50 (C-19), 20.24 (C-21), 21.09, 21.70, 23.91, 28.23, 37.66, 39.36, 39.37, 40.03, 40.48, 42.96, 46.93, 53.37, 53.82, 55.64, 56.71, 122.52 (q, $J = 3.6$ Hz), 123.26 (q, $J = 3.6$ Hz), 124.19 (q, $J = 272.3$ Hz), 124.48 (C-3), 124.96 (C-2), 126.27, 128.85, 129.14, 130.81 (q, $J = 32.4$ Hz), 138.71, 138.90, 211.92 (C-6). ^{19}F NMR (δ) -62.59 (s, 1F). HRMS: (API+) calculated for $\text{C}_{30}\text{H}_{38}\text{F}_3\text{O}$ ($[\text{M} + \text{H}]^+$) 471.2875, Found 471.2879.

2.1.17. (22E)-23-(4-trifluoromethylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8p)

The general procedure with *p*-trifluoromethylstyrene afforded 102 mg (71%) of the title compound **19a** as a colorless oil: IR ν (cm^{-1}) 2940, 1702, 1653, 1593, 1562, 1127, 963. ^1H NMR (CDCl_3) δ 0.72, 0.74 (both s, 3H, CH_3); 1.15 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.81 (m, 2H); 1.96–2.05 (m, 4H); 2.07 (dt, 1H, $J = 12.5$, $J' = 3.4$ Hz); 2.21–2.31 (m, 2H); 2.34–2.37 (m, 2H); 5.58 (m, 1H, H-3); 5.69 (m, 1H, H-2); 6.17 (dd, 1H, $J = 15.7$, $J' = 8.7$ Hz, H-22); 6.34 (d, 1H, $J = 15.7$ Hz, H-23); 7.41 (d, 2H, $J = 8.4$ Hz, $2 \times \text{Ar-H}$); 7.53 (d, 2H, $J = 8.4$ Hz, $2 \times \text{Ar-H}$). ^{13}C NMR δ 12.21 (C-18), 13.51 (C-19), 20.19 (C-21), 21.10, 21.70, 23.92, 28.19, 37.67, 39.34, 39.37, 40.04, 40.49, 42.99, 46.94, 53.38, 53.83, 55.85, 56.70, 124.28 (q, $J = 271.5$ Hz), 124.48 (C-3), 124.97 (C-

2), 125.38, 125.40 (q, $J = 3.6$ Hz, $2 \times \text{C}$), 126.04 ($2 \times \text{C}$), 126.35, 128.56 (q, $J = 32.4$ Hz), 139.66, 211.91 (C-6). ^{19}F NMR (δ) -62.25 (s, 1F). HRMS: (API+) calculated for $\text{C}_{30}\text{H}_{38}\text{F}_3\text{O}$ ($[\text{M} + \text{H}]^+$) 471.2875, Found 471.2879.

2.1.18. (22E)-23-(2,4-dimethylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8q)

The general procedure with 2,4-dimethylstyrene and purification on HPLC afforded 81 mg (61%) of the title compound **20a** as a colorless oil: IR ν (cm^{-1}) 2935, 1701, 1652, 1593, 1560, 964. ^1H NMR (CDCl_3) δ 0.72, 0.74 (both s, 3H, CH_3); 1.13 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.80 (m, 2H); 1.96–2.04 (m, 4H); 2.07 (dt, 1H, $J = 12.5$, $J' = 3.4$ Hz); 2.22–2.37 (m, 4H); 2.28 (s, 6H, $2 \times \text{Ar-CH}_3$); 5.57 (m, 1H, H-3); 5.69 (m, 1H, H-2); 5.85 (dd, 1H, $J = 15.6$, $J' = 8.9$ Hz, H-22); 6.45 (d, 1H, $J = 15.6$ Hz, H-23); 6.94 (s, 1H, Ar-H); 6.95 (d, 1H, $J = 8.3$ Hz, Ar-H); 7.27 (d, 1H, $J = 8.3$ Hz, Ar-H). ^{13}C NMR δ 12.20 (C-18), 13.48 (C-19), 19.76 (Ar- CH_3), 20.57 (C-21), 20.97 (Ar- CH_3), 21.08, 21.70, 23.92, 28.28, 37.66, 39.32, 39.35, 40.02, 40.68, 42.85, 46.93, 53.37, 53.79, 55.78, 56.75, 124.48 (C-3), 124.93 (C-2), 125.09, 125.38, 126.64, 130.84, 134.29, 134.77, 136.28, 137.58, 211.98 (C-6). HRMS: (API+) calculated for $\text{C}_{31}\text{H}_{43}\text{O}$ ($[\text{M} + \text{H}]^+$) 431.3314, Found 431.3318.

2.1.19. (22E)-23-(2,5-dimethylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8r)

The general procedure with 2,5-dimethylstyrene and purification on HPLC afforded 77 mg (58%) of the title compound **21a** as a colorless oil: IR ν (cm^{-1}) 2931, 1699, 1653, 1593, 1560, 963. ^1H NMR (CDCl_3) δ 0.72, 0.74 (both s, 3H, CH_3); 1.13 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.80 (m, 2H); 1.96–2.04 (m, 4H); 2.08 (dt, 1H, $J = 12.5$, $J' = 3.3$ Hz); 2.22–2.31 (m, 2H); 2.27, 2.30 (both s, 3H, Ar- CH_3); 2.32–2.37 (m, 2H); 5.57 (m, 1H, H-3); 5.69 (m, 1H, H-2); 5.88 (dd, 1H, $J = 15.6$, $J' = 8.9$ Hz, H-22); 6.46 (d, 1H, $J = 15.6$ Hz, H-23); 6.92 (dd, 1H, $J = 7.6$, $J' = 1.2$ Hz, Ar-H); 7.00 (d, 1H, $J = 7.6$ Hz, Ar-H); 7.18 (b s, 1H, Ar-H). ^{13}C NMR δ 12.20 (C-18), 13.48 (C-19), 19.36 (Ar- CH_3), 20.57 (C-21), 21.00 (Ar- CH_3), 21.09, 21.70, 23.92, 28.31, 37.66, 39.33, 39.37, 40.02, 40.73, 42.87, 46.93, 53.37, 53.80, 55.74, 56.77, 124.48 (C-3), 124.94 (C-2), 125.37, 126.07, 127.43, 130.01, 131.89, 135.23, 136.88, 138.12, 211.95 (C-6). HRMS: (API+) calculated for $\text{C}_{31}\text{H}_{43}\text{O}$ ($[\text{M} + \text{H}]^+$) 431.3314, Found 431.3318.

2.1.20. (22E)-23-(4-*t*-butylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8s)

The general procedure with *p*-*t*-butylstyrene and purification on HPLC afforded 84 mg (60%) of the title compound **22a** as a colorless oil: IR ν (cm^{-1}) 2939, 1702, 1651, 1593, 1564, 1372, 963. ^1H NMR (CDCl_3) δ 0.73, 0.74 (both s, 3H, CH_3); 1.13 (d, 3H, $J = 6.7$ Hz, CH_3); 1.32 (s, 9H, *t*-Bu); 1.71–1.80 (m, 2H); 1.97–2.05 (m, 4H); 2.08 (dt, 1H, $J = 12.7$, $J' = 3.4$ Hz); 2.23–2.32 (m, 2H); 2.35–2.38 (m, 2H); 5.59 (m, 1H, H-3); 5.70 (m, 1H, H-2); 6.03 (dd, 1H, $J = 15.7$, $J' = 8.7$ Hz, H-22); 6.30 (d, 1H, $J = 15.7$ Hz, H-23); 7.27–7.29 (m, 2H, $2 \times \text{Ar-H}$); 7.32–7.34 (m, 2H, $2 \times \text{Ar-H}$). ^{13}C NMR δ 12.18 (C-18), 13.48 (C-19), 20.46 (C-21), 21.08, 21.70, 23.90, 28.17, 31.28 ($3 \times \text{C}$, *t*-Bu), 37.66, 39.32, 39.35, 40.02, 40.39, 42.86, 46.94, 53.38, 53.80, 55.84, 56.74, 124.48 (C-3), 124.94 (C-2), 125.35 ($2 \times \text{C}$), 125.60 ($2 \times \text{C}$), 127.09, 135.17, 136.17, 149.72, 211.97 (C-6). HRMS: (API+) calculated for $\text{C}_{33}\text{H}_{47}\text{O}$ ($[\text{M} + \text{H}]^+$) 459.3627, Found 459.3629.

2.1.21. (22E)-23-(4-phenylphenyl)-24-nor-5 α -chola-2,22-dien-6-one (8t)

The general procedure with *p*-phenylstyrene afforded 47 mg (32%) of the title compound **23a** as a colorless oil: IR ν (cm^{-1}) 2949, 1700, 1654, 1593, 1565, 1465, 960. ^1H NMR (CDCl_3) δ 0.72, 0.75 (both s, 3H, CH_3); 1.15 (d, 3H, $J = 6.7$ Hz, CH_3); 1.71–1.80 (m, 2H); 1.96–2.04 (m, 4H); 2.08 (dt, 1H, $J = 12.5$, $J' = 3.1$ Hz); 2.22–2.31 (m, 2H); 2.33–2.37 (m, 2H); 5.57 (m, 1H, H-3); 5.69 (m, 1H, H-2); 6.11 (dd, 1H, $J = 15.7$, $J' = 8.7$ Hz, H-22); 6.34 (d, 1H, $J = 15.7$ Hz, H-23); 7.33 (m, 1H, Ar-H); 7.39–7.44 (m, 4H, $4 \times \text{Ar-H}$); 7.52–7.54 (m, 2H, $2 \times \text{Ar-H}$); 7.57–7.60 (m, 2H, $2 \times \text{Ar-H}$). ^{13}C NMR δ 12.20 (C-18), 13.50 (C-19),

20.39 (C-21), 21.10, 21.70, 23.92, 28.21, 37.68, 39.34, 39.37, 40.04, 40.47, 42.93, 46.95, 53.39, 53.82, 55.84, 56.74, 124.50 (C-3), 124.96 (C-2), 126.32 (2 × C), 126.86 (2 × C), 126.98, 127.12, 127.16 (2 × C), 128.73 (2 × C), 137.04, 137.15, 139.51, 140.83, 211.98 (C-6). HRMS: (API+) calculated for C₃₅H₄₃O ([M+H]⁺) 479.3314, Found 479.3317.

2.1.22. (22E)-23-(2-isopropoxyphenyl)-24-nor-5α-chola-2,22-dien-6-one (9)

The general procedure with different *bis-ortho* substituted styrene or without styrene afforded 20 mg (14%) of the title compound **9** as a colorless oil: IR ν (cm⁻¹) 2940, 1701, 1653, 1591, 1561, 963. ¹H NMR (CDCl₃) δ 0.72, 0.74 (both s, 3H, CH₃); 1.14 (d, 3H, J = 6.7 Hz, CH₃); 1.34 (d, 3H, J = 6.1 Hz, CH(CH₃)₂); 1.35 (d, 3H, J = 6.1 Hz, CH(CH₃)₂); 1.71–1.81 (m, 2H); 1.96–2.03 (m, 4H); 2.08 (dt, 1H, J = 12.5, J' = 3.2 Hz); 2.22–2.31 (m, 2H); 2.33–2.37 (m, 2H); 4.49 (septet, 1H, J = 6.1 Hz, OCH(CH₃)₂); 5.58 (m, 1H, H-3); 5.69 (m, 1H, H-2); 6.03 (dd, 1H, J = 15.9, J' = 8.9 Hz, H-22); 6.62 (d, 1H, J = 15.9 Hz, H-23); 6.85–6.90 (m, 2H); 7.13 (m, 1H); 7.40 (dd, 1H, J = 7.6, J' = 1.5 Hz). ¹³C NMR δ 12.21 (C-18), 13.50 (C-19), 20.51 (C-21), 21.10, 21.70, 22.25 (2 × C, CH(CH₃)₂), 23.94, 28.21, 37.69, 39.34, 39.37, 40.05, 40.70, 42.89, 46.97, 53.41, 53.82, 55.94, 56.77, 71.04 (OCH(CH₃)₂), 114.70, 120.71, 122.32, 124.51 (C-3), 124.94 (C-2), 126.31, 127.49, 128.43, 137.13, 154.79, 212.05 (C-6). HRMS: (API+) calculated for C₃₂H₄₅O₂ ([M+H]⁺) 461.3420, Found 461.3423.

2.2. General procedure for dihydroxylation of dienes

To a solution of diene (0.2 mmol); hydroquinidine 4-chlorobenzoate (23 mg; 0.049 mmol); methansulfonamide (33 mg; 0.35 mmol); potassium carbonate (140 mg; 1.01 mmol); and potassium ferricyanide (350 mg; 1.06 mmol) in the mixture of *t*-butanol and water (15 mL; 1:1 v/v) was added 0.1 mL of osmium tetroxide in *t*-butanol (1 g/20 mL; 0.02 mmol). Reaction mixture was stirred at room temperature for 24 h. A saturated solution of sodium sulfite (3 mL) was then added. After an additional 30 min of stirring, the reaction mixture was diluted with ethyl acetate (30 mL) and extracted with water (2 × 20 mL). The combined organic fractions were dried over anhydrous magnesium sulfate and evaporated under reduced pressure. Column chromatography on silica gel (mobile phase – MeOH/CHCl₃ – 1/16) gave the desired product.

Compounds **10a** (m. p. 269–271 °C, EtOH), **10d** (m. p. 276–278 °C, EtOH), **10g** (m. p. 251–253 °C, EtOH), **10j** (m. p. 248–249 °C, *i*-PrOH), and **10m** (m. p. 271–273 °C, EtOH) were prepared according to general procedure which is also described in literature [11]. All data for these compounds are in agreement with same literature.

2.2.1. (22R, 23R)-2α,3α,22,23-tetrahydroxy-23-(2-fluorophenyl)-24-nor-5α-cholan-6-one (10b)

The general procedure for dihydroxylation of **8b** (60 mg; 0.14 mmol) afforded 54 mg (77%) of the title compound **10b** as a white solid: m. p. 288–290 °C (EtOH); IR ν (cm⁻¹) 3348 vb, 2941, 1710, 1620, 1587, 1492, 757. ¹H NMR (DMSO-*d*₆) δ 0.34 (H-18), 0.60 (H-19) (both s, 3H, CH₃); 0.86 (d, 3H, J = 6.4 Hz, H-21, CH₃); 1.80 (m, 1H); 1.87 (m, 1H); 1.98 (m, 1H); 2.07 (t, 1H, J = 12.6 Hz); 2.58 (d, 1H, J = 12.4 Hz); 3.44 (m, 1H, H-2); 3.58 (m, 1H, H-3); 3.74 (m, 1H, H-22); 4.18 (br s, 1H, OH); 4.31 (d, 1H, J = 5.8 Hz, OH); 4.62 (d, 1H, J = 1.8 Hz, OH); 4.66 (d, 1H, J = 8.3 Hz, H-23); 5.25 (br s, 1H, OH); 7.12 (m, 1H); 7.18 (m, 1H); 7.29 (m, 1H); 7.38 (m, 1H). ¹³C NMR δ 11.32 (C-18), 12.51 (C-19), 13.30 (C-21), 20.73, 23.27, 26.73, 27.21, 36.77, 36.98, 39.06, 40.10, 41.79, 42.00, 45.94, 50.25, 51.89, 52.86, 55.93, 67.08 (C-2), 67.45 (C-3), 68.48 (C-23), 75.65 (C-22), 114.92 (d, J = 22.8 Hz), 124.31 (d, J = 3.6 Hz), 128.82 (d, J = 3.6 Hz), 128.87, 130.01 (d, J = 14.2 Hz), 159.61 (d, J = 243.5 Hz, C-F), 211.39 (C-6). ¹⁹F NMR {¹H} δ -118.49 (s, 1F). HRMS: (API+) calculated for C₂₉H₄₂FO₅ ([M+H]⁺) 489.3016, Found 489.3017. Anal. Calcd for C₂₉H₄₁FO₅: C, 71.28; H, 8.46. Found: C, 71.30; H, 8.45%.

2.2.2. (22R, 23R)-2α,3α,22,23-tetrahydroxy-23-(3-fluorophenyl)-24-nor-5α-cholan-6-one (10c)

The general procedure for dihydroxylation of **8c** (60 mg; 0.14 mmol) afforded 59 mg (85%) of the title compound **10c** as a white solid: m. p. 278–279 °C (EtOH); IR ν (cm⁻¹) 3334 vb, 2939, 1710, 1615, 1591, 1461, 763. ¹H NMR (DMSO-*d*₆) δ 0.34 (H-18), 0.59 (H-19) (both s, 3H, CH₃); 0.86 (d, 3H, J = 6.4 Hz, H-21, CH₃); 1.80 (m, 1H); 1.86 (m, 1H); 1.98 (dd, 1H, J = 12.8, J' = 4.4 Hz); 2.07 (t, 1H, J = 12.8 Hz); 2.57 (dd, 1H, J = 12.0, J' = 3.0 Hz); 3.43–3.48 (m, 2H, H-2, H-3); 3.74 (m, 1H, H-22); 4.18 (d, 1H, J = 2.4 Hz, OH); 4.32 (d, 1H, J = 6.1 Hz, OH); 4.39 (dd, 1H, J = 7.9, J' = 3.4 Hz, H-23); 4.57 (d, 1H, J = 4.3 Hz, OH); 5.24 (d, 1H, J = 4.0 Hz, OH); 7.04–7.12 (m, 3H); 7.35 (m, 1H). ¹³C NMR δ 11.38 (C-18), 12.53 (C-19), 13.35 (C-21), 20.78, 23.28, 26.81, 27.25, 36.48, 37.01, 39.08, 40.09, 41.83, 42.05, 45.94, 50.28, 51.88, 52.85, 55.94, 67.09 (C-2), 67.48 (C-3), 74.56 (C-23), 76.17 (C-22), 113.46 (d, J = 21.6 Hz), 113.88 (d, J = 21.6 Hz), 123.05, 129.96 (d, J = 8.4 Hz), 146.53 (d, J = 7.2 Hz), 162.04 (d, J = 242.3 Hz, C-F), 211.54 (C-6). ¹⁹F NMR {¹H} δ -113.41 (s, 1F). HRMS: (API+) calculated for C₂₉H₄₂FO₅ ([M+H]⁺) 489.3016, Found 489.3017. Anal. Calcd for C₂₉H₄₁FO₅: C, 71.28; H, 8.46. Found: C, 71.26; H, 8.49%.

2.2.3. (22R, 23R)-2α,3α,22,23-tetrahydroxy-23-(2-chlorophenyl)-24-nor-5α-cholan-6-one (10e)

The general procedure for dihydroxylation of **8e** (60 mg; 0.14 mmol) afforded 54 mg (78%) of the title compound **10e** as a white solid: m. p. 290–292 °C (EtOH); IR ν (cm⁻¹) 3307 vb, 2936, 1693, 1595, 1572, 1463, 760. ¹H NMR (DMSO-*d*₆) δ 0.38 (H-18), 0.60 (H-19) (both s, 3H, CH₃); 0.93 (d, 3H, J = 6.7 Hz, H-21, CH₃); 1.79 (m, 1H); 1.89 (m, 1H); 1.98 (dd, 1H, J = 13.4, J' = 4.7 Hz); 2.07 (t, 1H, J = 12.3 Hz); 2.58 (dd, 1H, J = 12.3, J' = 3.4 Hz); 3.45 (m, 1H, H-2); 3.58 (bd, 1H, J = 8.0 Hz, H-3); 3.74 (m, 1H, H-22); 4.18 (bd, 1H, J = 2.4 Hz, OH); 4.31 (d, 1H, J = 6.1 Hz, OH); 4.58 (bd, 1H, J = 3.4 Hz, H-23); 4.79 (d, 1H, J = 7.3 Hz, OH); 5.28 (bs, 1H, OH); 7.26 (m, 1H); 7.34 (td, 1H, J = 7.5, J' = 1.2 Hz); 7.39 (dd, 1H, J = 7.9, J' = 1.2 Hz); 7.44 (dd, 1H, J = 7.8, J' = 1.7 Hz). ¹³C NMR δ 11.37 (C-18), 13.36 (C-19), 13.39 (C-21), 20.76, 23.33, 26.81, 27.40, 37.01, 37.36, 39.04, 40.10, 41.82, 42.08, 45.93, 50.26, 52.13, 52.84, 55.93, 67.09 (C-2), 67.47 (C-3), 71.00 (C-23), 75.66 (C-22), 127.19, 128.65, 128.92, 129.13, 131.92, 141.13, 211.54 (C-6). HRMS: (API+) calculated for C₂₉H₄₂ClO₅ ([M+H]⁺) 505.2721, Found 505.2720. Anal. Calcd for C₂₉H₄₁ClO₅: C, 68.96; H, 8.18. Found: C, 68.93; H, 8.19%.

2.2.4. (22R, 23R)-2α,3α,22,23-tetrahydroxy-23-(3-chlorophenyl)-24-nor-5α-cholan-6-one (10f)

The general procedure for dihydroxylation of **8f** (60 mg; 0.14 mmol) afforded 56 mg (80%) of the title compound **10f** as a white solid: m. p. 269–270 °C (EtOH); IR ν (cm⁻¹) 3257 vb, 2936, 1709, 1600, 1576, 1461, 781, 699. ¹H NMR (DMSO-*d*₆) δ 0.35 (H-18), 0.60 (H-19) (both s, 3H, CH₃); 0.86 (d, 3H, J = 6.7 Hz, H-21, CH₃); 1.79 (m, 1H); 1.87 (m, 1H); 1.96 (dd, 1H, J = 13.4, J' = 4.6 Hz); 2.07 (t, 1H, J = 12.3 Hz); 2.58 (dd, 1H, J = 12.3, J' = 3.2 Hz); 3.43–3.46 (m, 2H, H-2, H-3); 3.74 (m, 1H, H-22); 4.18 (bd, 1H, J = 2.1 Hz, OH); 4.33 (d, 1H, J = 5.8 Hz, OH); 4.38 (bd, 1H, J = 8.3 Hz, H-23); 4.59 (d, 1H, J = 3.4 Hz, OH); 5.26 (bs, 1H, OH); 7.23 (m, 1H); 7.29–7.31 (m, 2H); 7.35 (t, 1H, J = 7.6 Hz). ¹³C NMR δ 11.44 (C-18), 12.58 (C-19), 13.39 (C-21), 20.81, 23.33, 26.85, 27.31, 36.59, 37.06, 39.12, 40.11, 41.88, 42.09, 45.98, 50.33, 51.93, 52.90, 55.98, 67.13 (C-2), 67.52 (C-3), 74.56 (C-23), 76.16 (C-22), 125.73, 126.77, 127.12, 130.01, 132.75, 146.12, 211.64 (C-6). HRMS: (API+) calculated for C₂₉H₄₂ClO₅ ([M+H]⁺) 505.2721, Found 505.2716. Anal. Calcd for C₂₉H₄₁ClO₅: C, 68.96; H, 8.18. Found: C, 68.95; H, 8.23%.

2.2.5. (22R, 23R)-2α,3α,22,23-tetrahydroxy-23-(2-bromophenyl)-24-nor-5α-cholan-6-one (10h)

The general procedure for dihydroxylation of **8h** (60 mg;

0.13 mmol) afforded 55 mg (81%) of the title compound **10h** as a white solid: m. p. 288–289 °C (EtOH); IR ν (cm⁻¹) 3313 vb, 2944, 1699, 1593, 1564, 1471, 786, 682. ¹H NMR (DMSO-*d*₆) δ 0.39 (H-18), 0.60 (H-19) (both s, 3H, CH₃); 0.95 (d, 3H, *J* = 6.7 Hz, H-21, CH₃); 1.78 (m, 1H); 1.89 (m, 1H); 1.98 (dd, 1H, *J* = 13.1, *J'* = 4.6 Hz); 2.07 (t, 1H, *J* = 12.3 Hz); 2.58 (dd, 1H, *J* = 12.3, *J'* = 3.2 Hz); 3.46 (m, 1H, H-2); 3.58 (bd, 1H, *J* = 7.0 Hz, H-3); 3.74 (m, 1H, H-22); 4.20 (d, 1H, *J* = 1.8 Hz, OH); 4.34 (d, 1H, *J* = 5.8 Hz, OH); 4.58 (m, 1H, H-23); 4.73 (d, 1H, *J* = 7.0 Hz, OH); 5.31 (bs, 1H, OH); 7.18 (m, 1H); 7.38 (m, 1H); 7.42 (dd, 1H, *J* = 7.9, *J'* = 1.8 Hz); 7.55 (dd, 1H, *J* = 7.9, *J'* = 1.2 Hz). ¹³C NMR δ 11.45 (C-18), 13.40 (C-19), 13.69 (C-21), 20.81, 23.40, 26.86, 27.47, 37.07, 37.60, 39.10, 40.10, 41.89, 42.15, 45.98, 50.34, 52.22, 52.89, 55.97, 67.13 (C-2), 67.52 (C-3), 73.47 (C-23), 75.65 (C-22), 122.72, 127.76, 129.08, 129.46, 132.26, 142.83, 211.68 (C-6). HRMS: (API+) calculated for C₂₉H₄₂⁷⁹BrO₅ ([M+H]⁺) 549.2216, Found 549.2211. Anal. Calcd for C₂₉H₄₁BrO₅: C, 63.38; H, 7.52. Found: C, 63.35; H, 7.56%.

2.2.6. (22*R*, 23*R*)-2 α ,3 α ,22,23-tetrahydroxy-23-(3-bromophenyl)-24-nor-5 α -cholan-6-one (**10i**)

The general procedure for dihydroxylation of **8i** (60 mg; 0.13 mmol) afforded 54 mg (79%) of the title compound **10i** as a white solid: m. p. 264–265 °C (EtOH); IR ν (cm⁻¹) 3242 vb, 2935, 1709, 1596, 1572, 1460, 782, 700. ¹H NMR (DMSO-*d*₆) δ 0.35 (H-18), 0.60 (H-19) (both s, 3H, CH₃); 0.86 (d, 3H, *J* = 6.7 Hz, H-21, CH₃); 1.79 (m, 1H); 1.87 (m, 1H); 1.98 (dd, 1H, *J* = 13.1, *J'* = 4.8 Hz); 2.07 (t, 1H, *J* = 12.3 Hz); 2.58 (dd, 1H, *J* = 12.3, *J'* = 3.2 Hz); 3.42–3.46 (m, 2H, H-2, H-3); 3.74 (m, 1H, H-22); 4.19 (bs, 1H, OH); 4.33 (bd, 1H, *J* = 5.5 Hz, OH); 4.37 (bd, 1H, *J* = 8.3 Hz, H-23); 4.59 (bs, 1H, OH); 5.26 (bs, 1H, OH); 7.26–7.30 (m, 2H); 7.43 (m, 1H); 7.45 (m, 1H). ¹³C NMR δ 11.44 (C-18), 12.58 (C-19), 13.38 (C-21), 20.80, 23.32, 26.84, 27.30, 36.60, 37.05, 39.11, 40.10, 41.87, 42.09, 45.97, 50.32, 51.93, 52.89, 55.97, 67.12 (C-2), 67.51 (C-3), 74.52 (C-23), 76.15 (C-22), 121.39, 126.10, 129.69, 130.00, 130.33, 146.38, 211.62 (C-6). HRMS: (API+) calculated for C₂₉H₄₂⁷⁹BrO₅ ([M+H]⁺) 549.2216, Found 549.2211. Anal. Calcd for C₂₉H₄₁BrO₅: C, 63.38; H, 7.52. Found: C, 63.32; H, 7.59%.

2.2.7. (22*R*, 23*R*)-2 α ,3 α ,22,23-tetrahydroxy-23-(2-methylphenyl)-24-nor-5 α -cholan-6-one (**10k**)

The general procedure for dihydroxylation of **8k** (60 mg; 0.14 mmol) afforded 53 mg (76%) of the title compound **10k** as a white solid: m. p. 269–271 °C (EtOH); IR ν (cm⁻¹) 3300 vb, 2942, 1706, 1631, 1604, 1461, 981, 759. ¹H NMR (DMSO-*d*₆) δ 0.29 (H-18), 0.55 (H-19) (both s, 3H, CH₃); 0.81 (d, 3H, *J* = 6.4 Hz, H-21, CH₃); 1.76 (m, 1H); 1.83 (m, 1H); 1.97 (dd, 1H, *J* = 13.1, *J'* = 4.9 Hz); 2.02 (t, 1H, *J* = 12.3 Hz); 2.26 (s, 3H, CH₃); 2.58 (dd, 1H, *J* = 12.3, *J'* = 3.2 Hz); 3.40 (m, 1H, H-2); 3.59 (dd, 1H, *J* = 8.1, *J'* = 2.6 Hz, H-3); 3.69 (m, 1H, H-22); 4.13 (d, 1H, *J* = 2.1 Hz, OH); 4.27 (d, 1H, *J* = 5.8 Hz, OH); 4.44 (d, 1H, *J* = 4.0 Hz, OH); 4.52 (dd, 1H, *J* = 8.3, *J'* = 1.8 Hz, H-23); 4.94 (d, 1H, *J* = 3.4 Hz, OH); 7.06–7.08 (m, 2H); 7.10 (m, 1H); 7.20 (d, 1H, *J* = 7.3 Hz). ¹³C NMR δ 11.42 (C-18), 13.25 (C-19), 13.37 (C-21), 19.32 (Ar-CH₃), 20.79, 23.35, 26.83, 27.35, 36.79, 37.04, 39.07, 40.10, 41.86, 42.09, 45.96, 50.31, 52.21, 52.87, 55.95, 67.11 (C-2), 67.50 (C-3), 71.41 (C-23), 75.50 (C-22), 125.74, 126.82, 127.38, 130.15, 135.06, 141.42, 211.61 (C-6). HRMS: (API+) calculated for C₃₀H₄₅O₅ ([M+H]⁺) 485.3267, Found 485.3268. Anal. Calcd for C₃₀H₄₄O₅: C, 74.34; H, 9.15. Found: C, 74.29; H, 9.19%.

2.2.8. (22*R*, 23*R*)-2 α ,3 α ,22,23-tetrahydroxy-23-(3-methylphenyl)-24-nor-5 α -cholan-6-one (**10l**)

The general procedure for dihydroxylation of **8l** (60 mg; 0.14 mmol) afforded 55 mg (79%) of the title compound **10l** as a white solid: m. p. 276–278 °C (EtOH); IR ν (cm⁻¹) 3315 vb, 2938, 1708, 1611, 1593, 1460, 988, 707. ¹H NMR (DMSO-*d*₆) δ 0.32 (H-18), 0.60 (H-19) (both s, 3H, CH₃); 0.85 (d, 3H, *J* = 6.4 Hz, H-21, CH₃); 1.80 (m, 1H); 1.87 (m, 1H); 1.98 (dd, 1H, *J* = 13.2, *J'* = 4.9 Hz); 2.07 (t, 1H, *J* = 12.6 Hz);

2.28 (s, 3H, CH₃); 2.58 (dd, 1H, *J* = 12.1, *J'* = 3.2 Hz); 3.44–3.49 (m, 2H, H-2, H-3); 3.74 (m, 1H, H-22); 4.18 (d, 1H, *J* = 2.1 Hz, OH); 4.30–4.33 (m, 2H, H-23, OH); 4.46 (d, 1H, *J* = 4.0 Hz, OH); 5.03 (d, 1H, *J* = 3.4 Hz, OH); 7.03–7.05 (m, 2H); 7.07 (s, 1H); 7.19 (t, 1H, *J* = 7.5 Hz). ¹³C NMR δ 11.45 (C-18), 12.51 (C-19), 13.37 (C-21), 18.59 (Ar-CH₃), 20.80, 23.32, 26.83, 27.23, 36.35, 37.04, 39.12, 40.10, 41.86, 42.08, 45.97, 50.31, 51.91, 52.88, 55.97, 67.11 (C-2), 67.50 (C-3), 75.15 (C-23), 76.21 (C-22), 124.15, 127.67, 127.83, 127.94, 136.99, 143.20, 211.62 (C-6). HRMS: (API+) calculated for C₃₀H₄₅O₅ ([M+H]⁺) 485.3267, Found 485.3269. Anal. Calcd for C₃₀H₄₄O₅: C, 74.34; H, 9.15. Found: C, 74.31; H, 9.16%.

2.2.9. (22*R*, 23*R*)-2 α ,3 α ,22,23-tetrahydroxy-23-(2-trifluoromethylphenyl)-24-nor-5 α -cholan-6-one (**10n**)

The general procedure for dihydroxylation of **8n** (60 mg; 0.13 mmol) afforded 50 mg (73%) of the title compound **10n** as a white solid: m. p. 268–270 °C (EtOH); IR ν (cm⁻¹) 3418 vb, 2939, 1701, 1613, 1589, 1315, 1114, 772. ¹H NMR (DMSO-*d*₆) δ 0.38 (H-18), 0.60 (H-19) (both s, 3H, CH₃); 0.87 (d, 3H, *J* = 6.4 Hz, H-21, CH₃); 1.82 (m, 1H); 1.88 (m, 1H); 1.99 (dd, 1H, *J* = 13.1, *J'* = 4.6 Hz); 2.07 (t, 1H, *J* = 12.6 Hz); 2.58 (dd, 1H, *J* = 12.2, *J'* = 3.1 Hz); 3.45 (m, 1H, H-2); 3.66 (m, 1H, H-3); 3.74 (bs, 1H, H-22); 4.18 (d, 1H, *J* = 2.1 Hz, OH); 4.31 (d, 1H, *J* = 6.1 Hz, OH); 4.63 (d, 1H, *J* = 4.3 Hz, H-23); 4.70 (d, 1H, *J* = 4.6 Hz, OH); 5.32 (bs, 1H, OH); 7.46 (m, 1H); 7.65–7.70 (m, 3H). ¹³C NMR δ 11.31 (C-18), 12.83 (C-19), 13.35 (C-21), 20.75, 23.34, 26.81, 27.38, 37.00, 37.77, 39.04, 40.10, 41.82, 42.11, 45.91, 50.28, 52.32, 52.84, 55.91, 67.09 (C-2), 67.47 (C-3), 69.76 (C-23), 75.19 (C-22), 124.49 (q, *J* = 274.7 Hz, CF₃), 126.26 (q, *J* = 28.8 Hz), 125.21 (q, *J* = 5.0 Hz), 127.59, 129.38, 132.39, 142.56, 211.54 (C-6). ¹⁹F NMR {¹H} δ -55.95 (s, 1F). HRMS: (API+) calculated for C₃₀H₄₂F₃O₅ ([M+H]⁺) 539.2984, Found 539.2986. Anal. Calcd for C₃₀H₄₁F₃O₅: C, 66.89; H, 7.67. Found: C, 66.88; H, 7.69%.

2.2.10. (22*R*, 23*R*)-2 α ,3 α ,22,23-tetrahydroxy-23-(3-trifluoromethylphenyl)-24-nor-5 α -cholan-6-one (**10o**)

The general procedure for dihydroxylation of **8o** (60 mg; 0.13 mmol) afforded 53 mg (77%) of the title compound **10o** as a white solid: m. p. 243–245 °C (EtOH); IR ν (cm⁻¹) 3257 vb, 2934, 1712, 1618, 1457, 1329, 1113, 706. ¹H NMR (DMSO-*d*₆) δ 0.32 (H-18), 0.59 (H-19) (both s, 3H, CH₃); 0.86 (d, 3H, *J* = 6.4 Hz, H-21, CH₃); 1.79 (m, 1H); 1.86 (m, 1H); 1.97 (dd, 1H, *J* = 13.1, *J'* = 4.6 Hz); 2.07 (t, 1H, *J* = 12.6 Hz); 2.57 (dd, 1H, *J* = 12.2, *J'* = 3.3 Hz); 3.42–3.49 (m, 2H, H-2, H-3); 3.74 (bs, 1H, H-22); 4.19 (d, 1H, *J* = 1.8 Hz, OH); 4.33 (d, 1H, *J* = 5.5 Hz, OH); 4.50 (d, 1H, *J* = 7.9 Hz, OH); 4.65 (bs, 1H, H-23); 5.35 (bs, 1H, OH); 7.54–7.62 (m, 4H). ¹³C NMR δ 11.30 (C-18), 12.59 (C-19), 13.36 (C-21), 20.78, 23.30, 26.83, 27.22, 36.54, 37.01, 39.08, 40.10, 41.84, 42.09, 45.94, 50.31, 51.93, 52.86, 55.94, 67.11 (C-2), 67.50 (C-3), 74.52 (C-23), 76.23 (C-22), 123.38 (q, *J* = 4.0 Hz), 123.84 (q, *J* = 4.0 Hz), 124.37 (q, *J* = 272.3 Hz, CF₃), 128.77 (q, *J* = 31.8 Hz), 129.14, 131.12, 145.01, 211.59 (C-6). ¹⁹F NMR {¹H} δ -60.97 (s, 1F). HRMS: (API+) calculated for C₃₀H₄₂F₃O₅ ([M+H]⁺) 539.2984, Found 539.2985. Anal. Calcd for C₃₀H₄₁F₃O₅: C, 66.89; H, 7.67. Found: C, 66.85; H, 7.71%.

2.2.11. (22*R*, 23*R*)-2 α ,3 α ,22,23-tetrahydroxy-23-(4-trifluoromethylphenyl)-24-nor-5 α -cholan-6-one (**10p**)

The general procedure for dihydroxylation of **8p** (60 mg; 0.13 mmol) afforded 54 mg (78%) of the title compound **10p** as a white solid: m. p. 265–266 °C (EtOH); IR ν (cm⁻¹) 3224 vb, 2937, 1710, 1621, 1461, 1331, 1121, 838. ¹H NMR (DMSO-*d*₆) δ 0.33 (H-18), 0.59 (H-19) (both s, 3H, CH₃); 0.86 (d, 3H, *J* = 6.4 Hz, H-21, CH₃); 1.80 (m, 1H); 1.86 (m, 1H); 1.98 (dd, 1H, *J* = 13.1, *J'* = 4.6 Hz); 2.07 (t, 1H, *J* = 12.6 Hz); 2.58 (dd, 1H, *J* = 12.2, *J'* = 3.1 Hz); 3.44 (m, 1H, H-2); 3.50 (d, 1H, *J* = 7.6 Hz, H-3); 3.74 (bs, 1H, H-22); 4.19 (bs, 1H, OH); 4.33 (d, 1H, *J* = 4.6 Hz, OH); 4.47 (d, 1H, *J* = 8.3 Hz, H-23); 4.66 (bs, 1H, OH); 5.34 (bs, 1H, OH); 7.50 (d, 2H, *J* = 8.1 Hz); 7.68 (d, 2H,

$J = 8.1$ Hz). ^{13}C NMR δ 11.41 (C-18), 12.57 (C-19), 13.36 (C-21), 20.80, 23.29, 26.83, 27.26, 36.61, 37.03, 39.11, 40.10, 41.85, 42.07, 45.96, 50.31, 51.92, 52.87, 55.95, 67.11 (C-2), 67.50 (C-3), 74.68 (C-23), 76.10 (C-22), 124.40 (q, $J = 271.9$ Hz, CF_3), 124.96 (q, $2 \times \text{C}$, $J = 3.6$ Hz), 127.70 (q, $J = 31.2$ Hz), 127.80 ($2 \times \text{C}$), 148.29, 211.59 (C-6). ^{19}F NMR $\{^1\text{H}\}$ δ -60.59 (s, 1F). HRMS: (API+) calculated for $\text{C}_{30}\text{H}_{42}\text{F}_3\text{O}_5$ ($[\text{M} + \text{H}]^+$) 539.2984, Found 539.2986. Anal. Calcd for $\text{C}_{30}\text{H}_{41}\text{F}_3\text{O}_5$: C, 66.89; H, 7.67. Found: C, 66.86; H, 7.70%.

2.2.12. (22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-(2,4-dimethylphenyl)-24-nor-5 α -cholan-6-one (10q)

The general procedure for dihydroxylation of **8q** (60 mg; 0.14 mmol) afforded 52 mg (75%) of the title compound **10q** as a white solid: m. p. 253–255 °C (EtOH); IR ν (cm^{-1}) 3350 vb, 2943, 1693, 1617, 1502, 1458, 987, 817. ^1H NMR (DMSO- d_6) δ 0.35 (H-18), 0.60 (H-19) (both s, 3H, CH_3); 0.85 (d, 3H, $J = 6.4$ Hz, H-21, CH_3); 1.78–1.88 (m, 2H); 1.98 (dd, 1H, $J = 13.1$, $J' = 4.6$ Hz); 2.07 (t, 1H, $J = 12.5$ Hz); 2.22 (s, 3H, Ar- CH_3); 2.26 (s, 3H, Ar- CH_3); 2.58 (dd, 1H, $J = 12.1$, $J' = 3.0$ Hz); 3.44 (m, 1H, H-2); 3.63 (dd, 1H, $J = 8.3$ Hz, H-3); 3.74 (m, 1H, H-22); 4.19 (bs, 1H, OH); 4.33 (bs, 1H, OH); 4.46 (bs, 1H, OH); 4.52 (d, 1H, $J = 8.3$ Hz, H-23); 4.89 (bs, 1H, OH); 6.93 (s, 1H); 6.96 (d, 1H, $J = 7.9$ Hz); 7.12 (d, 1H, $J = 7.9$ Hz). ^{13}C NMR δ 11.50 (C-18), 13.13 (C-19), 13.38 (C-21), 19.25 (Ar- CH_3), 20.65 (Ar- CH_3), 20.80, 23.36, 26.84, 27.36, 36.79, 37.05, 39.09, 40.10, 41.87, 42.09, 45.98, 50.31, 52.17, 52.88, 55.97, 67.12 (C-2), 67.51 (C-3), 71.26 (C-23), 75.44 (C-22), 126.41, 127.40, 130.84, 134.94, 135.62, 138.30, 211.64 (C-6). HRMS: (API+) calculated for $\text{C}_{31}\text{H}_{47}\text{O}_5$ ($[\text{M} + \text{H}]^+$) 499.3423, Found 499.3426. Anal. Calcd for $\text{C}_{31}\text{H}_{46}\text{O}_5$: C, 74.66; H, 9.30. Found: C, 74.63; H, 9.32%.

2.2.13. (22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-(2,5-dimethylphenyl)-24-nor-5 α -cholan-6-one (10r)

The general procedure for dihydroxylation of **8r** (60 mg; 0.14 mmol) afforded 49 mg (71%) of the title compound **10r** as a white solid: m. p. 281–283 °C (EtOH); IR ν (cm^{-1}) 3332 vb, 2933, 1707, 1616, 1506, 1461, 985. ^1H NMR (DMSO- d_6) δ 0.36 (H-18), 0.60 (H-19) (both s, 3H, CH_3); 0.86 (d, 3H, $J = 6.4$ Hz, H-21, CH_3); 1.82 (m, 1H); 1.87 (m, 1H); 1.98 (dd, 1H, $J = 13.1$, $J' = 4.6$ Hz); 2.07 (t, 1H, $J = 12.6$ Hz); 2.23 (s, 3H, Ar- CH_3); 2.24 (s, 3H, Ar- CH_3); 2.58 (dd, 1H, $J = 12.0$, $J' = 3.1$ Hz); 3.45 (m, 1H, H-2); 3.61 (bd, 1H, $J = 7.9$ Hz, H-3); 3.74 (bs, 1H, H-22); 4.20 (bs, 1H, OH); 4.34 (d, 1H, $J = 5.2$ Hz, OH); 4.47 (bs, 1H, OH); 4.52 (d, 1H, $J = 7.9$ Hz, H-23); 4.94 (bs, 1H, OH); 6.92 (dd, 1H, $J = 7.6$, $J' = 1.2$ Hz); 6.99 (d, 1H, $J = 7.6$ Hz); 7.06 (bs, 1H). ^{13}C NMR δ 11.47 (C-18), 13.29 (C-19), 13.40 (C-21), 18.92 (Ar- CH_3), 20.81, 20.85 (Ar- CH_3), 23.42, 26.86, 27.32, 36.91, 37.08, 39.10, 40.11, 41.88, 42.13, 45.98, 50.34, 52.25, 52.91, 55.98, 67.14 (C-2), 67.52 (C-3), 71.45 (C-23), 75.53 (C-22), 127.46, 127.99, 130.08, 131.91, 134.34, 141.26, 211.68 (C-6). HRMS: (API+) calculated for $\text{C}_{31}\text{H}_{47}\text{O}_5$ ($[\text{M} + \text{H}]^+$) 499.3423, Found 499.3425. Anal. Calcd for $\text{C}_{31}\text{H}_{46}\text{O}_5$: C, 74.66; H, 9.30. Found: C, 74.61; H, 9.33%.

2.2.14. (22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-(4-*t*-butylphenyl)-24-nor-5 α -cholan-6-one (10s)

The general procedure for dihydroxylation of **8s** (60 mg; 0.13 mmol) afforded 53 mg (77%) of the title compound **10s** as a white solid: m. p. 255–256 °C (EtOH); IR ν (cm^{-1}) 3234 vb, 2938, 1709, 1616, 1512, 1461, 1084, 991, 831. ^1H NMR (DMSO- d_6) δ 0.32 (H-18), 0.59 (H-19) (both s, 3H, CH_3); 0.85 (d, 3H, $J = 6.4$ Hz, H-21, CH_3); 1.26 (s, 3H, *t*-Bu); 1.77–1.88 (m, 2H); 1.98 (dd, 1H, $J = 13.1$, $J' = 4.6$ Hz); 2.07 (t, 1H, $J = 12.6$ Hz); 2.58 (dd, 1H, $J = 12.3$, $J' = 3.4$ Hz); 3.44 (m, 1H, H-2); 3.51 (d, 1H, $J = 8.3$ Hz, H-3); 3.74 (m, 1H, H-22); 4.20 (vb s, 2H, $2 \times \text{OH}$); 4.32 (d, 1H, $J = 8.3$ Hz, H-23); 4.42 (vb s, 1H, OH); 4.99 (vb s, 1H, OH); 7.18 (d, 2H, $J = 8.5$ Hz); 7.33 (d, 2H, $J = 8.5$ Hz). ^{13}C NMR δ 11.44 (C-18), 12.48 (C-19), 13.33 (C-21), 20.77, 23.29, 26.81, 27.13, 31.19 ($3 \times \text{C}$, *t*-Bu), 34.21, 36.33, 37.00, 39.10, 40.10, 41.81, 42.03, 45.94, 50.27, 51.95, 52.84, 55.92, 67.08 (C-

2), 67.47 (C-3), 74.95 (C-23), 76.05 (C-22), 124.72 ($2 \times \text{C}$), 126.75 ($2 \times \text{C}$), 140.23, 149.41, 211.53 (C-6). HRMS: (API+) calculated for $\text{C}_{33}\text{H}_{51}\text{O}_5$ ($[\text{M} + \text{H}]^+$) 527.3736, Found 527.3734. Anal. Calcd for $\text{C}_{33}\text{H}_{50}\text{O}_5$: C, 75.25; H, 9.57. Found: C, 75.22; H, 9.61%.

2.2.15. (22R, 23R)-2 α ,3 α ,22,23-tetrahydroxy-23-(4-phenylphenyl)-24-nor-5 α -cholan-6-one (10t)

The general procedure for dihydroxylation of **8t** (40 mg; 0.08 mmol) afforded 33 mg (72%) of the title compound **10t** as a white solid: m. p. 251–253 °C (EtOH); IR ν (cm^{-1}) 3328 vb, 2955, 1705, 1643, 1600, 1580, 1465, 967, 763. ^1H NMR (DMSO- d_6) δ 0.34 (H-18), 0.59 (H-19) (both s, 3H, CH_3); 0.89 (d, 3H, $J = 6.4$ Hz, H-21, CH_3); 1.78–1.89 (m, 2H); 1.97 (dd, 1H, $J = 13.1$, $J' = 4.6$ Hz); 2.07 (t, 1H, $J = 12.6$ Hz); 2.58 (dd, 1H, $J = 12.1$, $J' = 3.0$ Hz); 3.44 (m, 1H, H-2); 3.54 (d, 1H, $J = 8.6$ Hz, H-3); 3.74 (bs, 1H, H-22); 4.19 (bs, 1H, OH); 4.33 (bs, 1H, OH); 4.41 (d, 1H, $J = 8.9$ Hz, H-23); 4.55 (bs, 1H, OH); 5.15 (bs, 1H, OH); 7.33–7.37 (m, 3H); 7.45 (t, 2H, $J = 7.6$ Hz); 7.64 (d, 2H, $J = 8.3$ Hz); 7.68 (t, 2H, $J = 8.3$ Hz). ^{13}C NMR δ 11.49 (C-18), 12.52 (C-19), 13.35 (C-21), 20.79, 23.29, 26.81, 27.27, 36.52, 37.02, 39.12, 40.09, 41.83, 42.05, 45.95, 50.28, 51.92, 52.86, 55.95, 67.10 (C-2), 67.48 (C-3), 74.87 (C-23), 76.17 (C-22), 126.26 ($2 \times \text{C}$), 126.51 ($2 \times \text{C}$), 127.38, 127.67 ($2 \times \text{C}$), 128.96 ($2 \times \text{C}$), 138.74, 139.80, 142.58, 211.57 (C-6). HRMS: (API+) calculated for $\text{C}_{35}\text{H}_{47}\text{O}_5$ ($[\text{M} + \text{H}]^+$) 547.3423, Found 547.3421. Anal. Calcd for $\text{C}_{35}\text{H}_{46}\text{O}_5$: C, 76.89; H, 8.48. Found: C, 76.85; H, 8.53%.

3. Molecular docking

Docking was performed prediction of conformation and energy ranking between BRI1 receptor (PDBID: 3RGZ) and. Docking studies was carried out using AutoDock Vina 1.05 [13]. All 3D structures of BRI1 ligands were prepared with Marvin 5.10.3 [10], software which can be used for drawing, displaying and characterization of chemical structure, substructures and reactions. Ligands were prepared as derivatives of natural ligand brassinolide (BLD). Polar hydrogens were added to all ligands and proteins with the AutoDock Tools (ADT) [14] program prior to docking with Autodock Vina program. Grid box with size of 40 Å were centered on active site of protein. Exhaustiveness parameter was set to 20 (default 8). After docking we compared docked ligand with brassinolide crystal-like poses and the best crystal-like poses of each ligand were analyzed.

4. The pea inhibition biotest

Pea seedlings (*Pisum arvense* L. sort Arvica) germinating for 2 days were selected for uniformity from a large population and then transferred into pots containing perlite and 1/10 diluted Hoagland solution (half concentration, pH 5.7) After 24 h in a dark cultivation room (24 °C, humidity 75%) the seedlings were treated with different amounts of tested compounds in 5 μl fractionated lanolin. The substances were applied as microdrops to the scar left after the removal of bract. The control plants were treated with lanolin alone. At least seven plants were used for each experiment and the assays were repeated at least three times. The inhibition of etiolated pea stems were measured after 4 days and the difference in length between treated and control plants provided a measure of activity. For each treatment, 8 seedlings were analyzed in two biological replicates. The mean values were subjected to the statistical analysis using the Student's *t* test.

5. Determination of ethylene production

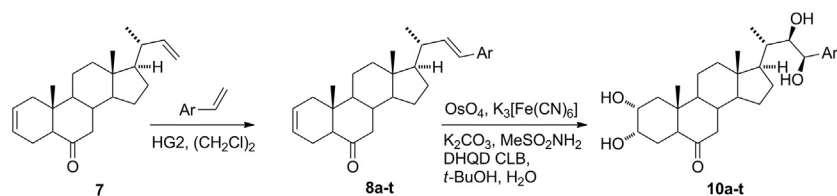
For measurement of ethylene production, pea seedlings (8 plants/ tested amount of substance) were placed in a 0.5 L glass container for 24 h in the dark. One milliliter of headspace gas from the chamber was withdrawn for each measurement and injected into a gas chromatograph (Agilent Technologies, GC System, USA) equipped with a flame

Table 1
Isolated yields of products in alkene cross-metathesis.

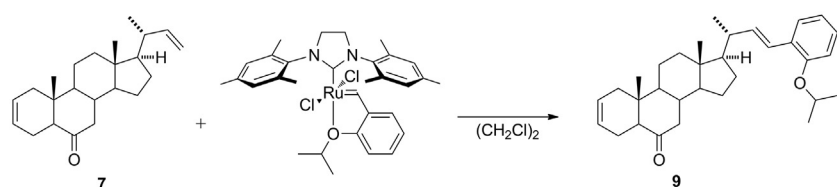
Entry	Ar	Yield (%)	Entry	Ar	Yield (%)
a		81	l		69
b		65	m		71
c		73	n		57
d		75	o		66
e		64	p		71
f		70	q		61
g		75	r		58
h		59	s		60
i		66	t		32
j		73	u		na
k		65	v		na

na – compound not isolated.

ionic detector (FID) and a capillary column (HP-AL/S stationary phase, 15 lm, i.d. = 0.535). The chromatographic analytical parameters were as follows: column temperature: 150 °C; detector temperature: 220 °C;



Scheme 1. Preparation of aryl-dienones and aryl-tetraols. For Ar see Table 1.



Scheme 2. Formation of isopropoxy derivative 9 if bis-ortho substituted or no styrenes are used in the reaction.

and helium was used as carrier gas. The area under the resultant peak (y -axis) versus sensitivity (x -axis; $nL \cdot mL^{-1}$) was representing a quantitative measure of ethylene concentration. The measurements were done in triplicates and data were statistically analyzed using the Student's t test.

6. Arabidopsis brassinosteroid sensitivity assay

Arabidopsis thaliana L. (Heyhn.) (Columbia ecotype, Col-0; referred to Arabidopsis) seedlings were stratified for 2 d at 4 °C and germinated on vertical half-strength Murashige and Skoog (1% w/v sucrose) agar plates with different concentrations of 24-epiBL and BR derivatives at 22 °C in a 16 h/8 h light-dark cycle for 5 d. Roots were then straightened on solid media plates, scanned with an Epson high-resolution scanner and the entire root length measured with ImageJ (<http://rsbweb.nih.gov/ij/>). For each treatment, more than 15 seedlings were analyzed in two biological repeats. P values were calculated with a two-tailed Student t -test using Excel software.

7. Cytotoxicity

Calcein AM cytotoxicity assay was performed exactly as described earlier [11].

8. Results and discussion

8.1. Chemistry

For the preparation of above mentioned compounds we started with the known [15] 24-nor-5 α -chola-2,22-dien-6-one (7). With this compound cross-metathesis were carried out using twenty commercially available styrenes substituted with fluorine, chlorine, bromine, trifluoromethyl, alkyls and phenyl group. The reaction was catalyzed by Hoveyda-Grubbs second generation catalyst in refluxing dichloroethane. This catalyst has proven to be efficient for different steroid side chain cross-metatheses [15–18]. Due to the fact that compound 7 and used styrenes belongs to the group of type I olefins (non-, meta- and para- substituted styrenes) and type II olefins (ortho- substituted styrenes) if Hoveyda-Grubbs second generation catalyst is used [19], the reactions proceed in two steps. Firstly, the homodimerization of styrene with strong release of ethylene takes place followed by the secondary cross-metathesis of steroid and *in-situ* formed stilbene. This is the main reason of long reaction time. In most cases, the cross-metathesis gave the corresponding 23-aryl products in good isolated yields 60–80% (see Scheme 1). Lower yield is typical for ortho-substituted styrenes as their homodimers are sparingly consumable during secondary cross-metathesis. In the case of 4-phenylstyrene the low yield (32%) is caused by formation of poorly soluble 4,4'-diphenylstilbene

Table 2

IC₅₀ (mol/L) values obtained from the pea inhibition biotest. Compounds **10a**, **10d**, **10g**, **10j**, and **10m** are presented for comparison; the results were published [11].

Compound	IC ₅₀ (mol/L)
24-epibrassinolide	1.66×10^{-5}
10f	1.80×10^{-5}
10b	2.15×10^{-5}
10n	2.30×10^{-5}
10a	2.52×10^{-6}
10d	2.0×10^{-6}
10g	1.8×10^{-5}
10j	2.7×10^{-4}
10m	1.8×10^{-4}

during homodimerization step. No desired products were observed if bis-*ortho* substituted styrenes were used. In these cases we were able to isolate only very small amount of 2-isopropoxyphenyl derivative **9** as a product of reaction between steroid olefin and Hoveyda-Grubbs catalyst (see Scheme 2). This was probably caused by steric hindrance near steroidal terminal double bond. All products **8a-t** were obtained as *trans* double bond isomers as indicated by NMR analysis.

Next, Sharpless simultaneous dihydroxylation of both double bonds was used to minimize formation of unnatural configuration of 22 and 23 hydroxy groups. Hydroquinidine 4-chlorobenzoate was used as chiral ligand. The reaction rate was increased by addition of methanesulfonamide [20]. Such reaction conditions led only to the desired 22*R*,23*R*-isomers **10a-t** in very good yields (75–83%).

All compounds were characterized by NMR, IR and MS techniques together with elemental analysis for all tetraols. Compounds **10a**, **10d**, **10g**, **10j**, and **10m** were also prepared by different synthetic strategy [11] (see Table 1).

8.2. Biological activity and docking

Biological activity of new BRs derivatives was monitored by pea inhibition biotest. This test is based on that BRs inhibit the growth of etiolated seedlings at high concentration and the inhibition is probably caused by ethylene production which is mediated by BRs. Dose response curves for all prepared BRs derivatives are shown in SFig. 1. The IC₅₀ values obtained from the pea inhibition biotest are summarized in Table 2. The most active BRs derivatives were **10f**, **10b**, and **10n** (IC₅₀ 1.8×10^{-5} – 2.3×10^{-5} mol.L⁻¹) compared to 24-epibrassinolide (IC₅₀ 1.6×10^{-5} mol.L⁻¹), used as a positive control (Fig. 2).

The crossstalk of BRs and ethylene regulates various aspects of plant growth and development. Hansen et al. [21] showed that exogenously

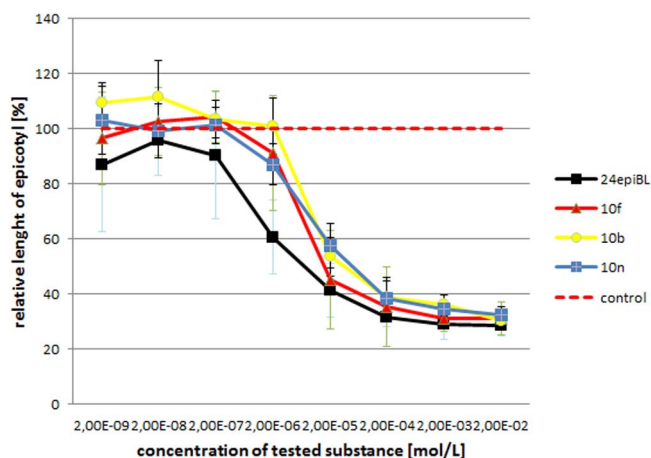


Fig. 2. Effect of selected brassinosteroid derivatives on the inhibition of etiolated pea seedlings. Error bars represent s.d.

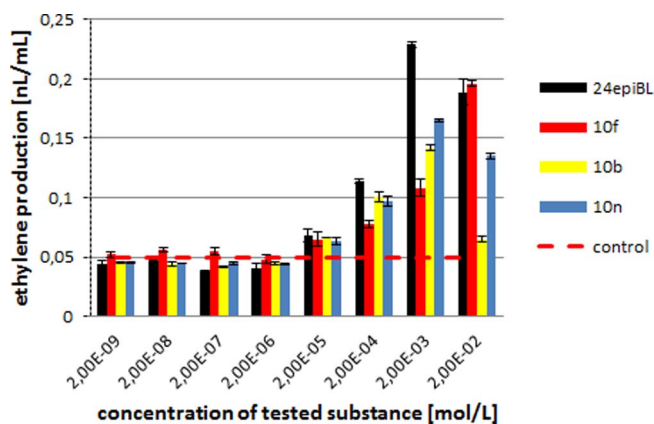


Fig. 3. Effect of selected brassinosteroid derivatives on ethylene production (nL.mL⁻¹) in etiolated pea seedlings determined by GC-FID 24 h after ventilation. Error bars represent s.d.

applied BRs stimulate ethylene production by stabilizing the ACS protein. When low levels of ethylene are applied to etiolated pea seedlings, the characteristic triple response can be observed: inhibition of stem elongation, radial swelling of the stem and the absence of a normal geotropic response [22]. Production of ethylene was measured in cultivation vessels during the incubation of etiolated pea plants after treatment of different BR derivatives (SFig. 2). The high concentrations of ethylene (196, 142 and 165 nL.L⁻¹) were determined after treatment of **10f**, **10b** and **10n** and compared to 229 nL.L⁻¹ for 24-epiBL treatment. While level of this gaseous plant hormone produced by untreated control pea plants, was found to be significantly lower (about 60 nL.L⁻¹, Fig. 3).

BRs derivatives were further assessed for biological activity based on their inhibitory effect on Arabidopsis root growth [23]. The effects of compounds **10f**, **10b** and **10n** on the Arabidopsis roots are shown in Fig. 4. Results of all prepared analogues are summarized in SFig. 3.

Molecular docking into BRI1 receptor shows that several compounds binds with very high binding energy. In some cases the binding energy was even better than for naturally occurring brassinolide (−10.6 kcal/mol) [11] (see SI for molecular docking of all new compounds). The high binding affinity of compounds **10f** (−11.0 kcal/mol), **10b** (−10.7 kcal/mol), and **10n** (−11.9 kcal/mol) also proves the good biological activities of these analogues (Fig. 5).

Relationship between biological activity and substitution pattern in the phenyl group shows that no substitution or substitution with one small group (fluorine, chlorine) led to promising compounds with high

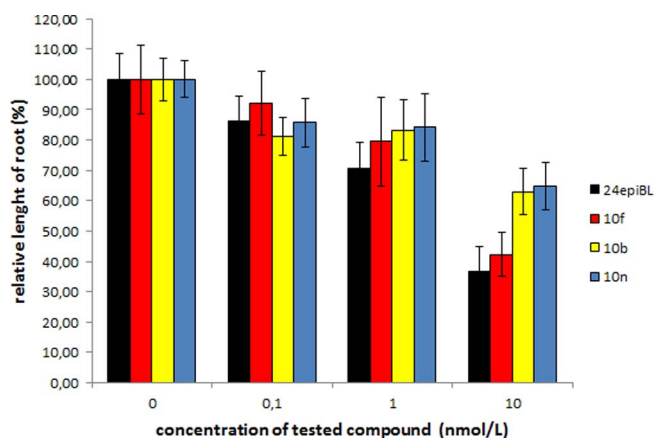


Fig. 4. Effect of selected brassinosteroid derivatives on the inhibition of Arabidopsis root length. 5 days old *Arabidopsis thaliana* seedlings (Columbia ecotype, Col-0) were treated by DMSO/24epiBL/BR analogues. For each treatment more than 25 seedlings were analyzed in two biological repeats.

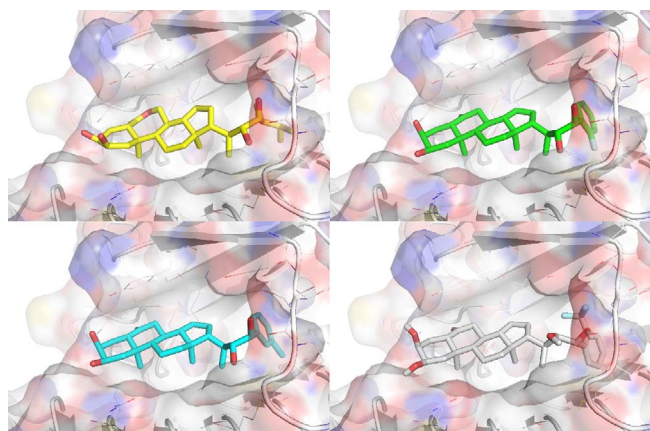


Fig. 5. Poses of 24-epibrassinolide (yellow) and three new analogues of brassinosteroids, **10b** (green), **10f** (cyan), and **10n** (grey) within BR11 binding site.

plant activities. The most active are compounds **10a** (no substitution), **10b** and **10c** (fluorine in *ortho* and *para* position), and **10f** (chlorine in *meta* position). On the other hand, substitution with bulky groups (phenyl, *t*-butyl, methyl or trifluoromethyl) or substitution with more than one group (e.g. **10q** and **10r**) causes significant decrease or complete loss of plant activity. The only exception from bulky groups is compound **10n** which showed good activity in two of the three assays. This may be explained by smaller interaction of substituents in *ortho* position with receptor cavity – the group can be oriented out of the cavity.

Antiproliferative activity of BRs derivatives was screened towards various tumor cell lines and normal cells, including T-lymphoblastic leukemia CEM, breast carcinoma (MCF7) and cervical carcinoma (HeLa) and human foreskin fibroblasts (BJ). All tested BR analogues had no detectable cytotoxic activity, even when tested in concentrations up to 50 μ M (data not shown).

9. Conclusions

Overall 15 novel and 5 known phenyl analogues of brassinosteroid were synthesized *via* alkene cross metathesis. The metathesis showed as an effective method for preparation of new brassinosteroid derivatives with plant growth promoting activities comparable with natural brassinosteroids. The results of biological screenings showed that molecular docking into BR11 is a powerful tool for prediction and design of new compounds with strong brassinosteroid activities. Very potent brassinosteroids analogues prepared by this synthesis can be also used for potential application in agriculture to improve growth and yield or to increase the resistance of plants against various biotic and abiotic stresses.

Acknowledgments

This work was supported by grant GJ15-08202Y of the Grant Agency of the Czech Republic and by the project of the Ministry of Education, Youth and Sports CR NPMU LO1204. This work was also supported by student project IGA_PrF_2017_013 and IGA_PrF_2017_028 of the Palacký University in Olomouc.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.steroids.2017.08.010>.

References

- [1] A. Bajguz, Brassinosteroids – occurrence and chemical structures in plants, in: S. Hayat, A. Ahmad (Eds.), *Brassinosteroids: A Class of Plant Hormone*, Springer, New York, 2011, pp. 1–27.
- [2] Y. Coll, F. Coll, A. Amoros, M. Pujol, Brassinosteroids roles and applications: an update, *Biologia* 70 (2015) 726–732.
- [3] J. Oklestkova, L. Rarova, M. Kvasnica, M. Strnad, Brassinosteroids: synthesis and biological activities, *Phytochem. Rev.* 14 (2015) 1053–1072.
- [4] S.D. Clouse, A History of Brassinosteroid research from through 2005: thirty-five years of phytochemistry, physiology, genes, and mutants, *J. Plant Growth Reg.* 34 (2015) (1970) 828–844.
- [5] D. Hla, O. Rothova, M. Kocova, L. Kohout, M. Kvasnica, The effect of brassinosteroids on the morphology, development and yield of field-grown maize, *Plant Growth Reg.* 61 (2010) 29–43.
- [6] J.M. Li, J. Chory, A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction, *Cell* 90 (1997) 929–938.
- [7] T. Kinoshita, A. Cano-Delgado, H. Seto, S. Hiranuma, S. Fujioka, S. Yoshida, J. Chory, Binding of brassinosteroids to the extracellular domain of plant receptor kinase BR11, *Nature* 433 (2005) 167–171.
- [8] J. She, Z. Han, B. Zhou, J. Chai, Structural basis for differential recognition of brassinolide by its receptors, *Protein Cell* 4 (2013) 475–482.
- [9] J. She, Z. Han, T.W. Kim, J. Wang, W. Cheng, J. Chang, S. Shi, J. Wang, M. Yang, Z.Y. Wang, J. Chai, Structural insight into brassinosteroid perception by BR11, *Nature* 474 (2011) 472–476.
- [10] M. Hothorn, Y. Belkhadir, M. Dreux, T. Dabi, J.P. Noel, I.A. Wilson, J. Chory, Structural basis of steroid hormone perception by the receptor kinase BR11, *Nature* 474 (2011) 467–471.
- [11] M. Kvasnica, J. Oklestkova, V. Bazgier, L. Rárová, P. Korinkova, J. Mikulík, M. Budesinsky, T. Béres, K. Berka, Q. Lu, E. Russinova, M. Strnad, Design, synthesis and biological activities of new brassinosteroid analogues with a phenyl group in the side chain, *Org. Biomol. Chem.* 14 (2016) 8691–8701.
- [12] M. Kvasnica, J. Oklestkova, V. Bazgier, L. Rarova, K. Berka, M. Strnad, Biological activities of new monohydroxylated brassinosteroid analogues with a carboxylic group in the side chain, *Steroids* 85 (2014) 58–64.
- [13] O. Trott, A.J. Olson, Software news and update autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *J. Comput. Chem.* 31 (2010) 455–461.
- [14] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility, *J. Comp. Chem.* 16 (2009) 2785–2791.
- [15] B. Eignerova, B. Slavikova, M. Budesinsky, M. Dracinsky, B. Klepetarova, E. Stastna, M. Katora, Synthesis of fluorinated brassinosteroids based on alkene cross-metathesis and preliminary biological assessment, *J. Med. Chem.* 52 (2009) 5753–5757.
- [16] D. Czajkowska, J.W. Morzycki, Metathesis reactions of Δ^{22} -steroids, *Tetrahedron Lett.* 50 (2009) 2004–2007.
- [17] B. Eignerová, M. Dračinský, M. Katora, Perfluoroalkylation through cross-metathesis between alkenes and (perfluoroalkyl)propenes, *Eur. J. Org. Chem.* 40 (2008) 4493–4499.
- [18] M. Bandini, M. Contento, A. Garelli, M. Monari, A. Tolomelli, A. Umani-Ronchi, E. Andriolo, M. Montorsi, A nonclassical stereoselective semi-synthesis of drospironone via cross-metathesis reaction, *Synthesis* 23 (2008) 3801–3804.
- [19] A.K. Chatterjee, T. Choi, D.P. Sanders, R.H. Grubbs, A general model for selectivity in olefin cross metathesis, *J. Am. Chem. Soc.* 125 (2003) 11360–11370.
- [20] K.B. Sharpless, W. Amberg, Y.L. Bennani, G.A. Criapino, J. Hartung, K.-S. Jeong, H.-L. Kwong, K. Morikawa, Z.-M. Wang, D. Xu, X.-L. Zhang, The osmium-catalyzed asymmetric dihydroxylation: a new ligand class and a process improvement, *J. Org. Chem.* 57 (1992) 2768–2771.
- [21] M. Hansen, H.S. Chae, J. Kieber, Regulation of ACS protein stability by cytokinin and brassinosteroids, *Plant J.* 57 (2009) 606–614.
- [22] P. Guzmán, J.R. Ecker, Exploiting the triple response of Arabidopsis to identify ethylene-related mutants, *Plant Cell* 2 (1990) 513–523.
- [23] S.D. Clouse, A.F. Hall, M. Langford, T.C. McMorris, M.E. Baker, Physiological and Molecular Effects of Brassinosteroids on Arabidopsis thaliana, *J. Plant Growth Regul.* 12 (1993) 61–66.

Appendix D

Mojzych M, Šubertová V, Bielawska A, Bielawski K, **Bazgier V**, Berka K, Gucký T,
Fornal E, Kryštof V:

Synthesis and kinase inhibitory activity of new sulfonamide derivatives of pyrazolo[4,3-
e][1,2,4]triazines.

Eur. J. Med. Chem., 78, 217-224, 2014.

DOI: 10.1016/j.ejmech.2014.03.054

IF = 4.519



Original article

Synthesis and kinase inhibitory activity of new sulfonamide derivatives of pyrazolo[4,3-*e*][1,2,4]triazines



Mariusz Mojzych^{a,*}, Veronika Šubertová^{b,1}, Anna Bielawska^c, Krzysztof Bielawski^c, Václav Bazgier^b, Karel Berka^d, Tomáš Gucký^b, Emilia Fornal^e, Vladimír Kryštof^{b,*}

^a Department of Chemistry, Siedlce University of Natural Sciences and Humanities, ul. 3 Maja 54, Siedlce 08-110, Poland

^b Centre of the Region Haná for Biotechnological and Agricultural Research, Laboratory of Growth Regulators, Faculty of Science, Palacký University, Šlechtitelů 11, 783 71 Olomouc, Czech Republic

^c Department of Medicinal Chemistry and Drug Technology, Medical University of Białystok, Białystok, Poland

^d Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacký University Olomouc, 17. listopadu 12, 77146 Olomouc, Czech Republic

^e Department of Chemistry, Laboratory of Separation and Spectroscopic Method Applications, Center for Interdisciplinary Research, The John Paul II Catholic University of Lublin, al. Krasnicka 102, 20-718 Lublin, Poland

ARTICLE INFO

Article history:

Received 2 December 2013

Received in revised form

13 March 2014

Accepted 16 March 2014

Available online 18 March 2014

Keywords:

Bcr-Abl

Kinase

Inhibitor

Synthesis

Cancer

Leukaemia

Pyrazolo[4,3-*e*][1,2,4]triazine

ABSTRACT

A new series of sulfonamide derivatives of pyrazolo[4,3-*e*][1,2,4]triazine has been synthesized and characterized. Their anticancer activity was tested *in vitro* against multiple human cancer cell lines and were found to have dose-dependent antiproliferative effects. Furthermore, some of the new compounds inhibited the Abl protein kinase with low micromolar IC₅₀ values and exhibited selective activity against the Bcr-Abl positive K562 and BV173 cell lines, providing starting points for the further development of this new kinase inhibitor scaffold.

© 2014 Elsevier Masson SAS. All rights reserved.

1. Introduction

Nitrogen-containing heterocycles have attracted huge interest over the past decades because of their diverse pharmacological activities, including protein kinase inhibition. Being involved in nearly all aspects of life at the cellular level, protein kinases have become the most important targets of drugs for various indications, such as cancers and inflammations [1–3]. Medicinal chemists have developed probes targeting a substantial fraction of the human kinome, and dozens of kinase inhibitors have been evaluated in clinical trials since the milestone approval of imatinib, the first anti-kinase drug targeting the fusion protein Bcr-Abl, a constitutively

active tyrosine kinase that initiates the onset of chronic myelogenous leukaemia [4].

Despite the clear clinical success of imatinib, many tumours become resistant to it, often due to point mutations in the kinase domain of Bcr-Abl. Several second-generation drugs have been developed that inhibit some imatinib-resistant Bcr-Abl mutants, but there are still significant numbers of resistant mutants for which no alternative drugs are available [5]. The same may well be true for other pharmacologically relevant kinases. Consequently, there is still a need for new compounds that will enrich the pool of known inhibitors, especially if they possess novel scaffolds, binding modes or mechanisms of action.

Pyrazolo[4,3-*e*][1,2,4]triazine is a scaffold that has not yet received significant attention in drug discovery although some of its derivatives are known to exhibit biological activity. Over the last few decades, seven naturally occurring pyrazolo[4,3-*e*][1,2,4]triazines have been isolated and characterized spectroscopically and

* Corresponding authors.

E-mail address: vladimir.krystof@upol.cz (V. Kryštof).

¹ These authors contributed equally.

structurally: pseudoiodinine, nostocine A, and fluviols A–E [6–9]. In three cases, the proposed structures were verified by total synthesis [10]. Importantly, synthetic 1,3,5-substituted pyrazolo[4,3-*e*] [1,2,4]triazines have been reported to exhibit significant activity against human cancer cell lines and to weakly inhibit cyclin-dependent kinase CDK2 [11].

The reported biological activities of pyrazolo[4,3-*e*] [1,2,4]triazines might stem, at least in part, from their physicochemical similarity to purine, another nitrogen-containing heterocycle that is widely distributed in nature [12]. To date, chemical modification of the purine core (via scaffold hopping) has led to the discovery of various purine bioisosteres that have been used as scaffolds for the synthesis and development of chemical probes and drugs [12,13], including cyclin-dependent kinase (CDK) inhibitors [14]. The CDKs are a family of Ser/Thr kinases which, in association with specific cyclins, play critical roles as regulators of the different phases of the cell cycle. These enzymes and their direct regulators are frequently mutated, amplified, or deleted in malignant cells, suggesting that pharmacological CDK inhibition may be an effective strategy for treating cancer [14]. We have previously prepared and characterized many CDK inhibitors based on purine scaffolds [15–17] and bioisosteric alternatives such as 8-azapurines and pyrazolo[4,3-*d*] pyrimidines. Many of the compounds based on these bioisosteres are comparable or superior to the corresponding purine derivatives in terms of anticancer activity and kinase inhibition [18,19]. The discovery of kinase inhibitors based on new scaffolds is a major goal of our current research, and this prompted us to investigate the structure–activity relationships of pyrazolo[4,3-*e*] [1,2,4]triazines. We have previously described the convenient synthesis and structural analysis of a series of 3,5,7-trisubstituted pyrazolo[4,3-*e*] [1,2,4]triazines [20–22]. Herein we describe the synthesis and *in vitro* biological activity of sulfonamides with a pyrazolo[4,3-*e*]

[1,2,4]triazine core. The new compounds are shown to exhibit selective toxicity towards Bcr-Abl positive cells.

2. Results and discussion

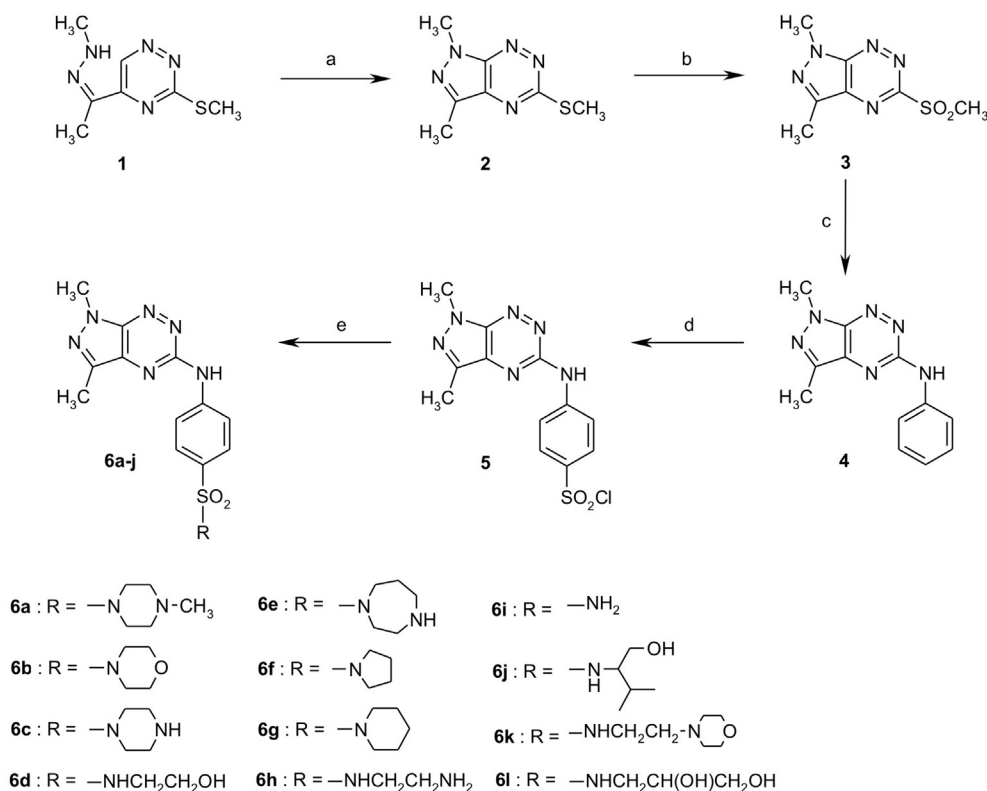
2.1. Chemistry

The synthetic pathway leading to the title compounds is outlined in Scheme 1. Methylhydrazone **1**, the precursor of pyrazolo[4,3-*e*] [1,2,4]triazine **2**, was prepared using a standard procedure. The cyclization of **1** to form **2** was then achieved under acidic conditions using a literature method [20,21]. Compound **2** was treated with potassium permanganate under phase transfer catalytic conditions at room temperature for 1 h to give sulfone **3** in nearly quantitative yield. Compound **3** was then reacted with anhydrous aniline in a sealed tube at 150 °C for 9 days. Chlorosulfonylation of the aniline-substituted pyrazolotriazine **4** in chlorosulfonic acid at 0 °C proceeded smoothly and selectively at the 4'-position of the phenyl ring to give the desired compound **5** in excellent yield. The chlorosulfonyl derivative **5** was readily coupled with amines to produce the target sulfonamides **6a–l** as shown in Scheme 1.

The structures and purity of the newly synthesized compounds were characterized using ¹H and ¹³C NMR spectroscopy, HPLC-MS, and elemental analysis. The spectral data confirmed that all of the new sulfonamides had the expected structures and were of high purity.

2.2. Antiproliferative activity *in vitro*

The antiproliferative activity of the new compounds against several breast carcinoma (MCF-7, MDA-MB-231) and leukaemia



Scheme 1. Reagents and conditions: (a) p-TSA, T; (b) KMnO_4 , Bu_4NBr , CH_3COOH , benzene- H_2O , rt, 2 h; (c) aniline, seal tube, 9 days, 150 °C; (d) ClSO_3H , 0 °C to rt, 2 h; (e) NH_3 or amine, MeCN, rt.

(K562, BV173, HL60, CCRF-CEM) cell lines was measured using the MTT assay after 24 h of incubation [23]. The results obtained are summarized in Table 1. Concentration-dependent activity was observed in all cases. The breast carcinoma lines were much less sensitive to all of the compounds (with IC₅₀ values ranging from 100 to 200 μM) than the leukaemia cell lines. However, in general, the compounds had similar rankings with respect to activity in all cell lines, with the most active being **6a**, **6c**, **6e** and **6h**. All of these species have similar side chains (*N*-methylpiperazine, piperazine, homopiperazine, 1,2-diaminoethane). Conversely, compounds **6d**, **6f**, **6i**, and **6j** exhibited very little activity.

The IC₅₀ values for the tested leukaemia cell lines ranged from 20 to 70 μM for CCRF-CEM, 24 to 58 μM for HL60, and 27 to 72 μM for BV173. The IC₅₀ range for the K562 line was even broader, with the most potent compound **6e** having a value of 21 μM while **6j** was inactive even at 200 μM. Surprisingly, the IC₅₀ values for the most potent compounds against K562 – **6e** and **6c** – were 7 and 5 times lower than their IC₅₀ values against breast carcinoma cell lines, respectively. This suggests that these compounds exhibit a degree of selective toxicity towards the K562 cell line. A similar activity pattern was observed in the BV173 cell line, with the most potent compounds again being **6e** and **6c**.

The compounds' antiproliferative activities are probably at least partly due to their cytostatic effects: preliminary investigations using breast carcinoma cell lines showed that they prevent the incorporation of [³H]thymidine into living cells. As in the MTT assay, all of the tested compounds exhibited concentration-dependent activity with different potencies (Supplementary Table S1). However, cytostatic effects alone cannot fully explain their antiproliferative activity because in several cases, the number of viable cells declined significantly after as little as 24 h of treatment. We therefore investigated the influence of the most potent compounds (**6c**, **6e** and **6h**) on the viability of K562 cells using the propidium iodide exclusion assay. Treatment with compound **6c** or **6e** caused sharp increases in the number of cells with compromised cell membranes (Fig. 1a). In parallel, we evaluated the test compounds' effects on cell cycling and found that **6c** and **6e** had no effect on the distribution of cell cycle phases (data not shown) but did cause pronounced increases in the sub-diploid population, which is indicative of ongoing apoptotic cell death (Fig. 1b). Compound **6h** had only marginal effects on viability and did not increase the sub-diploid cell population.

Table 1

In vitro antiproliferative activity of new sulfonamide derivatives of pyrazolo[4,3-*e*]-[1,2,4]triazine after 24 h incubation.

Compd.	MTT assay, IC ₅₀ (μM) ^a					
	MCF-7	MDA-MB-231	K562	BV173	HL60	CCRF-CEM
6a	102 ± 2	99 ± 2	66 ± 5	40 ± 8	49 ± 2	36 ± 2
6b	>200	>200	90 ± 8	58 ± 5	39 ± 1	69 ± 8
6c	150 ± 2	130 ± 2	27 ± 4	22 ± 6	55 ± 2	20 ± 2
6d	>200	>200	100 ± 4	41 ± 10	42 ± 5	49 ± 8
6e	140 ± 3	155 ± 2	21 ± 5	22 ± 4	38 ± 1	36 ± 12
6f	>200	>200	102 ± 1	47 ± 14	56 ± 6	30 ± 2
6g	200 ± 2	140 ± 1	98 ± 2	36 ± 9	24 ± 2	30 ± 2
6h	126 ± 1	120 ± 1	77 ± 7	39 ± 8	42 ± 6	56 ± 2
6i	>200	>200	106 ± 8	45 ± 11	58 ± 3	50 ± 1
6j	>200	>200	>200	58 ± 9	40 ± 2	54 ± 8
6k	146 ± 1	125 ± 2	96 ± 3	39 ± 8	41 ± 1	64 ± 6
6l	200 ± 2	140 ± 2	101 ± 2	42 ± 9	44 ± 5	57 ± 3
Chlorambucil	97 ± 2	93 ± 2	84 ± 6	34 ± 8	38 ± 2	21 ± 8
Imatinib	n.a.	n.a.	13 ± 2	20 ± 6	55 ± 7	45 ± 1

^a The reported values represent the mean ± S.D. for each compound based on four independent experiments.

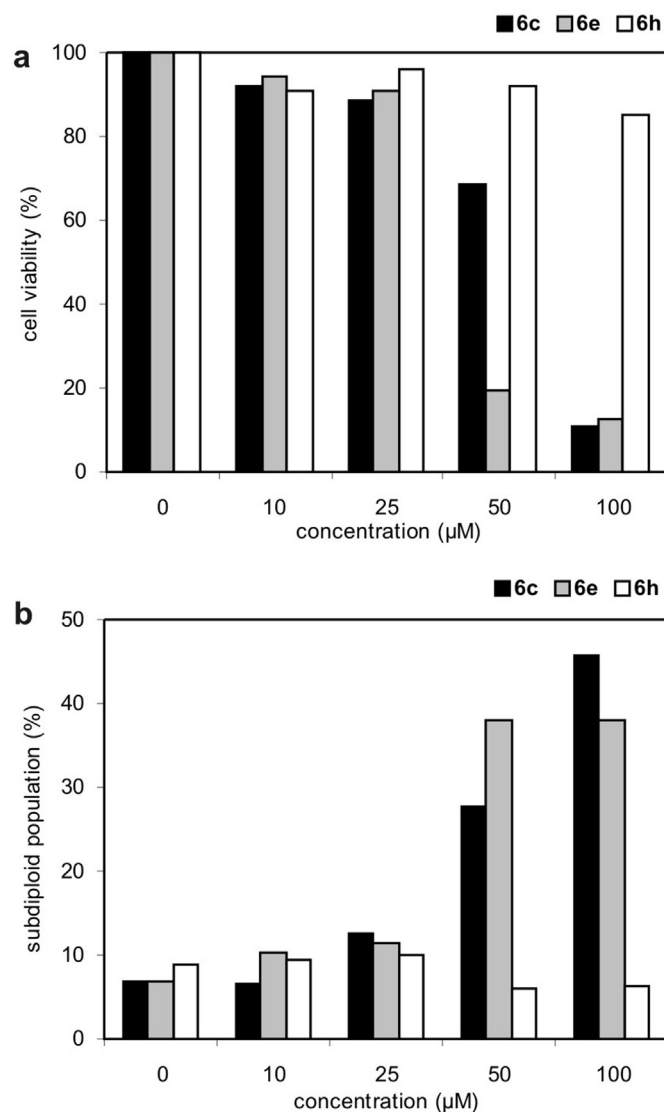


Fig. 1. Effect of test compounds on K562 cell viability after 24 h treatment. Propidium iodide exclusion (a) and the sub-diploid cell population (b) were quantified by flow cytometry in native or in ethanol-fixed cells, respectively.

2.3. Kinase inhibitory activity

The enhanced potency of the new pyrazolo[4,3-*e*][1,2,4]triazines against the K562 and BV173 cell lines (both of which are Bcr-Abl positive) relative to other cell lines prompted us to explore their inhibitory potency against a purified recombinant Abl kinase. The seven tested compounds significantly inhibited Abl activity, with IC₅₀ values in the micromolar range (Table 2). The most potent compounds **6c** and **6e** bear piperazine (IC₅₀ values of 5.8 μM) and homopiperazine (IC₅₀ values of 5.9 μM) appendages, respectively. Such rings, and especially the secondary amines within them, seem to be important for Abl inhibition: analogues bearing a methyl-substituted piperazine (**6a**) or deaza and oxo analogues (piperidine and morpholine derivatives **6g** and **6b**, respectively) exhibited dramatically lower activity (Table 2 and supplementary table S2). The acyclic 2-aminoethylamine derivative **6h** also exhibited somewhat lower activity (IC₅₀ = 15 μM). All other derivatives were either much weaker inhibitors (unsubstituted sulfonamide **6i** and hydroxyalkyl sulfonamides **6d** and **6l**) or entirely inactive at the tested concentrations (**6j** and **6f**).

Table 2
Inhibition of protein kinases by sulfonamide derivatives of pyrazolo[4,3-*e*][1,2,4]triazines.

Compd.	IC ₅₀ (μM)		
	Abl	CDK2	CK2
6a	>50	>50	>50
6b	>25	>25	>25
6c	5.8	>25	>25
6d	39	>50	>50
6e	5.9	>50	>50
6f	>50	>50	>50
6g	>25	>25	>25
6h	15.0	34.9	>50
6i	33.0	>50	>50
6j	>50	>50	>50
6k	20	>50	>50
6l	34.0	>50	>50
Roscovitine	>100	0.1	n.a.
Imatinib	>100	>100	n.a.
Silmitasertib	n.a.	1.8	0.1

To better understand the molecular basis for the inhibitory activity of the pyrazolo[4,3-*e*][1,2,4]triazine derivatives, we used molecular modelling to rationalize their binding to Abl. We hypothesized that compounds **6c** and **6e** might bind to the ATP-binding site of the Abl kinase in a manner similar to that of pyrido[2,3-*d*]pyrimidine PD180970, interacting with the protein primarily via non-polar interactions and H-bonds with the backbone NH group of the M318 residue [24]. Our studies confirmed that both **6c** and **6e** dock into the Abl kinase in this way. The piperazine ring of **6c** fits in to a pocket of the active site that is lined by the gate-keeper residue T315. This is the pocket occupied by the dichlorophenyl group of PD180970 and the methylbenzene moiety of imatinib. In addition, we observed a possible H-bond between I313 and the secondary amino groups on the sulfonamide appendages of **6c** and **6e** (Fig. 2). This H-bond has not previously been observed in any Abl co-crystal structure but its presence is consistent with the SAR data presented above: the analogous 2-aminoethylamine derivative **6h** is less active despite having a primary amine at about the right distance from I313 to form a similar H-bond, probably due to the greater flexibility of its side chain. The same is true for the even less active hydroxy derivative **6d**. The inactivity of compounds that lack a secondary amine or other H-bond acceptor capable of forming an interaction of this sort (e.g. the methyl-substituted compound **6a**, piperidine **6g**, and morpholine **6b**) further supports this hypothesis.

2.4. Kinase inhibition selectivity

We have recently demonstrated that certain pyrazolo[4,3-*e*][1,2,4]triazines inhibit CDK2/cyclin E [22]. However, most of the

compounds synthesized in this work did not inhibit either CDK2 or CK2 within the tested concentration range, showing that they have some degree of selectivity for Abl (Table 2). The only exception is compound **6h**, which exhibited weak activity towards CDK2 (IC₅₀ = 34.9 μM). This may contribute to its marginal anti-proliferative effects (Table 1). These preliminary selectivity data are however not sufficient to exclude the possibility that the new compounds may bind to and inhibit other targets, including other protein kinases.

To clarify the new compounds' lack of activity towards CDK2 and guide future optimization, we investigated the interactions of the pyrazolo[4,3-*e*][1,2,4]triazine derivatives with CDK2. Molecular docking experiments suggested that the mostly negative results obtained in the biochemical assays are due to the relatively unfavourable mode of binding adopted by the pyrazolo[4,3-*e*][1,2,4]triazines in the CDK2 active site (Fig. 3). Well-known purine CDK inhibitors such as olomoucine and roscovitine typically bear small hydrophobic chain at N9, an aromatic ring attached to the secondary amino group at C6, and a polar alkyl amine at C2. In addition, they have no substituents at N7 [13,14,16]. This substitution pattern, together with the donor–acceptor H-bond motif formed by the N7 group and the secondary amino group at C6 that interacts with the hinge region of CDK2, gives the known CDK2 inhibitors a well-defined binding orientation within the CDK2 active site. Because pyrazolo[4,3-*e*][1,2,4]triazines cannot be substituted at the corresponding positions, we anticipated, based on the well known structure–activity relationships for CDK inhibitors [15–17], that they might instead bind to CDK2 in a 'reversed' orientation, similar to that adopted by the guanine-based inhibitor NU6102 [25]. NU6102 binds to the hinge region via H-bonds involving the secondary amino group at C2 and purine nitrogens N3 and N9 (Fig. 3). However, it is likely that the methyl substitution of the pyrazole ring would both abolish one of these H-bonds and also form an unfavourable steric interaction that would discourage the binding of **6h** in the active site (Fig. 3a).

2.5. Cellular inhibition of tyrosine kinases

Compounds **6c** and **6e** emerged as most potent Abl inhibitors from this series. Because of their strong antiproliferative activity in Bcr-Abl positive cell lines (K562 and BV173), we decided to characterize their cellular effects in more detail. Derivative **6h** was also included in these experiments due to its lower activity and broader selectivity. We treated K652 cells with increasing doses of **6c**, **6e** and **6h** for 1 h and immunoblotted the cell lysates with phospho-specific antibodies towards phosphorylated forms of STAT5 and CrkL, known substrates of Bcr-Abl kinase that both contribute to the sustained proliferation of leukaemic cells (Fig. 4a). All three

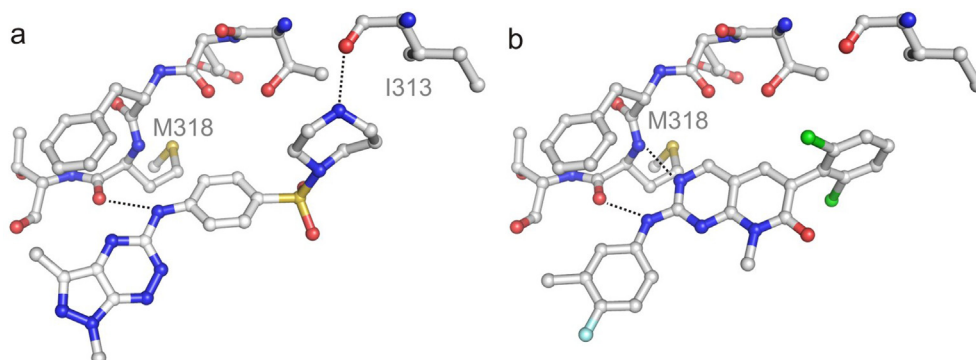


Fig. 2. The binding poses of compound **6e** (a) and PD180970 (b) in c-Abl (PDBID: 2hzi). The compounds' interactions with I313 and M318 are indicated by dotted lines.

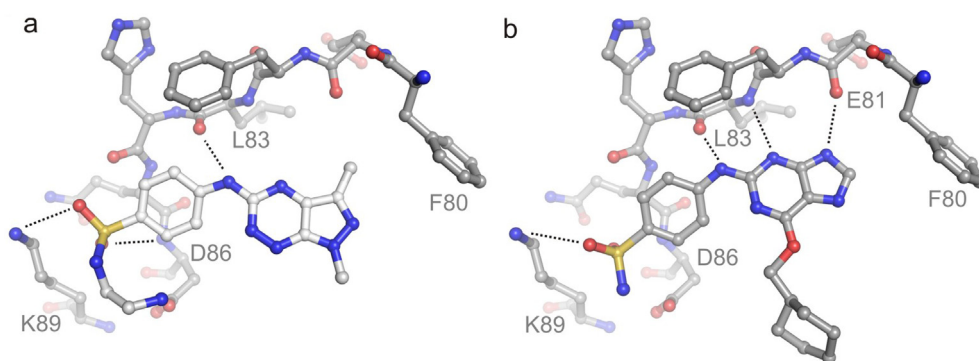


Fig. 3. The binding poses of compound **6h** (a) and NU6102 (b) in CDK2 (PDBID: 2c60). Interactions with D86, K89 and with the backbone of L83 and E81 are indicated by dotted lines.

derivatives caused dose–response decreases in the phosphorylation of STAT5, but only **6c** reduced the phosphorylation of CrkL.

We also evaluated the influence of the most potent compounds on the global tyrosine phosphorylation status of the proteins in K562 cells. This was done by immunoblotting their lysates with an anti-phosphotyrosine antibody (Fig. 4b). In untreated cells, several proteins phosphorylated on tyrosine residues were detected. Treatment with derivatives **6c** and **6e** reduced the phosphorylation of some of these proteins in a dose-dependent manner, including

some that formed intense bands with molecular weights between 150 and 250 kDa and probably corresponded to Abl and Bcr-Abl [26]. In addition, at least three additional phosphoproteins with molecular weights between 40 and 60 kDa (which yielded weaker bands) exhibited dose-dependent reductions in signal intensity. Importantly, treatment with the less active derivative **6h** did not reduce the phosphorylation of any of these proteins whereas imatinib (which was used as a positive control in both experiments) rapidly inhibited the phosphorylation of several proteins.

3. Conclusions

We have described an efficient method for the preparation of sulfonamide derivatives of pyrazolo[4,3-*e*][1,2,4]triazines (**6**). Approximately half of the prepared compounds exhibited modest activity towards cancer cell lines. In addition, our study identified two compounds, **6c** and **6e**, that exhibit micromolar inhibition of the Abl protein kinase and preferential antiproliferative activity in leukaemia cell lines that overexpress the oncogenic kinase Bcr-Abl. These compounds represent new scaffolds for protein kinase inhibitors, which are still needed in oncological drug discovery, especially because of the emerging resistance to existing drugs. Further studies aimed at optimizing the structures of the hits in order to increase their potency against the Bcr-Abl kinase and cancer cell lines *in vitro* are currently underway in our laboratories.

4. Experimental section

4.1. Chemistry

Melting points were determined on a Mel-Temp apparatus and are uncorrected. ^1H and ^{13}C NMR spectra were recorded on a Varian spectrometer (400 and 100 MHz, respectively). Chemical shift values are expressed in ppm (part per million) relative to tetramethylsilane as an internal standard. The relative integrals of peak areas agreed with those expected for the assigned structures. The molecular weights of the final compounds were determined by electrospray ionization mass spectrometry (ESI/MS) performed using an Agilent Technologies 6538 UHD Accurate Mass Q-TOF LC/MS. All elemental compositions were within $\pm 0.4\%$ of the calculated values.

4.2. Synthesis of precursors

4.2.1. 1,3-Dimethyl-5-phenylamino-1H-pyrazolo[4,3-*e*][1,2,4]triazine (**4**)

A mixture of **3** (1.68 g, 7.37 mmol) and freshly distilled aniline (10 ml, 110 mmol) was introduced into a sealed tube and heated at

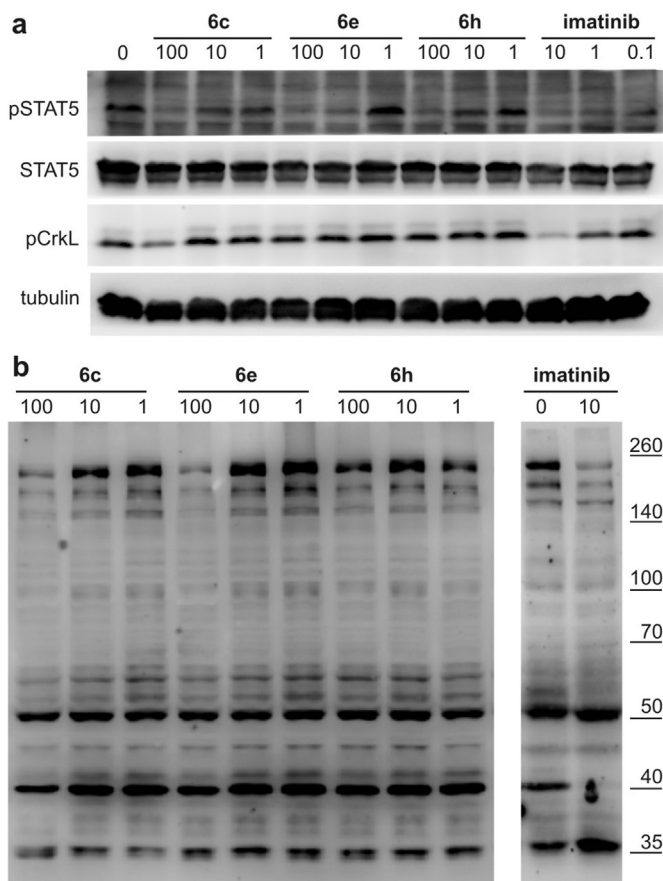


Fig. 4. Phosphoprotein analysis of K562 cells treated with test compounds. The cells were incubated for 1 h with the indicated doses of **6c**, **6e** and **6h** (in μM), lysed and immunoblotted with anti-phospho-STAT5, anti-STAT5, anti-phospho-CrkL and anti-tubulin antibodies (a). In a second experiment, the immunoblotting was performed with an anti-phosphotyrosine antibody (b). Imatinib (0.1, 1 or 10 μM) was used as a positive control in both experiments.

150 °C for 9 days. The excess aniline was then distilled out from the reaction mixture, and the resulting residue was purified by column chromatography on silica gel, eluting with a mixture of ethyl acetate/hexane (1:1). The combined fractions were evaporated, providing the crude product **4** as an orange solid (0.74 g, 3.07 mmol). Yield 42%, mp 215–217 °C. ¹H NMR (DMSO-*d*₆) δ: 2.49 (s, 3H), 4.09 (s, 3H), 6.98 (t, 1H, *J* = 7.6 Hz), 7.32 (t, 2H, *J* = 7.4 Hz), 7.81 (d, 2H, *J* = 9.0 Hz), 10.26 (s, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ: 10.72, 34.67, 118.54, 121.93, 128.95, 133.05, 138.33, 140.42, 146.17, 158.36. HRMS (ESI, *m/z*) calcd. for C₁₂H₁₃N₆ [M+H] 241.1196. Found 241.1195. Anal. Calcd. for C₁₂H₁₂N₆: C, 59.99; H, 5.03; N, 34.98. Found: C, 59.76; H, 5.06; N, 34.84.

4.2.2. 4-(1,3-Dimethyl-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-ylamino)benzene-1-sulfonyl chloride (**5**)

Compound **4** (1.03 g, 4.28 mmol) was added portionwise to stirred and ice-cooled chlorosulfonic acid (9.03 g, 77.48 mmol) under an argon atmosphere, after which the reaction mixture was gradually warmed to room temperature over 2 h. The reaction solution was cautiously poured over crushed ice (40 g) and the aqueous mixture was extracted with dichloromethane. The combined extracts were dried over anhydrous Na₂SO₄ and evaporated under reduced pressure to give the required sulfonyl chloride **5** as a yellow solid (1.37 g, 4.03 mmol). Yield 94%, mp 233 °C. ¹H NMR (DMSO-*d*₆) δ: 2.49 (s, 3H), 4.10 (s, 3H), 7.57 (d, 2H, *J* = 8.4 Hz), 7.79 (d, 2H, *J* = 8.8 Hz), 10.43 (s, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ: 10.56, 34.52, 117.17, 126.25, 132.86, 138.17, 140.69, 140.78, 146.04, 157.95.

4.3. General procedure for preparation of the sulfonamides (**6a–I**)

A mixture of chlorosulfonyl derivative (**5**) (100 mg, 0.29 mmol) and the appropriate amine (1.00 mmol) in anhydrous acetonitrile (10 ml) was stirred overnight at room temperature, after which the reaction mixture was concentrated under reduced pressure to afford crude sulfonamides, **6a–I** as yellow solids. The residues were purified by crystallization from EtOH or by column chromatography on silica using a mixture of CH₂Cl₂:EtOH (25/1, v/v) as the eluent.

4.3.1. 1,3-Dimethyl-N-[4-(4-methylpiperazin-1-ylsulfonyl)phenyl]-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-amine (**6a**)

Yield 71%, mp 246–251 °C. ¹H NMR (CDCl₃) δ: 2.27 (s, 3H), 2.50–2.52 (m, 4H), 2.63 (s, 3H), 3.07–3.08 (m, 4H), 4.25 (s, 3H), 7.77 (d, 2H, *J* = 8.8 Hz), 7.93 (d, 2H, *J* = 8.8 Hz), 8.12 (s, 1H, NH). ¹³C NMR (CDCl₃) δ: 10.87, 34.80, 45.68, 45.95, 54.07, 117.62, 127.98, 129.36, 133.97, 139.76, 143.53, 146.64, 157.14. HRMS (ESI, *m/z*) calcd. for C₁₇H₂₃N₈O₂S [M+H] 403.1659. Found 403.1661. Anal. Calcd. for C₁₇H₂₂N₈O₂S: C, 50.73; H, 5.51; N, 27.84. Found: C, 50.50; H, 5.53; N, 27.71.

4.3.2. 1,3-Dimethyl-N-[4-(morpholinosulfonyl)phenyl]-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-amine (**6b**)

Yield 72%, mp 280–285 °C. ¹H NMR (CDCl₃) δ: 2.64 (s, 3H), 3.03 (t, 4H, *J* = 4.8 Hz), 3.76 (t, 4H, *J* = 4.8 Hz), 4.26 (s, 3H), 7.77 (d, 2H, *J* = 6.8 Hz), 7.97 (d, 2H, *J* = 6.8 Hz), 8.19 (s, 1H, NH). ¹³C NMR (CDCl₃) δ: 10.87, 34.81, 46.05, 66.13, 117.65, 127.66, 129.43, 133.39, 139.75, 143.75, 146.65, 157.12. HRMS (ESI, *m/z*) calcd. for C₁₆H₂₀N₇O₃S [M+H] 390.1343. Found 390.1342. Anal. Calcd. for C₁₆H₁₉N₇O₃S: C, 49.35; H, 4.92; N, 25.18. Found: C, 49.09; H, 4.95; N, 25.04.

4.3.3. 1,3-Dimethyl-N-[4-(piperazin-1-ylsulfonyl)phenyl]-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-amine (**6c**)

Yield 86%, mp 300–307 °C. ¹H NMR (DMSO-*d*₆) δ: 2.52 (s, 3H), 2.79–2.81 (m, 8H), 4.15 (s, 3H), 7.68 (d, 2H, *J* = 8.8 Hz), 8.09 (d, 2H, *J* = 8.8 Hz), 10.92 (s, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ: 10.54, 34.56, 44.27, 46.14, 117.45, 126.04, 128.94, 132.78, 138.47, 144.76, 146.18,

157.32. HRMS (ESI, *m/z*) calcd. for C₁₆H₂₁N₈O₂S [M+H] 389.1503. Found 389.1502. Anal. Calcd. for C₁₆H₂₀N₈O₂S: C, 49.47; H, 5.19; N, 28.85. Found: C, 49.17; H, 5.23; N, 28.67.

4.3.4. 4-(1,3-Dimethyl-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-ylamino)-N-(2-hydroxyethyl)-benzenesulfonamide (**6d**)

Yield 65%, mp 242–244 °C. ¹H NMR (DMSO-*d*₆) δ: 2.53 (s, 3H), 2.78 (t, 2H, *J* = 6.0 Hz), 3.36 (q, 2H, *J* = 6.0 Hz), 4.16 (s, 3H), 4.64 (t, 1H, *J* = 6.0 Hz, OH), 7.37 (bs, 1H, NH), 7.75 (d, 2H, *J* = 8.4 Hz), 8.02 (d, 2H, *J* = 8.4 Hz), 10.82 (bs, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ: 10.53, 34.52, 45.10, 59.93, 117.45, 127.72, 132.26, 132.73, 138.38, 143.88, 146.08, 157.39. HRMS (ESI, *m/z*) calcd. for C₁₄H₁₈N₇O₃S [M+H] 364.1186. Found 364.1186. Anal. Calcd. for C₁₄H₁₇N₇O₃S: C, 46.27; H, 4.72; N, 26.98. Found: C, 46.04; H, 4.74; N, 26.84.

4.3.5. N-[4-(1,4-Diazepan-1-ylsulfonyl)phenyl]-1,3-dimethyl-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-amine (**6e**)

Yield 70%, mp 168–174 °C. ¹H NMR (DMSO-*d*₆) δ: 1.76–1.80 (m, 2H), 2.52 (s, 3H), 2.72 (t, 2H, *J* = 12.0 Hz), 2.77 (t, 2H, *J* = 10.0 Hz), 3.20 (t, 2H, *J* = 10.0 Hz), 3.25 (t, 2H, *J* = 12.0 Hz), 4.15 (s, 3H), 4.18 (bs, 1H, NH), 7.73 (d, 2H, *J* = 8.8 Hz), 8.05 (d, 2H, *J* = 8.8 Hz), 10.92 (bs, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ: 10.50, 27.87, 34.51, 45.74, 46.82, 47.89, 48.10, 117.52, 128.02, 129.89, 132.67, 138.36, 144.36, 146.07, 157.29. HRMS (ESI, *m/z*) calcd. for C₁₇H₂₃N₈O₂S [M+H] 403.1659. Found 403.1660. Anal. Calcd. for C₁₇H₂₂N₈O₂S: C, 50.73; H, 5.51; N, 27.84. Found: C, 50.52; H, 5.53; N, 27.72.

4.3.6. 1,3-Dimethyl-N-[4-(pyrrolidin-1-ylsulfonyl)phenyl]-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-amine (**6f**)

Yield 88%, mp 284–286 °C. ¹H NMR (DMSO-*d*₆) δ: 1.63–1.66 (m, 4H), 2.52 (s, 3H), 3.12–3.15 (m, 4H), 4.15 (s, 3H), 7.76 (d, 2H, *J* = 8.8 Hz), 8.07 (d, 2H, *J* = 8.8 Hz), 10.88 (s, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ: 10.67, 24.67, 34.65, 47.79, 117.40, 127.61, 128.54, 132.70, 138.39, 144.47, 146.10, 157.33. HRMS (ESI, *m/z*) calcd. for C₁₆H₂₀N₇O₂S [M+H] 374.1394. Found 374.1396. Anal. Calcd. for C₁₆H₁₉N₇O₂S: C, 51.46; H, 5.13; N, 26.26. Found: C, 51.30; H, 5.15; N, 26.19.

4.3.7. 1,3-Dimethyl-N-[4-(piperidin-1-ylsulfonyl)phenyl]-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-amine (**6g**)

Yield 85%, mp 241–248 °C. ¹H NMR (DMSO-*d*₆) δ: 1.29–1.38 (m, 2H), 1.60–1.66 (m, 4H), 2.53 (s, 3H), 2.95–2.98 (m, 4H), 4.13 (s, 3H), 7.66 (d, 2H, *J* = 8.4 Hz), 8.06 (d, 2H, *J* = 8.4 Hz), 10.89 (bs, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ: 10.70, 22.12, 24.70, 43.54, 46.66, 117.43, 126.80, 128.61, 132.70, 138.41, 144.53, 146.11, 157.32. HRMS (ESI, *m/z*) calcd. for C₁₇H₂₂N₇O₂S [M+H] 388.1550. Found 388.1510. Anal. Calcd. for C₁₇H₂₁N₇O₂S: C, 52.70; H, 5.46; N, 25.31. Found: C, 52.45; H, 5.49; N, 25.18.

4.3.8. N-(2-Aminoethyl)-4-(1,3-dimethyl-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-ylamino)benzenesulfonamide (**6h**)

Yield 78%, mp 238–242 °C. ¹H NMR (DMSO-*d*₆) δ: 2.51 (s, 3H), 2.65 (t, 2H, *J* = 6.6 Hz), 2.82 (t, 2H, *J* = 6.4 Hz), 4.08 (bs, 2H, NH₂), 4.14 (s, 3H), 7.75 (d, 2H, *J* = 8.8 Hz), 8.03 (d, 2H, *J* = 8.8 Hz), 10.86 (bs, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ: 10.53, 34.53, 43.30, 117.48, 127.79, 131.76, 132.78, 138.39, 144.03, 146.14, 157.38. HRMS (ESI, *m/z*) calcd. for C₁₄H₁₉N₈O₂S [M+H] 363.1346. Found 363.1346. Anal. Calcd. for C₁₄H₁₈N₈O₂S: C, 46.40; H, 5.01; N, 30.92. Found: C, 46.20; H, 5.03; N, 30.81.

4.3.9. 4-(1,3-Dimethyl-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-ylamino)benzenesulfonamide (**6i**)

Yield 82%, mp 317–320 °C. ¹H NMR (DMSO-*d*₆) δ: 2.49 (s, 3H), 4.12 (s, 3H), 7.18 (s, 2H, NH₂), 7.76 (d, 2H, *J* = 8.4 Hz), 7.96 (d, 2H, *J* = 8.4 Hz), 10.74 (s, 1H, NH). ¹³C NMR (DMSO-*d*₆) δ: 10.66, 34.65,

117.55, 126.87, 132.95, 136.38, 138.55, 143.52, 146.24, 157.60. HRMS (ESI, m/z) calcd. for $C_{12}H_{14}N_7O_2S$ [M+H] 320.0924. Found 320.0923. Anal. Calcd. for $C_{12}H_{13}N_7O_2S$: C, 45.13; H, 4.10; N, 30.70. Found: C, 44.83; H, 4.16; N, 30.53.

4.3.10. 4-(1,3-Dimethyl-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-ylamino)-N-(1-hydroxy-3-methylbutan-2-yl)benzenesulfonamide (6j)

Yield 75%, mp 306–307 °C. 1H NMR (DMSO- d_6) δ : 0.71 (d, 3H, $J = 6.8$ Hz), 0.75 (d, 3H, $J = 6.8$ Hz), 1.80–1.86 (m, 1H), 2.53 (s, 3H), 2.87–2.91 (m, 1H), 3.10–3.17 (m, 2H), 4.15 (s, 3H), 4.49 (t, 1H, $J = 5.4$ Hz, OH), 7.14 (d, 1H, $J = 8.4$ Hz, NH), 7.75 (d, 2H, $J = 9.2$ Hz), 8.01 (d, 2H, $J = 9.2$ Hz), 10.84 (bs, 1H, HN). HRMS (ESI, m/z) calcd. for $C_{17}H_{24}N_7O_3S$ [M+H] 406.1656. Found 406.1655. Anal. Calcd for $C_{17}H_{23}N_7O_3S$: C, 50.36; H, 5.72; N, 24.18. Found: C, 50.09; H, 5.74; N, 24.07.

4.3.11. (1,3-Dimethyl-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-ylamino)-N-(2-morpholinoethyl)-benzenesulfonamide (6k)

Yield 76%, mp 218–220 °C. 1H NMR (DMSO- d_6) δ : 2.26–2.31 (m, 6H), 2.49 (s, 3H), 2.84 (q, 2H, $J = 6.4$ Hz), 3.49 (t, 4H, $J = 4.4$ Hz), 4.13 (s, 3H), 7.29 (t, 2H, $J = 5.2$ Hz), 7.74–7.76 (d, 2H, $J = 8.8$ Hz), 8.00 (d, 2H, $J = 8.8$ Hz), 10.79 (s, 1H, NH). ^{13}C NMR (DMSO- d_6) δ : 10.51, 15.25, 34.52, 53.05, 57.09, 65.97, 117.44, 127.73, 131.76, 132.78, 138.36, 144.03, 146.14, 157.39. HRMS (ESI, m/z) calcd. for $C_{18}H_{25}N_8O_3S$ [M+H] 433.1765. Found 433.1763. Anal. Calcd. for $C_{18}H_{24}N_8O_3S$: C, 49.99; H, 5.59; N, 25.91. Found: C, 49.75; H, 5.62; N, 25.74.

4.3.12. N-(2,3-Dihydroxypropyl)-4-(1,3-dimethyl-1H-pyrazolo[4,3-*e*][1,2,4]triazin-5-ylamino)-benzenesulfonamide (6l)

Yield 87%, mp 235–236 °C. 1H NMR (DMSO- d_6) δ : 2.50 (s, 3H), 2.60 (dd, 1H, $J_1 = 7.2$ Hz, $J_2 = 1.32$ Hz), 2.85 (dd, 1H, $J_1 = 5.2$ Hz, $J_2 = 1.28$ Hz), 3.24–3.28 (m, 2H), 3.42–3.50 (m, 1H), 4.13 (s, 3H), 4.45 (t, 1H, $J = 3.6$ Hz, OH), 4.70 (d, 1H, $J = 4.4$ Hz, OH), 7.23 (bs, 1H, NH), 7.74 (d, 2H, $J = 8.4$ Hz), 8.00 (d, 2H, $J = 8.4$ Hz), 10.80 (bs, 1H, NH). ^{13}C NMR (DMSO- d_6) δ : 10.52, 34.51, 46.05, 63.52, 70.30, 117.40, 127.74, 132.15, 132.75, 138.37, 143.83, 146.11, 157.40. HRMS (ESI, m/z) calcd. for $C_{15}H_{19}N_7O_4S$ [M+H] 394.1292. Found 394.1292. Anal. Calcd. for $C_{15}H_{19}N_7O_4S$: C, 45.79; H, 4.87; N, 24.92. Found: C, 45.52; H, 5.03; N, 24.79.

4.4. Cell cultures

Human cancer cell lines MCF-7, MDA-MB-231, K562, CCRF-CEM, HL60 and BV173 (purchased from the American Type Culture Collection and the German Collection of Microorganisms and Cell Cultures) were maintained in DMEM supplemented with 10% foetal bovine serum (FBS), 50 U/ml penicillin, 50 μ g/ml streptomycin at 37 °C. Cells were cultured in Costar flasks and subconfluent cells were detached with trypsin (in case of breast cell lines), counted and plated at 5×10^5 cells per well in 6-well plates (Nunc) in 2 ml of growth medium. Cells reached 80% confluency on day 3 and in most cases such cells were used for the assays.

4.5. MTT assay

The assay was performed according to the literature method [23]. Cells exposed for 24 h to various concentrations of the studied compounds in microtitre plates were stained with MTT solution at 37 °C. The solution was then removed and 1 ml of 0.1 M HCl in absolute isopropanol was added to the attached cells. The absorbance of the converted dye in living cells was measured at a wavelength of 570 nm. Cell viability was calculated as a percentage of the viable control cells. All experiments were performed in triplicate. IC_{50} values (the drug concentrations that reduced the

number of viable cells to 50% of the control level) were determined from the dose–response curves.

4.6. Flow cytometry

K562 cells (5×10^5 cells/ml) were incubated with test compounds for 24 h. After incubation, cells were centrifuged, washed in PBS and split into two subsamples for viability and cell cycle analysis as described previously [17,19]. For viability analyses, the cells were directly stained with propidium iodide and the percentage of viable cells (*i.e.* those excluding propidium iodide) was determined by flow cytometry using a 488 nm laser (Cell Lab Quanta SC, Beckman Coulter). For cell cycle analysis, the washed cells were fixed with 70% ethanol, washed, stained with propidium iodide and analyzed by flow cytometry using a 488 nm laser (Cell Lab Quanta SC, Beckman Coulter). The sub-diploid fraction was designated as apoptotic.

4.7. Kinase inhibition assay

Kinase assays were performed according to established protocols [17,19]. Briefly, CDK2/cyclin E and Abl kinases were produced in Sf9 insect cells and purified on a NiNTA column (Qiagen). CK2 α 1 was purchased from ProQinase. CDK2 was assayed with 1 mg/ml histone H1 in the presence of 15 μ M ATP, 0.05 μ Ci [γ - ^{33}P]ATP and of the test compound in a final volume of 10 μ L, all in a reaction buffer (60 mM HEPES-NaOH, pH 7.5, 3 mM $MgCl_2$, 3 mM $MnCl_2$, 3 μ M Na-orthovanadate, 1.2 mM DTT, 2.5 μ g/50 μ L PEG_{20,000}). CK2 α 1 was assayed in the same buffer but with casein as a substrate (0.2 mg/ml). Abl was assayed with 500 μ M of a synthetic peptide (GGEAIYAAPFKK) 10 μ M of [γ - ^{33}P]ATP and the appropriate quantity of the test compound in a final volume of 10 μ L, all in a reaction buffer (25 mM Tris, pH 7.5, 5 mM $MgCl_2$, 0.5 mM EGTA, 1 mM DTT, 0.01% Brij35). The reactions were stopped by adding 5 μ L of 3% aq. H_3PO_4 . Aliquots were spotted onto P-81 phosphocellulose (Whatman), washed 3 \times with 0.5% aq. H_3PO_4 and finally air-dried. Kinase inhibition was quantified using a FLA-7000 digital image analyzer (Fujifilm). The concentration of the test compounds required to decrease the kinase activity by 50% was determined from dose–response curves and identified as the IC_{50} .

4.8. Immunoblotting

Immunoblotting analyses were performed as in our previous works [16–19]. Specific antibodies were purchased from Sigma–Aldrich (anti- α -tubulin and peroxidase-labelled secondary antibodies) and Cell Signaling Technology (STAT5, STAT5 phosphorylated at Y694, CrkL phosphorylated at Y207), and Millipore (anti-phosphotyrosine 4G10). Briefly, cellular lysates were prepared by harvesting cells in Laemmli sample buffer. Proteins were separated on SDS-polyacrylamide gels and electroblotted onto nitrocellulose membranes. After blocking, the membranes were incubated with specific primary antibodies overnight, washed and then incubated with peroxidase-conjugated secondary antibodies. Finally, peroxidase activity was detected with ECL+ reagents (AP Biotech) using a CCD camera LAS-4000 (Fujifilm).

4.9. Molecular modelling

3D structures of the studied compounds were prepared using Marvin 5.10.3, 2012, ChemAxon (<http://www.chemaxon.com>). The crystal structures of CDK2 with NU6102 (PDBID: 2c6o) and Abl with PD180970 (PDBID: 2hzi) were used as the protein docking templates. Polar hydrogens were added to all ligands and proteins with the AutoDock Tools program [27]. Molecular docking was

performed using the AutoDock Vina program [28] with a grid box centred on the protein's active site and a box size of 14 Å. The value of the exhaustiveness parameter was set to 20 (default 8). The two crystal-like poses with the lowest energies were selected for further analysis.

Acknowledgements

This research was funded by the National Science Centre, Poland (grant NN405 092340), the Czech Science Foundation (grant P305/12/0783), the Centre of the Region Haná for Biotechnological Agricultural Research (grant no. ED0007/01/01), Palacky University (Student Project PrF_2014_023), the Operational Program Research and Development for Innovations – European Regional Development Fund (CZ.1.05/2.1.00/03.0058). The authors are grateful to Prof. Robert Kawęcki and Dr. Konrad Zdanowski (Siedlce University of Natural Sciences and Humanities) for providing NMR spectra and Prof. W.T. Miller (Stony Brook University, NY) for providing the Abl construct.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.ejmech.2014.03.054>.

References

- [1] J. Zhang, P.L. Yang, N.S. Gray, *Nature Reviews Cancer* 9 (2009) 28–39.
- [2] S. Lapena, A. Giordano, *Nature Reviews Drug Discovery* 8 (2009) 547–566.
- [3] A. Levitzki, *Annual Review of Pharmacology and Toxicology* 53 (2013) 161–185.
- [4] R. Capdeville, E. Buchdunger, J. Zimmermann, A. Matter, *Nature Reviews Drug Discovery* 1 (2002) 493–502.
- [5] A. Quintás-Cardama, H. Kantarjian, J. Cortes, *Nature Reviews Drug Discovery* 6 (2007) 834–848.
- [6] H.J. Lindner, G. Schaden, *Chemische Berichte* 105 (1972) 1949–1955.
- [7] K. Hirata, H. Nakagami, J. Takashina, T. Mahmud, M. Kobayashi, Y. In, T. Ishida, K. Miyamoto, *Heterocycles* 43 (1996) 1513–1519.
- [8] V.V. Smirnov, E.A. Kiprianova, A.D. Garagulya, S.E. Esipov, S.A. Dovjenko, *FEMS Microbiology Letters* 153 (1997) 357–361.
- [9] V. Galasso, *Chemical Physics Letters* 472 (2009) 237–242.
- [10] T.R. Kelly, E.L. Elliott, R. Lebedev, J. Pagalday, *Journal of the American Chemical Society* 128 (2006) 5646–5647.
- [11] T. Gucky, I. Frysova, J. Slouka, M. Hajduch, P. Dzubak, *European Journal of Medicinal Chemistry* 44 (2009) 891–900.
- [12] M. Legraverend, D.S. Grierson, *Bioorganic & Medicinal Chemistry* 14 (2006) 3987–4006.
- [13] R. Jorda, K. Paruch, V. Krystof, *Current Pharmaceutical Design* 18 (2012) 2974–2980.
- [14] V. Krystof, S. Uldrijan, *Current Drug Targets* 11 (2010) 291–302.
- [15] V. Krystof, R. Lenobel, L. Havlicek, M. Kuzma, M. Strnad, *Bioorganic & Medicinal Chemistry Letters* 12 (2002) 3283–3286.
- [16] V. Krystof, I.W. McNae, M.D. Walkinshaw, P.M. Fischer, P. Muller, B. Vojtesek, M. Orsag, L. Havlicek, M. Strnad, *Cellular and Molecular Life Sciences* 62 (2005) 1763–1771.
- [17] M. Zatloukal, R. Jorda, T. Gucky, E. Reznickova, J. Voller, T. Pospisil, V. Malinkova, H. Adamcova, V. Krystof, M. Strnad, *European Journal of Medicinal Chemistry* 61 (2013) 61–72.
- [18] L. Havlicek, K. Fuksova, V. Krystof, M. Orsag, B. Vojtesek, M. Strnad, *Bioorganic & Medicinal Chemistry* 13 (2005) 5399–5407.
- [19] R. Jorda, L. Havlicek, I.W. McNae, M.D. Walkinshaw, J. Voller, A. Sturc, J. Navratilová, M. Kuzma, M. Mistrik, J. Bartek, M. Strnad, V. Krystof, *Journal of Medicinal Chemistry* 54 (2011) 2980–2993.
- [20] M. Mojzych, A. Rykowski, *Heterocycles* 53 (2000) 2175–2181.
- [21] M. Mojzych, A. Rykowski, *Journal of Heterocyclic Chemistry* 44 (2007) 1003–1007.
- [22] T. Gucky, E. Reznicková, P. Dzubak, M. Hajduch, V. Krystof, *Monatshefte für Chemie* 141 (2010) 709–714.
- [23] J. Carmichael, W. Degraff, A. Gazdar, J. Minna, J. Mitchell, *Cancer Research* 47 (1987) 936–942.
- [24] S.W. Cowan-Jacob, G. Zendrich, A. Floersheimer, P. Furet, J. Liebetanz, G. Rummel, P. Rheinberger, M. Centeleghe, D. Fabbro, P.W. Manley, *Acta Crystallographica Section D: Biological Crystallography* 63 (2007) 80–93.
- [25] I.R. Hardcastle, C.E. Arris, J. Bentley, F.T. Boyle, Y. Chen, N.J. Curtin, J.A. Endicott, A.E. Gibson, B.T. Golding, R.J. Griffin, P. Jewsbury, J. Menyerol, V. Mesguiche, D.R. Newell, M.E. Noble, D.J. Pratt, L.Z. Wang, H.J. Whitfield, *Journal of Medicinal Chemistry* 47 (2004) 3710–3722.
- [26] A. Jacquel, M. Herrant, L. Legros, N. Belhacene, F. Luciano, G. Pages, P. Hofman, P. Auberger, *FASEB Journal* 17 (2003) 2160–2162.
- [27] M.F. Sanner, *Journal of Molecular Graphics and Modelling* 17 (1999) 57–61.
- [28] O. Trott, A.J. Olson, *Journal of Computational Chemistry* 31 (2010) 455–461.

Appendix E

Řezníčková E, Popa A, Gucký T, Zatloukal M, Havlíček L, **Bazgier V**, Berka K, Jorda R,
Popa I, Nasereddin A, Jaffe CL, Kryštof V, Strnad M:

2,6,9-Trisubstituted purines as CRK3 kinase inhibitors with antileishmanial activity in
vitro. *Bioorg.*

Med. Chem. Lett., 25(11), 2298-2301, 2015.

DOI: 10.1016/j.bmcl.2015.04.030

IF = 3.746



Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

2,6,9-Trisubstituted purines as CRK3 kinase inhibitors with antileishmanial activity in vitro



Eva Řezníčková^a, Alexandr Popa^a, Tomáš Gucký^a, Marek Zatloukal^a, Libor Havlíček^a, Václav Bazgier^{a,b}, Karel Berka^{b,c}, Radek Jorda^{a,d}, Igor Popa^{a,e,f}, Abdelmajeed Nasereddin^g, Charles L. Jaffe^g, Vladimír Kryštof^{a,*}, Miroslav Strnad^a

^aLaboratory of Growth Regulators & Department of Chemical Biology and Genetics, Centre of the Region Haná for Biotechnological and Agricultural Research, Palacký University and Institute of Experimental Botany AS CR, Šlechtitelů 11, 783 71 Olomouc, Czech Republic

^bDepartment of Physical Chemistry, Faculty of Science, Palacký University, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic

^cRegional Centre of Advanced Technologies and Materials, Faculty of Science, Palacký University, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic

^dRegional Centre for Applied Molecular Oncology, Masaryk Memorial Cancer Institute, Žlutý kopec 7, Brno 656 53, Czech Republic

^eInstitute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University, Hněvotínská 5, 779 00 Olomouc, Czech Republic

^fDepartment of Organic Chemistry, Faculty of Science, Palacký University, 17. listopadu 12, 771 46 Olomouc, Czech Republic

^gDepartment of Microbiology and Molecular Genetics, IMRIC, Hebrew University—Hadassah Medical School, 9112102 Jerusalem, Israel

ARTICLE INFO

Article history:

Received 16 February 2015

Revised 7 April 2015

Accepted 9 April 2015

Available online 16 April 2015

Keywords:

Purine

Inhibitor

Cyclin-dependent kinase

Leishmania

ABSTRACT

Here we describe the leishmanicidal activities of a library of 2,6,9-trisubstituted purines that were screened for interaction with Cdc2-related protein kinase 3 (CRK3) and subsequently for activity against parasitic *Leishmania* species. The most active compound inhibited recombinant CRK3 with an IC₅₀ value of 162 nM and was active against *Leishmania major* and *Leishmania donovani* at low micromolar concentrations in vitro. Its mode of binding to CRK3 was investigated by molecular docking using a homology model.

© 2015 Elsevier Ltd. All rights reserved.

Leishmaniasis, a diverse group of diseases caused by protozoan parasites of the genus *Leishmania*, is a threat to more than 350 million people in 98 countries worldwide.¹ *Leishmania* are transmitted to the host via the bites of sand fly vectors belonging to the genus *Phlebotomus* or *Lutzomyia*, in which they exist as motile extracellular flagellated promastigotes. After transmission to humans or other mammalian hosts, the promastigotes invade macrophages, transform into oval nonflagellated amastigotes, and multiply. This is the stage of the parasite's lifecycle that causes the symptoms of leishmaniasis.²

Leishmaniasis can be divided into three major forms with different clinical manifestations: cutaneous, mucocutaneous and visceral. Cutaneous leishmaniasis is the most common form of the disease and is eventually defeated by the immune system in most cases. However it sometimes progress to the mucocutaneous form where the parasite metastasizes to mucosal tissues. This form is

usually refractory to therapy and can be fatal. The most serious form of the disease, visceral leishmaniasis or Kala-Azar (black fever in Hindi), is fatal if not rapidly diagnosed and treated, and is responsible for most leishmaniasis-related deaths. In visceral leishmaniasis, which is caused by *Leishmania donovani* and *Leishmania infantum* (*syn*=*Leishmania chagasi*), the parasite infects the liver, spleen and bone marrow causing patients to suffer symptoms including fevers, weight loss, and anaemia.²

Current leishmaniasis treatments are based on the administration of amphotericin B, pentavalent antimonials, paromomycin or miltefosine. Many of these drugs are expensive, highly toxic, and require long periods of hospitalization. In addition, parasite resistance to pentavalent antimonials precludes their use in a major endemic region, India.³ It follows that there is an urgent need for the development of new leishmaniasis therapies and the identification of new drug targets that could be exploited in the treatment of this disease. To ensure that new therapies exhibit good selectivity and effectiveness, they should have targets that are absent in the mammalian host or at least different in function and structure from the host homolog. Widely studied targets that are considered

Abbreviation: CRK, Cdc2-related protein kinase.

* Corresponding author. Tel.: +420 585634854; fax: +420 585634870.

E-mail address: vladimir.krystof@upol.cz (V. Kryštof).

<http://dx.doi.org/10.1016/j.bmcl.2015.04.030>

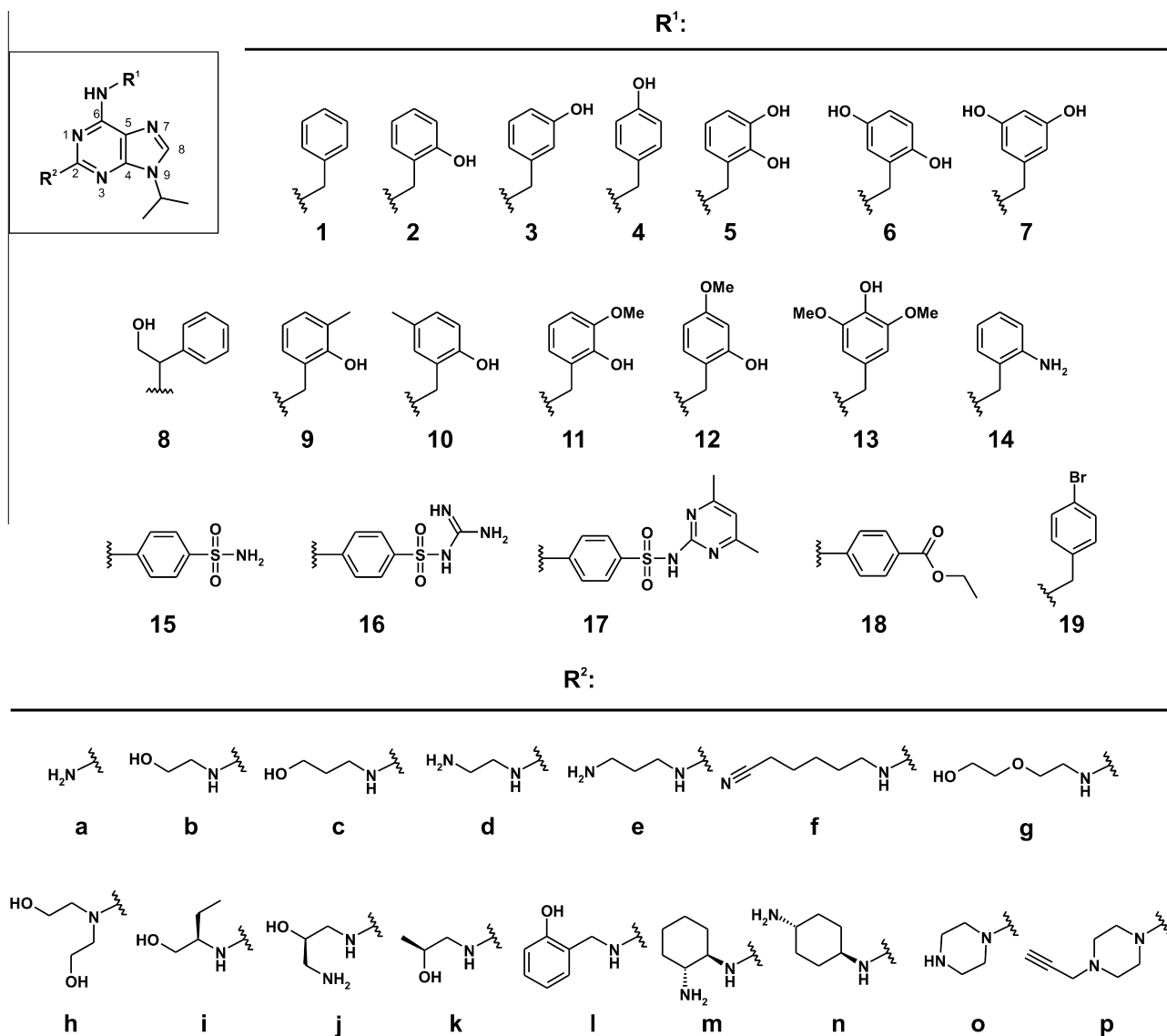
0960-894X/© 2015 Elsevier Ltd. All rights reserved.

particularly promising for this purpose include enzymes involved in essential parasitic metabolic pathways such as sterol biosynthesis or glycolysis; peptidases, which are important for parasite-host cell interaction; and protein kinases important for leishmanial proliferation and viability such as mitogen-activated protein kinases and Cdc2-related protein kinases (CRKs).^{3–5} CRKs belong to the group of serine/threonine protein kinases that play crucial roles in regulating the cell cycle and proliferation.⁶ The most extensively studied CRK kinase is CRK3, a *Leishmania* enzyme essential for parasite proliferation and viability,⁷ and a potential molecular target for drug development.

Several drug-like molecules have been screened for interaction with CRK3^{8–10} including known CDK inhibitors such as flavopiridol, indirubines, paullones, staurosporines, substituted pyrazolo[4,3-*d*]pyrimidines and roscovitine-like purines.^{7,9,11–14} Many of these compounds inhibit the parasitic CRK3, blocking the cell cycle and reducing parasite viability.

Here we report the leishmanicidal activities of a library of selected 2,6,9-trisubstituted purines that are moderate inhibitors of human CDK2, but exhibit low cytotoxicity in cancer cells.^{15–17} Their low cytotoxicity means that they are unlikely to have applications in cancer therapy but may be useful in other fields of

medicine where inactivity towards human cells is beneficial. The compounds were prepared using a previously reported three-step procedure^{15–18} starting from 2,6-dichloropurine, which is alkylated at N9 with isopropyl alcohol under Mitsunobu conditions. This is followed by a nucleophilic substitution at the C6 purine position with a substituted benzylamine, followed by a second nucleophilic substitution at C2 with an amino alcohol or diamine. The synthesis and characterization of the novel compounds in the tested series is presented in the [Supplementary material](#) of this manuscript. We evaluated the CRK3-inhibitory and antiproliferative activities of a library of different 2,6,9-trisubstituted purines ([Scheme 1](#)). All of the compounds share an isopropyl moiety at position 9 but have different substituents at positions 2 and 6. The C2 position was modified with diverse substituents ranging from simple aliphatic chains to branched and cyclic moieties, whereas most of the C6 substituents are substituted benzylamines. Kinase inhibition assays were performed using a complex of the leishmanial CRK3 kinase and CYC6,¹³ and cytotoxicity was determined against the promastigote stage of *Leishmania turanica*, an avirulent species found in the great gerbil (*Rhombomys opimus*) that is non-pathogenic towards humans. Promastigotes were treated for 72 h with tested compounds (30 μ M) and the compounds' cytotoxicity was



Scheme 1. Substituents of studied purines at positions 2 (substituents R²) and 6 (substituents R¹).

expressed in terms of the percentage of viable cells relative to untreated controls.

The kinase inhibition assays showed that many of the tested compounds reduced CRK3 activity to less than 20% of the control level when applied at a concentration of 30 μM (see [Supplementary data](#)). The active compounds exhibit greater diversity at C2 position. Compounds **2c** and **11n**, which have 3-hydroxypropyl and 4-aminocyclohexylamino C2 groups, respectively, reduced the CRK3 activity to 10.5% and 10.8% of its reference level, respectively. The strongest CRK3 inhibitors had a branched 1-(hydroxymethyl)propyl substituent at C2. Compounds in this group such as **2i**, **5i**, **6i**, **9i**, **10i**, **11i**, **15i**, **16i** and **17i** reduced the CRK3 kinase activity to less than 5% of its reference level at 30 μM , suggesting that this C2 substituent is optimal for CRK3 inhibition. Dose–response curves of kinase activity were constructed to obtain more detailed information on the CRK3/CYC6 inhibitory activity of these compounds. Their half maximal inhibitory concentrations (IC_{50} values) ranged from 0.1 to 1.6 μM for compounds **16i** and **9i**, respectively ([Table 1](#)), with the lowest values being achieved by compounds having an N6-benzyl substituent bearing an *o*-hydroxyl group.

To explain these biochemical results, we investigated the binding mode of the most potent compound, **10i**, to CRK3. Since the crystal structure of CRK3 has yet to be determined, we needed to create a model of *Leishmania major* CRK3 (LmCRK3; UNIPROT O96526) based on its homology to the human CDK2 kinase (HsCDK2; PDB: 4KD1),¹⁹ with which it exhibits 58% sequence

identity. This was achieved using an approach similar to one described previously.^{8,10} Interestingly, the CRK3 sequences for six human *Leishmania* pathogens show $\geq 96\%$ sequence identity over 311 amino acids, and their ATP binding pockets all have identical amino acid sequences (NCBI Blast—<http://blast.ncbi.nlm.nih.gov>, and TriTrypDB—<http://tritypdb.org/>) (see [Supplementary data](#)).

The resulting model showed that the LmCRK3 active site is similar to that of HsCDK2 but features several amino acid changes (see the [Supplementary data](#) for a more extensive comparison). The largest difference in the shapes of the active sites is due to the replacement of F82^{HsCDK2} by Y101^{LmCRK3}. It should be noted that while most amino acids in the LmCRK3 active site were conserved in all models, some amino acids (D105, K108) have multiple side-chain rotamers that differed between individual models. Salt bridge formation involving these two amino acids allowed the docked ligands to adopt a common pose in each case.

The common orientation of the active compounds in the LmCRK3 model is similar to that adopted by roscovitine in its complex with HsCDK2 ([Fig. 1](#)). Importantly, the hydrogen bonding D–A–D (donor–acceptor–donor) motif observed in the CDK complex is reproduced in CRK3 via the main chains of E100 and V102. However, the position occupied by the N6-benzyl ring of roscovitine in the HsCDK2 complex is blocked in the LmCRK3 model by the residue Y101^{LmCRK3}. Therefore, the N6 benzyl ring of the docked ligand is instead inclined towards the K108–D105 salt bridge, which interacts favorably with its *o*-hydroxyl group. Finally, the hydroxylated substituent at the ligand's C2 position forms a hydrogen bond with the main-chain carbonyl of A149. These interactions explain why an *o*-hydroxyl group on the N6 benzyl substituent and a 1-(hydroxymethyl)propylamino group at C2 are beneficial for activity against LmCRK3.

CRK3 is essential for the progression of the parasite's cell cycle⁷ and several studies have shown that specific inhibitors of CRK3 can reduce parasite proliferation and viability.^{9,10,12,13} We therefore tested the more active CRK3 inhibitors from the library to determine their cytotoxicity towards *L. turanica*. As expected, several of them (**2i**, **9i**, **10i**, **11i**, **19i**) strongly reduced leishmanial promastigote viability to <50% of the reference level at 30 μM . However, there was no clear correlation between CRK3 inhibitory activity and the degree of viability reduction ([Fig. 2](#)). It is possible that the differences in the compounds' leishmanicidal activity are due to different rates of membrane transport or intracellular metabolism, but interactions with other cellular targets cannot be ruled out either. IC_{50} values against three *Leishmania* species (*L.*

Table 1
CRK3/CYC6 inhibition and antileishmanial activity of the most active 2,6,9-trisubstituted purines

Compound	IC_{50}^a (μM)			
	CRK3/CYC6	<i>L. turanica</i>	<i>L. major</i>	<i>L. donovani</i>
2i	0.590	11.9	47.2	62.4
5i	0.314	>50.0	n.d.	n.d.
6i	0.384	>50.0	n.d.	n.d.
9i	1.648	23.0	19.8	12.4
10i	0.162	11.8	2.3	2.2
11i	0.641	35.4	98.5	95.9
15i	0.214	>50.0	n.d.	n.d.
16i	0.115	>50.0	n.d.	n.d.
19i	n.d.	20.1	19.4	16.8

n.d.—not determined.

^a All values were determined by duplicate assays.

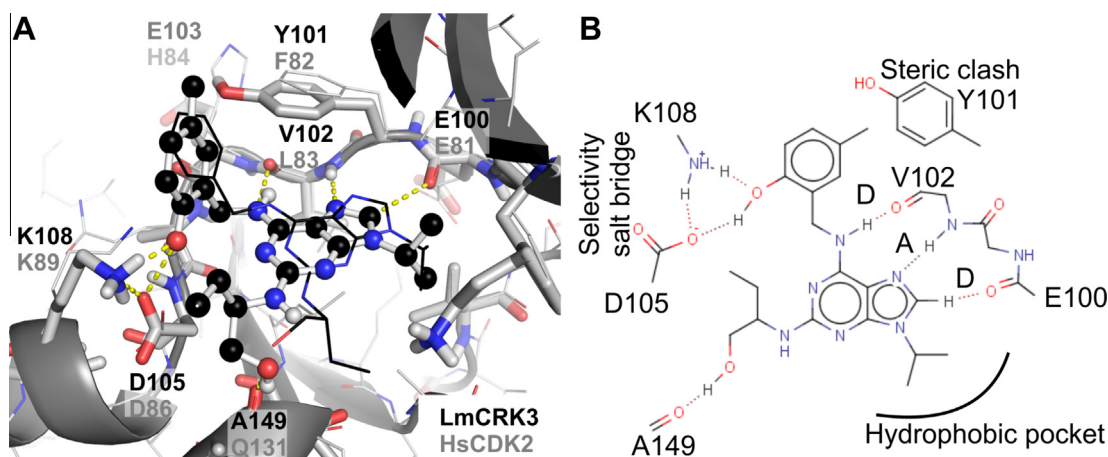


Figure 1. Docked binding pose of compound **10i** in LmCRK3 compared with that of roscovitine in HsCDK2.²⁰ (A) **10i** is shown with a ball-and-stick representation, LmCRK3 is shown as sticks, and the HsCDK2/roscovitine complex (PDB: 2A4L) is shown with lines. (B) **10i** interacts with the main chains of residues E100, V102 and A149 and with a salt bridge formed by K108 and D105. The N9 isopropyl group projects into the hydrophobic pocket of the binding site.

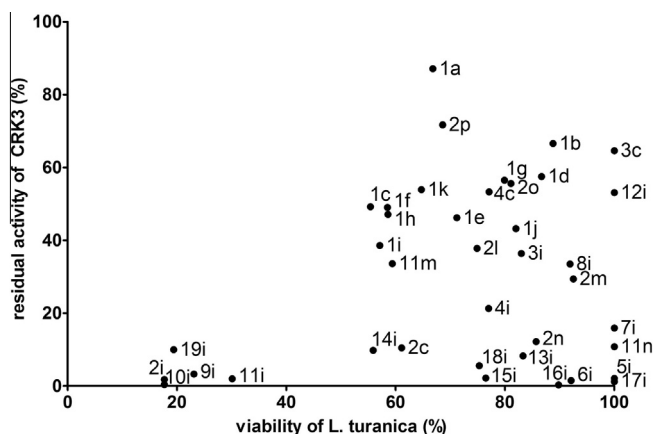


Figure 2. A plot showing the activity of selected compounds against *Leishmania turanica* and CRK3/CYC6 (expressed as percentage of inhibition). All reported measurements are based on at least duplicate experiments using the tested compounds at a concentration of 30 μ M (see [Supplementary data](#) for details).

turanica, and the human parasites *L. major* and *L. donovani*) were determined for the most active compounds. While the IC_{50} values observed for most such compounds against *L. turanica* were comparable to or lower than those for *L. major* and *L. donovani*, compound **10i** was more active against the human parasites, achieving low micromolar IC_{50} values of 2.3 and 2.2 μ M, respectively. We also tested the compounds' cytotoxicity towards normal human fibroblast cells (BJ). At 30 μ M, the most active compounds reduced leishmanial viability to 20% of the reference level seen in control samples while maintaining the viability of non-proliferating human fibroblasts above 75% (see [Supplementary data](#)).

In conclusion, a new group of submicromolar CRK inhibitors has been identified. Their affinity for CRK3 was rationalized by examining the binding of **10i** to a homology model. The compounds also suppressed the proliferation of three different *Leishmania* promastigote species, exhibiting low micromolar activities. Interestingly, some of the compounds also showed similar activity against *Trypanosoma rhodesiense* (data not shown). As such, these compounds could potentially serve as leads for antileishmanial and antitrypanosomal drug development.

Acknowledgements

The authors gratefully acknowledge support from the Ministry of Education, Youth and Sports of the Czech Republic (via projects LO1305 and LO1204), the European Regional Development Fund

and the State Budget of the Czech Republic (RECAMO, CZ.1.05/2.1.00/03.0101), the Operational Program Research and Development for Innovations (IMTM project CZ.1.05/2.1.00/01.0030), Palacký University in Olomouc (IGA_PrF_2015_027, IGA_PrF_2014_024 and IGA_PrF_2015_021) and the German Federal Ministry of Education and Research (Bio-Disc 5 project 315729). C.L.J. holds the Michael and Penny Feiweil Chair in Dermatology.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bmcl.2015.04.030>.

References and notes

- den Boer, M.; Argaw, D.; Jannin, J.; Alvar, J. *Clin. Microbiol. Infect.* **2011**, *17*, 1471.
- Sharlow, E. R.; Grogl, M.; Johnson, J.; Lazo, J. S. *Mol. Interventions* **2010**, *10*, 72.
- Singh, N.; Kumar, M.; Singh, R. K. *Trop. Med.* **2012**, *5*, 485.
- Chawla, B.; Madhubala, R. *J. Parasit. Dis.* **2010**, *34*, 1.
- Seifert, K. *Open Med. Chem. J.* **2011**, *5*, 31.
- Naula, C.; Parsons, M.; Mottram, J. C. *Biochim. Biophys. Acta* **2005**, *1754*, 151.
- Hassan, P.; Fergusson, D.; Grant, K. M.; Mottram, J. C. *Mol. Biochem. Parasitol.* **2001**, *113*, 189.
- Cleghorn, L. A.; Woodland, A.; Collie, I. T.; Torrie, L. S.; Norcross, N.; Luksch, T.; Mpamhanga, C.; Walker, R. G.; Mottram, J. C.; Brenk, R.; Frearson, J. A.; Gilbert, I. H.; Wyatt, P. G. *Chem. Med. Chem.* **2011**, *6*, 2214.
- Grant, K. M.; Dunion, M. H.; Yardley, V.; Skaltsounis, A. L.; Marko, D.; Eisenbrand, G.; Croft, S. L.; Meijer, L.; Mottram, J. C. *Antimicrob. Agents Chemother.* **2004**, *48*, 3033.
- Walker, R. G.; Thomson, G.; Malone, K.; Nowicki, M. W.; Brown, E.; Blake, D. G.; Turner, N. J.; Walkinshaw, M. D.; Grant, K. M.; Mottram, J. C. *PLoS Negl. Trop. Dis.* **2011**, *5*, e1033.
- Reichwald, C.; Shimony, O.; Dunkel, U.; Sacerdoti-Sierra, N.; Jaffe, C. L.; Kunick, C. *J. Med. Chem.* **2008**, *51*, 659.
- Xingi, E.; Smirlis, D.; Myrianthopoulos, V.; Magiatis, P.; Grant, K. M.; Meijer, L.; Mikros, E.; Skaltsounis, A. L.; Soteriadou, K. *Int. J. Parasitol.* **2009**, *39*, 1289.
- Jorda, R.; Sacerdoti-Sierra, N.; Voller, J.; Havlicek, L.; Krcalikova, K.; Nowicki, M. W.; Nasereddin, A.; Krystof, V.; Strnad, M.; Walkinshaw, M. D.; Jaffe, C. L. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4233.
- Houze, S.; Hoang, N. T.; Lozach, O.; Le, B. J.; Meijer, L.; Galons, H.; Demange, L. *Molecules* **2014**, *19*, 15237.
- Havlicek, L.; Hanus, J.; Vesely, J.; LeClerc, S.; Meijer, L.; Shaw, G.; Strnad, M. *J. Med. Chem.* **1997**, *40*, 408.
- Krystof, V.; Lenobel, R.; Havlicek, L.; Kuzma, M.; Strnad, M. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3283.
- Otyepka, M.; Krystof, V.; Havlicek, L.; Siglerova, V.; Strnad, M.; Koca, J. *J. Med. Chem.* **2000**, *43*, 2506.
- Zatloukal, M.; Jorda, R.; Gucky, T.; Reznickova, E.; Voller, J.; Pospisil, T.; Malinkova, V.; Adamcova, H.; Krystof, V.; Strnad, M. *Eur. J. Med. Chem.* **2013**, *61*, 61.
- Martin, M. P.; Olesen, S. H.; Georg, G. I.; Schonbrunn, E. *ACS Chem. Biol.* **2013**, *8*, 2360.
- de Azevedo, W. F. J.; Mueller-Dieckmann, H. J.; Schulze-Gahmen, U.; Worland, P. J.; Sausville, E.; Kim, S. H. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2735.

Appendix F

Řezníčková E, Weitensteiner S, Havlíček L, Jorda R, Gucký T, Berka K, **Bazgier V**, Zahler S, Kryštof V, Strnad M:

Characterization of a Pyrazolo[4,3- d]pyrimidine Inhibitor of Cyclin-Dependent Kinases 2 and 5 and Aurora A With Pro-Apoptotic and Anti-Angiogenic Activity In Vitro.

Chem. Biol. Drug. Des., 86(6), 1528–1540, 2015.

DOI: 10.1111/cbdd.12618

IF = 2.396



Characterization of a Pyrazolo[4,3-*d*]pyrimidine Inhibitor of Cyclin-Dependent Kinases 2 and 5 and Aurora A With Pro-Apoptotic and Anti-Angiogenic Activity *In Vitro*

Eva Řezníčková¹, Sabine Weitensteiner²,
Libor Havlíček³, Radek Jorda¹, Tomáš Gucký¹,
Karel Berka⁴, Václav Bazgier^{1,5}, Stefan Zahler²,
Vladimír Kryštof^{1,*} and Miroslav Strnad¹

¹Laboratory of Growth Regulators & Department of Chemical Biology and Genetics, Centre of the Region Haná for Biotechnological and Agricultural Research, Palacký University and Institute of Experimental Botany AS CR, Šlechtitelů 27, 78371 Olomouc, Czech Republic

²Department of Pharmacy, LMU Munich – Center for Drug Research – Pharmaceutical Biology, Butenandtstr. 5-13, 81377 Munich, Germany

³Isotope Laboratory, Institute of Experimental Botany ASCR, Vídeňská 1083, 14220 Prague, Czech Republic

⁴Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacký University, 17. listopadu 12, 77146 Olomouc, Czech Republic

⁵Department of Physical Chemistry, Faculty of Science, Palacký University, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic

*Corresponding author: Vladimír Kryštof, vladimir.krystof@upol.cz

Selective inhibitors of kinases that regulate the cell cycle, such as cyclin-dependent kinases (CDKs) and aurora kinases, could potentially become powerful tools for the treatment of cancer. We prepared and studied a series of 3,5,7-trisubstituted pyrazolo[4,3-*d*]pyrimidines, a new CDK inhibitor scaffold, to assess their CDK2 inhibitory and antiproliferative activities. A new compound, 2i, which preferentially inhibits CDK2, CDK5, and aurora A was identified. Both biochemical and cellular assays indicated that treatment with compound 2i caused the downregulation of cyclins A and B, the dephosphorylation of histone H3 at Ser10, and the induction of mitochondrial apoptosis in the HCT-116 colon cancer cell line. It also reduced migration as well as tube and lamellipodia formation in human endothelial cells. The kinase inhibitory profile of compound 2i suggests that its anti-angiogenic activity is linked to CDK5 inhibition. This dual mode of action involving apoptosis induction in cancer cells and the blocking of angiogenesis-like activity in endothelial cells offers possible therapeutic potential.

Key words: angiogenesis, apoptosis, aurora A, cyclin-dependent kinase, inhibitor, pyrazolo[4,3-*d*]pyrimidine

Abbreviations: BrdU, 5-bromo-2'-deoxyuridine; CDK, cyclin-dependent kinase; IC₅₀, half maximum inhibitory concentration; FITC, fluorescein isothiocyanate; HMEC, human microvascular epithelial cells; HUVEC, human umbilical vein endothelial cells; PARP-1, poly-(ADP-ribose)polymerase-1; PCNA, proliferating cell nuclear antigen; PDGFR, platelet-derived growth factor receptor; Plk-1, Polo-like kinase 1; PUMA, p53 upregulated modulator of apoptosis; Rb, retinoblastoma; VEGF, vascular endothelial growth factor.

Received 9 April 2015, revised 9 July 2015 and accepted for publication 13 July 2015

Established hallmarks of cancer cells include sustained proliferative signaling and evasion of growth suppressors (1). Both of these processes have been linked to genetic and epigenetic changes in cell cycle regulating genes. The selective targeting of cyclin-dependent kinases (CDKs) was one of the first proposed approaches to cancer therapy and has been studied extensively (2). While CDK2 is not regarded as a good target for cancer drugs (3,4), several synthetic CDK inhibitors have entered clinical trials as drug candidates for oncological indications (5,6). The negative perception of CDKs as drug targets was reinforced by the discovery that several CDKs have compensatory roles in the cell cycle. For example, CDK1 can compensate for the absence or inhibition of CDK2 and vice-versa (7). On the other hand, most inhibitors that have been tested in clinical trials target multiple CDKs and so may advantageously overcome this compensation.

While some existing inhibitors target CDKs that regulate cell cycling, many also interfere with the transcriptional CDKs 7 and 9. CDK9 in particular has been identified as a target of most of the CDK inhibitors that have entered clinical trials, e.g. flavopiridol, roscovitine, SNS-032, and AT7519 (5,6,8). CDK9 inhibition causes the downregulation of short half-life proteins connected to apoptosis (notably, Mcl-1, and XIAP) that are critical for the survival of cancer cells, especially those causing multiple myeloma and chronic lymphocytic leukemia (9–12).



Rather surprisingly, some CDK inhibitors that have been evaluated in clinical studies (SNS-032, flavopiridol, and ZK 304709) have also been shown to exert anticancer effects *in vivo* by interfering with tumor angiogenesis (13–15). This was ascribed to the down-regulation of both the mRNA and protein of vascular endothelial growth factor (VEGF), the most potent known tumor angiogenic factor, via CDK9 inhibition. However, we have recently shown that CDK5, previously identified as a regulator of neuronal processes, also regulates two processes that are essential for angiogenesis: endothelial cell migration and tube formation (16). Moreover, we found that the anti-angiogenic activity of several CDK inhibitors with different structures, including roscovitine, is at least partially due to interference with CDK5 (16,17).

We recently prepared a pyrazolo[4,3-*d*]pyrimidine bioisostere of the well-known CDK inhibitor roscovitine (18). The bioisostere proved to be a more potent CDK inhibitor than roscovitine itself, with greater anticancer activity *in vitro*. We therefore sought to build on this finding by exploiting our knowledge of structure-activity relationships in purine CDK inhibitors to prepare a small library of potential CDK inhibitors with pyrazolo[4,3-*d*]pyrimidine skeletons. Here, we describe the synthesis of these compounds and present data on their biological activity.

Methods and Materials

Synthesis

Melting points were determined on a Kofler block and are uncorrected. NMR spectra were measured on a UNITY Inova-400 or Gemini-300 spectrometer (Varian, Yarnton, UK) using deuterated solvents at 303 K. The residual solvent signal was used as an internal standard. Chemical shifts are given in the δ -scale [p.p.m.], and coupling constants in Hz. ESI or APCI mass spectra were determined using a Micromass ZMD mass spectrometer (Waters, Milford, MA, USA) (direct inlet, coin voltage 20 V). IR spectra were recorded on a Nicolet 200 FTIR instrument (Thermo Scientific, Waltham, MA, USA). Column chromatography was performed using silica gel Kieselgel 60 (Merck, Darmstadt, Germany) (230–400 mesh). The purity of all synthesized compounds was evaluated by HPLC-PDA (200–500 nm) and was $\geq 95\%$ in all cases other than compounds **1c** (82%) and **1e** (89%). Detailed information about synthesis and characterization of the compounds are described in Supporting information (Appendix S1).

Kinase inhibition

All prepared compounds were evaluated for the inhibition of CDK2/cyclin E as previously described (18,19). Briefly, CDK2/cyclin E was produced in Sf9 insect cells via baculoviral infection and purified on a NiNTA column (Qiagen, Hilden, Germany). The kinase reactions were assayed with 1 mg/mL histone H1 in the presence of 15 μM ATP, 0.05 μCi [γ - ^{32}P]ATP, and the tested compound in a final vol-

Pyrazolo[4,3-*d*]pyrimidine Kinase Inhibitors

ume of 10 μL , all in reaction buffer (60 mM HEPES-NaOH, pH 7.5, 3 mM MgCl_2 , 3 mM MnCl_2 , 3 μM Na-orthovanadate, 1.2 mM DTT, 2.5 $\mu\text{g}/50 \mu\text{L}$ PEG_{20,000}). The reaction was stopped by adding 3% aq. H_3PO_4 ; aliquots were spotted onto P-81 phosphocellulose (Whatman, Maidstone, UK), washed with 0.5% aq. H_3PO_4 , and air-dried.

The protein kinase selectivity of compound **2i** was tested against a panel of CDKs and aurora A (Table 2) using the ProKinase platform according to the manufacturer's protocol. Compound **2i** was also profiled at concentrations of 1 and 10 μM by screening against 88 enzymes (Table S1) as described previously (20).

Molecular docking

All 3D structures of studied compounds were prepared using Marvin 14.9.8, 2014, ChemAxon (<http://www.chemaxon.com>). The crystal structures of CDK2 with Cyclin A (PDBID: 3DDQ), CDK2 with Cyclin E (PDBID: 1W98), and aurora A (PDBID: 3O50) were used as the protein docking templates. Polar hydrogens were added to all ligands and proteins with the AutoDock Tools program (21). Molecular docking was performed using the AutoDock Vina program (22) with a grid box centered on the protein's active site and a box size of 18 Å. The value of the exhaustiveness parameter was set to 30 (default 8).

Cell maintenance and cytotoxicity assay

The cells were cultured in DMEM or RPMI-1640 (supplemented with 10% fetal calf serum, 4 mM glutamine, 100 IU/mL penicillin, 100 $\mu\text{g}/\text{mL}$ streptomycin) in a humidified CO_2 incubator at 37 °C and then seeded into 96-well microtiter plates at appropriate densities for their cell sizes and growth rates. After a preincubation period, the cells were treated with the test compound at six different concentrations, in triplicate, and incubated for 72 h. A solution of Calcein AM was then added and the fluorescence of the live cells was measured at 485 nm/538 nm (ex/em) with a Fluoroskan Ascent microplate reader (Labsystems, Helsinki, Finland). IC_{50} values, i.e. the compound concentrations required to reduce the number of viable cells by 50% relative to an untreated control, were determined from the dose-response curves.

Immunoblotting

Immunoblotting was performed as described previously (18,19). Total cellular protein lysates were prepared by harvesting treated cells in Laemmli sample buffer. Proteins were separated on SDS-polyacrylamide gels and electroblotted onto nitrocellulose membranes. The blotted membranes were stained with 0.2% Ponceau-S in 1% acetic acid to confirm equal protein loading, then destained and blocked using PBS and 0.1% Tween 20 (PBS-T) with 5% low fat milk or 3% bovine serum albumin (BSA). The membranes were then incubated with specific antibodies

overnight at 4 °C. After washing three times in PBS-T, the membranes were incubated with peroxidase-conjugated secondary antibodies. After another three washes in PBS-T, peroxidase activity was determined using ECL kit (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions.

Antibodies

Specific antibodies were purchased from Santa Cruz Biotechnology, Dallas, TX, USA (anti-ARK-1, clone N-20; anti-ARK-2, clone E-15; anti-Bcl-2, clone 100; anti-cyclin B1, clone GNS1; anti-cyclin E, clone HE12; anti-Mcl-1, clone S-19; anti-MDM-2, clone SMP14 and anti-PARP-1, clone F-2), Cell Signaling, Danvers, MA, USA (anti-phospho-Plk-1 Thr210; anti-PUMA; anti-phospho-Rb Ser807/811; anti-Rb, clone 4H1), DAKO Cytomation, Glostrup, Denmark (anti-caspase 3), Millipore, Billerica, MA, USA (anti-Histone H3 phosphorylated at Ser10), Beckman Coulter/Immunotech, Marseille, France (anti-cyclin A2-FITC), Roche Applied Science, Penzberg, Germany (anti-BrdU-FITC), Jackson Laboratories, West Grove, PA, USA (goat-anti-mouse FITC), Invitrogen, Carlsbad, CA, USA (goat-anti-rabbit-Alexa Fluor 488), Sigma Aldrich, St. Louis, MO, USA (peroxidase-labelled secondary antibodies) or were generously gifted by Dr. B. Vojtěšek (anti-p53, clone DO-1; anti-PCNA, clone PC-10).

Cell cycle analysis

Sub-confluent HCT-116 cells were treated with compound **2i** at different concentrations for 24 h. Cells treated with etoposide and nocodazole were analyzed simultaneously for comparative purposes. To study proliferation, the cultures were pulse-labeled with 10 μM 5-bromo-2'-deoxyuridine (BrdU) for 30 min before harvesting. The cells were then trypsinized, washed with PBS containing 1% BSA (PBS/BSA), fixed with ice-cold 70% ethanol, incubated on ice for 30 min, washed with PBS/BSA again, and resuspended in 2 M HCl containing 0.5% Triton-X100 for 30 min at room temperature to denature their DNA. After neutralization with 0.1 M $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$, the cells were harvested by centrifugation and washed with PBS/BSA containing 0.5% Tween-20. They were then stained with anti-BrdU fluorescein-labeled antibody for 1 hour at room temperature in the dark. To study histone H3 phosphorylation and the expression of cyclins A and B, cells were harvested by trypsinization, washed in PBS, fixed with ice-cold 90% methanol, incubated on ice for 30 min and washed with PBS/BSA containing 0.5% Tween-20. For cyclin A analysis, cells were incubated with an anti-cyclin A2 fluorescein-labeled antibody for 1 h in dark. Cells for the analysis of cyclin B expression and histone H3 phosphorylation were initially incubated with the appropriate primary antibody for 1 h at room temperature, then washed with PBS containing 1% BSA and incubated with the appropriate fluorescent-labeled secondary antibody for 1 h in the dark. After washing with PBS/BSA, each sample was

incubated with propidium iodide (final concentration 10 $\mu\text{g}/\text{mL}$) and RNase A (final concentration 200 $\mu\text{g}/\text{mL}$) for 30 min at room temperature in the dark. The cells were then analyzed by flow cytometry using a 488 nm laser (Cell Lab Quanta SC; Beckman Coulter, Brea, CA, USA).

Caspase-3/7 activity assay

The cells were harvested by centrifugation and homogenized in an extraction buffer (20 mM TRIS pH 7.4, 100 mM NaCl, 5 mM EDTA, 2 mM EGTA, 2 mM NaF, 0.2% Nonidet p40, inhibitors of proteases, pH 7.4) on ice for 30 min. The samples were centrifuged at 17 000 $\times g$ for 25 min at 4 °C; the proteins were quantified and diluted to equal concentrations. The lysates were then incubated for 4 h with 100 μM Ac-DEVD-AMC as a substrate in the assay buffer (25 mM PIPES, 2 mM EGTA, 2 mM MgCl_2 , 5 mM DTT, pH 7.3). The fluorescence of the product was measured using a Fluoroskan Ascent microplate reader (Labsystems, Helsinki, Finland) at 346/442 nm (ex/em).

Analysis of transcription

The cells were seeded in media containing [^{14}C]thymidine (62.5 Bq/mL in growth medium) for 36 h and then treated with various concentrations of compound **2i** in fresh cultivation medium for 24 h. During the final 30 min of the drug treatment period, nascent RNA was labeled by adding [^3H]uridine (0.75 MBq/mL) to the cultivation media. mRNA was isolated from cells using the Oligotex Direct mRNA kit (Qiagen, Hilden, Germany). For measurements of total RNA synthesis, cell lysates in 1% SDS were precipitated with 10% TCA after which the TCA-insoluble material was collected on spin filters (Invitek, Hayward, CA, USA) and washed. The nucleic acids were then eluted with 1 M NaOH. The ^3H and ^{14}C in the eluate were quantified simultaneously using an LS6500 liquid scintillation counter (Beckman Coulter, Brea, CA, USA). Relative total RNA synthesis and mRNA synthesis rates were determined by calculating the $^3\text{H}/^{14}\text{C}$ ratio for each sample and comparing it to that for an untreated control sample.

Scratch assay (Migration assay)

HUVECs were seeded into a 24-well plate. After reaching confluency, the cells were scratched using a 100 μL pipette tip. The wounded monolayers were washed twice with PBS containing $\text{Ca}^{2+}/\text{Mg}^{2+}$; then growth medium containing the appropriate concentration of the test compound was added. After 16 h of migration, the cells were washed and subsequently fixed with 4% paraformaldehyde. We have previously shown that under these conditions cell proliferation does not contribute substantially to scratch closure. Images were acquired using the TILLvisON system (TILL Photonics GmbH, Gräfelfing, Germany) and a CCD-camera connected to an Axiovert 200 microscope (Zeiss, Oberkochen, Germany). Quantitative image analyses of the cell-covered area were performed by Wimasis GmbH, Munich (23).

Quantification of lamellipodia and immunocytochemistry

Confluent layers of HUVECs were scratched and stimulated in eight-well μ -slides as described above. The cells were allowed to migrate for 8 h in the presence or absence of the appropriate concentrations of the tested compounds, until lamellipodia were clearly visible in the control. The actin cytoskeleton was then stained with rhodamine-phalloidin (Life Technologies, Darmstadt, Germany) and fluorescence images of the scratches were taken in 10 \times magnification with a Zeiss LSM 510 META instrument. For quantitative evaluation of lamellipodia formation, cells with prominent lamellipodia and ruffles were counted using IMAGEJ (Cell Counter plug-in by Kurt De Vos) in relation to the total number of cells at the scratch front. The ratio was calculated as the proportion of lamellipodia-positive cells per 100 cells at the scratch front. Immunofluorescence staining was performed using the same apparatus. Fixed cells were permeabilized for 2 min with 0.1 % Triton X-100 in PBS, after which the cells were washed and unspecific binding was blocked with 0.2% BSA in PBS for 15 min. The cells were then incubated with a primary antibody (cortactin, Cell Signaling, and Rac1, Upstate) in 0.2% BSA/PBS overnight at 4 °C. After three washing steps with PBS, the specimens were incubated with the appropriate AlexaFluor[®]-labeled secondary antibodies for 30 min at room temperature. Finally, the slides were again washed three times with PBS (5 min), embedded in FluorSave[™] Reagent mounting medium, and covered with 8 mm \times 8 mm glass cover slips (23).

Tube formation assay

Precooled BD Matrigel[™] Matrix Growth Factor Reduced (GFR) (BD Biosciences, Heidelberg, Germany) was placed in the lower compartment of μ -slide Angiogenesis wells (ibidi GmbH, Martinsried, Germany) on ice. The Matrigel[™] Matrix was polymerized by incubating the slides at 37 °C for

30 min. 12 000 HUVECs/well were then seeded onto the Matrigel[™] and stimulated for 16 h. The level of tube formation was determined by light microscopy using the TILLvisON system. Quantitative image analysis of tube length, number of branching points, and tube number was performed with a software tool from Wimasis GmbH, Munich (23).

Results

Synthesis

The tested compounds were synthesized using previously published general procedures (18,23,24).

A first series of 7-benzylamino substituted pyrazolo[4,3-*d*]pyrimidines having various side chains at the 5 position (compounds **1b–n**) was prepared by the nucleophilic substitution of 5-chloro- or 5-methylsulfonyl- $\{7$ -benzylamino-3-isopropyl-1*H*-pyrazolo[4,3-*d*]pyrimidin-5-yl $\}$ amine with appropriate amines at 120–150 °C (Figure 1) (18,23,24).

Compounds of the second series, i.e. 7-(2-aminobenzyl) amino-substituted pyrazolo[4,3-*d*]pyrimidines **2d–n**, were obtained in higher yields from 5-chloro-7-(2-aminobenzyl) amino-3-isopropyl-1*H*-pyrazolo[4,3-*d*]pyrimidine (**2b**). In particular, this route was the only viable way of preparing compounds **2i** and **2j**, which have bulky side chains. When 5-methylsulfonyl-7-(2-aminobenzyl)amino-3-isopropyl-1*H*-pyrazolo[4,3-*d*]pyrimidine was used as the starting material in reactions with sterically demanding amines, an elimination reaction occurred within the (2-aminobenzyl) side chain that primarily generated the corresponding 7-aminopyrazolo[4,3-*d*]pyrimidines. Compound **2b** was prepared by reacting the highly condensed heteroaromatic system 4,9-dichloro-2,7-diisopropyl-1,3,5,5b,6,8,10,10b-octaazacyclopenta[*h*,*i*]aceanthrylene with (2-aminobenzyl) amine (24).

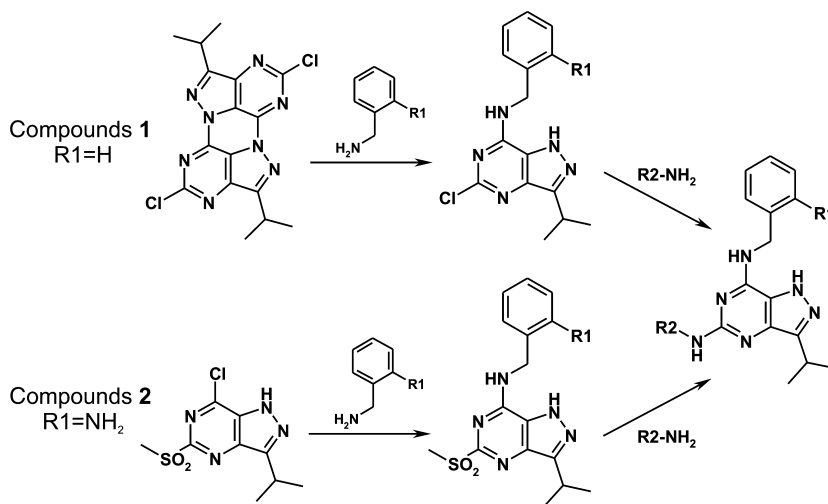


Figure 1: Synthesis of the new pyrazolo [4,3-*d*]pyrimidines. R1 and R2 substituents are listed in Table 1.

Structure-activity relationship for CDK inhibition and cytotoxicity

A series of 3,5,7-trisubstituted pyrazolo[4,3-*d*]pyrimidines was prepared based on our knowledge of the structure-activity relationships of CDK inhibitors with purine cores.

All of the prepared compounds have a common isopropyl moiety at the 3 position of the pyrazolo[4,3-*d*]pyrimidine

ring system because an isopropyl group was found to be the optimal substituent at the analogous position in purine CDK inhibitors (25–29). The 5 and 7 positions were also substituted with groups that conferred high activity in purines, i.e. benzylamines at the 7 position and alkyl- or cycloalkyl- amines at position 5. Table 1 summarizes the CDK2 inhibitory activities and antiproliferative properties of the new compounds.

Table 1: Substituents and biological activities of the tested pyrazolo[4,3-*d*]pyrimidine compounds

General structure	Cmpd	Substituents		IC ₅₀ (nM) ^a CDK2/E	GI ₅₀ (μM) ^a		
		R1	R2		MCF-7	K562	
	1b	H	Cl	130.2 ± 2.3	>100	>100	
	1c	H	H ₂ N-NH	690.0 ± 84.9	40.9 ± 15.1	53.1 ± 13.9	
	1d	H	H ₂ N-CH ₂ -CH ₂ -NH	128.3 ± 62.2	4.6 ± 1.1	10.8 ± 2.1	
	1e	H	H ₂ N-CH(OH)-CH ₂ -NH	111.0 ± 51.1	23.4 ± 5.1	30.9 ± 5.6	
	1f	H	CCCCCCCC-NH	>1000	21.4 ± 8.2	17.0 ± 3.4	
	1g	H	HO-CH(CH ₃)-CH ₂ -NH	29.3 ± 9.3	19.9 ± 4.4	29.7 ± 16.1	
	1i	H	HO-CH(CH ₂ CH ₃)-CH ₂ -NH	46.1 ± 26.9	5.9 ± 2.2	8.0 ± 1.7	
	1k	H	Cyclohexyl-NH ₂ -NH	100.7 ± 8.1	6.3 ± 1.3	4.5 ± 0.4	
	1l	H	Cyclohexyl-NH ₂ -NH	52.3 ± 6.8	2.7 ± 1.1	2.8 ± 0.8	
	1n	H	Cyclohexyl-NH ₂ -NH	186.7 ± 70.9	22.4 ± 19.0	20.9 ± 17.2	
	2a	NH ₂	H	H	367.5 ± 43.5	>100	>100
	2b	NH ₂	Cl	H	461.1 ± 72.0	>100	>100
	2d	NH ₂	H ₂ N-CH ₂ -CH ₂ -NH	H	230.0 ± 88.9	47.5 ± 31.1	63.6 ± 27.9
	2g	NH ₂	HO-CH(CH ₃)-CH ₂ -NH	H	16.5 ± 8.1	4.4 ± 0.5	15.5 ± 1.0
	2h	NH ₂	HO-CH(CH ₂ CH ₃)-CH ₂ -NH	H	30.5 ± 19.0	12.9 ± 2.0	41.0 ± 13.2
	2i	NH ₂	HO-CH(CH ₂ CH ₃)-CH ₂ -NH	H	10.5 ± 3.4	2.8 ± 0.4	9.2 ± 1.4

Table 1: continued

Cmpd	Substituents		IC ₅₀ (nM) ^a CDK2/E	GI ₅₀ (μM) ^a	
	R1	R2		MCF-7	K562
2j	NH ₂		15.3 ± 1.2	10.9 ± 4.0	22.3 ± 1.7
2m	NH ₂		313.3 ± 102.6	23.5 ± 16.1	21.2 ± 13.2
2n	NH ₂		93.3 ± 37.9	47.6 ± 25.7	35.6 ± 7.6

^aAverage values from at least three determinations ± SD.

Among prepared 7-benzylamino-derivatives, three compounds possess linear side chain at 5. The weakest CDK2 inhibitory activity was observed with compound **1f** having heptylamino at 5 (IC₅₀ value of 4.45 μM). Small substituents with terminal amino group increased activity of derivatives **1c** and **1d**. Interestingly, **1f** was still more potent in terms of cytotoxicity, probably due to its lower polarity and easier cellular uptake. Branched alkyl derivatives **1e**, **1g** and **1i** reached nanomolar potency in CDK2 assays. Introduction of substituted cyclohexyl at 5 further enhanced substantially not only CDK2 inhibition, but also cytotoxicity; **1k** and **1l** having (2-aminocyclohexyl)amino yielded single digit micromolar IC₅₀ values. Unexpectedly, closely related **1n** with (4-aminocyclohexyl)amino functions demonstrated 10-fold drop in cellular potency.

We also prepared compounds with a 2-aminobenzylamino moiety at position 7. The effects of varying the substituent at the 5 position in this series were almost identical to those observed in the 2-benzylamino series. Because purine derivatives bearing the *R* isomers of (2-hydroxypropyl)amine and [(2-hydroxymethyl)propyl]amine at the 5 position are known to be more active than their *S* counterparts (19), we prepared the corresponding pyrazolo[4,3-*d*]pyrimidine derivatives in both optically active isoforms (compounds **2g** and **2h**, and **2i** and **2j** respectively). In both cases, the *S* isomer was less active than the *R*. However, all of these compounds had lower IC₅₀ values (around 20 nM) than other members of the series. The most potent CDK2 inhibitor from the two series, compound **2i** (IC₅₀ = 10 nM), was subjected to further biological profiling.

Antiproliferative effects of compound **2i**

The antiproliferative activity of compound **2i** was tested against a panel of human cancer cell lines of different origins (Table 2). It exhibited strong activity in all of the tested lines, with an average IC₅₀ of 5 μM. It also reduced the

Table 2: Cytotoxicity of compound **2i** in human cancer cell lines

Cell line	Compound 2i IC ₅₀ (μM) ^a
MCF-7	2.8 ± 0.4
MDA-MB-468	4.0 ± 0.3
HeLa	4.8 ± 1.8
HCT-116	3.5 ± 0.3
HT-29	9.8 ± 3.6
CEM	3.5 ± 0.1
K562	9.2 ± 1.4
G-361	4.2 ± 0.4
RPMI-8226	4.4 ± 0.2
U266	3.0 ± 0.3
HOS	5.5 ± 0.4

^aAverage values from three determinations ± SD.

proliferation of normal HMEC-1 cells with an IC₅₀ value of 5.4 μM. However, compound **2i** had no adverse effect on the viability of these cells (data not shown). In addition, non-proliferating (confluent) human fibroblasts, treated with different concentrations of compound **2i** for 72 hours, showed no significant reduction of the cell viability up to the concentration of 100 μM.

Protein kinase selectivity

Due to the high structural similarity between roscovitine and compound **2i**, we initially explored its inhibitory potency against a panel of CDKs and other kinases. The compound proved to be selective toward CDKs, with sub-micromolar IC₅₀ values for CDK2 and CDK5 (70 and 260 nM, respectively) and low micromolar IC₅₀ values for CDK9 and CDK1 (1.91 and 3.13 μM respectively). Aside from aurora A, which it also inhibited at sub-micromolar concentrations (IC₅₀ = 0.40 μM), it was not active against any of the other tested kinases (Table 3, Table S1).

Table 3: Protein kinase selectivity of compound **2i**, roscovitine, and JNJ-7706621

Protein kinase	IC ₅₀ (μM)		
	Compound 2i ^a	Roscovitine ^b	JNJ-7706621 ^c
CDK1/cyclin B	3.134 ± 0.021	9.900	0.009
CDK2/cyclin A	0.125 ± 0.001	3.000	0.004
CDK2/cyclin E	0.072 ± 0.003	0.500	0.003
CDK4/cyclin D1	21.730 ± 2.970	29.000	0.253
CDK5/p25NCK	0.257 ± 0.002	3.900	<i>n.d.</i>
CDK6/cyclin D1	>100	27.000	0.175
CDK7/cyclin H/Mat1	>100	1.700	<i>n.d.</i>
CDK9/cyclin T	1.914 ± 0.129	1.600	<i>n.d.</i>
Aurora A	0.395 ± 0.006	>100	0.011

^aAverage values from two determinations ± SD; ^bData for roscovitine were previously published (50); ^cData for JNJ-7706621 were previously published (30).

To confirm the CDK2 inhibitory properties of **2i**, we evaluated its effect on the abundance of Ser807/811-phosphorylated retinoblastoma (Rb) protein in HCT-116 cells. The commercially available dual CDK/Aurora inhibitor JNJ-7706621 (30) and the pan-selective CDK inhibitor dinaciclib (31) were used as reference compounds. All three compounds reduced the phosphorylation of Rb at Ser807/811 after 16 h treatment (Figures 2A and S1A). In addition, studies on two downstream targets of Aurora A – Thr210 of Polo-like kinase 1 (Plk-1) (32,33) and Ser10 of histone H3 (34) – revealed that this kinase is also inhibited by **2i**. HCT-116 cells were synchronized in G2/M phase by treatment with nocodazole for 16 h and subsequently treated for 2 h with **2i** or one of the reference compounds. Both **2i** and JNJ-7706621 reduced the nocodazole-induced phosphorylation of Plk-1 at Thr210 and histone H3 at Ser10 but dinaciclib did not affect the level of either phosphorylated protein (Figures 2B and S1B).

Molecular modeling

The affinity of compound **2i** for aurora A was rather surprising given that roscovitine does not inhibit this kinase. We therefore used molecular modeling to study their interaction. An initial analysis with the Autodock Vina program (22) suggested that compound **2i** binds to the ATP-binding site of CDK2/cyclin A (PDBID: 3DDQ) in an orientation similar to that adopted by roscovitine in CDK2 (PDBID: 2A4L) (35), forming two hydrogen bonds with the main chain of Leu83 and one with the main chain of Glu81. Its binding is further stabilized by additional hydrogen bonds involving its 2-aminobenzylamino moiety, one with Asp86 as a donor and the other with Lys89 as an acceptor (Figure S2). This may be why compound **2i** binds more strongly than roscovitine to CDK2: the estimated binding energies for the two compounds obtained using Autodock Vina are −9.4 and −8.9 kcal/mol respectively. We then analyzed the interactions between

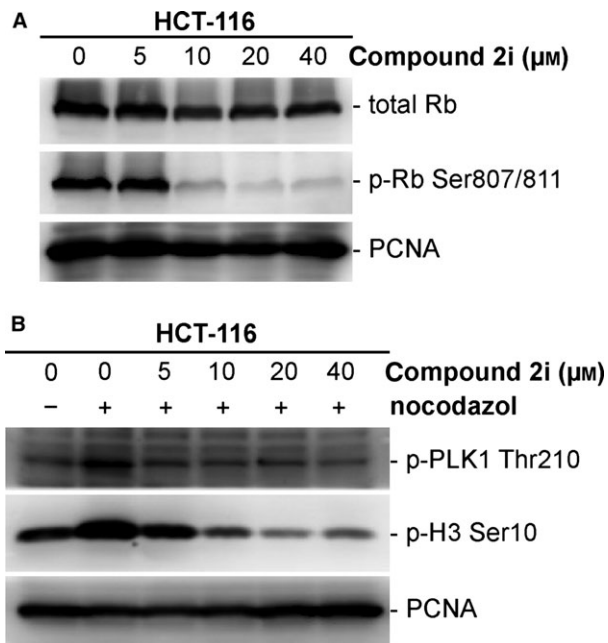


Figure 2: Compound **2i** reduces phosphorylation of Rb, Plk-1 and histone H3 proteins in HCT-116 cells. (A) Analysis of Rb phosphorylation at Ser807/811. Cells were treated with indicated concentrations of tested compound for 16 h. (B) Analysis of Plk-1 and histone H3 phosphorylation. Cells were synchronized using nocodazole (40 ng/mL) for 16 h and subsequently treated with indicated concentrations of tested compound for 2 h. Proliferating cell nuclear antigen (PCNA) levels were determined to confirm equal protein loading.

compound **2i** and aurora A, for which a binding energy of −7.9 kcal/mol was predicted. This revealed four hydrogen bonds that govern the binding of the compound in the kinase's active site: one between the Ala213 carbonyl group and N7, another between the Ala213 amino group and N1, a third between the Glu211 carbonyl group and N2, and a fourth between the Pro214 carbonyl group and the 2-aminobenzylamino substituent (Figure S2). The docking studies thus explained why compound **2i** binds more tightly than roscovitine to the ATP-binding site of CDK2 and suggest that it may bind to the aurora A kinase in a similar fashion.

Cell cycle analysis

The effects of compound **2i** on cell cycling were studied by using flow cytometry to determine the DNA content of asynchronously growing HCT-116 cells and their incorporation of BrdU after treatment with various concentrations of **2i** for 24 h (Figure 3A). In control experiments, cells were treated with etoposide, nocodazole and the dual CDK/aurora kinase inhibitor JNJ-7706621 (Figures S3 and S4). Treatment with **2i** reduced the population of cells in the G1 phase of the cell cycle and the number of cells undergoing active DNA replication (i.e. the number of BrdU-positive cells). While treatment with **2i** at 10 μM

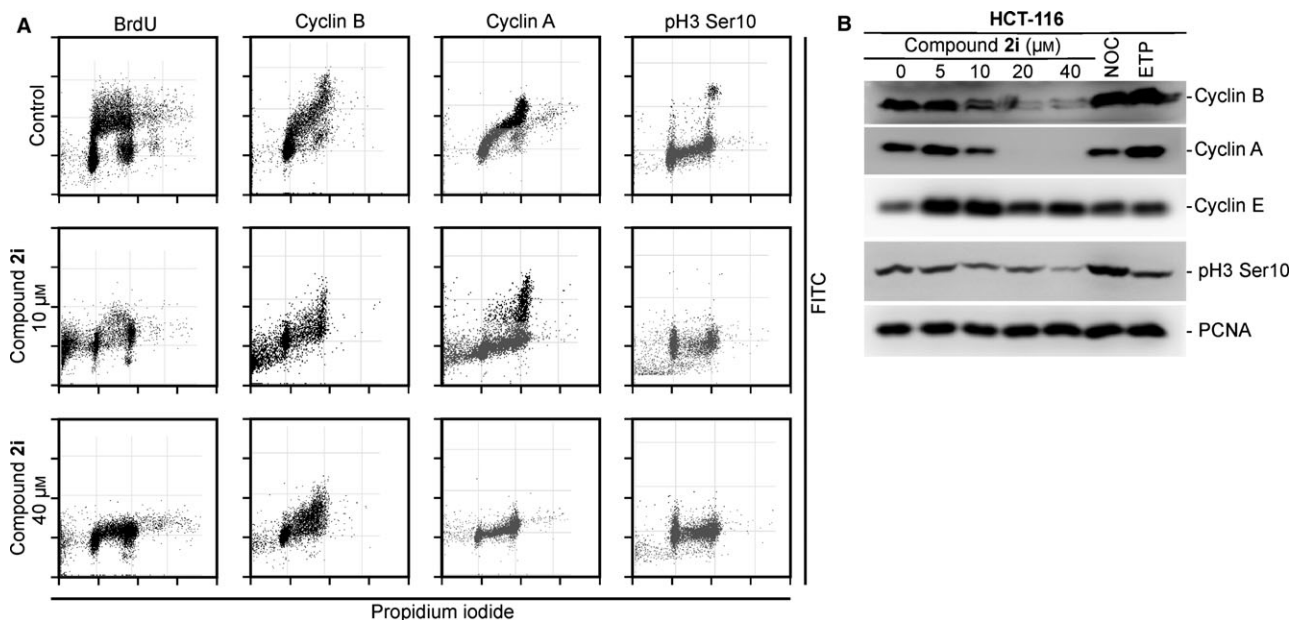


Figure 3: Cell cycle analysis of asynchronously growing HCT-116 cells treated with increasing concentrations of compound **2i** for 24 h. (A) Flow cytometric analysis of 5-bromo-2'-deoxyuridine (BrdU) incorporation, expression of cyclins B and A, and phosphorylation of histone H3 on Ser10 (pH3 Ser10). (B) Immunoblotting experiments to determine the levels of cyclins B, A, and E and the extent of phosphorylation of histone H3 at Ser10. Proliferating cell nuclear antigen (PCNA) levels were determined to confirm equal protein loading. Etoposide (ETP) and nocodazole (NOC)-treated cells were used as controls.

induced a strong increase in the sub-G1 cell population, higher concentrations had cytostatic effects and produced only a minimal number of apoptotic (sub-G1) cells. This unexpected result may be due to the parallel targeting of CDKs and Aurora kinases. The accumulation of cells in the G2 and M phases of the cell cycle is consistent with the inhibition of CDK2, CDK1, and Aurora (3,30,36,37).

Flow cytometry experiments revealed that cells treated with **2i** exhibited pronounced changes in the levels of important cell cycle regulators (cyclins A and B) and a common mitotic marker (the extent of phosphorylation of histone H3 at Ser10). Treatment with the highest tested concentration of compound **2i** (40 μM) completely suppressed cyclin expression and histone H3 phosphorylation. Similarly, treatment with **2i** at a concentration of 10 μM greatly reduced the size of the cell population expressing cyclins A and B relative to that seen in samples of untreated control cells (Figure 3A). The changes in the expression of cyclins A and B and the extent of histone H3 phosphorylation were confirmed by immunoblotting experiments and complemented with measurements of cyclin E expression (Figure 3B). The accumulation of BrdU-negative S phase cells, suppression of histone H3 phosphorylation, and reduced abundance of cyclins A and B all suggest that treatment with the higher dose of compound **2i** arrested the cell cycle of the HCT-116 cells in the S phase. Conversely, the lower dose allowed the cells to complete DNA replication and arrested them in the G2 phase.

Apoptosis and transcription

Flow cytometric analysis of HCT-116 cells treated with compound **2i** at 10 μM revealed a sub-diploid cell population, which suggested ongoing apoptosis. We therefore analyzed the levels of antiapoptotic (Bcl-2, Mcl-1) and proapoptotic (PUMA) proteins as well as the activation (fragmentation) of caspase-3 and the cleavage of its substrate, poly(ADP-ribose)polymerase-1 (PARP-1) in treated cells. Surprisingly, cells treated with **2i** at a concentration of 10 μM exhibited higher levels of cleaved caspase-3 and PARP-1 fragments than those treated with higher concentrations. The measured levels of Bcl-2 and PUMA in each case were consistent with this result. However, the expression of Mcl-1 declined in proportion to the applied concentration of compound **2i** (Figure 4A). The apoptosis-inducing effect of compound **2i** was confirmed by the results of a fluorimetric caspase-3/7 activity assay, which correlated with the immunoblotting analysis (Figure 4B). Control experiments were performed with dinaciclib and JNJ-7706621 (Figure S5).

Like roscovitine and other CDK inhibitors (38–40), compound **2i** also increased the expression of p53 at concentrations above 10 μM . This increase was accompanied by decreased expression of Mdm-2, the negative regulator of p53. Such changes are usually attributed to the interference of CDK inhibitors with transcription, so we measured the effect of compound **2i** on RNA synthesis in HCT-116 cells. Incubation with **2i** for 24 h caused a dose-dependent reduction in the levels of newly synthesized mRNA

and total RNA. Treatment with **2i** at a concentration of 5 μM reduced both mRNA synthesis and global transcription by 70% relative to untreated controls. The highest tested concentrations of compound **2i** (20 and 40 μM) completely suppressed RNA synthesis (Figure 4C).

Inhibition of migration and tube formation

We have recently shown that in addition to their cytotoxicity toward cancer cells, CDK inhibitors also influence the angiogenesis-related behavior of endothelial cells, probably by targeting CDK5 (16,41). Because compound **2i** is a

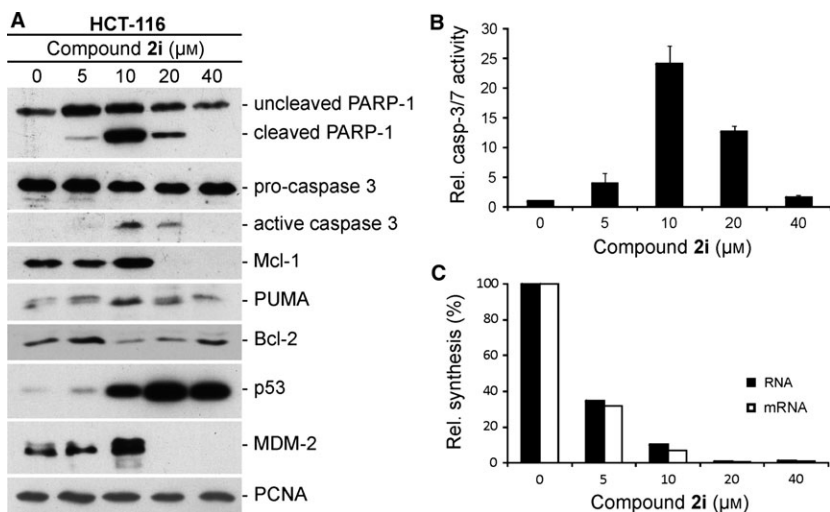


Figure 4: Compound **2i** blocks transcription and induces apoptosis in HCT-116 cells. (A) Immunoblotting of apoptotic markers; proliferating cell nuclear antigen (PCNA) levels were determined to confirm equal protein loading. (B) Results of a fluorimetric caspase-3/7 enzyme activity assay based on the cleavage of an Ac-DEVD-AMC peptide substrate. (C) The relationship between the applied concentration of compound **2i** and the mRNA and total RNA levels in HCT-116 cells.

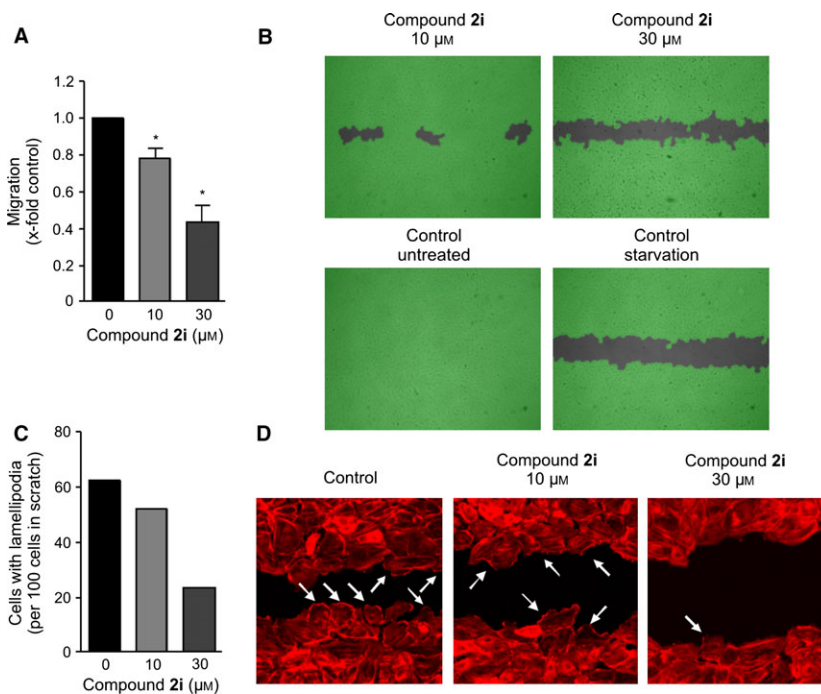


Figure 5: Compound **2i** blocks HUVEC migration and reduces the formation of lamellipodia. (A) Confluent layers of HUVECs were scratched and the cells were allowed to migrate for 16 h in the presence or absence of the specified concentrations of compound **2i**. The columns indicate the area re-covered by migrating cells ($n = 3$, mean \pm SEM, * $p < 0.05$, One Way ANOVA, Dunnett). (B) Scratches at the end-point of the experiment (representative images taken from one of the three experiments conducted for each set of tested conditions). (C) Confluent layers of HUVECs were scratched and the cells were allowed to migrate for 8 h in the presence or absence of 10 or 30 μM of compound **2i**, until clear lamellipodia formation was visible in the untreated control cells. The actin cytoskeleton was stained with rhodamin-phalloidin and fluorescence images of the scratches were acquired at 10 \times magnification. For quantitative evaluation of lamellipodia formation, the numbers of cells with prominent lamellipodia and ruffles were counted relative to the total number of cells at the scratch front ($n = 1$). (D) Representative images of the scratch front (F-actin, 40 \times magnification).

potent inhibitor of CDK5, we examined its effects on HUVEC migration and tube formation *in vitro*. Confluent HUVECs were scratched and the cells were allowed to migrate for 16 h in the presence or absence of compound **2i**. Compared to untreated controls, the area covered by migrating cells was 20 and 60% smaller, respectively, in cultures treated with compound **2i** at concentrations of 10 and 30 μM (Figure 5A,B). The influence of compound **2i** on the actin cytoskeleton in endothelial cells was examined by analyzing the F-actin distribution in migrating cells. Untreated HUVECs formed lamellipodia with a densely packed F-actin seam at the leading edge but these structures were largely absent in treated cells (Figure 5C,D). Similar results have previously been achieved by treatment with roscovitine or CDK5-targeting siRNA (16). Compound **2i** also reduced HUVEC tube formation. Treatment with compound **2i** at a concentration of 30 μM significantly reduced the numbers of branching points and tubes as well as the total and mean tube lengths (Figure 6). Control experiments were performed with roscovitine and JNJ-7706621. Whereas roscovitine blocked migration and tube formation at sub-toxic doses as described in our previous paper (16), JNJ-7706621 showed cytotoxicity but had no effect on angiogenesis (data not shown).

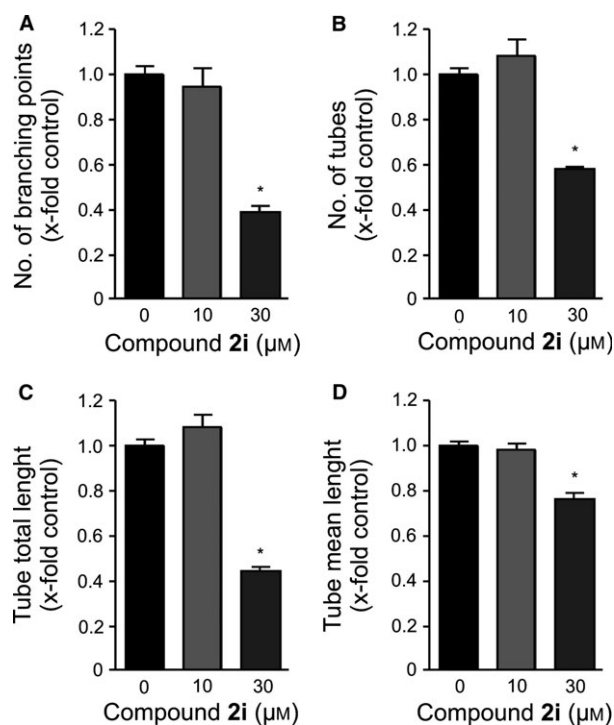


Figure 6: Compound **2i** reduces tube formation. HUVECs were seeded onto a growth factor containing matrix in the presence or absence of the indicated concentrations of compound **2i**. After 16 h of incubation, images were taken and the properties of the formed tubes were quantified ($n = 3$, mean \pm SEM, * $p < 0.05$, One Way ANOVA, Dunnett). (A) Number of branching points. (B) Number of tubes. (C) Total tube length. (D) Mean tube length.

Discussion

Selectively modulating the activity of protein kinases that regulate the cell cycle, such as CDKs and aurora kinases, could potentially be useful in the treatment of cancer. In addition to the selective CDK inhibitors that are currently undergoing clinical trials, some interesting broad-spectrum kinase inhibitors such as the 3,5-diamino-1,2,4-triazole JNJ-7706621 and the indirubin-like ZK 304709 have been developed (37,42). Both compounds are potent CDK inhibitors but also target other kinases: JNJ-7706621 also inhibits aurora kinases while ZK 304709 inhibits angiogenesis-related receptor kinases such as the VEGFRs and PDGFR β .

We unexpectedly identified a novel kinase inhibitor, compound **2i**, which targets CDK2, CDK5, and aurora A. Its structure is derived from the first-generation CDK inhibitor roscovitine. However, unlike roscovitine, **2i** has a pyrazolo [4,3-d]pyrimidine core; roscovitine analogs with this core structure are generally more potent than their purine-cored counterparts (18). Roscovitine inhibits the activities of CDK1/2/5/7/9 (43) but compound **2i** is a much stronger inhibitor of CDK5 and especially CDK2, which it inhibited at nanomolar concentrations. The distinctive selectivity of compound **2i** was confirmed by molecular modeling, which revealed that despite the very minor structural differences between the two compounds, it binds more tightly than roscovitine to both CDK2 and aurora A.

The *in vitro* cytotoxicity of compound **2i** was at least three times greater than that of roscovitine: the reported average IC_{50} values for roscovitine in a panel of cancer cell lines range from 15 μM to 28 μM (28,44–46), whereas the average IC_{50} for compound **2i** against the cell lines tested in this work was 5 μM . The death of HCT-116 cells treated with compound **2i** was accompanied by changes in the abundance of mitochondrial proteins (Mcl-1, Bcl-2, PUMA), the activation of caspases-3/7 and the fragmentation of PARP-1. Similar changes have been observed in cells undergoing apoptosis following treatment with roscovitine or other CDK inhibitors (11,18,47,48). Interestingly, whereas lower concentrations of compound **2i** had rather pro-apoptotic effects, higher concentrations apparently only had cytostatic effects, with little impact on markers of apoptosis – treatment with high concentrations of **2i** produced few sub-G1 cells and reduced caspase activity (Figures 3A and 4).

The greater cytotoxicity of compound **2i** relative to roscovitine was surprising because it is a less effective inhibitor of CDK7 and 9, both of which are regarded as critical targets of other potent CDK inhibitors. Inhibition of CDKs 7 and 9 blocks transcription, reduces the expression of Mcl-1, and increases the abundance and activity of p53, thereby inducing mitochondrial cell death (10,11,39,40). Cells treated with high (but not low or intermediate) doses

of compound **2i** also exhibited reduced levels of Mcl-1 level and transcription in general (Figure 4). Based on the results of the cellular assays and kinetic experiments using purified enzymes, we believe that the cytotoxicity of compound **2i** is due to the simultaneous inhibition of CDKs and aurora A. Cells treated with **2i** exhibited a pronounced downregulation of cyclins A and B along with reduced phosphorylation of histone H3 at Ser10, a site phosphorylated by aurora A (34). The suppression of histone H3 phosphorylation could be due to the direct inhibition of Aurora A or an indirect consequence of a reduction in its abundance (Figure S6). All these findings are consistent with the known anticancer activity of roscovitine, which has been found to reduce the transcription of many mitotic genes in HT29 cells including cyclin B, cdc25, and aurora A/B (49).

While the roles of CDK2 and aurora A in cancer development and the anticancer activity of their inhibitors are relatively well-explored, the link between CDK5 and tumor growth has emerged only recently. It has been shown that CDK5, a known regulator of neuronal processes, also plays a key role in the regulation of endothelial cell migration and tube formation, two essential steps in tumor-induced angiogenesis (16). Moreover, the inhibition of CDK5 by various CDK inhibitors was found to limit angiogenesis both *in vitro* and *in vivo* (17,41). We showed that compound **2i** can reduce cell migration as well as tube formation and the formation of lamellipodia in migrating HUVEC cells. Because it does not inhibit any VEGFR isoform, we assume that the molecular mechanism of its anti-angiogenic activity is linked to CDK5.

Conclusions

In conclusion, compound **2i** displays anticancer activity *in vitro* that appears to derive from the inhibition of CDKs and aurora A. This causes apoptosis in cancer cells and also suppresses angiogenesis-like activity in endothelial cells, a promising combination for therapeutic purposes. We are currently seeking to further optimize the pyrazolo [4,3-*d*]pyrimidine kinase inhibitors, particularly in terms of their selectivity and potency.

Acknowledgments

This work was supported by the Czech Science Foundation (grants P305/12/0783 and 14-19590S), the Ministry of Education, Youth and Sports of the Czech Republic (the National Program of Sustainability I, grant LO1204), Internal Grant Agency of Palacký University (projects IGA_PrF_2015_021, IGA_PrF_2014_024 and IGA_PrF_2014_023), German Research Foundation (grant DFG ZA 186/4-1), and the German Federal Ministry of Education and Research.

Conflict of Interest

The authors declare no competing financial interest.

References

- Hanahan D., Weinberg R.A. (2011) Hallmarks of cancer: the next generation. *Cell*;144:646–674.
- Malumbres M., Barbacid M. (2009) Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer*; 9:153–166.
- Payton M., Chung G., Yakowec P., Wong A., Powers D., Xiong L., Zhang N., Leal J., Bush T.L., Santora V., Askew B., Tasker A., Radinsky R., Kendall R., Coats S. (2006) Discovery and evaluation of dual CDK1 and CDK2 inhibitors. *Cancer Res*;66:4299–4308.
- Tetsu O., McCormick F. (2003) Proliferation of cancer cells despite CDK2 inhibition. *Cancer Cell*;3:233–245.
- Krystof V., Uldrijan S. (2010) Cyclin-dependent kinase inhibitors as anticancer drugs. *Curr Drug Targets*;11: 291–302.
- Lapenna S., Giordano A. (2009) Cell cycle kinases as therapeutic targets for cancer. *Nat Rev Drug Discov*;8:547–566.
- Santamaria D., Barriere C., Cerqueira A., Hunt S., Tardy C., Newton K., Caceres J.F., Dubus P., Malumbres M., Barbacid M. (2007) Cdk1 is sufficient to drive the mammalian cell cycle. *Nature*;448:811–815.
- Galons H., Oumata N., Meijer L. (2010) Cyclin-dependent kinase inhibitors: a survey of recent patent literature. *Expert Opin Ther Pat*;20:377–404.
- Chen R., Keating M.J., Gandhi V., Plunkett W. (2005) Transcription inhibition by flavopiridol: mechanism of chronic lymphocytic leukemia cell death. *Blood*;106: 2513–2519.
- MacCallum D.E., Melville J., Frame S., Watt K., Anderson S., Gianella-Borradori A., Lane D.P., Green S.R. (2005) Seliciclib (CYC202, R-Roscovitine) induces cell death in multiple myeloma cells by inhibition of RNA polymerase II-dependent transcription and down-regulation of Mcl-1. *Cancer Res*;65:5399–5407.
- Manohar S.M., Rathos M.J., Sonawane V., Rao S.V., Joshi K.S. (2011) Cyclin-dependent kinase inhibitor, P276-00 induces apoptosis in multiple myeloma cells by inhibition of Cdk9-T1 and RNA polymerase II-dependent transcription. *Leuk Res*;35:821–830.
- Canduri F., Peres P.C., Caceres R.A., de Azevedo W.F. (2008) CDK9 as a potential target for drug development. *Med Chem*;4:210–218.
- Ali M.A., Choy H., Habib A.A., Saha D. (2007) SNS-032 prevents tumor cell-induced angiogenesis by inhibiting vascular endothelial growth factor. *Neoplasia*;9:370–381.
- Melillo G., Sausville E.A., Cloud K., Lahusen T., Varesio L., Senderowicz A.M. (1999) Flavopiridol, a protein kinase inhibitor, down-regulates hypoxic induction of

- vascular endothelial growth factor expression in human monocytes. *Cancer Res*;59:5433–5437.
15. Scholz A., Wagner K., Welzel M., Remlinger F., Wiedenmann B., Siemeister G., Rosewicz S., Detjen K.M. (2009) The oral multitarget tumour growth inhibitor, ZK 304709, inhibits growth of pancreatic neuroendocrine tumours in an orthotopic mouse model. *Gut*;58:261–270.
 16. Liebl J., Weitensteiner S.B., Vereb G., Takacs L., Furst R., Vollmar A.M., Zahler S. (2010) Cyclin-dependent kinase 5 regulates endothelial cell migration and angiogenesis. *J Biol Chem*;285:35932–35943.
 17. Zahler S., Liebl J., Furst R., Vollmar A.M. (2010) Anti-angiogenic potential of small molecular inhibitors of cyclin dependent kinases *in vitro*. *Angiogenesis*;13:239–249.
 18. Jorda R., Havlíček L., McNae I.W., Walkinshaw M.D., Voller J., Šturc A., Navrátilová J., Kuzma M., Mistrík M., Bártek J., Strnad M., Kryštof V. (2011) Pyrazolo [4,3-d]pyrimidine bioisostere of roscovitine: evaluation of a novel selective inhibitor of cyclin-dependent kinases with antiproliferative activity. *J Med Chem*;54:2980–2993.
 19. Zatloukal M., Jorda R., Gucký T., Řezníčková E., Voller J., Pospíšil T., Malínková V., Adamcová H., Kryštof V., Strnad M. (2013) Synthesis and *in vitro* biological evaluation of 2,6,9-trisubstituted purines targeting multiple cyclin-dependent kinases. *Eur J Med Chem*;61: 61–72.
 20. Bain J., Plater L., Elliott M., Shpiro N., Hastie C.J., McLauchlan H., Klevernic I., Arthur J.S., Alessi D.R., Cohen P. (2007) The selectivity of protein kinase inhibitors: a further update. *Biochem J*;408:297–315.
 21. Morris G.M., Huey R., Lindstrom W., Sanner M.F., Belew R.K., Goodsell D.S., Olson A.J. (2009) AutoDock and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem*;30:2785–2791.
 22. Trott O., Olson A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*;31:455–461.
 23. Weitensteiner S.B., Liebl J., Kryštof V., Havlíček L., Gucký T., Strnad M., Furst T., Vollmar A.M., Zahler S. (2013) Trisubstituted pyrazolopyrimidines as novel angiogenesis inhibitors. *PLoS One*;8:e54607.
 24. Havlíček L., Moravcová D., Kryštof V., Strnad M. The identification of a novel highly condensed pentacyclic heteroaromatic ring system 1,3,5,5b,6,8,10,10b-octaazacyclopenta[h,i]aceanthrylene and its application in the synthesis of 5,7-substituted pyrazolo[4,3-d]pyrimidines. *J Heterocycl Chem*;52:669–673.
 25. Gray N.S., Wodicka L., Thunnissen A.M., Norman T.C., Kwon S., Espinoza F.H., Morgan D.O., Barnes G., LeClerc S., Meijer L., Kim S.H., Lockhart D.J., Schultz P.G. (1998) Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science*;281:533–538.
 26. Chang Y.T., Gray N.S., Rosania G.R., Sutherlin D.P., Kwon S., Norman T.C., Sarohia R., Leost M., Meijer L., Schultz P.G. (1999) Synthesis and application of functionally diverse 2,6,9-trisubstituted purine libraries as CDK inhibitors. *Chem Biol*;6:361–375.
 27. Bettayeb K., Oumata N., Echalié A., Ferandin Y., Endicott J.A., Galons H., Meijer L. (2008) CR8, a potent and selective, roscovitine-derived inhibitor of cyclin-dependent kinases. *Oncogene*;27:5797–5807.
 28. Bettayeb K., Sallam H., Ferandin Y., Popowycz F., Fournet G., Hassan M., Echalié A., Bernard P., Endicott J., Joseph B., Meijer L. (2008) N-&N, a new class of cell death-inducing kinase inhibitors derived from the purine roscovitine. *Mol Cancer Ther*;7:2713–2724.
 29. Oumata N., Bettayeb K., Ferandin Y., Demange L., Lopez-Giral A., Goddard M.L., Myriantopoulos V., Mikros E., Flajolet M., Greengard P., Meijer L., Galons H. (2008) Roscovitine-derived, dual-specificity inhibitors of cyclin-dependent kinases and casein kinases 1. *J Med Chem*;51:5229–5242.
 30. Emanuel S., Rugg C.A., Gruninger R.H., Lin R., Fuentes-Pesquera A., Connolly P.J., Wetter S.K., Hollister B., Kruger W.W., Napier C., Jolliffe L., Middleton S.A. (2005) The *in vitro* and *in vivo* effects of JNJ-7706621: a dual inhibitor of cyclin-dependent kinases and aurora kinases. *Cancer Res*;65:9038–9046.
 31. Parry D., Guzi T., Shanahan F., Davis N., Prabhavalkar D., Wiswell D., Seghezzi W. *et al.* (2010) Dinaciclib (SCH 727965), a novel and potent cyclin-dependent kinase inhibitor. *Mol Cancer Ther*;9:2344–2353.
 32. Macůrek L., Lindqvist A., Lim D., Lampson M.A., Klompaker R., Freire R., Clouin C., Taylor S.S., Yaffe M.B., Medema R.H. (2008) Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. *Nature*;455:119–123.
 33. Seki A., Coppinger J.A., Jang C.Y., Yates J.R., Fang G. (2008) Bora and the kinase Aurora a cooperatively activate the kinase Plk1 and control mitotic entry. *Science*;320:1655–1658.
 34. Crosio C., Fimia G.M., Loury R., Kimura M., Okano Y., Zhou H., Sen S., Allis C.D., Sassone-Corsi P. (2002) Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. *Mol Cell Biol*;22:874–885.
 35. de Azevedo W.F., Leclerc S., Meijer L., Havlicek L., Strnad M., Kim S.H. (1997) Inhibition of cyclin-dependent kinases by purine analogues: crystal structure of human cdk2 complexed with roscovitine. *Eur J Biochem*;243:518–526.
 36. Shapiro G.I. (2006) Cyclin-dependent kinase pathways as targets for cancer treatment. *J Clin Oncol*;24:1770–1783.
 37. Seamon J.A., Rugg C.A., Emanuel S., Calcagno A.M., Ambudkar S.V., Middleton S.A., Butler J., Borowski V., Greenberger L.M. (2006) Role of the ABCG2 drug transporter in the resistance and oral bioavailability of a potent cyclin-dependent kinase/Aurora kinase inhibitor. *Mol Cancer Ther*;5:2459–2467.

38. David-Pfeuty T., Nouvian-Dooghe Y., Sirri V., Roussel P., Hernandez-Verdun D. (2001) Common and reversible regulation of wild-type p53 function and of ribosomal biogenesis by protein kinases in human cells. *Oncogene*;20:5951–5963.
39. Demidenko Z.N., Blagosklonny M.V. (2004) Flavopiridol induces p53 via initial inhibition of Mdm2 and p21 and independently of p53, sensitizes apoptosis-reluctant cells to tumor necrosis factor. *Cancer Res*;64:3653–3660.
40. Kotala V., Uldrijan S., Horky M., Trbusek M., Strnad M., Vojtesek B. (2001) Potent induction of wild-type p53-dependent transcription in tumour cells by a synthetic inhibitor of cyclin-dependent kinases. *Cell Mol Life Sci*;58:1333–1339.
41. Liebl J., Krystof V., Vereb G., Takacs L., Strnad M., Pechan P., Havlicek L., Zatloukal M., Furst R., Vollmar A.M., Zahler S. (2011) Anti-angiogenic effects of purine inhibitors of cyclin dependent kinases. *Angiogenesis*;14:281–291.
42. Siemeister G., Luecking U., Wagner C., Detjen K., McCoy C., Bosslet K. (2006) Molecular and pharmacodynamic characteristics of the novel multi-target tumor growth inhibitor ZK 304709. *Biomed Pharmacother*;60:269–272.
43. Bach S., Knockaert M., Reinhardt J., Lozach O., Schmitt S., Baratte B., Koken M. *et al.* (2005) Roscovitine targets, protein kinases and pyridoxal kinase. *J Biol Chem*;280:31208–31219.
44. Meijer L., Bettayeb K., Galons H. (2006) (R)-Roscovitine (CYC202, Seliciclib). In: Smith P.J., Yue E.W., editors. *Inhibitors of Cyclin-Dependent Kinases as Anti-Tumor Agents*. Boca Raton, FL, USA: CRC Press: p. 187–226.
45. Krystof V., McNae I.W., Walkinshaw M.D., Fischer P.M., Muller P., Vojtesek B., Orsag M., Havlicek L., Strnad M. (2005) Antiproliferative activity of olomoucine II, a novel 2,6,9-trisubstituted purine cyclin-dependent kinase inhibitor. *Cell Mol Life Sci*;62:1763–1771.
46. McClue S.J., Blake D., Clarke R., Cowan A., Cummings L., Fischer P.M., MacKenzie M., Melville J., Stewart K., Wang S., Zhelev N., Zheleva D., Lane D.P. (2002) *In vitro* and *in vivo* antitumor properties of the cyclin dependent kinase inhibitor CYC202 (R-roscovitine). *Int J Cancer*;102:463–468.
47. Conroy A., Stockett D.E., Walker D., Arkin M.R., Hoch U., Fox J.A., Hawtin R.E. (2009) SNS-032 is a potent and selective CDK 2, 7 and 9 inhibitor that drives target modulation in patient samples. *Cancer Chemother Pharmacol*;64:723–732.
48. Santo L., Vallet S., Hideshima T., Cirstea D., Ikeda H., Pozzi S., Patel K. *et al.* (2010) AT7519, A novel small molecule multi-cyclin-dependent kinase inhibitor, induces apoptosis in multiple myeloma via GSK-3beta activation and RNA polymerase II inhibition. *Oncogene*;29:2325–2336.
49. Whittaker S.R., Te Poele R.H., Chan F., Linardopoulos S., Walton M.I., Garrett M.D., Workman P. (2007) The cyclin-dependent kinase inhibitor seliciclib (R-roscovitine; CYC202) decreases the expression of mitotic control genes and prevents entry into mitosis. *Cell Cycle*;6:3114–3131.
50. Sroka I.M., Heiss E.H., Havlicek L., Totzke F., Aristei Y., Pechan P., Kubbutat M.H., Strnad M., Dirsch V.M. (2010) A novel roscovitine derivative potently induces G1-phase arrest in platelet-derived growth factor-BB-activated vascular smooth muscle cells. *Mol Pharmacol*;77:255–261.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Synthesis and characterization of tested compounds.

Table S1. Protein kinase selectivity of compound **2i**.

Figure S1. Phosphorylation of the Rb, Plk-1, and histone H3 proteins in HCT-116 cells upon treatment with JNJ-7706621 and dinaciclib.

Figure S2. Molecular docking of compound **2i** to CDK2 (PDB ID: 3DDQ) and Aurora A (PDB ID: 3O50).

Figure S3. Cell cycle analysis of asynchronously growing HCT-116 cells treated with nocodazole (NOC) and etoposide (ETP) for 24 h.

Figure S4. Cell cycle analysis of asynchronously growing HCT-116 cells treated with JNJ-7706621 for 24 and 48 h.

Figure S5. Immunoblotting analysis of apoptotic markers in HCT-116 cells treated with dinaciclib and JNJ-7706621 for 24 h.

Figure S6. Expression of aurora A and B in HCT-116 cells treated with compound **2i** for 24 h.

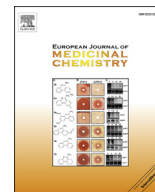
Appendix G

Baltus CB, Jorda R, Marot C, Berka K, **Bazgier V**, Kryštof V, Prié G, Viaud-Massuard M:
Synthesis, Biological Evaluation and Molecular Modeling of a Novel Series of 7-Azaindole
Based Tri-Heterocyclic Compounds as Potent CDK2/Cyclin E Inhibitors.

Eur. J. Med. Chem., 108(27), 701–719, 2016.

DOI: 10.1016/j.ejmech.2015.12.023

IF = 4.519



Research paper

Synthesis, biological evaluation and molecular modeling of a novel series of 7-azaindole based tri-heterocyclic compounds as potent CDK2/Cyclin E inhibitors



Christine B. Baltus^{a,1}, Radek Jorda^{b,**,1}, Christophe Marot^a, Karel Berka^{c,d},
Václav Bazgier^{b,d}, Vladimír Kryštof^b, Gildas Prié^a, Marie-Claude Viaud-Massuard^{a,*}

^a UMR 7292 GICC Equipe 4 Innovation Moléculaire et Thérapeutique, Labex SYNORG, University of Tours, Faculty of Pharmacy, 31 Avenue Monge, 37200 Tours, France

^b Laboratory of Growth Regulators & Department of Chemical Biology and Genetics, Centre of the Region Haná for Biotechnological and Agricultural Research, Palacký University and Institute of Experimental Botany AS CR, Šlechtitelů 27, CZ-78371 Olomouc, Czech Republic

^c Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacký University Olomouc, 17. Listopadu 12, 77146 Olomouc, Czech Republic

^d Department of Physical Chemistry, Faculty of Science, Palacký University, 17. Listopadu 1192/12, 771 46 Olomouc, Czech Republic

ARTICLE INFO

Article history:

Received 28 October 2015

Received in revised form

11 December 2015

Accepted 12 December 2015

Available online 19 December 2015

Keywords:

Cyclin-dependent kinase 2

Kinase inhibitors

Anti-tumor agent

1*H*-pyrrolo[2,3-*b*]pyridine

[3 + 2] cycloaddition

1,4-Triazole

1,5-Triazole

3D-QSAR CoMFA

ABSTRACT

From four molecules, inspired by the structural features of foscarnin, with an interesting potential to inhibit cyclin-dependent kinases (CDKs), we designed a new series of tri-heterocyclic derivatives based on 1*H*-pyrrolo[2,3-*b*]pyridine (7-azaindole) and triazole heterocycles. Using a Huisgen type [3 + 2] cycloaddition as the convergent key step, 24 derivatives were synthesized and their biological activities were evaluated. Comparative molecular field analysis (CoMFA), based on three-dimensional quantitative structure–activity relationship (3D-QSAR) studies, was conducted on a series of 30 compounds from the literature with high to low known inhibitory activity towards CDK2/cyclin E and was validated by a test set of 5 compounds giving satisfactory predictive r^2 value of 0.92. Remarkably, it also gave a good prediction of pIC_{50} for our tri-heterocyclic series which reinforce the validation of this model for the pIC_{50} prediction of external set compounds. The most promising compound, **43**, showed a micro-molar range inhibitory activity against CDK2/cyclin E and also an antiproliferative and proapoptotic activity against a panel of cancer cell lines.

© 2015 Elsevier Masson SAS. All rights reserved.

1. Introduction

Cyclin-dependent protein kinases (CDKs) are universal eukaryotic cell cycle regulators that promote the passage through the restriction point, initiation of DNA replication, and mitosis [1]. Given the fact that CDKs have oncogenic potential and are amenable to pharmacological inhibition, their inhibitors have therefore attracted great interest as potential anticancer agents [2,3].

As a part of our research on medicinal chemistry, we are interested in developing new CDK inhibitors (CKIs), in particular towards CDK4/cyclin D and CDK2/cyclin E, since they both participate in the phosphorylation of the retinoblastoma protein (pRb). Each member of the pRb pathway (such as CDK4(6)/cyclin D, p16 and CDK2/cyclin E), which activates the transcription factors at the G1-S transition phase, which in its turn regulates the expression of several genes involved in DNA replication, can be deregulated in cancers. In some human cancers (such as lung cancer or leukemia), cyclin E (E1/E2) is over-expressed and CDK2 is hyper-activated [4]. Therefore, many CKIs have been developed and some of them are undergoing clinical trials [5]. Most inhibitors that entered clinical trials belong to the group of pan-selective CDK inhibitors [6], but some highly selective CDK inhibitors like THZ1 [7] or PD-0332991 [8] were also described (Fig. 1). However, simultaneous targeting of multiple CDKs (particularly CDK1 and 2) seems to be more

Abbreviation: CCR2, CC chemokine receptor 2; CCL2, CC chemokine ligand 2; CCR5, CC chemokine receptor 5.

* Corresponding author.

** Corresponding author.

E-mail addresses: radek.jorda@upol.cz (R. Jorda), marie-claude.viaud-massuard@univ-tours.fr (M.-C. Viaud-Massuard).

¹ These authors contributed equally.

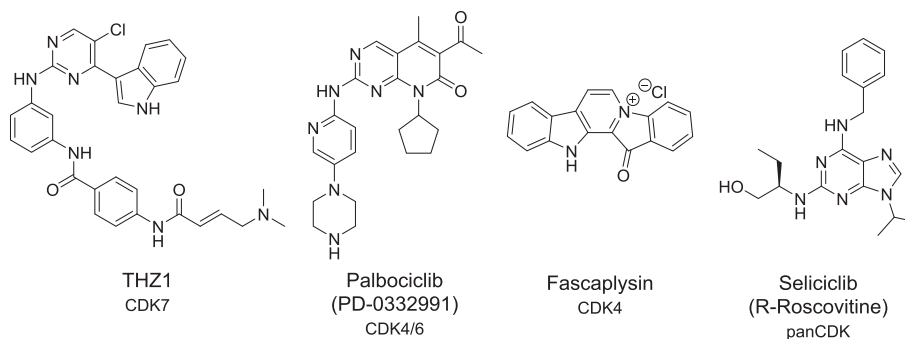


Fig. 1. Selected CDK inhibitors and their main CDK target.

advantageous in terms of anticancer activity, because many CDKs can compensate for the lack of others in cancer cells [9,10]. In addition, the knowledge about CDKs is growing constantly and it is still unclear whether observed effects are due to the activity of any single CDK or combinations of CDKs.

Since our laboratory has an expertise in the heterocyclic

chemistry domain and by analogy with the structural features of a known CKI, fascaplysin [11] (IC_{50} CDK4/D = 0.35 μ M) (Fig. 1), as other groups did with indole analogues [12–14], we designed and synthesized four novel heterocyclic molecules (21, 22, 24 and 37) (Figs. 2 and 3). 21 and 37 are composed of two 7-azaindole heterocycles linked each other by a 1,4- or 1,5-triazole type linker. The

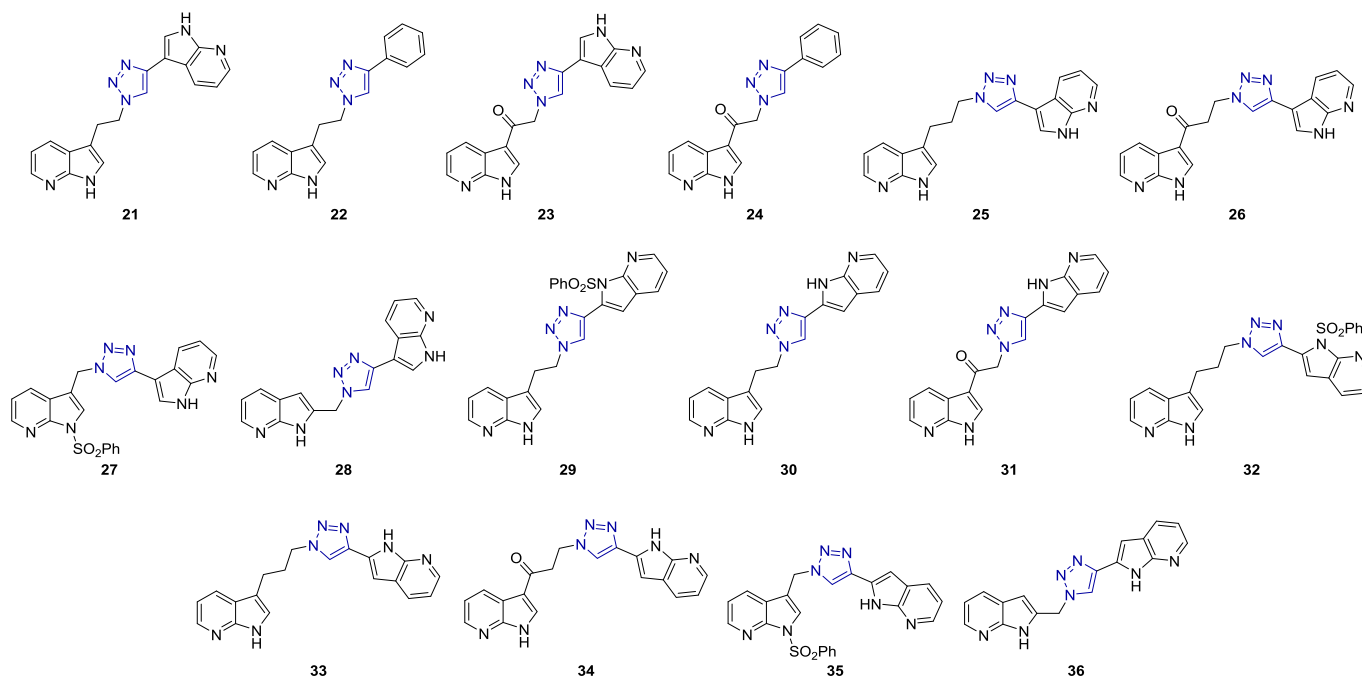


Fig. 2. 1,4-disubstituted triazole derivatives 21–36.

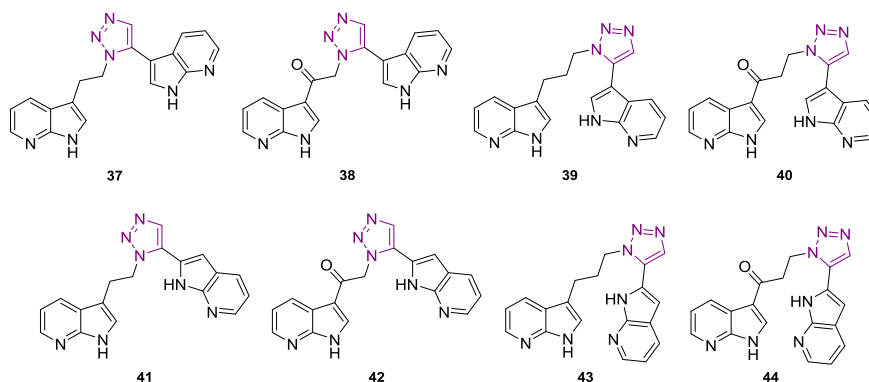
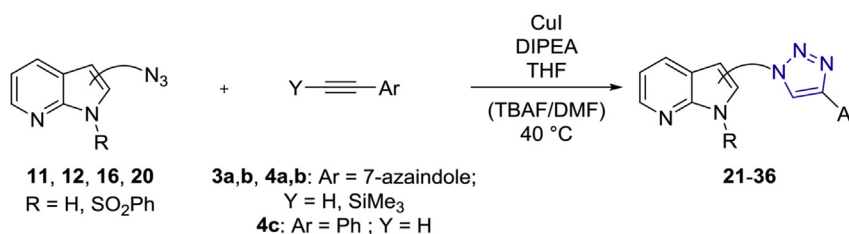


Fig. 3. 1,5-disubstituted triazoles derivatives 37–44.

Table 1
Copper-catalyzed cycloaddition reaction.2

Entry	Azide	Alkyne	Additive ^a	Time (h)	Product	Yield (%) ^b
1	11a	3a	TBAF	24	21	83
2	11a	4c	—	16	22	80
3	12a	3a	TBAF	24	23	—
4	12a	4c	—	26	24	72
5	12a	4a	DMF	21	23	26
6	11b	3a	TBAF	20	25	38
7	12b	3a	TBAF	16	26	—
8	12b	4a	DMF	72	26	57
9	16	4a	—	16	27	68
10	20	3a	TBAF	18	28	59
11 ^c	11a	3b	TBAF	18	29 and 30	43 and 39
12	12a	4b	DMF	16	31	84
13 ^c	11b	3b	TBAF	16	32 and 33	47 and 37
14	12b	4b	DMF	16	34	57
15	16	4b	—	16	35	89
16	20	3b	TBAF	16	36	53

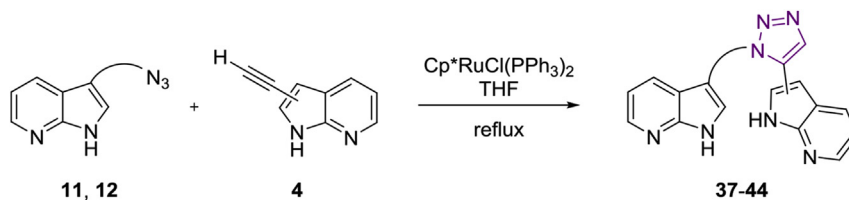
^a TBAF 3 equivalents; DMF 1 mL.^b Isolated yields.^c Products obtained as a mixture and separated by chromatography on silica gel column [30].

third ring in **22** and **24** was replaced by a phenyl group. Their skeleton can be viewed as an open and non-planar structure of faspaplysin, where the indole ring was replaced by a 7-azaindole. This open form design gives a more flexible molecule which could allow a better fitting into the binding site of the enzyme. Moreover, contrary to faspaplysin, free rotation potential should avoid DNA intercalation [15].

The inhibitory activities of these four molecules were tested on recombinant CDK2/cyclin E and CDK4/cyclin D complexes. Although the results obtained for CDK4/cyclin D with **21** and **22** were not satisfactory [16], the inhibitory activities on CDK2/cyclin E

of these four compounds were found to be interesting since the IC₅₀ values of **21** and **37** are in the micro molar range (experimental pIC₅₀: 5.1, 4.0, 4.3 and 5.1 respectively for **21**, **22**, **24** and **37**) (Table 4).

In parallel, 3D-QSAR CoMFA analyses on known CDK2/cyclin E inhibitors were performed and the results allowed us to create and validate a model, which will be presented later in this article. Notably, the CDK2/cyclin E activity predictions obtained by this model for **21**, **22**, **24** and **37** were in good correlation with the experimental ones with a maximal difference of about 1 log (predicted pIC₅₀: 5.2, 5.3, 5.3 and 5.8 respectively for **21**, **22**, **24** and **37**)

Table 2
Ruthenium-catalyzed cycloaddition reaction.3

Entry	Azide	Alkyne	Time (h)	Product	Yield (%) ^a
1	11a	4a	60	37	41
2	12a	4a	20	38	90
3	11b	4a	20	39	73
4	12b	4a	20	40	18
5	11a	4b	20	41	62
6	12a	4b	20	42	97
7	11b	4b	20	43	18
8	12b	4b	20	44	34

^a Isolated yields.

Table 3
Values of CoMFA models.

CoMFA	
Partial charge	Gasteiger–Huckel
q ^{2a}	0.49
r ^{2b}	0.99
S ^c	0.10
F value ^d	1021.002
N ^e	3
Contributions	
Steric	0.42
Electrostatic	0.58
Column filtering	2.0

^a Cross-validated correlation coefficient.^b Non-cross-validated correlation coefficient.^c Standard error of estimate.^d F test value.^e Optimum number of components.

(Table 4).

Following these preliminary results, we decided to design a series of 7-azaindole based tri-heterocyclic derivatives to evaluate them towards CDK2/cyclin E, to predict their activity by our 3D-QSAR CoMFA model and validate the model by comparison with the experimental values, in order to develop new CDK2/cyclin E inhibitors as potential anti-cancer agents. The synthesis of these derivatives was based on the pharmacomodulation of **21** and **37**: i) attachment of the triazole linker chain on both sides at the C-2 or C-3 position of the 7-azaindole moiety, ii) variation of the linker chain length, iii) carbonyl function group kept or not on the linker chain, iv) 1,4- or 1,5-triazoles.

Herein is presented the synthesis of a new series of 7-azaindole based tri-heterocyclic molecules as potent CDK2/cyclin E inhibitors, a 3D-QSAR CoMFA model, and their biological evaluation.

2. Results and discussion

2.1. Chemistry

Tri-heterocyclic derivatives were synthesized via a key reaction, a Huisgen type [3 + 2] cycloaddition [17] between an alkyne and an azide function substituted on a 7-azaindole heterocycle. The alkyne function was introduced on the C-2 or the C-3 position of the 7-azaindole moiety via an halogenation reaction followed by a palladium-catalysed cross-coupling reaction (Sonogashira coupling reaction) [18–20] (Scheme 1).

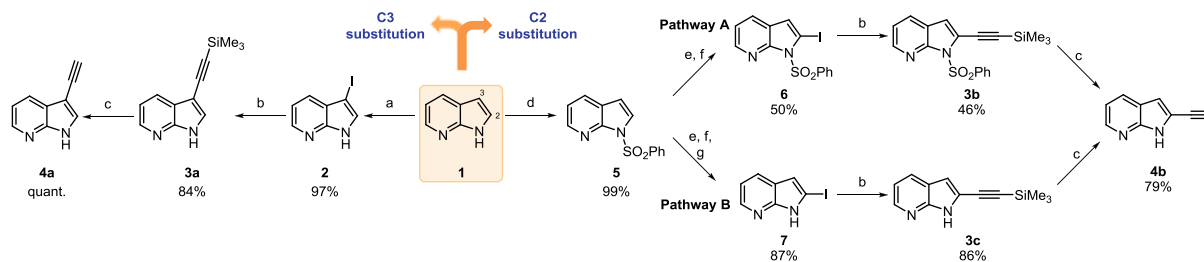
In one hand, a simple iodation reaction on the commercially available 7-azaindole **1** led to the derivative **2** bearing the iodine atom on the C-3 position. The latter was treated with ethynyltrimethylsilane in a Sonogashira cross-coupling reaction catalyzed by PdCl₂(PPh₃)₂ and copper iodide in the presence of triethylamine. The alkyne **3a** could be used in the cycloaddition reaction either as prepared, or after the removal of the trimethylsilyl group in the presence of tetrabutylammonium fluoride (TBAF) as **4a** (Scheme 1).

In the other hand, the introduction of the iodine atom at the C-2 position is described in Scheme 1. A benzenesulfonyl group was introduced at the N-1 position of **1** using a classical procedure of amine protection with benzenesulfonyl chloride [21]. The protected derivative **5** was halogenated by iodine on the C-2 position, after deprotonation with a strong base, here lithium diisopropylamide (LDA) leading to the expected compound **6** (in a mixture with a diiodinated derivative, see exp. part, ¹H NMR ratio 1:0.34), result also observed by Naka et al. [22] (Scheme 1, pathway A). Due to the low yield obtained for **6**, we then decided to perform a two steps reaction at once, halogenation followed by deprotection of the amine function, yielding **7** in a good yield for the two steps (Scheme 1, pathway B). To introduce the desired alkyne function, the Sonogashira cross-coupling reaction was performed on **6** and **7** with ethynyltrimethylsilane. **3c** was obtained with a much better

Table 4
Antiproliferative, CDK2/E activity and CoMFA prediction data for studied compounds.

Cmpds	IC ₅₀ (μM) ^a	pIC ₅₀	CoMFA prediction pIC ₅₀	Cytotoxic activity in cancer cell lines			Activation of caspases (fold) ^b
	CDK2/E			IC ₅₀ (μM) ^a	K562	MCF-7	
21	7.2	5.1	5.2	39.5	87.1	44.2	1.2
22	>50	4.0	5.3	>50	>50	33.6	3.1
23	29.3	4.5	6.4	>25	>25	>50	2.2
24	>50	4.0	5.3	>25	>25	>50	1.5
25	6.1	5.2	6.1	66.6	60.7	36.9	3.9
26	17.8	4.7	5.4	44.27	64.7	>50	4.5
27	25.8	4.6	5.9	>25	>25	>50	7.5
28	18.4	4.7	5.0	>50	47.4	>50	2.1
29	>25	4.6	6.8	>25	>25	>50	2.1
30	30.0	4.5	4.7	>50	58.5	>50	4.5
31	9.2	5.0	5.4	>25	>25	>50	1.7
32	18.0	4.7	5.7	>25	>25	>50	1.8
33	8.7	5.1	4.2	>25	>25	>50	1.9
34	3.6	5.4	4.7	>25	>25	>50	1.2
35	>25	4.6	6.1	>25	>25	>50	1.6
36	1.9	5.7	5.9	>25	>25	>50	1.9
37	9.1	5.0	5.8	>25	>25	>50	1.9
38	1.1	6.0	6.4	>25	>25	>50	2.9
39	2.7	5.6	6.0	>25	>25	33.4	4.2
40	13.4	4.9	6.2	>25	>25	>50	3.3
41	27.6	4.6	5.5	>25	>25	>50	2.0
42	3.0	5.5	5.8	>25	>25	>50	2.2
43	2.0	5.7	5.6	39.9	>25	31.5	13.8
44	36.1	4.5	5.8	>25	>25	>50	n.d.
Roscovitine	0.2	–	–	30.9	27.2	20.8	n.d.
Imatinib	>100	–	–	0.73	>10	n.d.	n.d.

^a Average values from at least three determinations.^b Relative caspase-3/7 activity in G361 melanoma cells after treatment with studied compounds at 50 μM concentration. Synthetic peptide Ac-DEVD-AMC was used as a substrate and caspase-3/7 activity was measured after 6 h by microplate reader at 346 nm/442 nm (excitation/emission).



Reagents and conditions: a) I_2 , KI, NaOH (1 M), EtOH, RT, 16 h. b) Ethynyltrimethylsilane, $PdCl_2(PPh_3)_2$, CuI, Et_3N , THF, RT, 20 h. c) TBAF, THF, RT, 2 h. d) Benzenesulfonyl chloride, TBAB, NaOH, CH_2Cl_2 , 0 °C, RT, 1 h. e) LDA, THF, -60 °C, 30 min. f) I_2 , -60 °C to RT, 16 h. g) NaOtBu, dioxane, reflux, 1 h.

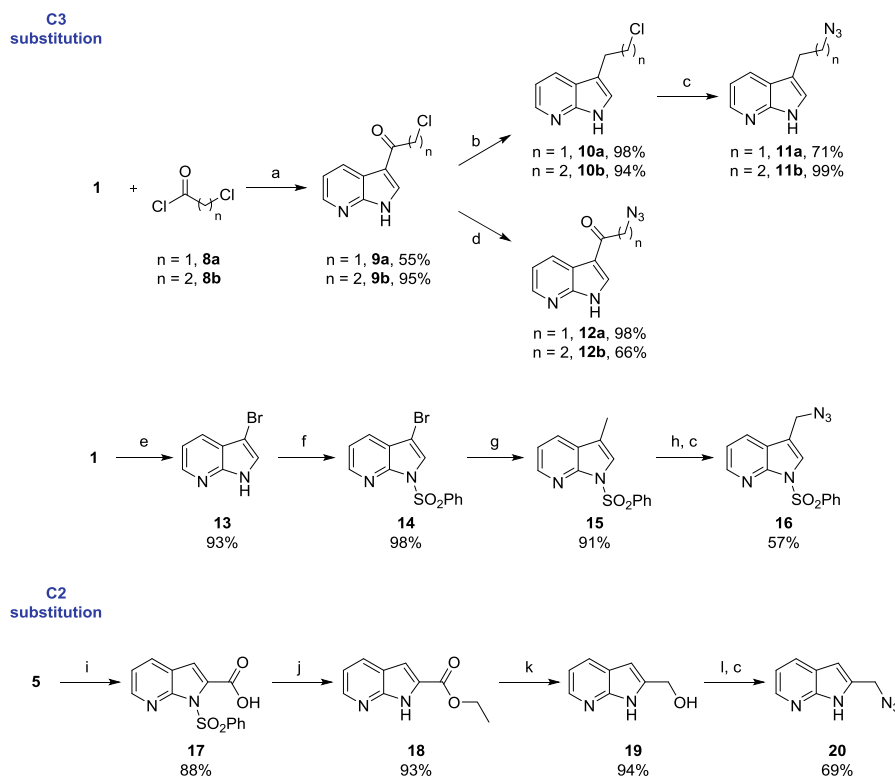
Scheme 1. Synthesis of the alkyne derivatives **3** and **4**.

yield than **3b** (86% and 46% respectively). The trimethylsilyl group of **3c** was removed with TBAF [23] leading to the derivative **4b**. The removal of the trimethylsilyl group in **3b** was attempted without success, as the reaction conditions led to the removal of both trimethylsilyl and benzenesulfonyl protecting groups, ending up to the unprotected derivative **4b** in 79% yield. The pathway B was found to be the pathway of choice for the synthesis of **4b** since it gave a better overall yield compared to pathway A, 59% and 18%

respectively.

In parallel, different length carbon chains bearing the desired azide function were introduced at the C-3 and C-2 (Scheme 2) positions of the 7-azaindole heterocycle.

First, concerning the compounds substituted at the C-3 position, **1** was reacted in a Friedel and Craft acylation reaction with selected acid chlorides (**8a** and **8b**) yielded the derivatives **9a** and **9b** containing a 2 or 3 carbons chain. The carbonyl group in **9** could be kept



Reagents and conditions: a) $AlCl_3$, CH_2Cl_2 , RT, 16 h. b) $LiAlH_4$, $AlCl_3$, DME, 0 °C, then RT, 16 h. c) NaN_3 , DMF, H_2O , 60 °C, 16 h. d) NaN_3 , DMF, H_2O , 60 °C, 16 h or RT, 36 h. e) NBS, CH_2Cl_2 , 0 °C, RT, 16 h. f) Benzenesulfonyl chloride, TBAB, NaOH, CH_2Cl_2 , 0 °C, RT, 1 h. g) $AlMe_3$, $Pd(PPh_3)_4$, THF, 60 °C, 24 h. h) NBS, AIBN, CCl_4 , reflux, 16 h. i) LDA, THF, -35 °C, then CO_2 , -55 °C, RT, 16 h; HCl. j) EtOH, H_2SO_4 , reflux, 20 h. k) $LiAlH_4$, DME, RT, 16 h. l) $SOCl_2$, CH_2Cl_2 , RT, 1 h.

Scheme 2. Synthesis of the azide derivatives.

or reduced with lithium aluminium hydride (LiAlH_4) in the presence of aluminium chloride (AlCl_3) [24] (leading to **10**) before the substitution of the chlorine atom by the azide group with sodium azide was set up. The azide derivatives **11** and **12** were obtained in moderate to good yields (Scheme 2).

The 7-azaindole derivative **16**, substituted in C-3 by a methylene azide group, was obtained in a 4 steps synthetic procedure from **1** (Scheme 2). Bromination of **1** at the C-3 position by *N*-bromosuccinimide (NBS) yielded **13**, which was protected with a benzenesulfonyl group. The derivative **14** was methylated at the C-3 position with AlMe_3 using tetrakis(triphenylphosphine)palladium as catalyst [25,26] to afford **15** in a good yield. Bromination of the methyl group with NBS and azobisisobutyronitrile (AIBN) [27] followed by a nucleophilic substitution with sodium azide led to **16** in moderate yield for the two steps.

Alternatively, the methylene azide scaffold was introduced on the C-2 position of the heterocycle (Scheme 2). The acid derivative **17** was obtained by deprotonation of **5** by LDA followed by the addition of carbon dioxide. The compound **17** was esterified with ethanol yielding **18**, which was reduced to the alcohol **19** by using LiAlH_4 in dimethoxyethane (DME). The hydroxyl group newly formed was substituted by a chlorine, in the presence of thionyl chloride, which was itself substituted by the azide group leading to the derivative **20** in good yield.

The alkyne (**3** and **4**) and azide (**11**, **12**, **16** and **20**) derivatives were then engaged in [3 + 2] cycloaddition reactions. Depending on the chosen catalyst, copper or ruthenium, the reaction led to 1,4- or 1,5-disubstituted triazole respectively.

First, copper (I) iodide was used as catalyst for the copper-catalyzed [3 + 2] Huisgen cycloaddition reaction in the presence of diisopropylethylamine (DIPEA) in THF (Table 1) [28]. An attempt using copper sulfate and sodium ascorbate was made but it did not give as good results as with copper iodide (results not shown). The corresponding 1,4-disubstituted triazoles **21–36** (Fig. 2) were obtained with moderate to good yields in 16 h–72 h. When the non-terminal alkyne derivatives **3a** and **3b** were used as starting materials, this reaction was carried out in a one pot procedure together with the deprotection of the trimethylsilyl group by TBAF (Table 1, entries 1, 3, 6, 7, 10, 11, 13 and 16) [29]. However, when this one pot reaction was tested on the azide derivatives containing a carbonyl function **12**, the reaction did not occur (Table 1, entries 3 and 7). In this case, terminal alkyne derivatives (**4a**, **4b** or phenylacetylene **4c**) had to be used in order to obtain the expected tri-heterocycle derivatives (Table 1, entries 4, 5, 8, 12 and 14). Moreover, due to the poor solubility of **12a** and **12b** in THF, small amount of dimethylformamide (DMF) was added in the reaction mixture. This copper-catalyzed reaction led to the obtention of a series of 1,4-disubstituted triazole derivatives shown in Fig. 2.

Secondly, terminal alkyne (**4**) and azide (**10** and **12**) derivatives were engaged in a ruthenium-catalyzed [3 + 2] cycloaddition reaction with pentamethylcyclopentadienylbis(triphenylphosphine) ruthenium(II) chloride ($\text{Cp}^*\text{RuCl}(\text{PPh}_3)_2$) in THF (Table 2) [31]. The 1,5-disubstituted triazole derivatives (**37–44**), obtained in low to good yields in 20 h–60 h, are presented in Fig. 3.

To conclude, all 24 synthesized compounds have consisted of three parts: 7-azaindole, triazole and aryl part. The first heterocycle, 7-azaindole, is linked by carbon C2 or C3 to nitrogen N1 of triazole in all derivatives, third aryl is linked to triazole at position 4 or 5 resulting two series: 1,4-disubstituted (**21–36**) and 1,5-disubstituted triazoles (**37–44**). Both series are subdivided to two another libraries differing by binding of third aryl (mainly 7-azaindole) to triazole. While compounds **29–36** and **41–44** bind by carbon C2 of 7-azaindole to triazole, the rest of compounds containing 7-azaindole (**21**, **23**, **25–28**, **37–40**) bind through carbon C3 to triazole. Only two derivatives (**22**, **24**) have varied from the

whole series by presence of phenyl group instead of 7-azaindole ring.

2.2. 3D-QSAR model

To rationalize the design and, thus, to understand physicochemical properties and structural parameters of the pharmacophore, a three-dimensional quantitative structure activity relationship (3D-QSAR) study of numerous CKIs derivatives by comparative molecular field analysis (CoMFA) [32] was performed. Traditional QSAR models are unable to explain complex structure activity data because of the extreme specificity of biological activity described by 3D intermolecular forces and predicted on 3D molecular structures. Consequently, the most relevant QSAR model would be shape-dependent and would describe steric and electrostatic interactions with sufficient accuracy. Indeed, the comparative molecular field analysis (CoMFA) method meets these requirements and has become a powerful tool for studying 3D-QSAR.

A CoMFA study starts by the examination of the differences in targeted properties which are related to changes in the shape of the steric and electrostatic fields surrounding the molecules. A QSAR table is then used to accommodate the details of the shape of each field by sampling their magnitudes at regular intervals throughout a specified region of space [32]. Similarity indices are calculated at regularly spaced grid points for the prealigned molecules. Instead of the direct measurement of the similarity between all mutual pairs of a molecule, indirect evaluation of the similarity of each molecule in the data set with a common probe atom is calculated [33]. A linear regression equation of similarity with biological activities is finally derived.

This indirect ligand-based approach can assist in understanding structure–activity relationships (SARs) and can also serve as a tool in designing more potent anticancer agents. We believe this study provides useful information about the structural requirements of CDK2 inhibitors, and expect the results will aid in the design of new synthetic drugs.

2.3. Computational methods

2.3.1. Data sets

A total of 30 ligands from the literature, 25 for the training set (**L1–L25**) and 5 for the test set (**L26–L30**), with known structures and biological activities towards CDK2/cyclin E, were used in the study (See Supp. Info. Tables S1 and S2) [34]. The IC_{50} values were converted to the corresponding pIC_{50} ($-\log\text{IC}_{50}$) and used as dependent variables in CoMFA analyses. The pIC_{50} values span a range of 4.5 log units, providing a broad and homogenous data set for the 3D-QSAR study. The initial structures of 30 compounds were constructed using the Sketch Builder module on Sybyl X 2.1 (TRIPOS) [35]. Conformations of compounds in the training set and test set were generated using the multisearch utility in Sybyl X 2.1. The conformer with the lowest energy was extracted and energy minimization was performed using the Tripos force field [36], with a distance-dependent dielectric constant, and the Powell conjugate gradient algorithm, with a convergence criterion of $0.01 \text{ kcal mol}^{-1}$.

2.3.2. Molecular alignment

Structural alignment is one of the most sensitive parameters in 3D-QSAR analyses. The accuracy of the prediction of CoMFA models and the reliability of the contour models depend strongly on the structural alignment of the molecules [37]. The molecular alignment of the training set was achieved by the Sybyl X 2.1 routine docking (FlexX) with X-ray crystal structure of CDK2 (PDB: 2A4L) (Fig. 4a). In order to investigate the interaction mechanism

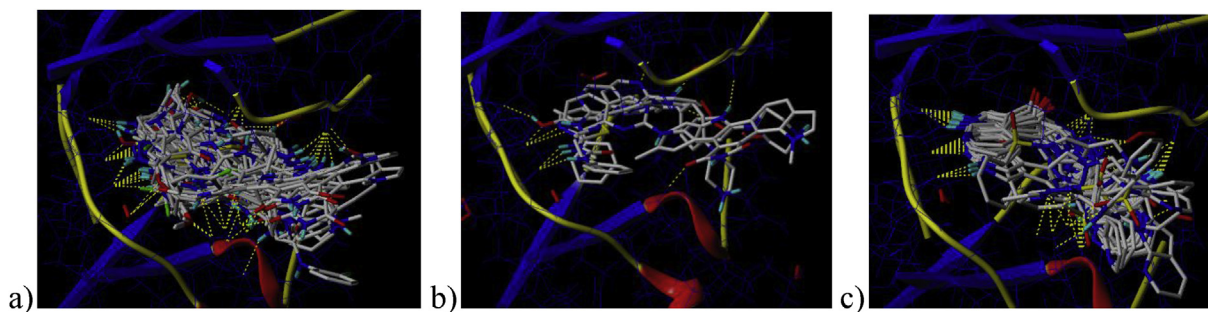


Fig. 4. Alignment in the active site of CDK2 (PDB: 2A4L) of a) training set ligands L1–L25, b) test set ligands L26–L30, c) our designed ligands 21–44.

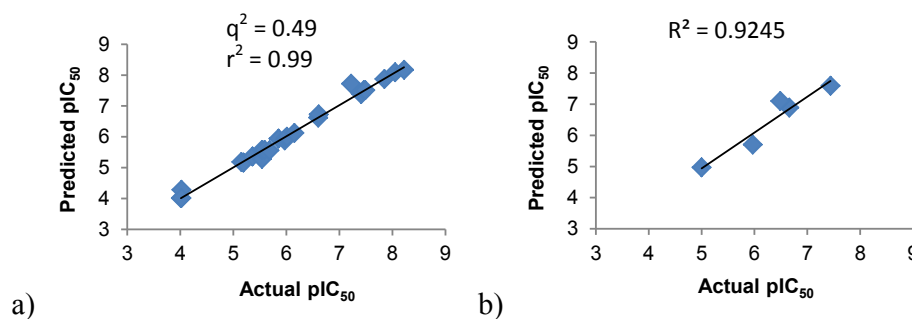


Fig. 5. CoMFA predicted and actual activities of ligands from a) the training set, b) the test sets.

between the ligand and the targeted enzyme, the molecular docking analysis was performed based on the active site of the built 3D model of 2A4L. The FlexX docking methods [38] were used to explore the reasonable binding mode of ligands in the active site of 2A4L. Once the molecules from the training set are docked, the poses with the best docking score values were selected as the

aligned molecules used for the CoMFA study.

2.3.3. CoMFA studies

Sybyl X 2.1 was used for molecular modeling and 3D-QSAR analysis on a PC workstation equipped with an Intel(R) Xeon(R) CPU 2.67 GHz processor. Steric and electrostatic CoMFA fields were calculated using the Lennard-Jones and the Coulomb potentials.³² The partial atomic charge was calculated by the Gasteiger-Hückel method [39]. Default parameters (Tripos force field, dielectric distance $1/r^2$, steric and electrostatic cut-off 30 kcal mol^{-1} , positively charged sp^3 hybridized carbon atom, grid spacing 2 Å) were used unless stated otherwise. With standard options for scaling of variables, the regression analysis was carried out using the full cross-validated partial least squares (PLS) method (leave one out) [40]. To derive 3D-QSAR models, the CoMFA descriptors were used as independent variables and the pIC_{50} as the dependent variable. Partial least squares (PLS) regression analysis was conducted with the standard implementation in the Sybyl X 2.1 package. The predicted values of the models were evaluated by leave-one-out cross-validation [41] (Table 3).

2.3.4. Predictive power of CoMFA models

To test the predictability of the analyses, the activities of training set compounds were calculated from the best CoMFA model considering Gasteiger-Hückel partial charges, steric and electrostatic field combinations. The correlation between experimental and predicted training set activities is shown in Fig. 5a and indicates a good agreement between experimental and predicted values (See Supp. Info. Table S1).

A test set made of 5 compounds was used to verify the efficacy of the CoMFA model (See Supp. Info. Table S2). These 5 molecules were aligned in the active site of the enzyme in the same way as the training set (Fig. 4b). Fig. 5b shows that the test set pIC_{50} predictions and the actual pIC_{50} were in a similar range of values (See Supp. Info. Table S2), giving a predictive correlation coefficient r^2 of

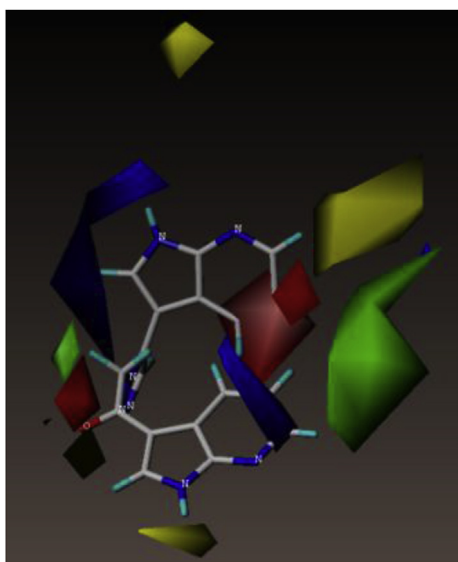


Fig. 6. Contour plot using CoMFA steric and electrostatic fields for CDK2 inhibitors with 38. Green indicates regions where bulky groups increase activity; yellow indicates regions where bulky groups decrease activity; blue indicates regions where positive charges increase activity or negative charges decrease it; red indicates regions where negative charges increase activity or positive charges decrease it. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

0.92, which validates this model.

Once the model was validated with the test set, our 24 tri-heterocyclic derivatives were aligned in the active site of the enzyme (Fig. 4c) and the CoMFA allowed us to obtain a prediction for their inhibitory activities (pIC_{50}) (Table 4).

2.3.5. CoMFA contour maps

The CoMFA steric and electrostatic fields from the final non-cross-validated analysis were plotted as 3D colored contour maps. In these contour maps, each colored contour represents particular properties (Fig. 6): green for regions of high steric tolerance (80% contribution), yellow for low steric tolerance (20% contribution), red for regions of decreased tolerance for positive charge (20% contribution) and, blue for regions of decreased tolerance for negative charge (80% contribution). The larger size of the green–yellow region compared to the red–blue region indicates a greater contribution of steric fields than electrostatic ones in determining the biological activity. The derivative **38** was selected to feature in the CoMFA contour map (Fig. 6), since it was the tri-heterocyclic derivative having the best IC_{50} value on the recombinant CDK2/cyclin E complex (Table 4). Indeed, the carbonyl group in **38** was located in a region of negatively charged electrostatic field and the whole molecule seemed to meet the requirement of the steric constraints.

In this study, 3D-CoMFA QSAR analyses were used to predict the CDK2/cyclin E inhibitory activity of a set of tri-heterocyclic compounds. The QSAR model gave good statistical results in terms of q^2 and r^2 values, respectively 0.495 and 0.99. The CoMFA model provided the most significant correlation of steric and electrostatic fields with the biological activities. Overall, the CoMFA method provided a better statistical model, which implies the significance of steric and electrostatic fields in the selectivity and activity of these compounds. The statistical significance and robustness of the 3D-QSAR models generated were confirmed using a test set ($r^2 = 0.92$). The effects of the steric, electrostatic fields around the aligned molecules on their activities were clarified by analyzing the CoMFA contour maps. Most of our compounds (~70%) were predicted with a maximum deviation of 1 log unit. The micromolar range activity of our best compounds were correctly predicted (**36**, **38**, **39**, **42** and **43**), since the predictive deviation is only about 0.5 log unit. Only a few inactive compounds were predicted with a micromolar activity (however, this shows that our model do not miss potent compounds).

2.4. CDK inhibition

The 24 tri-heterocyclic derivatives synthesized were tested for CDK2/cyclin E kinase inhibition. The obtained data are presented in Table 4.

The majority of these compounds displayed submicromolar activities toward CDK2, what we consider as a good starting point for further improvement of molecules. The structural diversity of library allowed us to evaluate structure–activity relationships. The first observed trend is related to substitution of 7-azaindole at nitrogen N1 by bulky phenylsulphonyl moiety that leads to a decrease of CDK2 activity of unmodified molecules (compared pairs **27–28**, **29–30**, **32–33**, **35–36**). No measurable CDK inhibition of two compounds bearing phenyl group as a third aryl (compounds **22**, **24**) showed us that 7-azaindole as a third heterocycle is beneficial (probably due to the presence of nitrogen N1 as a hydrogen donor). In fact, CDK2 inhibition was also reduced mostly with the shortening of the alkyl chain between first 7-azaindole and triazole (compared pairs **30–33**, **37–39**, **41–43**). In addition, modification of alkyl chain can also influence the activity; while the presence of a keto group leads to increase of CDK2 inhibitory activity in

comparison with unmodified structures (pairs **30–31**, **33–34**, **37–38**, **41–42**), in some pairs the activity decreased (pairs **21–23**, **25–26**, **39–40**, **43–44**). Observed differences are probably related to the orientation of the whole molecules in the active site of kinase and with the possibility to use the carbonyl as a hydrogen bond acceptor. The most active compounds from this library belong to the group of 1,5-disubstituted triazoles (structural isomers **38–42**, **39–43**) and their IC_{50} (CDK2) reaches 1.1 μ M (Table 4).

2.5. Flexible docking

To further explore SAR of novel compounds we decided to use our flexible docking procedure in conjunction with the Autodock Vina molecular docking program to model binding interactions with CDK2 (see the Experimental section for details). Crystal structure (PDBID: 2A4L) reveals that the ATP binding groove is formed by a planar, mainly hydrophobic, pocket between Ile10 and Leu134. On the edge of this pocket are situated main chains of Glu81 and Leu83 forming hydrogen bond acceptor-donor-acceptor pattern for interaction with purine bases of roscovitine or ATP. The pocket further tilts and extends towards the bay between two polar residues His84 and Lys89. This bay is filled with phenyl moiety of roscovitine in the crystal structure, but it can also accommodate larger ligands, such as biaryls [42].

Our molecular docking showed the binding mode of molecule **43** within the binding pocket of CDK2 kinase (Fig. 7; binding energy -9.5 kcal/mol). The ligand forms a U-shape with two branches formed by 7-azaindoles and turning loop formed by triazole moiety. Triazole shows two interactions – directly with Thr14 and more distantly with Asp145 (possibly through water molecule). First 7-azaindole branch interacts with the typical main chain interactions – Glu81 and Leu83. Second branch seems to be more movable as it has just interaction with main chain of Glu12, whereas it is more exposed to the water.

2.6. Anticancer activity in vitro

Table 4 and Table S4 (See Supp. Inf.) summarized also anticancer activity of prepared compounds towards three cancer cell lines.

Inhibition of CDKs in cells either by chemical or genetic depletion lead to the i) reduction of cell proliferation, ii) changes in cell cycle profile and iii) induction of apoptosis [43,44].

We therefore analyzed the antiproliferative and proapoptotic activity of all derivatives to confirm reported inhibition of CDKs. First, we treated a panel of cancer cell lines with increasing concentrations and determined the compounds cytotoxicity (IC_{50}) after 72 h (Table 4). Unexpectedly, majority of compounds did not display measurable IC_{50} , due to the limited solubility in culture medium. Otherwise, we determined the IC_{50} values of the most active derivatives reaching up to mid-micromolar ranges (Table 4).

The most active compound **43** was further profiled on a additional cancer cell lines (See Supp. Info. Table S3).

In parallel, we tested the ability of our compounds to activate apoptosis in treated G361 cells after 24 h at a single dose of 50 μ M. Using the one-step cellular caspase-3/7 activity assay [45] we determined that many compounds increased caspases activity over untreated control cells (Table 4); compound **43** more than ten-times and therefore we chose **43** as a candidate and evaluated its biological effects on melanoma G361 cell line in more detail.

In addition, the selectivity profile of the compound **43** was further characterized by assays on additional CDKs, including CDK1, CDK4, CDK5, CDK7 and CDK9. The compound yielded IC_{50} values in a mid-micromolar range (Table 5) and clearly showed that **43** is a pan-specific inhibitor of CDKs.

Asynchronously growing cells were treated by increasing

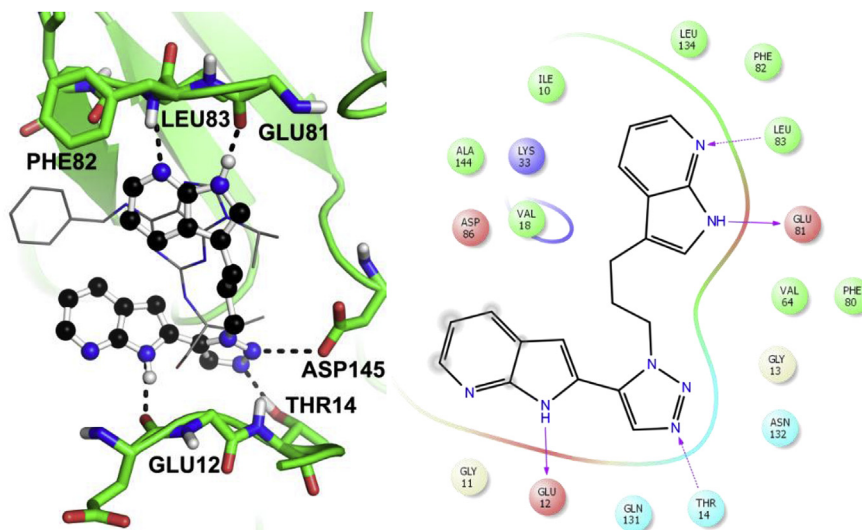


Fig. 7. Molecular docking of **43** into CDK2 (PDBID: 2A4L). Docking pose of **43** is shown on the left with ball and stick representation with annotated hydrogen bonds and roscovitine moiety in line representation. 2D interaction pattern of **43** is shown on the right. Amino acids within 3.5 Å from the ligand are shown whereas direct hydrogen bonds are shown with arrows. Solvent exposed part of the ligand is shown with circles.

Table 5
CDK selectivity profile for compound **43** and some known CDK inhibitors assayed as a control.

Kinases	IC ₅₀ (μM) ^a				
	43	Dinacliclib	BS-181	LY2835219	LDC000067
CDK1	11.5	0.072	14.0	0.371	3.95
CDK2	2.0	0.002	1.8	0.347	2.33
CDK4	9.0	0.127	44.7	0.005	3.16
CDK5	10.7	0.045	3.7	0.405	4.96
CDK7	>50	0.170	0.134	3.112	>20
CDK9	9.0	0.178	1.790	0.101	0.23

^a Average values from at least three determinations.

concentrations of **43** for 24 h and then analyzed by flow cytometry. Compound **43** potently reduced population of actively-replicating cells (BrdU positive) in G361 (Fig. 8A) and MCF-7 (data not shown) cells and accumulated cells in G1 and G2/M phases. We also observed a dose-dependent increase in sub G1/apoptotic cells that reached up 34% for **43** (80 μM).

The previous analysis was complemented by the monitoring of the cellular expression of some cell cycle regulators in G361 cells treated with **43** for 24 h. We observed a rapid dose-dependent decrease of Rb level starting already after of 20 μM treatment. This inhibitory pattern corresponds to reduction of number of cells in S and G2/M phases upon treatment documented by decrease of protein levels of cyclins A and B (Fig. 8B).

In addition, we evaluated the effect of **43** on the

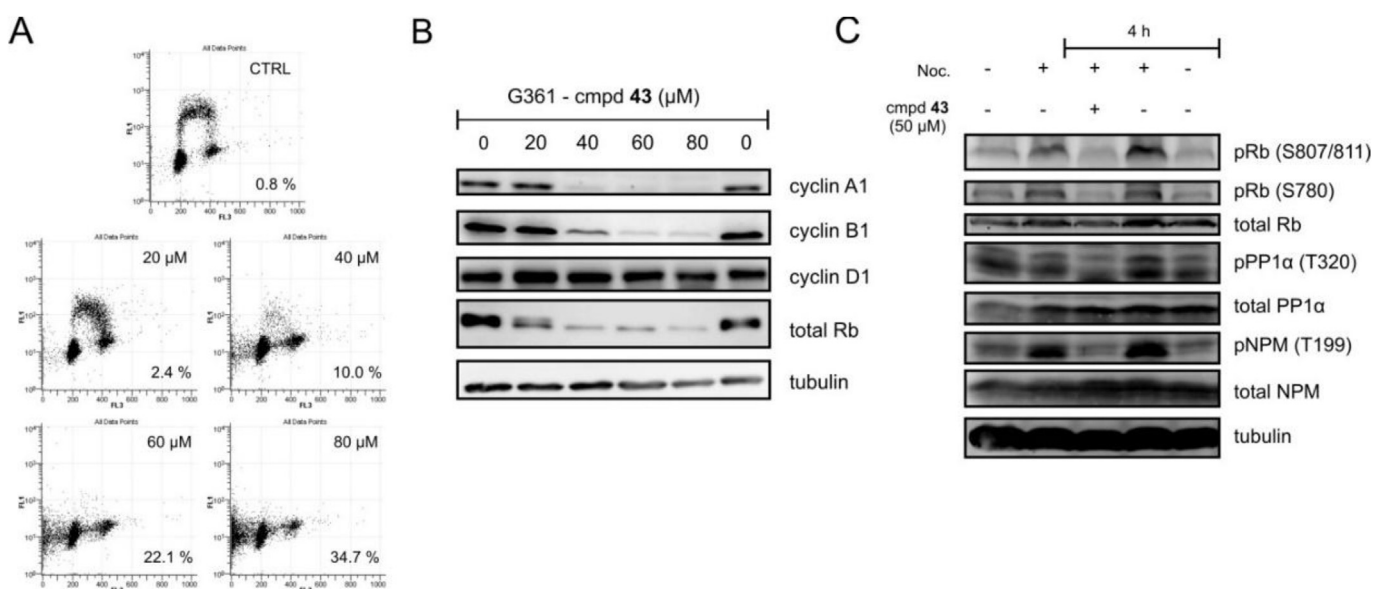


Fig. 8. A) Effect of compound **43** on cell cycle in G361 cells treated for 24 h. Flow cytometric analysis (10 000 counts) of DNA stained by propidium iodide (FL3 signal) and 5-bromo-2'-deoxyuridine (FL1 signal). The number in the lower right corner refers to the percentage of apoptotic cells (subG1 population). B) Immunoblotting analysis of cell cycle-related proteins in G361 cells treated by compound **43** for 24 h. C) Compound **43** inhibits the phosphorylation of CDKs substrates. G361 cells were first synchronized with nocodazole (3 ng/mL for 16 h) and then treated with **43** as indicated. Substrates of CDKs were detected by immunoblotting analysis and tubulin was included as a control for equal protein loading.

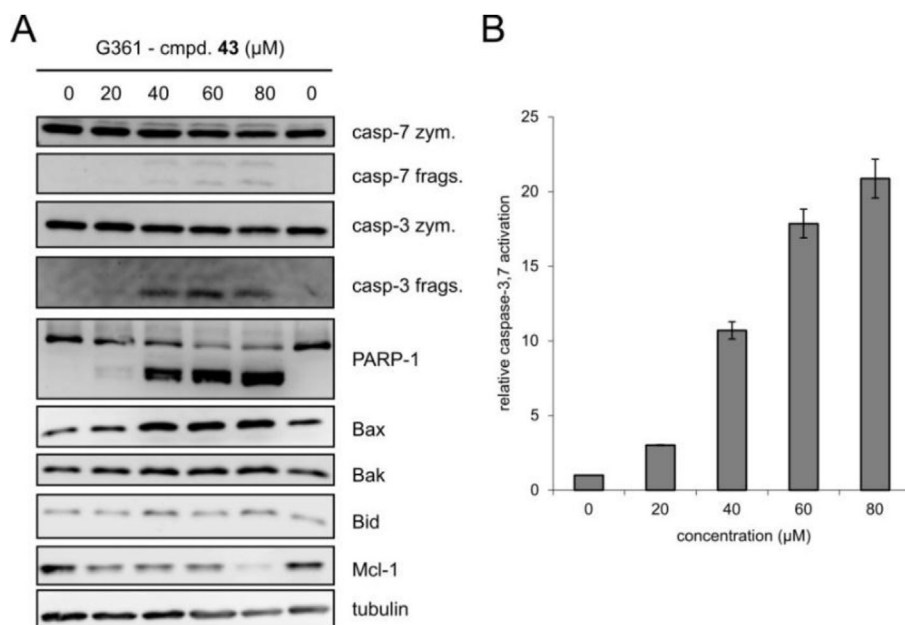


Fig. 9. A) Immunoblotting analysis of apoptosis-related proteins in G361 cells treated by compound **43** for 24 h. Tubulin level is included as a control for equal loading. B) Fluorimetric caspase-3,7 activity assay in G361 cell lysates treated with varying doses of **43** for 24 h. The activities of caspases were measured using the fluorogenic substrate Ac-DEVD-AMC and normalized against an untreated control.

phosphorylation of known CDK substrates. When G361 cells were treated with **43** for 4 h, the abundance of the phospho-forms of Rb, protein phosphatase PP1alpha and nucleophosmin (NPM) decreased (Fig. 8c).

2.7. Induction of apoptosis

Besides measuring antiproliferative activity of compounds **43** we also evaluated in a detail the ability of **43** to induce apoptosis in G361 cells by immunoblotting and caspase activity assay. As shown in Fig. 9A, compound **43** influenced mitochondrial pathway of apoptosis in treated cells as documented by a decrease of pro-survival protein Mcl-1 and an increase of proapoptotic proteins Bax, Bak and Bid. Dose-dependent response has clearly correlated with observations of i) cellular level of zymogenes of effector caspases 3 and 7, ii) their active fragments and iii) cleavage of substrate, PARP (poly-ADP ribose polymerase). In addition, the activation of caspases was confirmed by an enzyme activity assay using fluorescently labeled substrates of caspases 3 and 7 (Fig. 9B).

3. Conclusion

In summary, thanks to a convergent synthesis pattern, 24 pyrrolo[2,3-*b*]pyridine based derivatives (**21–44**) were synthesized in 5–10 steps, with moderate to very good yields. The key step [3 + 2] cycloaddition reaction allowed us to enlarge our library of tricyclics by giving us the opportunity to make 1,4- and 1,5-disubstituted triazoles. All the compounds were evaluated for *in vitro* inhibitory activity against CDK2/cyclin E and antiproliferative activity on cancer cell lines. Interestingly, some of them showed a single digit micro-molar inhibitory activity towards CDK2/cyclin E. From known CKIs from the literature, a 3D-QSAR CoMFA model was built and validated with good statistical results. This model gave us a correct prediction of the biological activity of our potent compounds towards CDK2/cyclin E and the contour map obtained will be of great use for improving our CKIs through a pharmacomodulation process.

Among the most potent compounds, the derivative **43**, containing a 1,5-disubstituted triazole, was found to be a pan-specific inhibitor of CDKs. The binding mode of **43** in CDK2 cavity have been described by molecular docking study and revealed donor-acceptor pattern of interaction of 7-azaindole ring with the main chain of Glu81 and Leu83. Cellular inhibition of CDK2 was confirmed by the monitoring of changes in the phosphorylation of known CDK substrates, Rb, PP1alpha and NPM.

Antiproliferative activity of **43** was studied in detail on G361 and MCF-7 cell lines and revealed a decrease of DNA replication and proliferation and a cell accumulation in G1 and G2/M phases in treated cells as it was previously described for numerous other pan-selective CDK inhibitors. In parallel, compound **43** showed to induce apoptosis in treated cells as it was documented by flow cytometry (subG1 population), immunoblot analysis of cleaved PARP and caspase activity assay.

4. Experimental section

4.1. Material and methods

All commercial materials were used without further purification. For anhydrous and inert reactions, the glassware was dried in an oven and several vacuum–nitrogen cycles were performed beforehand. The thin layer chromatography (TLC) studies were performed using commercial pre-coated aluminium sheets silica gel (60 Å, F₂₅₄) marketed by Merck. The purifications by chromatography on silica gel columns were carried out on an ISCO purification unit, Combi Flash RF 75 PSI, with Redisepp flash silica gel columns (60 Å, 230–400 mesh, grade 9385). Purities of compounds were assessed by inspection of their proton and carbon nuclear magnetic resonance spectroscopy (¹H and ¹³C NMR) spectra and high resolution mass spectrometry (HRMS) spectra. NMR spectra were measured on a Brücker Ultrashield 300 spectrometer, 300 MHz (¹H) and 75 MHz (¹³C) in deuterated chloroform (CDCl₃) or dimethylsulfoxide (DMSO-*d*₆). The notations used are: δ: chemical shift (ppm), s: singlet; d: doublet; dd: doublet of doublet; t: triplet; m:

multiplet; br: broad signal; *J*: coupling constant (Hz). The HRMS was performed by the mass spectrometry service on a Q-Exactive spectrometer from Thermo Scientific using the electrospray ionisation (ESI) technique. From known CKIs from the literature, a 3D-QSAR CoMFA model was built and validated with good statistical results. This model gave us a correct prediction of the biological activity of our potent compounds towards CDK2/cyclin E and the contour map obtained will be of great use for improving our CKIs through a pharmacomodulation process.

4.2. Synthesis of compounds 2–20

4.2.1. 3-Iodo-1*H*-pyrrolo[2,3-*b*]pyridine (**2**)

1*H*-Pyrrolo[2,3-*b*]pyridine (**1**) (1.00 g, 8.47 mmol) was dissolved in ethanol (40 mL) then iodine (3.23 g, 12.71 mmol), potassium iodide (2.11 g, 12.71 mmol) and a sodium hydroxide solution (12.71 mL, 12.71 mmol, 1 M) were added. The mixture was stirred overnight at room temperature. The reaction mixture was diluted with water and extracted with ethyl acetate (3 times). The organic layers were gathered, washed with a thiosulfate solution (w/w 5%) (twice) and a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated to give an orange solid. The crude product was recrystallised in methanol and water to give 2.00 g of the pure expected product as an orange solid in 97% yield. ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm): 12.11 (br, 1H), 8.26 (dd, *J* = 1.5, 4.8 Hz, 1H), 7.72 (d, *J* = 2.7 Hz, 1H), 7.69 (dd, *J* = 1.5, 7.8 Hz, 1H), 7.16 (dd, *J* = 4.8, 7.8 Hz, 1H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ (ppm): 148.5, 144.2, 130.9, 128.5, 122.4, 116.9, 54.8.

4.2.2. 3-((trimethylsilyl)ethynyl)-1*H*-pyrrolo[2,3-*b*]pyridine (**3a**)

General procedure for the Sonogashira cross-coupling reaction. Compound **2** (1.5 g, 6.15 mmol), copper iodide (118 mg, 0.62 mmol), bis(triphenylphosphine)palladium(II) dichloride (217 mg, 0.31 mmol), triethylamine (1.28 mL, 9.23 mmol), ethynyl(trimethyl) silane (1.30 mL, 9.23 mmol) and THF (35 mL) were introduced in an oven dried round bottom flask under inert atmosphere. The mixture was stirred at room temperature for 20 h. The reaction mixture was diluted with EtOAc and filtered on Dicalite. The filtrate was washed with water and a saturated solution of sodium chloride, dried over magnesium sulfate, filtered and concentrated. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 7:3, to give 1.099 g of the clean expected product as a beige solid in 84% yield. ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm): 12.12 (br, 1H), 8.30 (dd, *J* = 1.5, 4.7 Hz, 1H), 7.93 (dd, *J* = 1.5, 7.8 Hz, 1H), 7.88 (d, *J* = 2.7 Hz, 1H), 7.18 (dd, *J* = 4.7, 7.8 Hz, 1H), 0.24 (s, 9H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ (ppm): 148.0, 144.4, 131.4, 127.6, 120.8, 117.0, 99.6, 95.5, 95.4, 0.7 (3C).

4.2.3. 3-Ethynyl-1*H*-pyrrolo[2,3-*b*]pyridine (**4a**)

Compound **3a** (181 mg, 0.85 mmol) was dissolved in anhydrous THF, TBAF in solution in THF (1.7 mL, 1.70 mmol, 1 M) was added and the mixture was stirred at room temperature for 2 h. The reaction mixture was diluted with EtOAc, washed with water and a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure to give 125 mg of the expected product as a pale brown solid in quantitative yield. ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm): 12.08 (br, 1H), 8.29 (dd, *J* = 1.5, 4.9 Hz, 1H), 7.96 (dd, *J* = 1.5, 7.8 Hz, 1H), 7.86 (br s, 1H), 7.17 (dd, *J* = 4.9, 7.8 Hz, 1H), 4.12 (s, 1H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ (ppm): 148.0, 144.3, 131.3, 127.6, 120.8, 117.0, 94.8, 82.0, 77.8.

4.2.4. 1-(Phenylsulfonyl)-1*H*-pyrrolo[2,3-*b*]pyridine (**5**)

In a heat dried and nitrogen purged round bottom flask, 1*H*-pyrrolo[2,3-*b*]pyridine (**1**) (1.01 g, 8.58 mmol), tetrabutylammonium bromide (81 mg, 0.25 mmol), finely grounded sodium hydroxide

(1.02 g, 25.41 mmol) and CH₂Cl₂ (20 mL) were mixed, stirred and cooled down to 0 °C in an ice bath, then benzene sulfonylchloride (1.35 mL, 10.59 mmol) was added slowly. The mixture was left to warm up to room temperature and stirred at this temperature for 1 h. The reaction was hydrolysed with water (20 mL) and extracted by CH₂Cl₂ (twice). The organic layer was washed with a saturated sodium chloride solution, dried over magnesium sulfate and concentrated under reduced pressure to give 2.38 g of a beige solid. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 7:3, to give 2.17 g of the pure expected product as a white solid in 99% yield. ¹H NMR (300 MHz, CDCl₃) δ (ppm): 8.45 (dd, *J* = 1.5, 4.8 Hz, 1H), 8.20 (d, *J* = 7.5 Hz, 2H), 7.86 (dd, *J* = 1.5, 7.8 Hz, 1H), 7.73 (d, *J* = 4.0 Hz, 1H), 7.57 (dd, *J* = 7.4 Hz, 1H), 7.52–7.47 (m, 2H), 7.19 (dd, *J* = 4.8, 7.8 Hz, 1H), 6.61 (d, *J* = 4.0 Hz, 1H). ¹³C NMR (75 MHz, CDCl₃) δ (ppm): 144.7 (2C), 138.3, 134.1, 129.9, 129.1 (2C), 128.0 (2C), 126.5, 123.0, 119.0, 105.5.

4.2.5. 2-Iodo-1-(phenylsulfonyl)-1*H*-pyrrolo[2,3-*b*]pyridine (**6**)

In a heat dried and nitrogen purged tricol round bottom flask, **5** (1.00 g, 3.88 mmol) was dissolved in THF (40 mL). The solution was cooled down to –60 °C and LDA (2.33 mL, 4.66 mmol, 2 M in THF) was added. The mixture was left to stir at –60 °C for 15 min. Diiodine (1.97 g, 7.76 mmol) in solution in THF (10 mL) was added slowly. The mixture was left to warm up to room temperature overnight and the reaction was quenched with water (25 mL). The mixture was diluted with EtOAc, washed with a thiosulfate solution (w/w 5%), water and a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure to give 1.8 g of a brown solid. The crude product, a mixture of **6** and a diiodinated derivative **6'**, was purified by flash chromatography on silica gel column, cyclohexane/EtOAc 85:15, to give 750 mg of the clean expected product **6** as a pale yellow solid in 50% yield, **6'** described below and a mixed fraction, both not quantified. ¹H NMR (300 MHz, CDCl₃) δ (ppm): 8.55 (dd, *J* = 1.5, 4.8 Hz, 1H), 8.22 (d, *J* = 7.2 Hz, 2H), 7.79 (dd, *J* = 1.6, 7.8 Hz, 1H), 7.58 (dd, *J* = 7.3 Hz, 1H), 7.52–7.43 (m, 2H), 7.20 (dd, *J* = 4.8, 7.8 Hz, 1H), 6.83 (s, 1H). ¹³C NMR (75 MHz, CDCl₃) δ (ppm): 149.7, 144.8, 138.8, 134.3, 129.2 (2C), 128.3 (2C), 127.9, 124.0, 120.4, 119.5, 76.4.

4.2.6. 2-Iodo-1-((2-iodophenyl)sulfonyl)-1*H*-pyrrolo[2,3-*b*]pyridine (**6'**)

¹H NMR (300 MHz, CDCl₃) δ (ppm): 8.55 (dd, *J* = 1.6, 8.0 Hz, 1H), 8.07 (dd, *J* = 1.6, 4.8 Hz, 1H), 7.93 (dd, *J* = 1.2, 7.8 Hz, 1H), 7.74 (dd, *J* = 1.6, 7.8 Hz, 1H), 7.58 (ddd, *J* = 1.2, 7.4, 8.0 Hz, 1H), 7.22 (ddd, *J* = 1.6, 7.8, 7.8 Hz, 1H), 7.08 (dd, *J* = 4.8, 7.8 Hz, 1H), 7.05 (s, 1H). ¹³C NMR (75 MHz, CDCl₃) δ (ppm): 153.7, 149.3, 144.2, 142.6, 141.9, 134.6, 134.2, 128.1, 127.8, 123.5, 119.5, 119.3, 92.1.

4.2.7. 2-Iodo-1*H*-pyrrolo[2,3-*b*]pyridine (**7**)

In a heat dried and nitrogen purged round bottom flask, **5** (509 mg, 1.97 mmol) was dissolved in THF (20 mL). The solution was cooled down to –60 °C and LDA (1.97 mL, 3.95 mmol, 2 M in THF) was added. The mixture was left to stir at –60 °C for 15 min the diiodine (752 mg, 2.96 mmol) in solution in THF (10 mL) was added. The mixture was left to warm up to room temperature overnight and the reaction was quenched with water (5 mL). The mixture was diluted with EtOAc, washed with a thiosulfate solution (w/w 5%), water and a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure. The crude product was purified by flash chromatography on silica gel column, cyclohexane/EtOAc 85:15. A mixture of **6** and **6'** was obtained. This mixture was solubilised in dioxane (10 mL) and sodium *tert*-butoxide (380 mg, 3.95 mmol) was added. The mixture was left under stirring and reflux for 2 h. After being cooled down to room temperature, the reaction was quenched with water

(20 mL) and extracted with EtOAc. The organic layer was washed with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure to give 418 mg of the expected product as a pale orange solid in 87% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.19 (br, 1H), 8.28 (dd, $J = 1.6, 4.7$ Hz, 1H), 7.95 (dd, $J = 1.5, 7.9$ Hz, 1H), 7.09 (dd, $J = 4.7, 7.9$ Hz, 1H), 6.77 (sd, $J = 1.6$ Hz, 1H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 155.7, 147.7, 131.7, 126.9, 121.1, 114.8, 85.9.

4.2.8. 1-(Phenylsulfonyl)-2-((trimethylsilyl)ethynyl)-1H-pyrrolo[2,3-b]pyridine (**3b**)

Compound **3b** was synthesized following the general procedure for the Sonogashira coupling reaction from **6**. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 8:2, to give 285 mg of the expected product as a yellow oil (purity 70%), 46% calculated yield. ^1H NMR (300 MHz, CDCl_3) δ (ppm): 8.53 (dd, $J = 1.5, 4.7$ Hz, 1H), 8.21 (d, $J = 7.3$ Hz, 2H), 7.79 (dd, $J = 1.5, 7.9$ Hz, 1H), 7.58 (dd, $J = 7.3$ Hz, 1H), 7.53–7.43 (m, 2H), 7.20 (dd, $J = 4.8, 7.9$ Hz, 1H), 6.83 (s, 1H), 0.35 (s, 9H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 148.3, 146.4, 139.3, 134.1, 129.4, 129.1 (2C), 128.0 (2C), 121.4, 121.3, 119.7, 114.2, 105.0, 94.6, -0.25 (3C).

4.2.9. 2-((Trimethylsilyl)ethynyl)-1H-pyrrolo[2,3-b]pyridine (**3c**)

The derivative **3c** was synthesized following the general procedure for the Sonogashira coupling reaction from **7**. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 98:2, to give 381 mg of the clean expected product as an orange solid in 86% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.14 (brs, 1H), 8.27 (dd, $J = 1.5, 4.7$ Hz, 1H), 7.93 (ddd, $J = 0.6, 1.5, 7.9$ Hz, 1H), 7.09 (dd, $J = 4.7, 7.9$ Hz, 1H), 6.77 (sd, $J = 2.0$ Hz, 1H), 0.27 (s, 9H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 148.5, 145.2, 129.1, 119.6, 119.2, 116.9, 107.4, 99.2, 98.0, 0.2 (3C).

4.2.10. 2-Ethynyl-1H-pyrrolo[2,3-b]pyridine (**4b**)

Compound **3c** (331 mg, 1.55 mmol) was dissolved in THF (5 mL) and tetrabutylammonium fluoride (3.1 mL, 3.10 mmol, 1 M) was added. The mixture was stirred at room temperature for 2 h. The reaction mixture was diluted with EtOAc, washed with water and a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 98:2–7:3, to give 149 mg of the clean expected product as a beige solid in 68% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.20 (br, 1H), 8.28 (dd, $J = 1.6, 4.7$ Hz, 1H), 7.95 (dd, $J = 1.6, 7.9$ Hz, 1H), 7.09 (dd, $J = 4.7, 7.9$ Hz, 1H), 6.77 (s, 1H), 4.58 (s, 1H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 148.5, 145.0, 129.0, 119.6, 118.8, 116.8, 107.0, 85.4, 76.9.

4.2.11. 2-Chloro-1-(1H-pyrrolo[2,3-b]pyridin-3-yl)ethanone (**9a**)

In a heat dried and nitrogen purged round bottom flask, aluminium chloride (3.99 g, 30.0 mmol) and CH_2Cl_2 (300 mL) were mixed and stirred at room temperature for 5 min then **1** (714 mg, 6.05 mmol) was added. The mixture was stirred for 1.5 h then cooled in an ice bath and chloroacetyl chloride (2.39 mL, 30.0 mmol) was added slowly. The reaction mixture was left to warm up to room temperature and left to stir overnight (16 h) at room temperature under nitrogen atmosphere. The mixture was cooled in an ice bath and quenched by MeOH (60 mL). The solvents were removed under reduced pressure and a saturated sodium bicarbonate solution (120 mL) was added to the resulting mixture which was then extracted three times by EtOAc. The gathered organic layers were washed with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure to give 1.01 g of a beige solid. The solid was triturated in CH_2Cl_2 and filtered to give 648 mg of the pure expected product as a pale beige solid in 55% yield. ^1H NMR (300 MHz,

DMSO- d_6) δ (ppm): 12.68 (br, 1H), 8.60 (s, 1H), 8.46 (dd, $J = 1.8, 7.8$ Hz, 1H), 8.36 (dd, $J = 1.8, 4.8$ Hz, 1H), 7.29 (dd, $J = 4.8, 7.8$ Hz, 1H), 4.94 (s, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 186.9, 149.4, 145.1, 135.6, 130.0, 118.9, 118.2, 112.8, 46.7.

4.2.12. 3-Chloro-1-(1H-pyrrolo[2,3-b]pyridin-3-yl)propan-1-one (**9b**)

In a heat dried and nitrogen purged round bottom flask, aluminium chloride (3.225 g, 25.40 mmol) and CH_2Cl_2 (100 mL) were mixed and stirred at room temperature for 5 min then **1** (600 mg, 5.08 mmol) was added. The mixture was stirred for 1.5 h then cooled in an ice bath and 3-chloropropionyl chloride (2.42 mL, 25.40 mmol) was added slowly. The reaction mixture was left to warm up to room temperature and left to stir overnight (16 h) at rt under nitrogen atmosphere. The mixture was cooled in an ice bath and quenched by MeOH (15 mL). The solvents were removed under reduced pressure and a saturated sodium bicarbonate solution (120 mL) was added to the resulting mixture which was then extracted three times by EtOAc. The gathered organic layers were washed with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure to give 1.01 g of the expected product as a beige solid 95% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.57 (br, 1H), 8.58 (br s, 1H), 8.49 (dd, $J = 1.8, 7.8$ Hz, 1H), 8.34 (dd, $J = 1.8, 4.8$ Hz, 1H), 7.26 (dd, $J = 4.8, 7.8$ Hz, 1H), 3.96 (t, $J = 6.3$ Hz, 2H), 3.40 (t, $J = 6.3$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 192.2, 149.5, 144.8, 135.4, 130.0, 118.7, 118.1, 115.5, 41.3, 40.6.

4.2.13. 3-(2-Chloroethyl)-1H-pyrrolo[2,3-b]pyridine (**10a**)

DME (10 mL) was put in a heat dried and nitrogen purged round bottom flask and cooled in an ice bath to 0 °C, then aluminium chloride (685 mg, 5.15 mmol) and lithium aluminium hydride (1.08 mL, 2.58 mmol, 2.4 mol L $^{-1}$) are added slowly. **9a** (200 mg, 1.03 mmol) in solution in DME (10 mL) was added and the mixture was stirred and left to warm up to room temperature then stirred at room temperature for 18 h. The reaction is quenched with MeOH (10 mL), concentrated under reduced pressure, mixed with a saturated sodium hydrogenocarbonate solution and extracted with EtOAc. The organic layer was washed with a saturated sodium chloride solution, dried over magnesium sulfate and concentrated under reduced pressure to give 182 mg of the expected product as an orange solid in 98% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 11.47 (br, 1H), 8.19 (dd, $J = 1.5, 4.8$ Hz, 1H), 8.01 (dd, $J = 1.5, 7.8$ Hz, 1H), 7.37 (d, $J = 2.4$ Hz, 1H), 7.04 (dd, $J = 4.8, 7.8$ Hz, 1H), 3.86 (t, $J = 7.4$ Hz, 2H), 3.15 (t, $J = 7.4$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 148.9, 143.0, 127.1, 124.6, 119.6, 115.4, 110.2, 45.5, 28.9.

4.2.14. 3-(3-Chloropropyl)-1H-pyrrolo[2,3-b]pyridine (**10b**)

DME (20 mL) was put in a heat dried and nitrogen purged round bottom flask and cooled in an ice bath to 0 °C, then aluminium chloride (798 mg, 6.00 mmol) and lithium aluminium hydride (1.25 mL, 3.00 mmol, 2.4 mol L $^{-1}$) are added slowly. **9b** (250 mg, 1.20 mmol) in solution in DME (20 mL) was added and the mixture was stirred and left to warm up to room temperature then stirred at room temperature for 20 h. The reaction is quenched with MeOH (30 mL), concentrated under reduced pressure, mixed with a saturated sodium hydrogenocarbonate solution and extracted with EtOAc. The organic layer was washed with a saturated sodium chloride solution, dried over magnesium sulfate and concentrated under reduced pressure to give 237 mg of an orange oil. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 6:4, to give 218 mg of the expected product as a pale yellow solid in 94% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 11.38 (br, 1H), 8.19 (dd, $J = 1.5, 4.5$ Hz, 1H), 7.95 (dd, $J = 1.5, 7.8$ Hz, 1H), 7.27 (br s, 1H), 7.03 (dd, $J = 4.5, 7.8$ Hz, 1H), 3.66 (t,

$J = 6.5$ Hz, 2H), 2.82 (dd, $J = 6.9, 7.5$ Hz, 2H), 2.03–2.12 (m, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 149.1, 142.9, 126.9, 123.5, 119.7, 115.3, 112.4, 45.5, 33.1, 22.3.

4.2.15. 3-(2-Azidoethyl)-1H-pyrrolo[2,3-b]pyridine (**11a**)

Compound **10a** (178 mg, 0.99 mmol), sodium azide (322 mg, 4.95 mmol), DMF (5 mL) and water (2.5 mL) were mixed and stirred at 60 °C for 16 h. The mixture was cooled to room temperature, diluted with water and extracted three times with EtOAc. The organic layer was washed three times with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated to give 190 mg of a brown oil. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 1:1, to give 141 mg of the expected product as a beige solid in 71% yield. ^1H NMR (300 MHz, DMSO) δ (ppm): 11.46 (br, 1H), 8.20 (dd, $J = 1.8, 4.5$ Hz, 1H), 8.01 (dd, $J = 1.8, 7.8$ Hz, 1H), 7.35 (d, $J = 2.4$ Hz, 1H), 7.04 (dd, $J = 4.5, 7.8$ Hz, 1H), 3.60 (t, $J = 7.0$ Hz, 2H), 2.97 (t, $J = 7.0$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 149.2, 143.0, 127.1, 124.3, 119.7, 115.4, 110.2, 51.4, 25.0.

4.2.16. 3-(3-Azidopropyl)-1H-pyrrolo[2,3-b]pyridine (**11b**)

Compound **10b** (195 mg, 1.00 mmol), sodium azide (325 mg, 5.00 mmol), DMF (5 mL) and water (2.5 mL) were mixed and stirred at 60 °C for 16 h. The mixture was cooled to room temperature, diluted with water and extracted three times with EtOAc. The organic layer was washed three times with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated to give 198 mg of the expected product as a yellow oil 99% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 11.37 (br, 1H), 8.18 (dd, $J = 1.5, 4.7$ Hz, 1H), 7.94 (dd, $J = 1.5, 7.8$ Hz, 1H), 7.27 (br s, 1H), 7.03 (dd, $J = 4.7, 7.8$ Hz, 1H), 3.37 (t, $J = 6.9$ Hz, 2H), 2.75 (dd, $J = 7.2, 7.8$ Hz, 2H), 1.85–1.95 (ddd, $J = 6.6, 6.9, 7.2, 7.8$ Hz, 2H). ^{13}C NMR (75 MHz DMSO- d_6) δ (ppm): 149.1, 142.9, 126.9, 123.4, 119.7, 115.3, 112.7, 50.8, 29.4, 22.3.

4.2.17. 2-Azido-1-(1H-pyrrolo[2,3-b]pyridin-3-yl)ethanone (**12a**)

9a (50 mg, 0.26 mmol), sodium azide (85 mg, 1.30 mmol), DMF (2 mL) and water (1 mL) were mixed and stirred at room temperature for 36 h. The mixture was diluted with water (20 mL) and extracted three times by EtOAc. The gathered organic layers were washed twice with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure to give 51 mg of the pure expected product as a white solid in 98% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.66 (br, 1H), 8.53 (s, 1H), 8.47 (dd, $J = 1.8, 7.8$ Hz, 1H), 8.36 (dd, $J = 1.8, 4.8$ Hz, 1H), 7.29 (dd, $J = 4.8, 7.8$ Hz, 1H), 4.66 (s, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 189.8, 149.4, 145.0, 135.2, 129.9, 118.9, 118.0, 112.7, 54.2.

4.2.18. 3-Azido-1-(1H-pyrrolo[2,3-b]pyridin-3-yl)propan-1-one (**12b**)

Compound **9b** (238 mg, 1.14 mmol) was dissolved in DMF (5 mL) then sodium azide (371 mg, 5.70 mmol) in solution in water (2.5 mL) was added. The mixture was stirred at 60 °C overnight (16 h) then poured onto ice and extracted three times with EtOAc. The gathered organic layers were washed twice with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure to give 282 mg a yellow solid. The crude product was purified by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{EtOAc}$ 7:3, to give 146 mg of the expected product as a white solid in 60% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.56 (br, 1H), 8.57 (s, 1H), 8.48 (dd, $J = 1.8, 7.8$ Hz, 1H), 8.33 (dd, $J = 1.8, 4.8$ Hz, 1H), 7.26 (dd, $J = 4.8, 7.8$ Hz, 1H), 3.68 (t, $J = 6.3$ Hz, 2H), 3.22 (t, $J = 6.3$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 193.1, 149.4, 144.8, 135.2, 130.0, 118.6, 118.0, 115.3, 46.7, 37.9.

4.2.19. 3-Bromo-1H-pyrrolo[2,3-b]pyridine (**13**)

In a heat dried and nitrogen purged round bottom flask, **1** (1.00 g, 8.46 mmol) was dissolved in CH_2Cl_2 (20 mL) and cooled in an ice bath, then *N*-bromosuccinimide (1.66 g, 9.31 mmol) was added. The mixture was left to warm up to room temperature overnight under nitrogen atmosphere. The reaction mixture was diluted with CH_2Cl_2 and washed with a saturated sodium hydrogencarbonate solution. The aqueous layer was extracted three times with EtOAc and the gathered organic layers were washed with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated to give 1.79 g of a brown solid. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 1:1, to give 1.56 g of the expected product as a light brown solid in 93% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 11.45 (br, 1H), 8.38 (dd, $J = 1.2, 5.1$ Hz, 1H), 8.06 (dd, $J = 1.2, 7.8$ Hz, 1H), 7.48 (s, 1H), 7.28 (dd, $J = 5.1, 7.8$ Hz, 1H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 145.0, 140.9, 130.2, 125.6, 121.5, 116.3, 89.9.

4.2.20. 3-Bromo-1-(phenylsulfonyl)-1H-pyrrolo[2,3-b]pyridine (**14**)

In a heat dried and nitrogen purged round bottom flask, **13** (175 mg, 0.89 mmol), tetrabutylammonium bromide (10 mg, 0.03 mmol), finely grounded sodium hydroxide (107 mg, 2.67 mmol) and CH_2Cl_2 (5 mL) were mixed, stirred and cooled down to 0 °C in an ice bath, then benzene sulfonyl chloride (0.142 mL, 1.11 mmol) was added slowly. The mixture was left to warm up to room temperature and was stirred at this temperature for 1 h. The reaction was hydrolyzed with water (6 mL) and extracted by CH_2Cl_2 (twice). The organic layer was washed with water and a saturated sodium chloride solution, dried over magnesium sulfate and concentrated under reduced pressure to give an orange solid. The crude product was triturated in pentane, filtered, washed with pentane and dried to give 294 mg of the expected product as an orange solid in 98% yield. ^1H NMR (300 MHz, CDCl_3) δ (ppm): 8.48 (dd, $J = 1.5, 4.8$ Hz, 1H), 8.24–8.17 (m, 2H), 7.82 (dd, $J = 1.5, 8.0$ Hz, 1H), 7.79 (s, 1H), 7.60 (dd, $J = 7.5$ Hz, 1H), 7.54–7.46 (m, 2H), 7.27 (dd, $J = 4.8, 8.0$ Hz, 1H). ^{13}C NMR (75 MHz, CDCl_3) δ (ppm): 146.0 (2C), 138.0, 134.3, 129.2 (2C), 128.6, 128.1 (2C), 125.0, 122.5, 119.5, 95.6.

4.2.21. 3-Methyl-1-(phenylsulfonyl)-1H-pyrrolo[2,3-b]pyridine (**15**)

In a heat dried and nitrogen purged round bottom flask, **14** (1.064 g, 3.16 mmol) and tetrakis triphenylphosphine palladium (183 mg, 0.16 mmol) were dissolved in anhydrous THF (20 mL) then trimethylaluminum (3.17 mL, 6.32 mmol, 2.0 mol L^{-1}) was added slowly. The mixture was stirred at 60 °C for 16 h then cooled down to room temperature, diluted with a saturated sodium hydrogencarbonate solution and extracted by EtOAc three times. The organic layers were gathered, washed with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated to give 1.157 g of a red solid. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 8:2, to give 782 mg of the expected product as a yellow solid in 91% yield. ^1H NMR (300 MHz, CDCl_3) δ (ppm): 8.43 (dd, $J = 1.5, 4.8$ Hz, 1H), 8.20–8.12 (m, 2H), 7.78 (dd, $J = 1.5, 7.8$ Hz, 1H), 7.55 (dd, $J = 7.1, 7.5$ Hz, 1H), 7.50–7.42 (m, 4H), 7.18 (dd, $J = 4.8, 7.8$ Hz, 1H), 2.25 (sd, $J = 1.3$ Hz, 3H). ^{13}C NMR (75 MHz, CDCl_3) δ (ppm): 144.7 (2C), 138.6, 133.8, 129.0 (2C), 128.0, 127.8 (2C), 124.1, 123.2, 118.6, 115.1, 9.8.

4.2.22. 3-(Azidomethyl)-1-(phenylsulfonyl)-1H-pyrrolo[2,3-b]pyridine (**16**)

In a heat dried and nitrogen purged round bottom flask, **15**

(574 mg, 2.11 mmol), *N*-bromosuccinimide (413 mg, 2.32 mg), azobisisobutyronitrile (13 mg, 0.08 mmol) and carbon tetrachloride (10 mL) were mixed and stirred under reflux for 16 h. The reaction mixture was cooled to room temperature, and concentrated under reduced pressure. The crude product (brown oil), sodium azide (686 mg, 10.55 mmol), DMF (15 mL) and water (7.5 mL) were mixed and stirred at 60 °C for 16 h. The mixture was cooled to room temperature, diluted with water and extracted three times with EtOAc. The organic layer was washed three times with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated to give a yellow oil. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 8:2 to give 377 mg of the expected product as a white solid in 57% yield. ¹H NMR (300 MHz, CDCl₃) δ (ppm): 8.48 (dd, *J* = 1.6, 4.8 Hz, 1H), 8.23–8.16 (m, 2H), 7.91 (dd, *J* = 1.6, 7.9 Hz, 1H), 7.76 (s, 1H), 7.59 (dd, *J* = 7.5 Hz, 1H), 7.54–7.45 (m, 2H), 7.23 (dd, *J* = 4.8, 7.9 Hz, 1H), 4.46 (s, 2H). ¹³C NMR (75 MHz, CDCl₃) δ (ppm): 147.4, 145.7, 138.1, 134.2, 129.1 (2C), 128.3, 128.1 (2C), 124.9, 121.6, 119.4, 113.5, 46.2.

4.2.23. 1-(Phenylsulfonyl)-1*H*-pyrrolo[2,3-*b*]pyridine-2-carboxylic acid (**17**)

In a heat dried and nitrogen purged tricol round bottom flask, **5** (1.05 g, 4.07 mmol) was solubilized in dry THF (15 mL), cooled to –35 °C then lithium diisopropylamide (5.1 mL, 10.16 mmol, 2 M) was added slowly. The mixture was then cooled to –55 °C and dry ice was added slowly. The mixture was left to warm up to room temperature under stirring overnight. The reaction was quenched with water and extracted with EtOAc. The aqueous layer was acidified by the addition of a 6 N hydrochloric acid solution to pH = 1 and extracted with EtOAc. The latter organic layer was washed with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure to give 1.09 g of the expected product as a beige solid in 88% yield. ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm): 13.87 (br, 1H), 8.51 (dd, *J* = 1.6, 4.7 Hz, 1H), 8.26 (d, *J* = 7.0 Hz, 2H), 8.13 (dd, *J* = 1.6, 7.9 Hz, 1H), 7.77 (dd, *J* = 7.2 Hz, 1H), 7.72–7.65 (m, 2H), 7.39 (dd, *J* = 4.7, 7.9 Hz, 1H), 7.28 (s, 1H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ (ppm): 164.4, 149.5, 147.3, 138.5, 135.1, 133.3, 132.1, 129.8 (2C), 128.3 (2C), 121.1, 120.7, 112.5.

4.2.24. Ethyl 1*H*-pyrrolo[2,3-*b*]pyridine-2-carboxylate (**18**)

Compound **17** (302 mg, 1.00 mmol), EtOH (20 mL) and concentrated sulfuric acid (0.2 mL) were mixed in a round bottom flask and stirred under reflux for 20 h. The reaction mixture was cooled to room temperature and concentrated under reduced pressure. The residue was solubilized in EtOAc and washed three times with a saturated sodium hydrogenocarbonate solution, washed with water and a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure. The crude product was purified by chromatography on silica gel column, CH₂Cl₂/EtOAc 8:2, to give 178 mg of the expected product as a white solid in 94% yield. ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm): 12.49 (br, 1H), 8.41 (dd, *J* = 1.7, 4.6 Hz, 1H), 8.11 (dd, *J* = 1.6, 8.0 Hz, 1H), 7.17 (s, 1H), 7.16 (dd, *J* = 4.6, 8.0 Hz, 1H), 3.34 (q, *J* = 7.1 Hz, 2H), 1.34 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ (ppm): 161.5, 149.3, 147.2, 131.2, 128.3, 119.5, 117.3, 106.9, 61.1, 14.7.

4.2.25. (1*H*-Pyrrolo[2,3-*b*]pyridin-2-yl)methanol (**19**)

Dimethoxyethane (10 mL) was placed in a dry round bottom flask under nitrogen atmosphere and cooled to 0 °C. Lithium aluminium hydride (1.53 mL, 3.66 mmol, 2.4 M) then **18** (348 mg, 1.83 mmol) in solution in dimethoxyethane (10 mL) were added. The mixture was left to slowly warm up to room temperature under stirring overnight. The reaction was quenched with MeOH (15 mL) and the mixture was concentrated under reduced pressure. The

residue was solubilized in EtOAc, washed with a saturated sodium hydrogenocarbonate solution, water and a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated under reduced pressure to give 113 mg of the expected product as a pale yellow solid in 94% yield. ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm): 11.51 (br, 1H), 8.13 (dd, *J* = 1.5, 4.7 Hz, 1H), 7.85 (dd, *J* = 1.5, 7.8 Hz, 1H), 7.00 (dd, *J* = 4.7, 7.8 Hz, 1H), 6.29 (s, 1H), 5.28 (t, *J* = 5.0 Hz, 1H), 4.60 (d, *J* = 5.0 Hz, 2H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ (ppm): 149.2, 142.3, 141.6, 127.9, 120.5, 115.8, 97.5, 57.4.

4.2.26. 2-(Azidomethyl)-1*H*-pyrrolo[2,3-*b*]pyridine (**20**)

In a heat dried and nitrogen purged round bottom flask, **19** (311 mg, 2.10 mmol) and CH₂Cl₂ (10 mL) were mixed then thionyl chloride (0.184 mL, 2.52 mmol) was added. The mixture was stirred at room temperature for 1 h then concentrated under reduced pressure. The crude product was mixed with sodium azide (683 mg, 10.50 mmol), DMF (10 mL) and water (5 mL). The mixture was stirred at 60 °C overnight. The reaction mixture was poured onto ice and extracted three times with EtOAc. The organic layers were gathered and washed three times with a saturated sodium chloride solution, dried over magnesium sulfate, filtered and concentrated. The crude product was purified by chromatography on silica gel column, cyclohexane/EtOAc 7:3, to give 252 mg of the expected product as an off white solid in 69%. ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm): 11.86 (br, 1H), 8.21 (dd, *J* = 1.6, 4.7 Hz, 1H), 7.94 (dd, *J* = 1.5, 7.8 Hz, 1H), 7.06 (dd, *J* = 4.7, 7.8 Hz, 1H), 6.49 (sd, *J* = 1.9 Hz, 1H), 4.58 (s, 2H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ (ppm): 149.3, 143.5, 134.2, 128.7, 120.2, 116.3, 100.6, 47.6.

4.3. Synthesis of compounds 21–33. General procedure for the copper-catalyzed [3 + 2] cycloaddition reaction

4.3.1. 3-(2-(4-(1*H*-pyrrolo[2,3-*b*]pyridin-3-yl)-1*H*-1,2,3-triazol-1-yl)ethyl)-1*H*-pyrrolo[2,3-*b*]pyridine (**21**)

In a heat dried and nitrogen purged round bottom flask, **3a** (58 mg, 0.27 mmol), **11a** (50 mg, 0.27 mmol), copper iodide (1.4 mg, 0.03 mmol) and THF (5 mL) were mixed and stirred then tetrabutylammonium fluoride (0.81 mL, 0.81 mmol, 1 M) and diisopropylethylamine (0.092 mL, 0.54 mmol) were added. The mixture was stirred at 40 °C for 16 h. The reaction mixture was cooled to room temperature, concentrated and purified by chromatography on silica gel column, CH₂Cl₂/MeOH 95:5, to give 125 mg of a brown oil which was triturated in MeOH/diethylether to give 61 mg of the expected product as a beige solid in 69% yield. ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm): 11.87 (br, 1H), 11.42 (br, 1H), 8.47 (s, 1H), 8.37 (dd, *J* = 1.2, 8.0 Hz, 1H), 8.28 (dd, *J* = 1.5, 4.7 Hz, 1H), 8.18 (dd, *J* = 1.5, 4.7 Hz, 1H), 7.99 (dd, *J* = 1.1, 7.9 Hz, 1H), 7.85 (sd, *J* = 2.6 Hz, 1H), 7.25 (d, *J* = 2.4 Hz, 1H), 7.16 (dd, *J* = 4.7, 8.0 Hz, 1H), 7.03 (dd, *J* = 4.7, 7.9 Hz, 1H), 4.70 (t, *J* = 7.3 Hz, 2H), 2.34–2.21 (m, 2H). ¹³C NMR (75 MHz, DMSO-*d*₆) δ (ppm): 149.0, 148.9, 143.6, 143.0, 142.2, 128.7, 126.9, 124.2, 123.5, 120.2, 119.7, 117.5, 116.5, 115.5, 109.6, 105.7, 50.4, 26.4. HRMS-ESI (*m/z*): found 330.14636, calcd for C₁₈H₁₆N₇ [M + H]⁺ 330.14617.

Compounds **22–36** were synthesized following the general procedure for the copper-catalysed cycloaddition reaction (with or without additives).

4.3.2. 3-(2-(4-Phenyl-1*H*-1,2,3-triazol-1-yl)ethyl)-1*H*-pyrrolo[2,3-*b*]pyridine (**22**)

Purification by chromatography on silica gel column, CH₂Cl₂/MeOH 98:2, to give 123 mg of the expected product as an off white solid in 80% yield. ¹H NMR (300 MHz, DMSO-*d*₆) δ (ppm): 11.42 (br, 1H), 8.57 (s, 1H), 8.20 (br, 1H), 7.97 (d, *J* = 7.2 Hz, 1H), 7.80 (d, *J* = 7.2 Hz, 2H), 7.44 (dd, *J* = 7.2 Hz, 1H), 7.32 (dd, *J* = 7.2 Hz, 1H), 7.24 (d, 1H, *J* = 2.1 Hz), 7.03 (dd, *J* = 4.8, 7.8 Hz, 1H), 4.69 (t, *J* = 7.2 Hz,

2H), 3.33 (t, $J = 7.2$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 148.9, 145.6, 143.0, 131.3, 129.4 (3C), 128.4, 126.9, 125.5 (2C), 124.3, 121.8, 109.5, 50.6, 26.3. HRMS-ESI (m/z): found 290.14018, calcd for $\text{C}_{17}\text{H}_{16}\text{N}_5$ $[\text{M} + \text{H}]^+$ 290.14002.

4.3.3. 2-(4-(1H-Pyrrolo[2,3-*b*]pyridin-3-yl)-1H-1,2,3-triazol-1-yl)-1-(1H-pyrrolo[2,3-*b*]pyridin-3-yl)ethanone (**23**)

Purification by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 9:1–8:2, to give 123 mg of the expected product as a brown solid not pure. The solid was triturated in water, filtered, washed with diethyl ether and dried to give 44 mg of the expected product as a brown solid in 26% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.80 (br, 1H), 11.93 (br, 1H), 8.80 (d, $J = 3.0$ Hz, 1H), 8.50 (s, 1H), 8.59–8.42 (m, 2H), 8.42–8.20 (m, 2H), 7.97 (d, $J = 2.1$ Hz, 1H), 7.30 (dd, $J = 4.8, 7.8$ Hz, 1H), 7.19 (dd, $J = 4.2, 7.8$ Hz, 1H), 6.05 (s, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 187.1, 149.4, 149.1, 145.2, 143.6, 142.5, 135.7, 134.5, 129.9, 128.8, 123.7, 122.1, 119.0, 118.1, 112.7, 105.7, 55.6. HRMS-ESI (m/z): found 344.12582, calcd for $\text{C}_{18}\text{H}_{14}\text{N}_7\text{O}$ $[\text{M} + \text{H}]^+$ 344.12543.

4.3.4. 2-(4-Phenyl-1H-1,2,3-triazol-1-yl)-1-(1H-pyrrolo[2,3-*b*]pyridin-3-yl)ethanone (**24**)

Purification by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 98:2, to give 52 mg of the expected product as a white solid in 68% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.80 (br, 1H), 8.79 (sd, $J = 3.0$ Hz, 1H), 8.61 (s, 1H), 8.43 (dd, $J = 1.5, 7.8$ Hz, 1H), 8.38 (dd, $J = 1.5, 4.8$ Hz, 1H), 7.89 (d, $J = 7.6$ Hz, 2H), 7.48 (dd, $J = 7.3, 7.6$ Hz, 2H), 7.35 (dd, $J = 7.6$ Hz, 1H), 7.30 (dd, $J = 4.8, 7.8$ Hz, 1H), 6.05 (s, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 186.9, 149.4, 146.7, 145.2, 135.8, 131.3, 129.9, 129.4 (2C), 128.3, 125.6 (2C), 123.7, 119.0, 118.1, 112.6, 55.7. HRMS-ESI (m/z): found 304.11961, calcd for $\text{C}_{17}\text{H}_{14}\text{N}_5\text{O}$ $[\text{M} + \text{H}]^+$ 304.11929.

4.3.5. 3-(3-(4-(1H-Pyrrolo[2,3-*b*]pyridin-3-yl)-1H-1,2,3-triazol-1-yl)propyl)-1H-pyrrolo[2,3-*b*]pyridine (**25**)

Purification by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 9:1, to give 49 mg of the expected product as a beige solid not pure. The solid was triturated in THF, filtered, washed with THF and dried to give 30 mg of the expected product as a brown solid in 38% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 11.88 (br, 1H), 11.39 (br, 1H), 8.52 (s, 1H), 8.43 (dd, $J = 1.2, 7.8$ Hz, 1H), 8.28 (dd, $J = 1.5, 4.8$ Hz, 1H), 8.19 (dd, $J = 1.5, 4.5$ Hz, 1H), 7.95 (dd, $J = 1.2, 7.8$ Hz, 1H), 7.89 (d, $J = 2.8$ Hz, 1H), 7.32 (d, $J = 2.1$ Hz, 1H), 7.17 (dd, $J = 4.8, 8.1$ Hz, 1H), 7.03 (dd, $J = 4.8, 7.8$ Hz, 1H), 4.46 (t, $J = 7.3$ Hz, 2H), 2.74 (t, $J = 7.3$ Hz, 2H), 2.28 (td, $J = 7.3$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 149.2, 149.1, 143.6, 142.9, 142.4, 128.8, 127.0, 123.6 (2C), 120.2, 119.7, 117.5, 116.4, 115.3, 112.5, 105.7, 49.6, 30.7, 22.1. HRMS-ESI (m/z): found 344.16204, calcd for $\text{C}_{19}\text{H}_{18}\text{N}_7$ $[\text{M} + \text{H}]^+$ 344.16182.

4.3.6. 3-(4-(1H-Pyrrolo[2,3-*b*]pyridin-3-yl)-1H-1,2,3-triazol-1-yl)-1-(1H-pyrrolo[2,3-*b*]pyridin-3-yl)propan-1-one (**26**)

Purification by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 9:1–85:15, to give 93 mg of the expected product as an orange solid not pure. The solid was triturated in CH_2Cl_2 , filtered, washed with diethyl ether and dried to give 77 mg of the expected product as a beige solid in 57% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.56 (br, 1H), 11.87 (br, 1H), 8.58 (s, 1H), 8.50 (s, 1H), 8.47 (dd, $J = 1.5, 7.8$ Hz, 1H), 8.39 (dd, $J = 1.5, 7.8$ Hz, 1H), 8.32 (dd, $J = 1.5, 4.8$ Hz, 1H), 8.27 (dd, $J = 1.5, 4.8$ Hz, 1H), 7.88 (sd, $J = 2.8$ Hz, 1H), 7.25 (dd, $J = 4.8, 7.8$ Hz, 1H), 7.16 (dd, $J = 4.2, 7.8$ Hz, 1H), 4.84–4.74 (m, 2H), 3.72–3.60 (m, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 192.5, 149.4, 149.0, 144.8, 143.6, 142.3, 135.3, 130.0, 128.7, 123.6, 120.5, 118.7, 118.0, 117.5, 116.4, 115.2, 105.6, 45.5, (1C in the DMSO peak). HRMS-ESI (m/z): found 358.14171, calcd for $\text{C}_{19}\text{H}_{16}\text{N}_7\text{O}$

$[\text{M} + \text{H}]^+$ 358.14108.

4.3.7. 3-((4-(1H-Pyrrolo[2,3-*b*]pyridin-3-yl)-1H-1,2,3-triazol-1-yl)methyl)-1-(phenylsulfonyl)-1H-pyrrolo[2,3-*b*]pyridine (**27**)

Purification by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 98:2, to give 71 mg of the expected product as a white solid in 68% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 11.88 (br, 1H), 8.54 (s, 1H), 8.45–8.35 (m, 2H), 8.27 (dd, $J = 1.6, 4.8$ Hz, 1H), 8.19 (s, 1H), 8.15 (d, $J = 7.2$ Hz, 2H), 8.02 (dd, $J = 1.6, 8.0$ Hz, 1H), 7.88 (d, $J = 2.6$ Hz, 1H), 7.73 (dd, $J = 7.5$ Hz, 1H), 7.67–7.59 (m, 2H), 7.33 (dd, $J = 4.8, 8.0$ Hz, 1H), 7.16 (dd, $J = 4.8, 7.9$ Hz, 1H), 5.84 (s, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 149.0, 147.1, 145.7, 143.7, 142.9, 137.8, 135.4, 130.1 (2C), 129.4, 128.8, 128.1 (2C), 126.5, 123.9, 121.7, 120.2, 120.0, 117.4, 116.5, 114.64, 105.4, 44.8. HRMS-ESI (m/z): found 456.12390, calcd for $\text{C}_{23}\text{H}_{18}\text{N}_7\text{O}_2\text{S}$ $[\text{M} + \text{H}]^+$ 456.12372.

4.3.8. 3-(1-((1H-Pyrrolo[2,3-*b*]pyridin-2-yl)methyl)-1H-1,2,3-triazol-4-yl)-1H-pyrrolo[2,3-*b*]pyridine (**28**)

Purification by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 95:5, to give 110 mg of the expected product as an orange solid not pure. The solid was triturated in water, filtered, washed with diethyl ether and dried to give 74 mg of the clean expected product as a brown solid in 59% yield. ^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 11.96 (br, 1H), 11.89 (br, 1H), 8.52 (s, 1H), 8.44 (dd, $J = 1.3, 7.9$ Hz, 1H), 8.34–8.16 (m, 2H), 7.97–7.89 (m, 2H), 7.18 (dd, $J = 4.7, 7.9$ Hz, 1H), 7.06 (dd, $J = 4.7, 7.8$ Hz, 1H), 6.50 (sd, $J = 1.7$ Hz, 1H), 5.82 (s, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 149.3, 149.1, 143.6, 143.5, 142.8, 134.3, 128.8, 123.8, 120.3, 117.5, 116.5, 116.4, 105.5, 100.6, 47.2. HRMS-ESI (m/z): found 316.13090, calcd for $\text{C}_{17}\text{H}_{14}\text{N}_7$ $[\text{M} + \text{H}]^+$ 316.13052.

4.3.9. 2-(1-(2-(1H-Pyrrolo[2,3-*b*]pyridin-3-yl)ethyl)-1H-1,2,3-triazol-4-yl)-1-(phenylsulfonyl)-1H-pyrrolo[2,3-*b*]pyridine (**29**) and 3-(2-(4-(1H-pyrrolo[2,3-*b*]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)ethyl)-1H-pyrrolo[2,3-*b*]pyridine (**30**)

Purification by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 95:5, to give two fractions. Fraction 1: 87 mg of a brown solid not pure. The solid was triturated in diethyl ether, filtered, washed with diethyl ether and dried to give 80 mg of the expected product **29** as a brown solid in 43% yield. Fraction 2: 176 mg of a brown solid/oil not pure. The solid/oil was triturated in EtOAc/ $\text{CH}_2\text{Cl}_2/\text{cyclohexane}$, filtered, washed with CH_2Cl_2 and dried to give 51 mg of the product **30** as a pale pink solid in 39% yield.

4.3.10. 2-(1-(2-(1H-Pyrrolo[2,3-*b*]pyridin-3-yl)ethyl)-1H-1,2,3-triazol-4-yl)-1-(phenylsulfonyl)-1H-pyrrolo[2,3-*b*]pyridine (**29**)

^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.19 (br, 1H), 8.45 (s, 1H), 8.34 (dd, $J = 1.5, 4.8$ Hz, 1H), 8.29–8.15 (br, 1H), 8.10 (dd, $J = 1.5, 7.9$ Hz, 1H), 7.99–7.89 (m, 3H), 7.72 (s, 1H), 7.60 (dd, $J = 7.5$ Hz, 1H), 7.47 (dd, $J = 7.5, 8.0$ Hz, 2H), 7.30 (dd, $J = 4.8, 7.9$ Hz, 1H), 7.08 (dd, $J = 4.8, 7.9$ Hz, 1H), 6.80 (s, 1H), 4.82 (t, $J = 6.8$ Hz, 2H), 3.40–3.30 (m, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 149.9, 149.4, 147.1, 145.3, 143.3, 140.5, 137.8, 135.0, 130.4, 129.9 (2C), 129.2, 128.4, 128.3, 127.6 (2C), 124.8, 122.8, 122.3, 119.7, 116.0, 97.3, 49.4, 25.7. HRMS-ESI (m/z): found 470.13994, calcd for $\text{C}_{24}\text{H}_{20}\text{N}_7\text{O}_2\text{S}$ $[\text{M} + \text{H}]^+$ 470.13937.

4.3.11. 3-(2-(4-(1H-Pyrrolo[2,3-*b*]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)ethyl)-1H-pyrrolo[2,3-*b*]pyridine (**30**)

^1H NMR (300 MHz, DMSO- d_6) δ (ppm): 12.13 (br, 1H), 11.43 (br, 1H), 8.41 (s, 1H), 8.19 (dd, $J = 1.5, 4.8$ Hz, 2H), 7.97 (dd, $J = 1.5, 7.9$ Hz, 1H), 7.93 (dd, $J = 1.5, 7.9$ Hz, 1H), 7.24 (sd, $J = 2.4$ Hz, 1H), 7.09–7.00 (m, 2H), 6.78 (sd, $J = 2.0$ Hz, 1H), 4.75 (t, $J = 7.0$ Hz, 2H), 3.39–3.29 (m, 2H). ^{13}C NMR (75 MHz, DMSO- d_6) δ (ppm): 149.5, 149.0, 143.2, 143.0, 140.4, 130.6, 128.3, 126.9, 124.4, 122.2, 120.9, 119.6, 116.4, 115.5, 109.4, 97.2, 50.6, 26.4. HRMS-ESI (m/z): found 330.14637,

calcd for C₁₈H₁₆N₇ [M + H]⁺ 330.14617.

4.3.12. 2-(4-(1H-Pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)-1-(1H-pyrrolo[2,3-b]pyridin-3-yl)ethanone (**31**)

The reaction mixture was cooled to room temperature and filtered. The solid was washed with THF and dried under reduced pressure to give 115 mg of the expected product as an off white solid in 84% yield. ¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 12.81 (br, 1H), 12.26 (br, 1H), 8.80 (sd, *J* = 3.1 Hz, 1H), 8.55 (s, 1H), 8.44 (dd, *J* = 1.2, 7.9 Hz, 1H), 8.39 (d, *J* = 3.9 Hz, 1H), 8.32–8.15 (br, 1H), 7.97 (d, *J* = 7.9 Hz, 1H), 7.31 (dd, *J* = 4.8, 7.9 Hz, 1H), 7.15–7.02 (m, 1H), 6.89 (s, 1H), 6.13 (s, 2H). ¹³C NMR (75 MHz, DMSO-d₆) δ (ppm): 186.9, 149.4 (2C), 145.2 (2C), 140.5, 135.8 (2C), 130.6, 129.9, 128.3, 124.0 (2C), 119.0, 118.1, 112.6, 97.4, 55.7. HRMS-ESI (*m/z*): found 344.12578, calcd for C₁₈H₁₄N₇O [M + H]⁺ 344.12543.

4.3.13. 2-(1-(3-(1H-Pyrrolo[2,3-b]pyridin-3-yl)propyl)-1H-1,2,3-triazol-4-yl)-1-(phenylsulfonyl)-1H-pyrrolo[2,3-b]pyridine (**32**) and 3-(3-(4-(1H-pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)propyl)-1H-pyrrolo[2,3-b]pyridine (**33**)

The reaction mixture was cooled to room temperature and concentrated. The residue was purified by chromatography on silica gel column, CH₂Cl₂/MeOH 95:5, to give two fractions. Fraction 1: 146 mg of an orange solid which was triturated in EtOAc/CH₂Cl₂/cyclohexane, filtered, washed with water and diethyl ether, and dried to give 107 mg of the clean expected product **32** as a light brown solid in 47% yield. Fraction 2: 188 mg of a brown solid/oil which was triturated in EtOAc/CH₂Cl₂/cyclohexane, filtered, washed with water and diethylether, and dried to give 59 mg of the clean expected product **33** as a light brown solid in 37% yield.

4.3.14. 2-(1-(3-(1H-Pyrrolo[2,3-b]pyridin-3-yl)propyl)-1H-1,2,3-triazol-4-yl)-1-(phenylsulfonyl)-1H-pyrrolo[2,3-b]pyridine (**32**)

¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 12.19 (br, 1H), 8.56 (s, 1H), 8.37 (dd, *J* = 1.5, 4.8 Hz, 1H), 8.26–8.15 (br, 1H), 8.12–8.05 (m, 3H), 7.95 (d, *J* = 7.8 Hz, 1H), 7.79 (s, 1H), 7.70 (dd, *J* = 7.5 Hz, 1H), 7.66–7.56 (m, 2H), 7.32 (dd, *J* = 4.8, 7.8 Hz, 1H), 7.07 (dd, *J* = 4.8, 7.8 Hz, 1H), 6.83 (s, 1H), 4.52 (t, *J* = 7.0 Hz, 2H), 2.79–2.70 (m, 2H), 2.35–2.22 (m, 2H). ¹³C NMR (75 MHz, DMSO-d₆) δ (ppm): 147.4, 145.2 (2C), 143.2, 140.7, 138.1, 137.1, 135.1, 130.6, 130.0 (2C), 129.2, 128.3, 127.8 (2C), 123.8, 123.0, 122.2, 119.6, 119.0 (2C), 97.3, 49.6, 29.5, 21.7. HRMS-ESI (*m/z*): found 484.15571, calcd for C₂₅H₂₂N₇O₂S [M + H]⁺ 484.15502.

4.3.15. 3-(3-(4-(1H-Pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)propyl)-1H-pyrrolo[2,3-b]pyridine (**33**)

¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 12.18 (br, 1H), 11.41 (br, 1H), 8.55 (s, 1H), 8.23–8.17 (m, 2H), 7.97 (dd, *J* = 1.2, 3.5 Hz, 1H), 7.94 (dd, *J* = 1.2, 3.5 Hz, 1H), 7.33 (sd, *J* = 2.3 Hz, 1H), 7.07 (dd, *J* = 4.8, 7.9 Hz, 1H), 7.04 (dd, *J* = 4.8, 7.9 Hz, 1H), 6.83 (sd, *J* = 2.0 Hz, 1H), 4.50 (t, *J* = 7.0 Hz, 2H), 2.74 (t, *J* = 7.2 Hz, 2H), 2.26 (td, *J* = 7.0, 7.2 Hz, 2H). ¹³C NMR (75 MHz, DMSO-d₆) δ (ppm): 149.5, 149.2, 143.2, 142.9, 140.6, 130.6, 128.3, 127.0, 123.6, 122.2, 121.0, 119.7, 116.5, 115.3, 112.4, 97.3, 49.8, 30.7, 22.1. HRMS-ESI (*m/z*): found 344.16199, calcd for C₁₉H₁₈N₇ [M + H]⁺ 344.16182.

4.3.16. 3-(4-(1H-Pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)-1-(1H-pyrrolo[2,3-b]pyridin-3-yl)propan-1-one (**34**)

The reaction mixture was cooled to room temperature and filtered. The filtrate was concentrated and triturated with cyclohexane/EtOAc. The two solids were gathered and washed with water and diethylether to give 78 mg of the expected product as a brown solid in 57% yield. ¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 12.57 (br, 1H), 12.18 (br, 1H), 8.58 (d, *J* = 2.8 Hz, 1H), 8.55 (s, 1H), 8.48 (dd, *J* = 0.8, 7.8 Hz, 1H), 8.40–8.27 (br, 1H), 8.25–8.10 (br, 1H), 7.93

(d, *J* = 7.8 Hz, 1H), 7.26 (dd, *J* = 4.8, 7.8 Hz, 1H), 7.13–6.99 (m, 1H), 6.81 (s, 1H), 4.83 (t, *J* = 6.4 Hz, 2H), 3.65 (t, *J* = 6.4 Hz, 2H). ¹³C NMR (75 MHz, DMSO-d₆) δ (ppm): 192.4, 149.4, 144.8, 140.4, 135.7, 135.3, 130.6, 130.0, 122.6, 118.9, 115.1, 97.3, 45.5, 38.8. HRMS-ESI (*m/z*): found 358.14135, calcd for C₁₉H₁₆N₇O [M + H]⁺ 358.14108.

4.3.17. 3-((4-(1H-Pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)methyl)-1-(phenylsulfonyl)-1H-pyrrolo[2,3-b]pyridine (**35**)

The reaction mixture was cooled to room temperature and filtered. The solid was washed with MeOH and diethylether, and dried to give 72 mg of the expected product as a beige solid in 89% yield. ¹H NMR (300 MHz, DMSO-d₆) δ (ppm): 12.09 (br, 1H), 8.54 (s, 1H), 8.41 (dd, *J* = 1.2, 4.6 Hz, 1H), 8.23 (s, 1H), 8.21–8.11 (m, 3H), 8.03 (dd, *J* = 1.4, 7.9 Hz, 1H), 7.93 (d, *J* = 7.9 Hz, 1H), 7.74 (dd, *J* = 7.4 Hz, 1H), 7.64 (dd, *J* = 7.4, 7.8 Hz, 2H), 7.35 (dd, *J* = 4.8, 7.9 Hz, 1H), 7.05 (dd, *J* = 4.8, 7.9 Hz, 1H), 6.82 (sd, *J* = 1.8 Hz, 1H), 5.90 (s, 2H). ¹³C NMR (75 MHz, DMSO-d₆) δ (ppm): 149.5, 147.1, 145.8, 143.3, 140.9, 137.8, 135.4, 130.3, 130.2 (2C), 129.4, 128.4, 128.1 (2C), 126.8, 122.1, 121.7, 120.9, 120.0, 116.5, 114.2, 97.5, 44.9. HRMS-ESI (*m/z*): found 456.12450, calcd for C₂₃H₁₈N₇O₂S [M + H]⁺ 456.12372.

4.3.18. 2-((4-(1H-Pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)methyl)-1H-pyrrolo[2,3-b]pyridine (**36**)

The reaction mixture was cooled to room temperature and concentrated. The residue was purified by chromatography on silica gel column, CH₂Cl₂/MeOH 95:5, to give 97 mg of a light brown solid. The solid was triturated in MeOH and diethylether, filtered and dried. 64 mg of the expected product was obtained as a light brown solid in 53% yield. ¹H NMR (300 MHz, DMSO) δ (ppm): 12.14 (br, 1H), 11.98 (br, 1H), 8.51 (s, 1H), 8.30–8.12 (br, 2H), 8.00–7.90 (m, 2H), 7.13–7.02 (m, 2H), 6.85 (s, 1H), 6.55 (s, 1H), 5.85 (s, 2H). ¹³C NMR (75 MHz, DMSO) δ (ppm): 149.4, 149.3, 143.6, 143.2, 140.7, 133.7, 130.4, 128.9, 128.3, 122.2 (2C), 116.8, 116.7, 116.5, 101.1, 97.4, 47.4. HRMS-ESI (*m/z*): found 316.13073, calcd for C₁₇H₁₄N₇ [M + H]⁺ 316.13052.

4.4. Synthesis of compounds 38–44. General procedure for the ruthenium-catalyzed [3 + 2] cycloaddition reaction

4.4.1. 3-(2-(5-(1H-pyrrolo[2,3-b]pyridin-3-yl)-1H-1,2,3-triazol-1-yl)ethyl)-1H-pyrrolo[2,3-b]pyridine (**37**)

In a heat dried and nitrogen purged round bottom flask, **11a** (64 mg, 0.34 mmol), **4a** (61 mg, 0.43 mmol), Cp^{*}RuCl(PPh₃)₂ (20 mg, 0.03 mmol) and THF (4 mL) were mixed and stirred under reflux for 60 h. The reaction mixture was cooled to room temperature, concentrated and purified by chromatography on silica gel column, CH₂Cl₂/MeOH 95:5–9:1, to give 46 mg of the expected product as an orange solid in 41% yield. ¹H NMR (300 MHz, DMSO) δ (ppm): 12.22 (br, 1H), 11.36 (br, 1H), 8.31 (dd, *J* = 1.5, 4.7 Hz, 1H), 8.12 (dd, *J* = 1.5, 4.7 Hz, 1H), 7.93 (s, 1H), 7.87 (dd, *J* = 1.2, 7.8 Hz, 1H), 7.68 (d, *J* = 2.7 Hz, 1H), 7.50 (dd, *J* = 1.2, 7.8 Hz, 1H), 7.11–7.15 (m, 2H), 6.89 (dd, *J* = 4.7, 7.8 Hz, 1H), 4.70 (t, *J* = 7.2 Hz, 2H), 3.20 (t, *J* = 7.2 Hz, 2H). ¹³C NMR (75 MHz, DMSO) δ (ppm): 148.8, 148.7, 144.2, 142.9, 132.8, 131.6, 127.6, 126.4, 126.1, 124.2, 119.5, 118.3, 116.9, 115.3, 109.5, 100.0, 49.0, 25.9. HRMS-ESI (*m/z*): found 330.14622, calcd for C₁₈H₁₆N₇ [M + H]⁺ 330.14617.

Compounds **38–44** were synthesized following the general procedure for the ruthenium-catalysed [3 + 2] cycloaddition reaction.

4.4.2. 2-(5-(1H-Pyrrolo[2,3-b]pyridin-3-yl)-1H-1,2,3-triazol-1-yl)-1-(1H-pyrrolo[2,3-b]pyridin-3-yl)ethan-1-one (**38**)

The reaction mixture was cooled to room temperature and filtered. The solid was washed with CH₂Cl₂, EtOAc and diethylether, and dried. 106 mg of the expected product was obtained as a brown

solid in 90% yield. ^1H NMR (300 MHz, DMSO) δ (ppm): 12.79 (br, 1H), 12.13 (br, 1H), 8.75 (sd, $J = 2.0$ Hz, 1H), 8.38 (dd, $J = 1.5, 4.2$ Hz, 1H), 8.36 (s, 1H), 8.31 (dd, $J = 1.2, 4.6$ Hz, 1H), 8.15 (s, 1H), 8.13 (d, $J = 7.8$ Hz, 1H), 7.69 (sd, $J = 2.6$ Hz, 1H), 7.26 (dd, $J = 5.1, 7.5$ Hz, 1H), 7.19 (dd, $J = 4.7, 7.8$ Hz, 1H), 6.12 (s, 2H). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 187.4, 149.4, 148.7, 145.2, 144.3, 135.9, 132.8, 132.1, 129.9, 128.0, 125.4, 119.0, 118.2, 118.0, 117.0, 112.7, 100.2, 54.4. HRMS-ESI (m/z): found 344.12590, calcd for $\text{C}_{18}\text{H}_{15}\text{N}_7\text{O}$ [$\text{M} + \text{H}$] $^+$ 344.12543.

4.4.3. 3-(3-(5-(1H-Pyrrolo[2,3-b]pyridin-3-yl)-1H-1,2,3-triazol-1-yl)propyl)-1H-pyrrolo[2,3-b]pyridine (39)

The reaction mixture was cooled to room temperature and concentrated. The residu was purified by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 95:5–9:1, to give 80 mg of the expected product as an orange solid in 73% yield. ^1H NMR (300 MHz, DMSO) δ (ppm): 12.29 (br, 1H), 11.30 (br, 1H), 8.34 (dd, $J = 1.5, 4.7$ Hz, 1H), 8.14 (dd, $J = 1.5, 4.7$ Hz, 1H), 8.03 (dd, $J = 1.5, 7.8$ Hz, 1H), 8.01 (s, 1H), 7.91 (sd, $J = 2.7$ Hz, 1H), 7.72 (dd, $J = 1.2, 7.8$ Hz, 1H), 7.16 (dd, $J = 4.7, 7.8$ Hz, 1H), 7.16 (s, 1H), 6.93 (dd, $J = 4.7, 7.8$ Hz, 1H), 4.52 (t, $J = 7.0$ Hz, 2H), 2.63 (t, $J = 7.4$ Hz, 2H), 2.12 (td, $J = 7.0, 7.4$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 149.1, 148.8, 144.3, 142.8, 132.8, 131.5, 127.7, 126.7, 126.3, 123.3, 119.5, 118.3, 117.0, 115.2, 112.5, 100.0, 48.1, 30.1, 22.1. HRMS-ESI (m/z): found 344.16184, calcd for $\text{C}_{19}\text{H}_{18}\text{N}_7$ [$\text{M} + \text{H}$] $^+$ 344.16182.

4.4.4. 3-(5-(1H-Pyrrolo[2,3-b]pyridin-3-yl)-1H-1,2,3-triazol-1-yl)-1-(1H-pyrrolo [2,3-b]pyridin-3-yl)propan-1-one (40)

The reaction mixture was cooled to room temperature and concentrated. The residu was purified by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 95:5–9:1, to give 24 mg of the expected product as a yellow solid in 18% yield. ^1H NMR (300 MHz, DMSO) δ (ppm): 12.53 (br, 1H), 12.33 (br, 1H), 8.52 (s, 1H), 8.40 (dd, $J = 1.4, 8.0$ Hz, 1H), 8.34 (dd, $J = 1.1, 4.6$ Hz, 1H), 8.31 (dd, $J = 1.2, 4.7$ Hz, 1H), 8.08 (dd, $J = 1.0, 7.8$ Hz, 1H), 8.05 (sd, $J = 2.5$ Hz, 1H), 8.00 (s, 1H), 7.23 (dd, $J = 4.8, 7.9$ Hz, 1H), 7.19 (dd, $J = 4.8, 8.0$ Hz, 1H), 4.78 (t, $J = 6.7$ Hz, 2H), 3.64 (t, $J = 6.7$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 192.5, 149.4, 148.8, 144.7, 144.3, 135.2, 132.5, 131.8, 130.0, 127.8, 126.4, 118.6, 118.4, 118.0, 117.0, 115.1, 100.1, 44.0, 38.1. HRMS-ESI (m/z): found 358.14135, calcd for $\text{C}_{19}\text{H}_{16}\text{N}_7\text{O}$ [$\text{M} + \text{H}$] $^+$ 358.14108.

4.4.5. 3-(2-(5-(1H-Pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)ethyl)-1H-pyrrolo[2,3-b]pyridine (41)

The reaction mixture was cooled to room temperature and concentrated. The residu was purified by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 95:5, to give 64 mg of the expected product as a brown solid in 62% yield. ^1H NMR (300 MHz, DMSO) δ (ppm): 12.30 (br, 1H), 11.41 (br, 1H), 8.31 (dd, $J = 1.6, 4.7$ Hz, 1H), 8.14 (dd, $J = 1.6, 4.7$ Hz, 1H), 8.13 (s, 1H), 8.00 (dd, $J = 1.6, 7.9$ Hz, 1H), 7.73 (dd, $J = 1.2, 7.9$ Hz, 1H), 7.21 (sd, $J = 2.4$ Hz, 1H), 7.14 (dd, $J = 4.7, 7.9$ Hz, 1H), 6.94 (dd, $J = 4.6, 7.9$ Hz, 1H), 6.77 (s, 1H), 4.89 (t, $J = 7.2$ Hz, 2H), 3.26 (t, $J = 7.2$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 149.5, 148.9, 144.7, 143.0, 133.5, 130.4, 129.2, 126.6, 124.7, 124.3, 120.6, 119.6, 116.9, 115.4, 109.2, 101.1, 49.8, 25.6. HRMS-ESI (m/z): found 330.14608, calcd for $\text{C}_{18}\text{H}_{16}\text{N}_7$ [$\text{M} + \text{H}$] $^+$ 330.14617.

4.4.6. 2-(5-(1H-Pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)-1-(1H-pyrrolo [2,3-b]pyridin-3-yl)ethan-1-one (42)

The reaction mixture was cooled to room temperature and filtered. The solid was washed with CH_2Cl_2 , EtOAc and diethylether, and dried. 113 mg of the expected product was obtained as a brown solid 97% yield. ^1H NMR (300 MHz, DMSO) δ (ppm): 12.83 (br, 1H), 12.37 (br, 1H), 8.84 (sd, $J = 2.8$ Hz, 1H), 8.37 (s, 1H), 8.35 (dd, $J = 1.6, 4.7$ Hz, 1H), 8.29 (s, 1H), 8.27 (dd, $J = 1.6, 4.7$ Hz, 1H), 7.91 (dd, $J = 1.5, 7.9$ Hz, 1H), 7.26 (dd, $J = 5.0, 7.7$ Hz, 1H), 7.05 (dd, $J = 4.7, 7.9$ Hz, 1H),

6.71 (sd, $J = 2.0$ Hz, 1H), 6.31 (s, 2H). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 186.6, 149.4, 149.3, 145.2, 144.8, 136.0, 133.1, 131.8, 129.9, 129.3, 124.9, 120.5, 119.0, 118.0, 116.8, 112.5, 100.2, 55.3. HRMS-ESI (m/z): found 344.12580, calcd for $\text{C}_{18}\text{H}_{14}\text{N}_7\text{O}$ [$\text{M} + \text{H}$] $^+$ 344.12543.

4.4.7. 3-(3-(5-(1H-Pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)propyl)-1H-pyrrolo[2,3-b]pyridine (43)

The reaction mixture was cooled to room temperature and concentrated. The residu was purified by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 95:5, to give 20 mg of the expected product as a brown solid in 18% yield. ^1H NMR (300 MHz, DMSO) δ (ppm): 12.30 (br, 1H), 11.41 (br, 1H), 8.29 (dd, $J = 1.6, 4.7$ Hz, 1H), 8.21–8.17 (m, 2H), 7.91–7.83 (m, 2H), 7.29 (sd, $J = 2.4$ Hz, 1H), 7.12 (dd, $J = 4.7, 7.9$ Hz, 1H), 6.98 (dd, $J = 4.7, 7.9$ Hz, 1H), 6.54 (dd, $J = 2.0$ Hz, 1H), 4.65 (t, $J = 7.3$ Hz, 2H), 2.78 (t, $J = 7.3$ Hz, 2H), 2.19 (td, $J = 7.3$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 149.4, 149.2, 144.8, 142.9, 133.3, 130.2, 129.17, 126.9, 124.7, 123.8, 120.5, 119.7, 116.8, 115.3, 112.3, 100.7, 48.7, 30.1, 22.0. HRMS-ESI (m/z): found 344.16226, calcd for $\text{C}_{19}\text{H}_{18}\text{N}_7$ [$\text{M} + \text{H}$] $^+$ 344.16182.

4.4.8. 3-(5-(1H-Pyrrolo[2,3-b]pyridin-2-yl)-1H-1,2,3-triazol-1-yl)-1-(1H-pyrrolo [2,3-b]pyridin-3-yl)propan-1-one (44)

The reaction mixture was cooled to room temperature and concentrated. The residu was purified by chromatography on silica gel column, $\text{CH}_2\text{Cl}_2/\text{MeOH}$ 95:5, to give 52 mg of the expected product as a beige solid in 34% yield. ^1H NMR (300 MHz, DMSO) δ (ppm): 12.55 (br, 1H), 12.36 (br, 1H), 8.56 (s, 1H), 8.45 (d, $J = 7.9$ Hz, 1H), 8.36–8.30 (m, 2H), 8.17 (s, 1H), 8.06 (d, $J = 7.9$ Hz, 1H), 7.24 (dd, $J = 4.8, 7.9$ Hz, 1H), 7.15 (dd, $J = 4.8, 7.9$ Hz, 1H), 7.02 (s, 1H), 4.95 (t, $J = 6.6$ Hz, 2H), 3.70 (t, $J = 6.6$ Hz, 2H). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 192.3, 149.5, 149.4, 144.8, 135.3, 133.3, 130.7, 130.0, 129.3, 124.8, 120.7, 118.6, 118.1, 117.9, 116.9, 115.0, 101.3, 44.8, 37.8. HRMS-ESI (m/z): found 358.14136, calcd for $\text{C}_{19}\text{H}_{16}\text{N}_7\text{O}$ [$\text{M} + \text{H}$] $^+$ 358.14108.

4.5. Cell maintenance and cytotoxicity assays

The cytotoxicity of the studied compounds was determined using cell lines of different histological origin as described earlier [46]. Briefly, compounds in three-fold dilutions were added to the cells in triplicate. The treatment lasted for 72 h, after which Calcein AM solution was added, and the fluorescence of live cells at 485 nm/538 nm (excitation/emission) was measured with a Fluoroskan Ascent microplate reader (Labsystems). IC_{50} (the drug concentration that reduced the number of viable cells to 50%) values were determined from the dose–response curves.

4.6. One-step cellular caspase-3/7 activity assay

The activity of cellular caspases-3/7 was measured according to published procedures [45]. Briefly, G361 cells were incubated in a density of 20,000 cells/well in a 96-well plate overnight. Next day, the compounds were added (the final concentration reached 2 μM) and the cells were incubated for the next 24 h. After incubation, 3x caspase-3/7 assay buffer (150 mM HEPES pH 7.4, 450 mM NaCl, 150 mM KCl, 30 mM MgCl_2 , 1.2 mM EGTA, 1.5% Nonidet P40, 0.3% CHAPS, 30% sucrose, 30 mM DTT, 3 mM PMSF) with 37.5 μM Ac-DEVD-AMC as a substrate (Sigma–Aldrich) was added to the wells and plates were incubated at 37 $^\circ\text{C}$. The caspase-3/7 activity was measured after 4 h using Fluoroskan Ascent microplate reader (Labsystems) at 346 nm/442 nm (excitation/emission).

4.7. Immunoblotting and antibodies

Immunoblotting analysis was performed as described earlier

[46]. Briefly, cellular lysates were prepared by harvesting cells in Laemmli sample buffer. Proteins were separated on SDS-polyacrylamide gels and electroblotted onto nitrocellulose membranes. After blocking, the membranes were incubated with specific primary antibodies overnight, washed and then incubated with peroxidase-conjugated secondary antibodies. Finally, peroxidase activity was detected with ECL + reagents (AP Biotech) using a CCD camera LAS-4000 (Fujifilm). Specific antibodies were purchased from Sigma–Aldrich (anti- α -tubulin, clone DM1A; peroxidase-labeled secondary antibodies), Santa Cruz Biotechnology (anti-cyclin B, clone GNS1; anti-Mcl-1, clone S-19; anti-PARP, clone F-2), Cell Signaling (anti-cyclin A, clone BF683; anti-Rb, clone 4H1; anti-pRb antibodies phosphorylated at S780 and S807/811; anti-pPP1 α (T320); anti-PP1 α ; anti-pNPM (T199); anti-NPM; anti-caspase-7; anti-caspase-3, clone 3G2; anti-Bax, clone D2E11; anti-Bak, clone D4E4; anti-Bid), eBioscience (anti-5-bromo-2'-deoxyuridine-fluorescein, clone BU20A) or were a generous gift from prof. Bártek from Danish Cancer Society Research Center, Copenhagen, Denmark (anti-cyclin D1).

4.8. Kinase inhibition assays

CDK2/Cyclin E and CDK1/Cyclin B kinases were produced in Sf9 insect cells via baculoviral infection and purified on a Ni-NTA column (Qiagen). CDK4/Cyclin D1, CDK5/p35NCK, CDK7/Cyclin H/MAT1 and CDK9/Cyclin T1 were purchased from ProQinase GmbH. The kinases were assayed with 1 mg/mL histone H1 (for CDK1/2/5) or (YSPTSPS)₂ KK peptide (for CDK7/CDK9) or RPPTLSPIPIHR peptide (for CDK4) in the presence of 15/15/15/0.15/1.5/1.5 μ M ATP (for CDK1/2/4/5/7/9), 0.05 μ Ci [γ -³³P]ATP and of the test compound in a final volume of 10 μ L, all in a reaction buffer (60 mM HEPES-NaOH, pH 7.5, 3 mM MgCl₂, 3 mM MnCl₂, 3 μ M Na-orthovanadate, 1.2 mM DTT, 2.5 μ g/50 μ L PEG_{20,000}). The reactions were stopped by adding 5 μ L of 3% aq. H₃PO₄. Aliquots were spotted onto P-81 phosphocellulose (Whatman), washed 3 times with 0.5% aq. H₃PO₄ and finally air-dried. Kinase inhibition was quantified using a FLA-7000 digital image analyzer (Fujifilm). The concentration of the test compounds required to decrease the CDK activity by 50% was determined from dose–response curves and designated as IC₅₀.

4.9. Cell cycle analysis

The cultures were pulse-labeled with 10 μ M 5-bromo-2'-deoxyuridine (BrdU) for 30 min at 37 °C prior to harvesting. The cells were then washed in PBS, fixed with 70% ethanol, and denatured in 2 M HCl. Following neutralization, the cells were stained with anti-BrdU fluorescein-labeled antibodies, washed, stained with propidium iodide and analyzed by flow cytometry using a 488 nm laser (Cell Lab Quanta SC, Beckman Coulter) as described previously [46].

4.10. Caspase activity assay

The cells were homogenized in an extraction buffer (10 mM KCl, 5 mM HEPES, 1 mM EDTA, 1 mM EGTA, 0.2% CHAPS, inhibitors of proteases, pH 7.4) on ice for 20 min. The homogenates were clarified by centrifugation at 10 000 \times g for 30 min at 4 °C, and then the proteins were quantified and diluted to equal concentrations. Lysates were then incubated for 3 h with 100 μ M Ac-DEVD-AMC as a substrate of caspases 3 and 7 in the assay buffer (25 mM PIPES, 2 mM EGTA, 2 mM MgCl₂, 5 mM DTT, pH 7.3). The fluorescence of the product was measured using a Fluoroskan Ascent microplate reader (Labsystems) at 355/460 nm (excitation/emission) as described previously [46].

4.11. Molecular docking study

3D structures of compounds were prepared with Marvin, a software used for drawing, displaying, and characterizing chemical structures, substructures, and reactions [Marvin 14.7.7, 2014, ChemAxon (<http://www.chemaxon.com>)]. 3D structures of ligands were optimized and all hydrogens were added within the MarvinSketch 14.7.7 program. All nonaromatic and nonring bonds were set as rotatable within AutoDock Tools 1.5.4 program [47]. The crystal structure for CDK2 with roscovitin (PDBID: 2A4L) was used as the protein docking template with a docking grid of 16 Å around the center of the ligand in the crystal structure, which was deleted prior to docking. Polar hydrogens were added to receptor or selected for all ligands with the AutoDock Tools program prior to docking with the Autodock Vina 1.1.2 program [48].

Author contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Acknowledgments

This work was supported by the European funds for the economical and regional development (FEDER), the Laboratory of excellence Labex SYNORG, ANR-11-LABX-0029 and the Medical Research Fondation (FRM: SPF20130526815). The authors also gratefully acknowledge support through the projects LO1305 and LO1204 from the Ministry of Education, Youth and Sports of the Czech Republic. The work was also supported by the Czech Science Foundation (grant 15-17282Y) and by student projects IGA_PrF_2015_027 and IGA_PrF_2015_024 of Palacky University in Olomouc.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.ejmech.2015.12.023>.

References

- [1] M. Malumbres, M. Barbacid, Mammalian cyclin-dependent kinases, *Trends Biochem. Sci.* 30 (2005) 630–641.
- [2] A.M. Senderowicz, Small molecule modulators of cyclin-dependent kinases for cancer therapy, *Oncogene* 19 (2000) 6600–6606.
- [3] S. Lim, P. Kaldis, Cdk, cyclins and CKIs: roles beyond cell cycle regulation, *Development* 140 (2013) 3079–3093.
- [4] A. Deshpande, P. Sicinski, P.W. Hinds, Cyclins and Cdk in development and cancer: a perspective, *Oncogene* 24 (2005) 2909–2915.
- [5] J.S. Blachly, J.C. Byrd, Emerging drug profile: cyclin-dependent kinase inhibitors, *Leukemia Lymphoma* 54 (2013) 2133–2143.
- [6] U. Asghar, A.K. Witkiewicz, N.C. Turner, E.S. Knudsen, The history and future of targeting cyclin-dependent kinases in cancer therapy, *Nat. Rev. Drug Discov.* 14 (2015) 130–146.
- [7] N. Kwiatkowski, T. Zhang, P.B. Rahl, B.J. Abraham, J. Reddy, S.B. Ficarro, A. Dastur, A. Amzallag, S. Ramaswamy, B. Tesar, C.E. Jenkins, N.M. Hannett, D. McMillin, T. Sanda, T. Sim, N.D. Kim, T. Look, C.S. Mitsiades, A.P. Weng, J.R. Brown, C.H. Benes, J.A. Marto, R.A. Young, N.S. Gray, Targeting transcription regulation in cancer with a covalent CDK7 inhibitor, *Nature* 511 (2014) 616–620.
- [8] D.W. Fry, P.J. Harvey, P.R. Keller, W.L. Elliott, M. Meade, E. Trachet, M. Albassam, X. Zheng, W.R. Leopold, N.K. Pryer, P.L. Toogood, Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated anti-tumor activity in human tumor xenografts, *Mol. Cancer Ther.* 3 (2004) 1427–1438.
- [9] D. Cai, V.M. Latham Jr., X. Zhang, G.I. Shapiro, Combined depletion of cell cycle and transcriptional cyclin-dependent kinase activities induces apoptosis in cancer cells, *Cancer Res.* 66 (2006) 9270–9280.
- [10] A. Echalié, E. Cot, A. Camasses, E. Hodimont, F. Hoh, P. Jay, F. Sheinerman, L. Krasinska, D. Fisher, An integrated chemical biology approach provides insight into Cdk2 functional redundancy and inhibitor sensitivity, *Chem. Biol.*

- 19 (2012) 1028–1040.
- [11] R. Soni, L. Muller, P. Furet, J. Schoepfer, C. Stephan, S. Zumstein-Mecker, H. Fretz, B. Chaudhuri, Inhibition of cyclin-dependent kinase 4 (Cdk4) by faspaplysin, a marine natural product, *Biochem. Biophys. Res. Commun.* 275 (2000) 877–884.
- [12] P.R. Jenkins, J. Wilson, D. Emmerson, M.D. Garcia, M.R. Smith, S.J. Gray, R.G. Britton, S. Mahale, B. Chaudhuri, Design, synthesis and biological evaluation of new tryptamine and tetrahydro- β -carboline-based selective inhibitors of CDK4, *Bioorg. Med. Chem.* 16 (2008) 7728–7739.
- [13] C. Aubry, A.J. Wilson, D. Emmerson, E. Murphy, Y.Y. Chan, M.P. Dickens, M.D. Garcia, P.R. Jenkins, S. Mahale, B. Chaudhuri, Faspaplysin-inspired diindolyls as selective inhibitors of CDK4/cyclin D1, *Bioorg. Med. Chem.* 17 (2009) 6073–6084.
- [14] S. Mahale, C. Aubry, A.J. Wilson, P.R. Jenkins, J.-D. Maréchal, M.J. Sutcliffe, B. Chaudhuri, CA224, a non-planar analogue of faspaplysin, inhibits Cdk4 but not Cdk2 and arrests cells at G0/G1 inhibiting pRB phosphorylation, *Bioorg. Med. Chem. Lett.* 16 (2006) 4272–4278.
- [15] A. Hörmann, B. Chaudhuri, H. Fretz, DNA binding properties of the marine sponge pigment faspaplysin, *Bioorg. Med. Chem.* 9 (2001) 917–921.
- [16] CDK4/cyclin D inhibitory activities of 21 and 22 were performed by the Cerep company. The results are presented in the supporting informations.
- [17] R. Huisgen, 1,3-Dipolar cycloadditions. Past and future, *Angew. Chem. Int. Ed.* 2 (1963) 565–632.
- [18] K. Sonogashira, Y. Tohda, N. Hagihara, A convenient synthesis of acetylenes: catalytic substitutions of acetylenic hydrogen with bromoalkenes, iodoarenes and bromopyridines, *Tetrahedron Lett.* 16 (1975) 4467–4470.
- [19] S. Thorand, N. Krause, Improved procedures for the palladium-catalyzed coupling of terminal Alkynes with aryl bromides (sonogashira coupling), *J. Org. Chem.* 63 (1998) 8551–8553.
- [20] R. Chinchilla, C. Nájera, The sonogashira reaction: a booming methodology in synthetic organic chemistry, *Chem. Rev.* 107 (2007) 874–922.
- [21] C. Chaulet, C. Croix, D. Alagille, S. Normand, A. Delwail, L. Favot, J.-C. Lecron, M.-C. Viaud-Massuard, Design, synthesis and biological evaluation of new thalidomide analogues as TNF- α and IL-6 production Inhibitors, *Bioorg. Med. Chem. Lett.* 21 (2011) 1019–1022.
- [22] H. Naka, Y. Akagi, K. Yamada, T. Imahori, T. Kasahara, Y. Kondo, Fluorous synthesis of yuehchukene by α -lithiation of perfluoroalkyl-tagged 1-(aryl-sulfonyl)indole with mesityllithium, *Eur. J. Org. Chem.* (2007) 4635–4637.
- [23] T. Greene, P.G.M. Wuts, *Protective Groups in Organic Synthesis*, third ed., John Wiley & Sons, New York, 1991.
- [24] C. Galvez, P. Viladoms, Reactivity of 1H-pyrrolo[2,3-b]pyridine. II. Synthesis of 3-(β -haloethyl)-7-azaindole, *J. Heterocycl. Chem.* 21 (1984) 421–425.
- [25] J. Blum, D. Gelman, W. Baidossi, E. Shakh, A. Rosenfeld, Z. Aizenshtat, Palladium-catalyzed methylation of Aryl and vinyl halides by stabilized methyl-aluminum and methylgallium complexes, *J. Org. Chem.* 62 (1997) 8681–8686.
- [26] N. Joubert, R. Pohl, B. Klepetářová, M. Hocek, Modular and practical synthesis of 6-substituted pyridin-3-yl C-nucleosides, *J. Org. Chem.* 72 (2007) 6797–6805.
- [27] R. Liu, P. Zhang, T. Gan, J.M. Cook, Regiospecific bromination of 3-methylindoles with NBS and its application to the concise synthesis of optically active unusual tryptophans present in marine cyclic peptides¹, *J. Org. Chem.* 62 (1997) 7447–7456.
- [28] V. D'Elia, L. Yinghao, H. Zipse, Immobilized DMAP derivatives rivaling homogeneous DMAP, *Eur. J. Org. Chem.* (2011) 1527–1533.
- [29] V. Aucagne, D.A. Leigh, Chemoselective formation of successive triazole linkages in one pot: “click–click” chemistry, *Org. Lett.* 8 (2006) 4505–4507.
- [30] Regarding entries 11 and 13, it should be noted that even if compounds 30 and 33 could have been obtained directly by cycloaddition between the azide derivatives 11a or 11b and the alkyne derivative 4b, we did not attempt these reactions since 30 and 33 were obtained in a good yield for two steps and in enough quantity for their characterization and biological evaluation.
- [31] L. Zhang, X. Chen, P. Xue, H.H.Y. Sun, I.D. Williams, K.B. Sharpless, V.V. Fokin, G. Jia, Ruthenium-catalyzed cycloaddition of alkynes and organic azides, *J. Am. Chem. Soc.* 127 (2005) 15998–15999.
- [32] R.D. Cramer, D.E. Patterson, J.D. Bunce, Comparative molecular field analysis (CoMFA). 1. effect of shape on binding of steroids to carrier proteins, *J. Am. Chem. Soc.* 110 (1988) 5959–5967.
- [33] G. Klebe, U.J. Abraham, Comparative molecular similarity index analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries, *J. Comput. Aided Mol. Des.* 13 (1999) 1–10.
- [34] The references for the molecules used in the training and the test sets can be found in the supplementary data.
- [35] TRIPOS, 1699 South Hanley Road, St. Louis, MO 63144, USA.
- [36] M.C. Clark, R.D. Cramer III, N. van Opden Bosch, Validation of the general purpose tripos 5.2 force field, *J. Comput. Chem.* 10 (1989) 982–1012.
- [37] S.J. Cho, A. Tropsha, Cross-validated R2-guided region selection for comparative molecular field analysis: a simple method to achieve consistent results, *J. Med. Chem.* 38 (1995) 1060–1066.
- [38] B. Kramer, M. Rarey, T. Lengauer, Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking, *Proteins Struct. Funct. Genet.* 37 (1999) 228–241.
- [39] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges, *Tetrahedron* 36 (1980) 3219–3288.
- [40] B.L. Bush, R.B. Nachbar Jr., Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA, *J. Comput. Aided Mol. Des.* 7 (1993) 587–619.
- [41] A.R. Leach, *Molecular Modelling: Principles and Applications*, Henry Ling, London, 2001, p. 695.
- [42] T. Gucky, R. Jorda, M. Zatloukal, V. Bazgier, K. Berka, E. Reznickova, T. Beres, M. Strnad, V. Krystof, A novel series of highly potent 2,6,9-trisubstituted purine cyclin-dependent kinase inhibitors, *J. Med. Chem.* 56 (2013) 6234–6247.
- [43] M. Payton, G. Chung, P. Yakowec, A. Wong, D. Powers, L. Xiong, N. Zhang, J. Leal, T.L. Bush, V. Santora, B. Askew, A. Tasker, R. Radinsky, R. Kendall, S. Coats, Discovery and evaluation of dual CDK1 and CDK2 inhibitors, *Cancer Res.* 66 (2006) 4299–4308.
- [44] M. Zatloukal, R. Jorda, T. Gucky, E. Reznickova, J. Voller, T. Pospisil, V. Malinkova, H. Adamcova, V. Krystof, M. Strnad, Synthesis in vitro biological evaluation of 2,6,9-trisubstituted purines targeting multiple cyclin-dependent kinases, *Eur. J. Med. Chem.* 61 (2013) 61–72.
- [45] R.A. Carrasco, N.B. Stamm, B.K.R. Patel, One-step cellular caspase-3/7 assay, *Biotechniques* 34 (2003) 1064–1067.
- [46] V. Krystof, I.W. McNae, M.D. Walkinshaw, P.M. Fischer, P. Muller, B. Vojtesek, M. Orsag, L. Havlicek, M. Strnad, Antiproliferative activity of olomoucine II, a novel 2,6,9-trisubstituted purine cyclin-dependent kinase inhibitor, *Cell Mol. Life Sci.* 62 (2005) 1763–1771.
- [47] M.F. Sanner, Python: a programming language for software integration and development, *J. Mol. Graph. Model.* 17 (1999) 57–61.
- [48] O. Trott, A.J. Olson, Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multi-threading, *J. Comput. Chem.* 31 (2010) 455–461.

Appendix H

Gucký T, Jorda R, Zatloukal M, **Bazgier V**, Berka K, Řezníčková E, Béres T, Strnad M, Kryštof V:

A Novel Series of Highly Potent 2,6,9-Trisubstituted Purine Cyclin-Dependent Kinase Inhibitors.

J. Med. Chem., 56(15), 6234-6247, 2013.

DOI: 10.1021/jm4006884

IF=6.259

Novel thidiazuron-derived inhibitors of cytokinin oxidase/dehydrogenase

Jaroslav Nisler^{1,2,3} · David Kopečný⁴ · Radka Končítíková⁴ · Marek Zatloukal^{1,2} · Václav Bazgier^{1,5} · Karel Berka^{5,6} · David Zalabák⁷ · Pierre Briozzo⁸ · Miroslav Strnad¹ · Lukáš Spíchal^{1,2}

Received: 21 April 2016 / Accepted: 29 June 2016 / Published online: 15 July 2016
© Springer Science+Business Media Dordrecht 2016

Abstract

Key message Two new TDZ derivatives (HETDZ and 3FMTDZ) are very potent inhibitors of CKX and are promising candidates for in vivo studies.

Abstract Cytokinin hormones regulate a wide range of essential processes in plants. Thidiazuron (*N*-phenyl-*N'*-1,2,3-thiadiazol-5-yl urea, TDZ), formerly registered as a cotton defoliant, is a well known inhibitor of cytokinin oxidase/dehydrogenase (CKX), an enzyme catalyzing the degradation of cytokinins. TDZ thus increases the lifetime

of cytokinins and their effects in plants. We used in silico modeling to design, synthesize and characterize twenty new TDZ derivatives with improved inhibitory properties. Two compounds, namely 1-[1,2,3]thiadiazol-5-yl-3-(3-trifluoromethoxy-phenyl)urea (3FMTDZ) and 1-[2-(2-hydroxyethyl)phenyl]-3-(1,2,3-thiadiazol-5-yl)urea (HETDZ), displayed up to 15-fold lower *IC*₅₀ values compared with TDZ for AtCKX2 from *Arabidopsis thaliana* and ZmCKX1 and ZmCKX4a from *Zea mays*. Binding modes of 3FMTDZ and HETDZ were analyzed by X-ray crystallography. Crystal structure complexes, solved at 2.0 Å resolution, revealed that HETDZ and 3FMTDZ bound differently in the active site of ZmCKX4a: the thiadiazolyl ring of 3FMTDZ was positioned over the isoalloxazine ring of FAD, whereas that of HETDZ had the opposite orientation, pointing toward

Electronic supplementary material The online version of this article (doi:10.1007/s11103-016-0509-0) contains supplementary material, which is available to authorized users.

✉ Jaroslav Nisler
jaroslav.nisler@gmail.com

✉ David Kopečný
david.kopecny@upol.cz

- ¹ Laboratory of Growth Regulators, Centre of the Region Haná for Biotechnological and Agricultural Research (CRH), Institute of Experimental Botany AS CR, Palacký University, Šlechtitelů 27, 783 71 Olomouc, Czech Republic
- ² Department of Chemical Biology and Genetics, Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacký University, Šlechtitelů 27, 783 71 Olomouc, Czech Republic
- ³ Department of Chemistry of Natural Compounds, Faculty of Food and Biochemical Technology, University of Chemistry and Technology in Prague, Technická 5, 166 28 Prague, Czech Republic
- ⁴ Department of Protein Biochemistry and Proteomics, Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacký University, Šlechtitelů 27, 783 71 Olomouc, Czech Republic

- ⁵ Department of Physical Chemistry, Faculty of Science, Palacký University, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic
- ⁶ Regional Centre of Advanced Technologies and Materials, Faculty of Science, Palacký University, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic
- ⁷ Department of Molecular Biology, Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacký University, Šlechtitelů 27, 783 71 Olomouc, Czech Republic
- ⁸ Institut Jean-Pierre Bourgin, INRA, AgroParisTech, Université Paris-Saclay, 78026 Versailles, France

the entrance of the active site. The compounds were further tested for cytokinin activity in several cytokinin bioassays. We suggest that the combination of simple synthesis, lowered cytokinin activity, and enhanced inhibitory effects on CKX isoforms, makes 3FMTDZ and HETDZ suitable candidates for *in vivo* studies.

Keywords Cytokinin oxidase/dehydrogenase · Crystal structure · Molecular docking · Organic synthesis · Thidiazuron · Cytokinin

Introduction

Cytokinins are hormones that influence a wide range of essential processes in plants, including cell division, shoot and root development, seed and fruit development, germination, senescence and response to environmental stresses (Mok and Mok 2001). They are adenine derivatives with a distinct N^6 side-chain. Nowadays, use of plant growth regulators with cytokinin-like activity is a common practice in agriculture and horticulture, and thus much effort has been devoted to the development of new compounds with improved cytokinin-like properties. Thidiazuron (TDZ) is one of the most widely used synthetic compounds in agriculture, showing a strong cytokinin effect based on a combination of two activities: TDZ possesses strong cytokinin activity (Mok et al. 1982; Yamada et al. 2001; Spichal et al. 2004) and also inhibits the enzyme cytokinin oxidase/dehydrogenase (abbreviated as CKO or CKX, EC 1.5.99.12, Chatfield and Armstrong 1986; Hare and Van Staden 1994). CKX catalyzes the irreversible oxidative breakdown of cytokinins to form adenine/adenosine and the corresponding aldehyde (Whitty and Hall 1974; Brownlee et al. 1975; Chatfield and Armstrong 1986). Therefore, in addition to a direct effect at the receptor level, TDZ increases endogenous cytokinin concentrations in plants.

CKXs are encoded by small multigene families. The model plant *Arabidopsis thaliana* contains seven CKX isoforms, which are involved in the regulation of endogenous cytokinin levels (Werner et al. 2003), of which AtCKX2 is the most active and well-studied isoform (Galuszka et al. 2007). In contrast, the model plant *Zea mays* contains 13 CKX isoforms, which differ in substrate specificity, spatial and temporal expression patterns and subcellular localization (Massonneau et al. 2004; Vyroubalová et al. 2009; Šmehilová et al. 2009; Zalabák et al. 2014). Among these, ZmCKX1 is the best studied isoform, which has been shown to play a crucial role in cytokinin degradation in maize (Houba-Hérin et al. 1999; Morris et al. 1999). ZmCKX1 preferentially oxidizes cytokinin bases and ribosides (R), with N^6 -(Δ^2 -isopentenyl)adenine (iP) being the best substrate (Bilyeu et al. 2001; Kopečný et al. 2005). In

contrast, other cytokinins, including cytokinin glucosides (G) and monophosphates (MP), have been reported to be poor substrates (Zalabák et al. 2014; Kopečný et al. 2015). ZmCKX2, 3, 4a, 4b and 5, expressed mainly in reproductive organs, have been shown to preferentially oxidize iPRMP and iP9G as substrates, followed by iP. Such a change in substrate specificity is caused by substitution of the Glu381 residue in ZmCKX1 by a serine, alanine or glycine residue in the other isoforms. Moreover, the crystal structures of ZmCKX2 and ZmCKX4a display a mobile domain, which is well defined in the ZmCKX1 structure (Kopečný et al. 2015). The domain includes two helices and a loop belonging to the substrate binding domain and likely contributes to binding of the N^9 substituted cytokinins. ZmCKX1, ZmCKX4a and AtCKX2 have been expressed in *Yarrowia lipolytica*, *Escherichia coli* and *Saccharomyces cerevisiae* (Kopečný et al. 2005; Zalabák et al. 2014; Frébortová et al. 2007), allowing production of large amounts of recombinant CKX.

To date, several groups of compounds are known to inhibit CKX. One such group includes urea compounds, such as diphenyl urea, TDZ and 1-phenyl-3-(2-chloropyridin-4-yl)urea (CPPU) (Chatfield and Armstrong 1986; Laloue and Fox 1989; Burch and Horgan 1989; Hare and van; Staden 1994). TDZ and CPPU are important promoters of growth of various horticultural crops, such as kiwi, melons, watermelons and grapes (Arima et al. 1995). TDZ has also been extensively used as a cotton defoliant in the USA (Arndt et al. 1976). A new series of CPPU derivatives were described recently and tested on ZmCKX1 (Kopečný et al. 2010). The two best compounds were found to display IC_{50} values 25 times lower than that for TDZ. However, the synthesis of these compounds was complicated. Another group of compounds known to inhibit CKX is the suicide substrates N^6 -(buta-2,3-dienyl)adenine and N^6 -(penta-2,3-dienyl)adenine, which are strong CKX inhibitors with significant cytokinin activity (Suttle and Mornet 2005; Kopečný et al. 2008). Yet another group includes anilino-purine derivatives (Zatloukal et al. 2008). However, their synthesis includes the use of hydrogen fluoride, which is highly corrosive and toxic. The strongest inhibitor from anilino-purine group was successfully used to support *in vitro* organogenesis (Aremu et al. 2015) and to enhance plant stress tolerance toward salinity (Aremu et al. 2014) and cadmium (Gemrotová et al. 2013). Therefore, the development of new urea derivatives that are easy to synthesize and inhibit various CKX isoforms is desired for potential applications in horticulture, agriculture and biotechnology.

In the present study, we employed *in silico* drug design to identify new TDZ derivatives with improved CKX inhibitory properties. The designed inhibitors were synthesized in a two-step reaction by fusing 1,2,3-thiadiazol-5-isocyanate with phenylamine or benzylamine derivatives and

the products were characterized. The binding properties of the TDZ derivatives obtained *in silico* were compared to IC_{50} values (concentration at which the enzyme activity is inhibited by 50%) determined *in vitro* using AtCKX2. A subset of these derivatives was further screened using ZmCKX1 and ZmCKX4a and binding modes were analyzed by X-ray crystallography using ZmCKX4a. TDZ is also known to strongly activate all three cytokinin receptors in *A. thaliana*: ARABIDOPSIS HISTIDINE KINASE2 (AHK2), AHK3, and CYTOKININ RESPONSE1/AHK4 as well as primary cytokinin-sensitive genes including the *Arabidopsis Response Regulator 5* (*ARR5*; Spichal et al. 2004; Stolz et al. 2011; Lomin et al. 2015). *AHK* genes are expressed in almost all cells of different organs although with different strength (Nishimura et al. 2004). While AHK2 shows a lower sensitivity in planta, apparent affinity constant (K_D) values for major natural cytokinins and synthetic TDZ for all three receptors are in the low nanomolar range (Yamada et al. 2001; Romanov et al. 2006; Stolz et al. 2011). We were able to screen TDZ derivatives for binding to AHK3 and CRE1/AHK4 receptors and for their cytokinin activity in *Arabidopsis* using transgenic plants harboring the *ARR5::GUS* reporter gene ($GUS = \beta$ -glucuronidase; D'Agostino et al. 2000). Finally, the cytokinin activity of the compounds was also assessed in three classical cytokinin bioassays, i.e., wheat leaf senescence, tobacco callus and *Amaranthus* bioassays.

Results and discussion

Active sites in ZmCKX1, ZmCKX4a and AtCKX2

The active site of ZmCKX1 is composed of two pockets connected by a narrow tunnel bordered by Asp169. The crystal structure in complex with CPPU (PDB ID: 2QKN, Kopečný et al. 2010) shows that Asp169 forms hydrogen bonds to both urea nitrogen atoms of CPPU. The smaller of the two pockets is occupied by the isoalloxazine ring of FAD, whereas the larger pocket is exposed to solvent and is guarded by Glu381. CPPU can adopt two possible orientations, i.e., with either the phenyl ring or 2-chloropyridin-4-yl ring in direct π -stacking interaction with the isoalloxazine ring of FAD. The crystal structure of ZmCKX4a shows that CPPU is bound only in the orientation with the 2-chloropyridin-4-yl ring pointing towards the entrance of the active site and phenyl ring in direct π -stacking interaction with the isoalloxazine ring (PDB ID: 4O95 and 4OAL, Kopečný et al. 2015). ZmCKX4a has an alanine in place of Glu381 found in ZmCKX1 and this residue apparently affects the binding of urea-derived cytokinins.

A homology model of AtCKX2, prepared based on the ZmCKX1 structure, revealed that it shares a similar shape

of active site. Amino acids in the active site of AtCKX2 and ZmCKX1 are almost identical and include Asp150/Asp169, Val345/Val378, Trp358/Trp391, Asn366/Asn399, Tyr460/Tyr491, Leu461/Leu492 and especially Glu348/Glu381. A small difference is replacement of Leu397 by Ile429. To confirm the ligand binding, CPPU and TDZ were docked into the active site of ZmCKX1 and AtCKX2 using AutoDock Vina. The results were nearly identical and agreed with positions observed for CPPU in known crystal structures. CPPU was found in two orientations with binding energies of -6.6 and -8.2 kcal mol $^{-1}$. The latter corresponded to preferred binding with the phenyl ring in π -stacking interaction with the isoalloxazine ring. In contrast, when docking TDZ, the thiadiazolyl ring was located over the isoalloxazine ring.

Design and chemistry of TDZ derivatives

Molecular docking to the active site of AtCKX2 was performed in parallel to preparation of new *N,N'*-bis substituted urea derivatives possessing the 1,2,3-thiadiazolyl ring of TDZ. Twenty of the most promising compounds, based on the obtained ligand conformations, binding energies (Table 1) and availability of phenyl- and benzyl-amines for the synthesis, were selected and synthesized (Table 1; Fig. 1). Commercially available 1,2,3-thiadiazol-5-ylamine was a starting compound for the synthesis of the intermediate 5-isocyanato-1,2,3-thiadiazole, which was prepared using diphosgene (Kurita et al. 1976). Although, many TDZ derivatives have been synthesized before (Mok et al. 1982; Abad et al. 2004), use of 5-isocyanato-1,2,3-thiadiazole has not yet been reported. Final compounds were prepared by mild heating of this isocyanate and the corresponding amine in the presence of a catalytic amount of triethylamine in tetrahydrofuran (THF) (Goldschmidt and Bardach 1892). The compounds were characterized by LC-PDA-MS and 1H NMR (Supplementary material and Figure S3 and S4). Two compounds HETDZ and 3FMTDZ were also characterized in MS/MS experiments (Figure S5).

Inhibition of AtCKX2 and correlation with molecular docking

TDZ derivatives all docked into the active site of the AtCKX2 model with apparent preferred orientation of the 1,2,3-thiadiazolyl ring towards the isoalloxazine ring of FAD. Both urea nitrogen atoms bound to Asp150. The polar functional groups attached to the phenyl ring interacted with Glu348 at the entrance, whereas those attached to the benzyl ring were too large to be accommodated easily. This finding is in line with the calculated binding energies, which were more favorable for phenyl derivatives (ΔG from -8.1 to -9.3 kcal mol $^{-1}$) than those for benzyl derivatives (ΔG from -6.3 to -8.2 kcal mol $^{-1}$). However,

Table 1 Structures and abbreviations of prepared compounds and overview of their binding energies and comparison to IC_{50} values measured with AtCKX2, ZmCreKX1 and ZmCKX4a

	R1	R2	R3	R4	ΔG (kcal/mol)	IC_{50} (μM)		
						AtCKX2	ZmCKX1	ZmCKX4a
Compounds of series I								
TDZ	H	H	H	H	-8.8	62±6	40±7	273±18
1	OCH ₃	H	H	H	-8.4	36±3*	35±5	>200
2	OH	H	H	H	-8.3	51±5	37±4	>200
3	H	OCH ₃	H	H	-8.8	22±4*	9±3*	>200
4	H	OH	H	H	-8.8	66±5	43±4	>200
5	OCH ₃	H	H	OCH ₃	-8.9	7.6±1*	24±4*	>200
6	OH	H	H	OH	-8.9	–	–	–
7	H	OH	CH ₃	H	-9.2	33±2*	26±3*	117±12*
8	CH ₂ OH	H	H	H	-8.3	18±2*	14±2*	91±10*
9 HETDZ	CH ₂ CH ₂ OH	H	H	H	-8.5	3.9±0.6*	2.8±0.5*	120±15*
10 3FMTDZ	H	OCF ₃	H	H	-9.3	5.5±0.6*	4.8±0.3*	35±4*
11	H	CF ₃	H	H	-8.1	45±7	48±5	>200
12	Cl	H	Cl	H	-9.3	>200	>200	>200
Compounds of series II								
13	OH	H	H	H	–	>200	–	–
14	CH ₃	H	H	H	-6.9	n.a.	–	–
15	H	OH	H	H	-7.2	106±17*	–	–
16	H	Cl	H	H	-6.3	>200	–	–
17	H	F	H	H	-7.3	n.a.	–	–
18	H	OCH ₃	H	H	–	n.a.	–	–
19	OH	CH ₃	H	H	-6.3	n.a.	–	–
20	OCH ₃	H	H	OCH ₃	-8.2	n.a.	–	–

Binding energies (ΔG) were obtained by docking TDZ derivatives into the active site of AtCKX2. CKX activity was measured using the PMS/MTT method. *n.a.* Compound not active as inhibitor; – data not determined. Errors show SD of two parallel assays, each consisting from three replicates. An asterisk means statistically significant difference from TDZ in an ANOVA analysis (*t* test) at $p < 0.05$

ΔG differences among phenyl derivative series are too small to predict their real in vivo binding and thus data correlate rather roughly with IC_{50} values measured with AtCKX2 (Table 1). No correlation was found for compound 12 despite it displayed together with FMTDZ the best binding energy (ΔG value of -9.3 kcal mol⁻¹) found by docking. As the compound 12 is more hydrophobic due to the presence of chlorine atoms, it is likely that lower solubility or non-specific binding may have negatively affected its properties.

The ability of the prepared derivatives to inhibit the activity of AtCKX2 was compared with that of TDZ (Fig. 2). IC_{50} values were determined using a saturating concentration of the substrate, i.e., 45 μM iP (Table 1). An IC_{50} value of 62 μM was determined for TDZ. Similar IC_{50} values were determined for compounds 2 and 4, which were *ortho*- and *meta*-hydroxylated TDZ derivatives, respectively. No significant change was observed for compounds 1 (*ortho*-methoxy TDZ), 7 (*ortho*-methyl, *meta*-hydroxy TDZ) and 11 (*meta*-trifluoromethyl TDZ). Lower IC_{50} values were observed for compounds 8 (IC_{50} =18 μM) and 3

(IC_{50} =22 μM) bearing *ortho*-methylhydroxy and *meta*-methoxy substituents, respectively. Compounds with IC_{50} values below 10 μM were compounds 5 (IC_{50} =7.6 μM), 10 (3FMTDZ, IC_{50} =5.5 μM) and 9 (HETDZ, IC_{50} =3.9 μM) bearing 2,5-dimethoxy-, *meta*-trifluoro-methoxy- and *ortho*-ethylhydroxy groups. Compounds bearing a dichlorophenyl ring or substituted benzyl ring were less active or even inactive. 3FMTDZ was one of the top two compounds with the best binding energy found by docking (ΔG = -9.3 kcal mol⁻¹), whereas the other compound, 12, was not effective at all. As the latter compound is more hydrophobic due to the presence of chlorine atoms, it is likely that lower solubility or non-specific binding may have negatively affected its properties. Zatloukal et al. (2008) reported an IC_{50} value for TDZ of 29 μM toward AtCKX2. This is about half that of our result (62 μM). This difference can be explained by the use of lower substrate concentration in the previous study (30 μM iP). Zatloukal's best inhibitors reached an IC_{50} of about 1–2 μM , which is very close to the inhibitory strength of HETDZ (IC_{50} =3.9 μM). In contrast, our IC_{50} was obtained with 45 μM iP as a substrate.

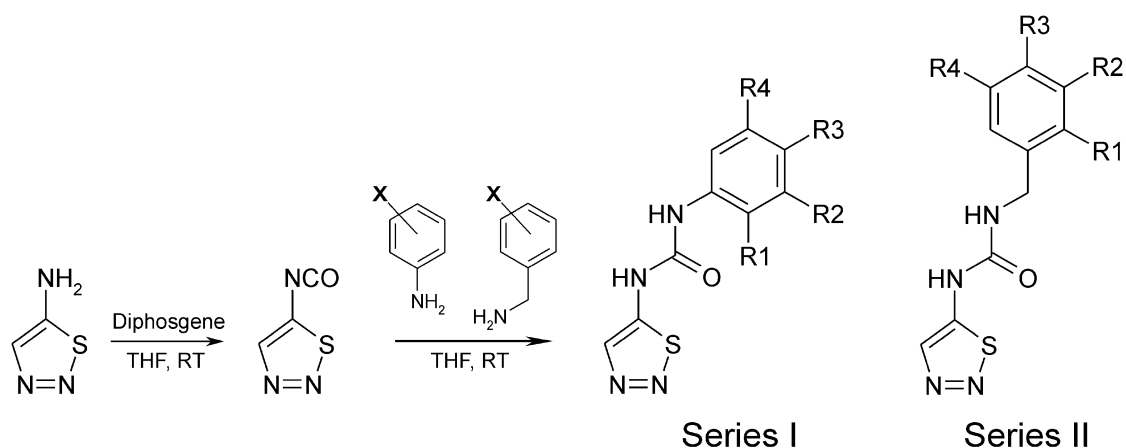


Fig. 1 Reaction scheme of the synthesis of new TDZ derivatives. In the first step, a 1,2,3-thiadiazol-5-ylamine was dissolved in THF and added dropwise into a solution of diphosgene to form an intermediate

5-isocyanato-1,2,3-thiadiazole. The isocyanate was further reacted with a phenyl- or benzyl-amine derivative in the presence of a catalytic amount of triethylamine to form the final product

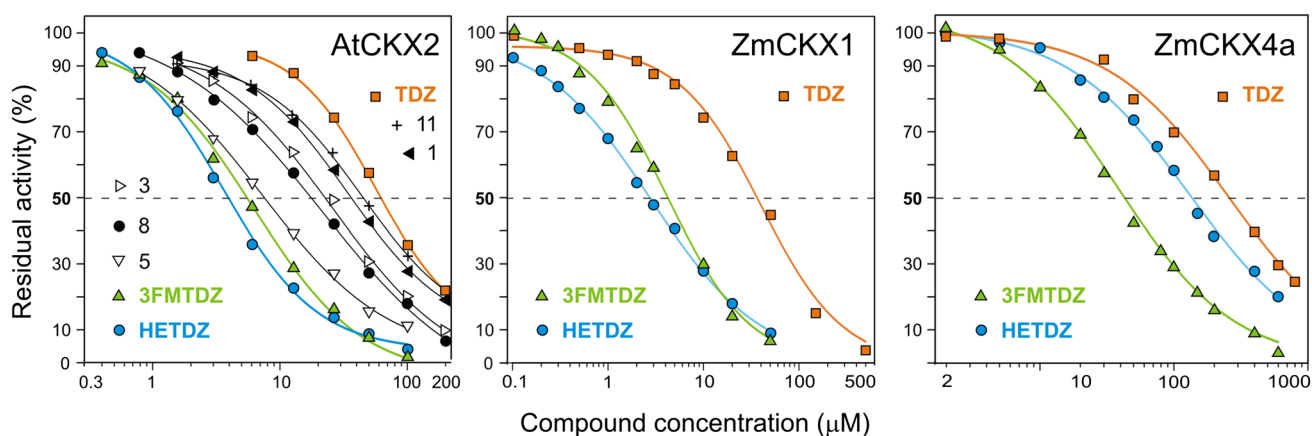


Fig. 2 Inhibitor strength of selected TDZ derivatives against AtCKX2, ZmCKX1 and ZmCKX4a. CKX activity was measured by PMS/MTT kinetic assay (Frébort et al. 2002). The standard deviation of three

replicates did not exceed 10%. The whole assays were performed twice and presented graphs are representative examples

Inhibition of ZmCKX1 and ZmCKX4a

Based on the results obtained with AtCKX2, selected TDZ derivatives were further tested using ZmCKX1 and ZmCKX4a. Compounds 1–12 displayed similar properties toward ZmCKX1, probably because it has a similar active site composition to that of AtCKX2. HETDZ and 3FMTDZ were the strongest inhibitors of ZmCKX1, with IC_{50} values of 2.8 and 4.8 μM , respectively. Only compounds 7, 8, HETDZ and 3FMTDZ inhibited ZmCKX4a better than TDZ itself: the IC_{50} value for TDZ was slightly higher than 200 μM , whereas it was 117, 91 and 120 μM for compounds 7, 8 and HETDZ, respectively. The best inhibitor was 3FMTDZ, with IC_{50} of 35 μM (Table 1; Fig. 2). Previously, the two strongest CPPU derivatives toward the ZmCKX1 isoform were reported to have IC_{50} values of 1.8 and 2 μM (Kopečný et al. 2010). However,

as mentioned above, the synthesis of these compounds is much more complicated than that of HETDZ and 3FMTDZ.

Although earlier work suggested that TDZ was a non-competitive inhibitor (Hare and Van Staden 1994), more recent studies have shown that it is a competitive inhibitor of CKX when 2,6-dichlorophenol indophenol (DCPIP) is used as an electron acceptor (Bilyeu et al. 2001; Kopečný et al. 2010). Using a method based on the coupled redox reaction of PMS and MTT (Frébort et al. 2002), we showed that TDZ and the two strongest inhibitors HETDZ and 3FMTDZ inhibit AtCKX2 as well as ZmCKX4a in a competitive manner (Fig. 3). With AtCKX2, K_i values for HETDZ and 3FMTDZ were 1.25 and 1.8 μM , respectively, whereas that for TDZ was one order of magnitude higher ($K_i=21$ μM). With ZmCKX4a, K_i values for TDZ and 3FMTDZ were 7.2 and 3.6 μM , respectively.

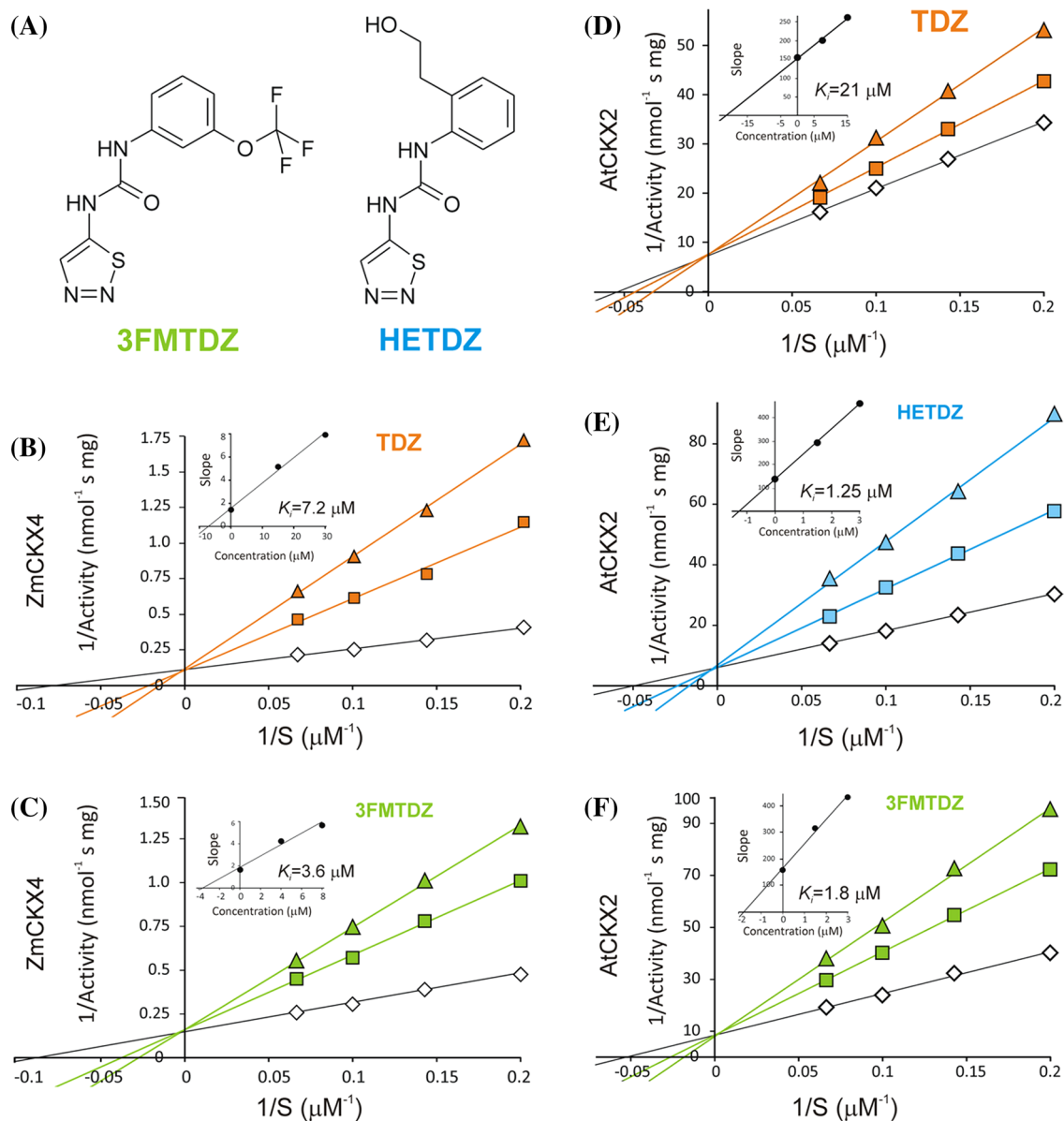


Fig. 3 Inhibition of ZmCKX4a and AtCKX2 with selected TDZ derivatives. **a** Structures of the two best inhibitors HETDZ and 3FMTDZ. **b, c** Double-reciprocal plots of competitive inhibition of ZmCKX4a by TDZ and 3FMTDZ. Lineweaver–Burk plots were constructed using iP as a substrate with concentration range 5–15 μM . The respective inhibitor concentrations were 0 μM (open diamond), 15 μM (filled square) and 30 μM (filled triangle) for TDZ and 0 μM (open diamond), 4 μM (filled square) and 8 μM (filled triangle) for

3FMTDZ. **d–f** Double-reciprocal plots showing competitive inhibition of AtCKX2 by TDZ, HETDZ and 3FMTDZ. The respective inhibitor concentrations were 0 μM (open diamond), 7.5 μM (filled square) and 15 μM (filled triangle) for TDZ and 0 μM (open diamond), 1.5 μM (filled square) and 3 μM (filled triangle) for HETDZ and 3FMTDZ. The insets show secondary plots of slope against inhibitor concentration, which were used for K_i determination

Crystal structures of ZmCKX4a in complex with 3FMTDZ and HETDZ

Two crystal structures of ZmCKX4a in complex with two urea inhibitors 3FMTDZ and HETDZ were solved at 2.0 \AA resolution (PDB IDs 5HMR and 5HQX). A summary of the refinement results together with model statistics is given in Table 2. The crystal structure complexes revealed that

HETDZ and 3FMTDZ occupy the same place in the active site as cytokinin substrates. Surprisingly, HETDZ and 3FMTDZ bind in opposite orientations (Fig. 4). Their urea nitrogen atoms are at a distance of 2.7–3.1 \AA from the side chain oxygen atom of catalytic aspartate D170, allowing hydrogen bond interactions. The urea's carbonyl oxygen is 3.3 \AA distant from the N5 atom of the isoalloxazine plane of FAD. The thiadiazolyl ring of 3FMTDZ is positioned over

Table 2 Data collection and refinement statistics

	ZmCKO4a+3FMTDZ	ZmCKO4a+HETDZ
PDB ID	5HMR	5HQX
Space group	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2
Asymmetric unit	Monomer	Monomer
Unit cell (Å)		
a	79.7	74.5
b	79.7	74.5
c	203.7	208.1
$\alpha=\beta=\gamma$ (°)	90.0	90.0
Resolution (Å)	49.3–2.00	47.0–2.05
Observed reflections	325,424 (52,088) ^a	542,798 (83,013)
Unique reflections	45,506 (7123)	38,064 (5908)
Completeness (%)	99.7 (98.6)	99.6 (97.5)
<i>I</i> / σ (<i>I</i>)	14.7 (1.4)	23.0 (2.3)
<i>CC</i> _{1/2} ^b	99.9 (59.8)	100.0 (88.6)
<i>R</i> _{sym} (%)	8.6 (152.2)	8.0 (109.9)
<i>R</i> _{cryst} (%)	18.2	19.2
<i>R</i> _{free} (%)	20.6	22.4
RMSD bond lengths (Å)	0.010	0.010
RMSD bond angles (°)	0.96	0.99
B average value (Å ²)		
Protein	48.8	52.6
Ligand	45.0	53.4
Solvent	50.1	54.1

^aNumbers in parentheses represent values in the highest resolution shell 2.00–2.12 Å

^bPercentage of correlation between intensities from random half-dataset (Karplus and Diederichs 2012)

the flavin ring and one of its two nitrogen atoms can establish a hydrogen bond to Asn391 directly (3.3 Å distance) or via a water molecule. One of the urea nitrogens forms an H-bond to a water molecule, which is further bound to the side-chain oxygen atom of D170, O4 atom of isoalloxazine, hydroxyl group of Tyr425 and the main-chain oxygen of Ile185. The 3-trifluoromethoxy group is oriented toward the hydrophobic pocket composed of side chains of Trp389, Trp383, Phe57, Leu377 and His387. In contrast, in the case of HETDZ, the 2-(2-hydroxyethyl)phenyl ring is positioned over the flavin ring, whereas the thiadiazolyl ring points toward the entrance of the active site. The oxygen atom of the hydroxyethyl side chain replaces a water molecule present in the ZmCKX4a-3FMTDZ structure and makes a H-bond to the side-chain oxygen atom of Asp170 and hydroxyl group of Tyr425. This water molecule is present in most of the CKX structures, including substrate and inhibitor complexes, and is H-bonded to a catalytic aspartate residue. Orientation and position of 3FMTDZ found

in the crystal structure was very similar to that predicted by molecular docking for AtCKX2. Interestingly, orientation of the docked HETDZ was similar to that of 3FMTDZ with thiadiazolyl ring positioned over the flavin ring and hydroxyethyl side chain pointing towards and interacting with Arg344 (Arg369 in ZmCKX4a). However, this interaction does not appear in the crystal structure of ZmCKX4a and further crystallization experiments are needed to verify whether HETDZ binding may differ among CKX isoforms.

Activity in classical cytokinin bioassays

Three classical cytokinin bioassays, i.e., wheat leaf senescence, *Amaranth* and tobacco callus assay, were used to investigate the cytokinin activity of the selected compounds using *N*⁶-benzyladenine (BA) and TDZ as controls (Table 3). TDZ is a very potent cytokinin and none of the tested derivatives possessed higher activity than TDZ itself in any bioassay. In the senescence assay, where the degradation of chlorophyll is assessed, eight compounds were completely inactive and only three had an *IC*₅₀ (concentration at which the chlorophyll degradation is inhibited by 50%) lower than 100 μM, although the *IC*₅₀ of TDZ in this assay was about 10 μM. The most active compounds were those bearing benzyl substituted in the *meta*-position by fluorine or hydroxyl group (compounds 17 and 15). In the *Amaranth* assay, which is based on the dark induction of betacyanin synthesis in *Amaranth* cotyledons by cytokinins, *EC*₅₀ values (concentration at which the activation response reaches 50%) were determined. Compound 8 was the most active in this assay, slightly exceeding the activity of BA. This compound had an *ortho*-methylhydroxy phenyl ring. Moderately less active were compounds 1, 2 and 3, which had an *ortho*-methoxy, *ortho*-hydroxy and *meta*-methoxy substituted phenyl ring, respectively, and compound 17, which had a *meta*-fluorobenzyl group. In the tobacco callus assay, where the growth of callus is cytokinin dependent, none of the compounds reached the TDZ activity. The most active compounds (8, 15 and 17) showed similar activity to that of BA. In contrast, substances 3, 11, 14, 16 and 3FMTDZ showed maximum activity for a concentration one order of magnitude higher.

Activation of cytokinin primary response gene *ARR5*

We employed transgenic *Arabidopsis* plants (*A. thaliana*) harboring the *ARR5::GUS* reporter gene (D'Agostino et al. 2000) to monitor complex in vivo cytokinin activity of each compound. *ARR5* is the primary response gene with a cytokinin-dependent promoter, and its activation integrates the responses of several putative cytokinin signaling pathways as well as the contribution of individual cytokinin receptors. Compounds were tested at 0.5, 1.0 and 10 μM concentrations

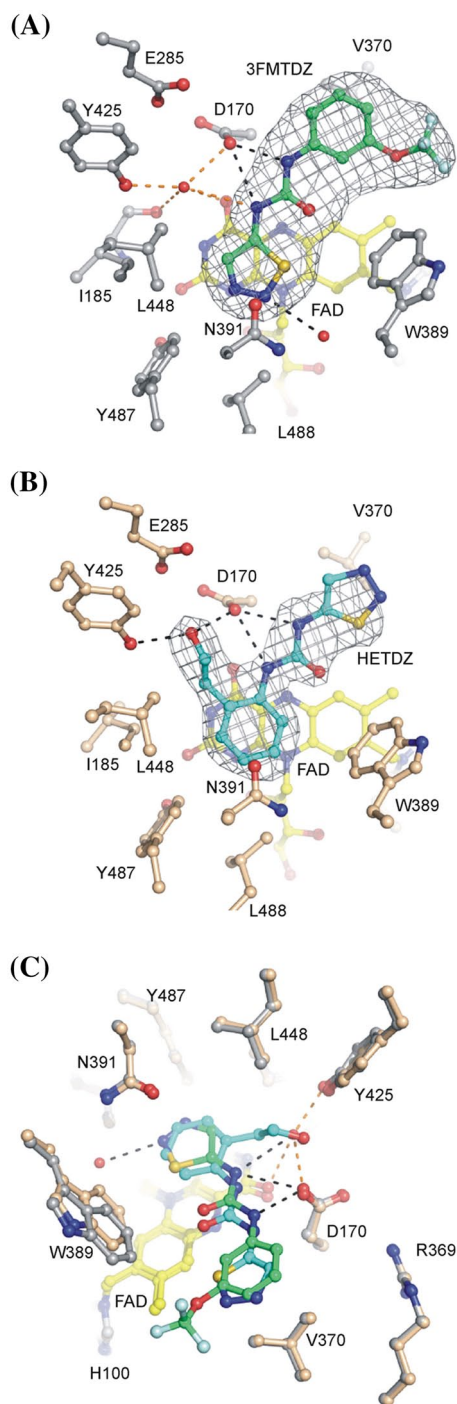


Fig. 4 Binding of the two urea inhibitors 3FMTDZ and HETDZ in the active site of ZmCKX4a. **a** Binding of 3FMTDZ (PDB ID: 5HMR). The inhibitor is colored green and shown in a Fo–Fc omit map contoured at 4.0 σ . **b** Binding of HETDZ (PDB ID: 5HQX). The inhibitor is colored blue and shown in a Fo–Fc omit map contoured at 3.0 σ . **c** Superposition of bound 3FMTDZ and HETDZ. Both ligands form a hydrogen bond to the oxygen atom of D170 via two nitrogen atoms of the urea backbone. The oxygen atom of the hydroxyethyl side chain of HETDZ replaces a water molecule observed in the ZmCKX4a-3FMTDZ structure. The FAD cofactor is colored yellow and neighboring residues are labeled

(Table 3) and their ability to activate the *ARR5::GUS* expression is presented for their optimal concentration. TDZ had highest activity in the nanomolar range, whereas BA reached its maximum activity at 1 μM . The most active TDZ derivatives, namely compounds 1, 8 and HETDZ (9), possessed a similar activity to that of BA. For these compounds, there was a clear structure–activity relationship, with the activity decreasing in the order *ortho*-methoxy group (1) > *ortho*-methylhydroxy group (8) > *ortho*-ethylhydroxy group (HETDZ). Slightly lower activities were exhibited by compounds 3, 4 and 10 (3FMTDZ) bearing methoxy, hydroxy and trifluoromethoxy groups, respectively, in the *meta* position of the phenyl ring. Compounds 2, 5, 11, 12 and 15 caused maximal activation only at 10 μM concentration. Other compounds, which did not activate the *ARR5::GUS* expression even at 10 μM concentration, were categorized as very poor cytokinins. These included TDZ derivatives with a substituted benzyl ring rather than phenyl ring.

Interaction with Arabidopsis cytokinin receptors AHK3 and CRE1/AHK4

Transformed *E. coli* cells expressing the genes of cytokinin receptor AHK3 or CRE1/AHK4 and the cytokinin-activated reporter gene *cps::lacZ* (Suzuki et al. 2001; Yamada et al. 2001) were employed to examine the ability of the synthesized compounds to activate the receptor (in case of CRE1/AHK4 receptor only) or bind to the cytokinin binding site. In the CRE1/AHK4 receptor activation assay, EC_{50} values (half maximal effective concentration for receptor activation) of the compounds were determined and compared with those for *trans*-zeatin (*tZ*) and TDZ (see Figure S1). The EC_{50} values for *tZ* and TDZ were 0.25 and 0.72 μM , respectively. This is in line with previous results reporting an EC_{50} value for *tZ* of 0.2 μM (Nisler et al. 2010). The receptor was activated by several compounds at the two highest concentrations tested (10 and 50 μM). All of the derivatives showed decreased cytokinin activity when compared to TDZ. The most active was compound 2 bearing an *ortho*-hydroxyphenyl ring (EC_{50} = 5.6 μM), followed by compounds 15 (EC_{50} = 17 μM) > 4 (EC_{50} = 22 μM) \cong 17 (EC_{50} = 25 μM) > 8 (EC_{50} = 32 μM) > 3 (EC_{50} > 50 μM) bearing a *meta*-hydroxybenzyl, *meta*-hydroxyphenyl, *meta*-fluorobenzyl, *ortho*-hydroxymethylphenyl and *meta*-methoxyphenyl ring, respectively. Interestingly, except for the fluorinated derivative (17), only compounds possessing hydroxyl group or methoxy group were capable of the activation. None of the other compounds, including HETDZ and 3FMTDZ, activated the CRE1/AHK4 receptor (Figure S1).

The same *E. coli* cells were also used to determine whether 3FMTDZ or HETDZ possessed affinity toward the

Table 3 Cytokinin activity of synthesized TDZ derivatives in classical cytokinin bioassays

Compound	Senescence bioassay	Amaranthus bioassay	Tobacco callus bioassay		Arabidopsis <i>ARR5::GUS</i> assay	
	IC_{50} (μ M)	EC_{50} (μ M)	Optimal conc. (mol l^{-1})	Relative activity (%)	Optimal conc. (mol l^{-1})	Relative activity (%)
BA	155±22	0.55±0.13	10 ⁻⁶	100	10 ⁻⁶	100
TDZ	13±4*	0.004±0.002*	10 ⁻⁷	102±5	10 ⁻⁷	100±10
1	127±13	1.5±0.4	10 ⁻⁴	103±2	5×10 ⁻⁷	74±6
2	113±14	3.5±1.0*	10 ⁻⁴	94±3	10 ⁻⁵	105±8
3	82±7*	3.3±1.6	10 ⁻⁵	93±3	10 ⁻⁶	110±9
4	152±12	29±5.2*	10 ⁻⁴	108±2	10 ⁻⁶	106±5
5	n.a	>100	10 ⁻⁴	75±8	10 ⁻⁶	47±6
6	n.a	>100	10 ⁻⁴	15±7	n.a	n.a
7	n.a	n.a	10 ⁻⁴	48±6	10 ⁻⁵	41±3
8	115±12	0.42±0.2	10 ⁻⁶	105±5	5×10 ⁻⁷	71±4
9-HETDZ	127±14	10.5±2.1*	10 ⁻⁵	95±5	5×10 ⁻⁷	62±6
10-3FMTDZ	n.a	>100	10 ⁻⁵	96±4	10 ⁻⁶	102±1
11	n.a	25±8*	10 ⁻⁵	104±6	10 ⁻⁵	105±5
12	105±10	>100	10 ⁻⁵	77±4	10 ⁻⁵	112±13
13	n.a	>100	10 ⁻⁴	90±4	10 ⁻⁵	77±5
14	157±11	65±7*	10 ⁻⁵	91±3	10 ⁻⁵	53±5
15	65±6*	9.0±1.2*	10 ⁻⁶	100±4	10 ⁻⁵	104±9
16	n.a	83±7	10 ⁻⁵	102±2	10 ⁻⁵	84±8
17	32±5*	2.5±0.8	10 ⁻⁶	95±7	10 ⁻⁵	64±7
18	n.a	n.a	n.a	n.a	n.a	n.a
19	n.a	>100	10 ⁻⁴	32±5	10 ⁻⁵	14±2
20	160±12	>100	–	–	10 ⁻⁵	62±4

In the tobacco callus bioassay and Arabidopsis *ARR5::GUS* assay, the effect at optimal concentration is compared with the activity of BA: 100 % corresponds to 10⁻⁶ M BA. *n.a.* Compound not active; – data not determined. Errors show SD of at least two parallel assays, each consisting from at least three replicates. An asterisk means statistically significant difference from BA in an ANOVA analysis (*t* test) at *p*<0.05 (for Senescence and Amaranthus assays)

CRE1/AHK4 receptor. Their affinities were compared with those of *tZ*, TDZ and adenine (Figure S1). While *tZ* at 10 nM concentration (used as a positive control) displaced 66 % of [2-³H]*tZ* from the receptor, TDZ was inactive at this concentration. TDZ displaced 85 % of tritiated *tZ* from the receptor at much higher 1 μ M concentration whereas both 3FMTDZ and HETDZ were inactive at this concentration. They both displaced only about 15 % of tritiated *tZ* from the receptor at 10 μ M concentration. The same result was achieved with the negative control, adenine. This confirms the results of the receptor activation assay and demonstrates limited interaction of 3FMTDZ and HETDZ with this receptor.

The AHK3 receptor is known to display better affinity (lower K_D values) and lower selectivity towards natural and synthetic cytokinins than the AHK4 receptor (Spíchal et al. 2004; Romanov et al. 2006). The K_D values for TDZ were reported as 13 nM (AHK3) and 40 nM (AHK4). In our measurements (Figure S1), 10 nM TDZ displaced about 80 % of tritiated *tZ* from the AHK3 receptor while

all phenyl derivatives (series I) and compound 15 (series II) displaced tritiated *tZ* at 20 μ M concentration. Remaining compounds of the series II did not bind at the concentration tested. Compounds 2, 3, 4, 6, 8, HETDZ and 15 displayed stronger binding as they displaced more than 70 % of tritiated *tZ*. Compounds 5, 7, 3FMTDZ, 11 and 12 displaced less than 50 % of tritiated *tZ* from the AHK3 receptor. In summary, both compounds HETDZ and 3FMTDZ bind to AHK3 receptor at much higher concentrations (and much lower affinity) compared with TDZ.

Conclusion—cytokinin activity versus CKX inhibitory activity

Previously described CKX inhibitors, including TDZ, CPPU and its derivatives, suicide substrates or anilino-purines, have been shown to activate *Arabidopsis* cytokinin receptors AHK3 and/or AHK4 (Spíchal et al. 2004;

Kopečný et al. 2010, 2008; Zatloukal et al. 2008). These compounds therefore possess intrinsic cytokinin activity including TDZ, the urea-derived compound with the highest cytokinin activity. In contrast, the new TDZ inhibitors 3FMTDZ or HETDZ described in this work do not interact with the CRE1/AHK4 receptor. They both possess lower intrinsic cytokinin activity than the cytokinins TDZ and *tZ* and display wide and increased inhibitory properties to various CKX isoforms. Their lower affinity to cytokinin receptors AHK3 and AHK4 (AHK2 receptors were not tested in this work) makes both attractive candidates for manipulations with endogenous cytokinins in planta and for further development of new derivatives. It has been proposed that enhancing the content of endogenous cytokinins in plants via inhibiting CKX is more beneficial to plant growth and development than using the high intrinsic cytokinin activity of exogenously applied compounds (Gemrotová et al. 2013). This indicates that plants favor very gentle regulation of cytokinin biosynthesis, metabolism and signaling. Therefore, CKX inhibitors might affect different cytokinin functions in plants, thereby having possible positive effects on seed filling, delayed senescence and stress tolerance toward biotic and abiotic stresses, thus improving crop yields.

Materials and methods

Molecular modeling and docking

A model of AtCKX2 was built in Modeller 9.8 (Eswar et al. 2006) within UCSF Chimera 1.9 (Pettersen et al. 2004) by aligning the sequence (UNIPROT ID: Q9FUJ3) to a template structure of ZmCKX1 (PDB ID: 2QKN, Kopečný et al. 2010). Sequence alignment between both proteins can be found in the supplement (Figure S2). Modeller was instructed to build 5 models of AtCKX2 without explicit hydrogens with non-protein residues treated as rigid bodies. We have selected the best performing model in the zDOPE value with rotamer of Asp150 in the same interaction position as Asp169 in ZmCKX1. Model also did not suffer from sterical clashes and RMSD for all C α atoms between final model and template was just 0.176 Å. Ligands for molecular docking were constructed using Marvin 14.9.8 software (ChemAxon). Polar hydrogens were added to all ligands and proteins using AutoDock Tools (ADT) software (Morris et al. 2009) prior to docking using Autodock Vina 1.05 (Trott and Olson 2010). A grid box with size of 16 Å was centered on the active site around Asp150. The exhaustiveness parameter was set to 20 (default 8). After docking, the best docked ligand was selected and the best crystal-like poses of each ligand were analyzed with Pymol 1.7.4 (The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC.) and Maestro 2014-3 (Schrödinger Release 2014-3:

Maestro, version 9.9, Schrödinger, LLC, New York, NY, 2014). RMSD cutoff for crystal-like pose is default defined below 2 Å.

Chemicals

1,2,3-Thiadiazol-5-ylamine was supplied by TCI Europe (Zwijndrecht, Belgium). TDZ, *tZ* and BA were supplied by Olchemim (Olomouc, Czech Republic). Other compounds, including 2-methoxy-phenylamine, 2-aminophenol, 3-methoxy-phenylamine, 3-amino-phenol, 2,5-dimethoxy-phenylamine, 2-hydroxy-3-methyl-phenylamine, (2-amino-phenyl)-methanol, (2-amino-phenyl)-ethanol, 3-trifluoromethoxy-phenylamine, 3-trifluoromethyl-phenylamine, 2,4-dichloro-phenylamine, 2-hydroxy-benzylamine, 2-methyl-benzylamine, 3-hydroxy-benzylamine, 3-methoxy-benzylamine, 3-fluoro-benzylamine, 3-chloro-benzylamine, 2,5-dimethoxy-benzylamine and *tert*-butyl-dimethyl-silyl chloride, were purchased from Sigma Aldrich (Germany). All solvents and chemicals used were of standard p.a. quality.

Synthesis of 5-isocyanato-1,2,3-thiadiazole and final products

The synthesis of isocyanates from amines has been described elsewhere (Kurita et al. 1976). Briefly, 1,2,3-thiadiazol-5-ylamine (1.01 g, 10 mmol) was dissolved in THF (40 mL) and added dropwise into a solution of diphosgene (2.6 g, 13 mmol) in THF (100 mL). The reaction mixture was stirred for 40 min at 0 °C and then heated to 30 °C. After another 30 min, the solvent and excess of diphosgene were evaporated. The yellow solid residue of 5-isocyanato-1,2,3-thiadiazole was re-suspended in diethyl ether and filtered. Yield: 95, ¹H NMR (δ , ppm, DMSO-*d*₆): 7.75(1H, s, CH).

All the prepared compounds were synthesized according to common protocols for the synthesis of disubstituted urea derivatives, using amine and isocyanate as starting materials in the presence of a catalytic amount of triethylamine (Goldschmidt and Bardach 1892). Firstly, 5-isocyanato-1,2,3-thiadiazole was mixed with THF (1:100) in a high pressure tube. Then, a catalytic amount of triethylamine was added and the suspension was allowed to dissolve. The starting amine was added in an equimolar amount into the solution and the reaction mixture was stirred for 5–12 h at 60 °C. The conversion was monitored by TLC (CHCl₃:MeOH, 9:1). After reaction, the solvent was evaporated to a solid/semi-solid residue, which was purified by flash silica column chromatography (mobile phase CHCl₃:MeOH, 9:1). Compounds were usually prepared in milligram quantities and yields varied between 30 and 80%. If the starting amine contained a free hydroxyl group, it was protected by *tert*-butyl-dimethylsilyl chloride prior to the condensation with 5-isocyanato-1,2,3-thiadiazole and then de-protected in 2-propanolic HCl (Wuts

and Greene 1991). All products (Table 1) were characterized by ^1H NMR and LC-PDA-MS (Supplementary material). The MS/MS spectra of HETDZ and 3FMTDZ were also recorded using tandem mass analyser Quatro Micro API with MassLynx data system software (Waters, Manchester, UK; see Supplementary material for details).

General experimental procedures

The purity and mass spectra of the synthesized compounds were recorded using method LC-PDA-MS. Compounds (1 mg) were dissolved in 1 ml of 10% methanol and injected (10 μL) onto a reverse-phased column (Symmetry C18, 5 μm , 150 mm \times 2.1 mm; Waters, Milford, MA, USA) incubated at 25 $^\circ\text{C}$. Solvent (A) consisted of 15 mM ammonium formate adjusted to pH 4.0. Solvent (B) consisted of methanol. At flow-rate of 200 $\mu\text{L}/\text{min}$, following binary gradient was used: 0 min, 10% B; 0–24 min; linear gradient to 90% B; 25–34 min; isocratic elution of 90% B; 35–45 min; linear gradient to 10% B. The effluent was introduced then to PDA detector (scanning range 210–700 nm with 1.2 nm resolution) and a tandem mass analyser Quatro Micro API (Waters, Manchester, UK) with an electrospray source (source temperature 120 $^\circ\text{C}$, desolvation temperature 300 $^\circ\text{C}$, capillary voltage 3 kV). Nitrogen was used as well as cone gas (50 L/h) and desolvation gas (500 L/h). Data acquisition was performed in the full scan mode (50–1000 Da), scan time of 0.5 s and cone voltage 20 V. Analyses were performed in positive mode (ESI+) therefore data were collected as protonated ions $[\text{M}+\text{H}]^+$. Analytical thin-layer chromatography (TLC) was carried out using silica gel 60 WF254 plates (Merck). If necessary, compounds were separated on a flash chromatography column using 63 μm chromatographic silica (Davisil) and eluted with a mobile phase containing $\text{CHCl}_3/\text{MeOH}$ (9:1, v/v). ^1H NMR spectra were measured on a Jeol 500 SS spectrometer operating at a temperature of 300 K and a frequency of 500.13 MHz (^1H). Samples were prepared by dissolving the compounds in $\text{DMSO}-d_6$ or CDCl_3 . Tetramethylsilane (TMS) was used as an internal standard.

CKX inhibition measurements

The ability of the prepared compounds to inhibit AtCKX2, ZmCKX1 and ZmCKX4a was evaluated by determining their IC_{50} values in the PMS/MTT (phenazine methosulfate/3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide) activity assay previously adapted for screening in ELISA microtitre plates (Table 1, Frébert et al. 2002). Each well contained 100 μL of a reaction mixture consisting of 0.1 M KH_2PO_4 (pH adjusted to 7.4 with KOH), 1.5 mM PMS, 0.3 mM MTT, the tested compound

(at various concentrations) and substrate iP at saturating concentration (45 μM for AtCKX2 and ZmCKX4a and 10 μM for ZmCKX1). Cell-free growth medium of *S. cerevisiae* strain 23344c ura- harbouring the plasmid pYES2-AtCKX2 (50 μL) was used directly as a source of AtCKX2 (Frébertová et al. 2007). Pure enzymes were used in the case of ZmCKX1 and ZmCKX4a. ZmCKX1 was expressed in yeast *Y. lipolytica* and purified according to a published protocol (Kopečný et al. 2005). *E. coli* BL21 STAR (DE3) cells carrying pTYB12-ZmCKX4a plasmid were grown for 16 h at 18 $^\circ\text{C}$ to produce ZmCKX4a, which was further purified by affinity chromatography on a column filled with a chitin resin, followed by ion-exchange chromatography (Zalabák et al. 2014). Plates were incubated in the dark for 20 min at 37 $^\circ\text{C}$ and the enzymatic reaction was monitored every 2 min by measuring the absorbance of the mixture in each well at 578 nm (spectrophotometer Synergy H4 Hybrid Reader). The absorbance of samples without substrate was subtracted as a baseline. Compounds were tested with two replicates and the entire test was repeated at least twice. Gene 5 software was used to calculate the residual CKX activity for each compound and concentration. IC_{50} values for each compound were determined in Origin Pro software. Lineweaver–Burk plots for K_i determination were constructed using iP as a substrate with concentration range from 5 to 15 μM .

Crystallization and structure determination of 3FMTDZ and HETDZ complexes with ZmCKX4a

Preliminary crystallization conditions were identified in sitting drops using Crystal Screen and Crystal Screen 2 (Hampton Research). ZmCKX4a was co-crystallized with 3FMTDZ in drops containing 100 mM HEPES pH 7.5, 80% MPD, 0.16 mM ZmCKX4a and 10 mM 3FMTDZ in DMSO. For HETDZ, crystals of ZmCKX4a were grown in 22% ethylene-glycol and then infiltrated by 5 mM HETDZ in DMSO for 1 h. Crystals were directly flash frozen in liquid nitrogen. Diffraction data were collected at 100 K on the Proxima 1 beamline at the SOLEIL synchrotron (Saint-Aubin, France) at 2.0 Å resolution. Intensities were integrated using the XDS program (Kabsch 2010) and data quality was assessed using the correlation coefficient $CC_{1/2}$ (Karplus and Diederichs 2012). Crystal structures were determined by performing molecular replacement with Phaser (Storoni et al. 2004) using a monomer of ZmCKX1 (PDB 2QKN) as search model. Model refinement was achieved with BUSTER-TNT (Bricogne et al. 2011). Electron density maps were evaluated using COOT (Emsley and Cowtan 2004). Refinement statistics are presented in Table 2. Molecular graphics images were generated using PYMOL (www.pymol.org).

Cytokinin bioassays and Arabidopsis *ARR5::GUS* reporter gene assay

Classical cytokinin bioassays, i.e., tobacco callus assay, *Amaranthus* assay and wheat senescence assay with excised wheat leaves in the dark, were performed as described in Holub et al. (1998). The tobacco callus bioassay was performed in 6-well microtiter plates (3 mL of MS medium in each well, into which 0.1 g of callus was placed). In the wheat leaf senescence and *Amaranthus* assays, IC_{50} values (concentration at which chlorophyll degradation is inhibited by 50%) and EC_{50} values (half maximal effective concentration) for each compound were determined (Table 3). In the tobacco callus bioassay, the effect at optimal concentration was compared with BA activity (100% corresponded to 10^{-6} M BA). Quantitative determination of GUS activity was performed by measuring fluorescence on a Synergy H4 hybrid reader (Biotek, USA) at excitation and emission wavelengths of 365 and 445 nm, respectively. The assay was carried out as described previously (Romanov et al. 2002). GUS specific activity was expressed in relative % and was calculated as RFU (relative fluorescence units) divided by protein content. Determination of protein content was done using the bicinchoninic acid assay (Smith et al. 1985).

Live-cell cytokinin-binding assays and CRE1/AHK4 receptor activation assay

Escherichia coli strain KMI001, harboring either the plasmid pIN-III-AHK4 or pSTV28-AHK3, which express the Arabidopsis histidine kinases CRE1/AHK4 or AHK3 (Suzuki et al. 2001; Yamada et al. 2001) was used in the experiments. Bacterial strains were kindly provided by T. Mizuno (Nagoya, Japan). The CRE1/AHK4 receptor activation assay was performed as previously described in Spíchal et al. (2004). The live-cell cytokinin-binding assay was performed essentially as described in Romanov et al. (2005). The competition reaction was allowed to proceed with 3 nM [$2\text{-}^3\text{H}$]tZ and various concentrations of the tested compounds for 30 min at 4 °C. When a binding equilibrium was reached, the suspension was centrifuged (6000g), the supernatant was removed and the bacterial pellet was re-suspended in scintillation cocktail (Beckman, Ramsey, MN, USA). Radioactivity was measured by scintillation counting on a Beckman LS 6500 scintillation counter (Beckman, Ramsey, MN, USA).

Accession numbers

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.rcsb.org) under the accession codes 5HMR for ZmCKO4a in complex

with 3FMTDZ and 5HQX for ZmCKO4a in complex with HETDZ.

Acknowledgments The authors gratefully acknowledge support through the projects LO1204 and LO1305 from the Ministry of Education, Youth and Sports of the Czech Republic and the grant 15-22322S and 15-19266S from the Czech Science Foundation. The work was also supported by student projects IGA_PrF_2016_028 and IGA_PrF_2016_018 of Palacký University, Olomouc.

Author contributions J. Nisler, L. Spíchal and M. Strnad designed the research, V. Bazgier and K. Berka performed molecular modeling and docking, J. Nisler and M. Zatloukal synthesized the compounds, J. Nisler, R. Končítiková, D. Zalabák contributed to enzyme kinetics, R. Končítiková, D. Kopečný and P. Briozzo performed X-ray crystallographic study, J. Nisler and D. Kopečný wrote the paper.

References

- Abad A, Agulló C, Cuñat AC, Jiménez R, Vilanova C (2004) Preparation and promotion of fruit growth in kiwifruit of fluorinated *N*-phenyl-*N'*-1,2,3-thiadiazol-5-yl ureas. *J Agr Food Chem* 52:4675–4683
- Aremu AO, Masondo NA, Sunmonu TO, Kulkarni MG, Zatloukal M, Spíchal L, Doležal K, Van Staden J (2014) A novel inhibitor of cytokinin degradation (INCYDE) influences the biochemical parameters and photosynthetic apparatus in NaCl-stressed tomato plants. *Planta* 240:877–889
- Aremu AO, Stirk WA, Masondo NA, Plačková L, Novák O, Pěňčík A, Zatloukal M, Nisler J, Spíchal L, Doležal K, Finnica JF, Van Staden J (2015) Dissecting the role of two cytokinin analogues (INCYDE and PI-55) on in vitro organogenesis, phytohormone accumulation, phytochemical content and antioxidant activity. *Plant Sci* 238:81–94
- Arima Y, Oshima K, Shudo K (1995) Evolution of a novel urea-type cytokinin: horticultural uses of forchlorofenuron. *Acta Hort* 394:75–83
- Arndt F, Rusch R, Stillfried HV (1976) SN49537, a new cotton defoliant. *Plant Physiol* 57:599
- Bilyeu KD, Cole JL, Laskey JG, Riekhof WR, Esparza TJ, Kramer MD, Morris RO (2001) Molecular and biochemical characterization of a cytokinin oxidase from maize. *Plant Physiol* 125:378–386
- Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart OS, Vornrhein C, Womack TO (2011) BUSTER version 2.1.0 Cambridge. Global Phasing Ltd, UK
- Brownlee BG, Hall RH, Whitty CD (1975) 3-Methyl-2-butenal: an enzymatic degradation product of the cytokinin, N-6-(delta-2 isopentenyl) adenine. *Can J Biochem* 53:37–41
- Burch LR, Horgan R (1989) The purification of cytokinin oxidase from *Zea mays* kernels. *Phytochemistry* 28:1313–1319
- Chatfield JM, Armstrong DJ (1986) Regulation of cytokinin oxidase activity in callus tissues of *Phaseolus vulgaris* L. cv Great Northern. *Plant Physiol* 80:493–499
- D'Agostino IB, Deruere J, Kieber JJ (2000) Characterization of the response of the Arabidopsis response regulator gene family to cytokinin. *Plant Physiol* 124(1):1706–1717
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132
- Eswar N, Webb B, Marti-Renom M, Madhusudhan M, Eramian D, Shen M, Pieper U, Sali A (2006) Comparative protein structure modeling using Modeller. *Curr Protoc Bioinform*, Chap 5: Unit 5

- Frébort I, Šebela M, Galuszka P, Werner T, Schmülling T, Peč P (2002) Cytokinin oxidase/cytokinin dehydrogenase assay: optimized procedures and applications. *Anal Biochem* 306:1–7
- Frébortová J, Galuszka P, Werner T, Schmülling T, Frébort I (2007) Functional expression and purification of cytokinin dehydrogenase from *Arabidopsis thaliana* (AtCKX2) in *Saccharomyces cerevisiae*. *Biol Plantarum* 51:673–682
- Galuszka P, Popelková H, Werner T, Frébortová J, Pospíšilová H, Mik V, Köllmer I, Schmülling T, Frébort I (2007) Biochemical characterization and histochemical localization of cytokinin oxidases/dehydrogenases from *Arabidopsis thaliana* expressed in *Nicotiana tabacum* L. *J Plant Growth Regul* 26:255–267
- Gemrotová M, Kulkarni MG, Stirk WA, Strnad M, Van Staden J, Spíchal L (2013) Seedlings of medicinal plants treated with either a cytokinin antagonist (PI-55) or an inhibitor of cytokinin degradation (INCYDE) are protected against the negative effects of cadmium. *Plant Growth Regul* 71:137–145
- Goldschmidt H, Bardach B (1892) Zur Kenntniss der Diazoamidokörper. *Chem Ber* 25:1347–1378
- Hare PD, Van Staden J (1994) Inhibitory effect of thidiazuron on the activity of cytokinin oxidase isolated from soybean callus. *Plant Cell Physiol* 35:1121–1125
- Holub J, Hanuš J, Hanke DE, Strnad M (1998) Biological activity of cytokinins derived from *ortho*- and *meta*-hydroxybenzyladenine. *Plant Growth Regul* 26:109–115
- Houba-Hérin N, Pethe C, d'Alayer J, Laloue M (1999) Cytokinin oxidase from *Zea mays*: purification, cDNA cloning and expression in moss protoplasts. *Plant J* 17:615–626
- Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66:125–132
- Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336:1030–1033
- Kopečný D, Pethe C, Šebela M, Houba-Hérin N, Madzak C, Majira A, Laloue M (2005) High-level expression and characterization of *Zea mays* cytokinin oxidase/dehydrogenase in *Yarrowia lipolytica*. *Biochimie* 87:1011–1022
- Kopečný D, Šebela M, Briozzo P, Spíchal L, Houba-Hérin N, Mašek V, Joly N, Madzak C, Anzenbacher P, Laloue M (2008) Mechanism-based inhibitors of cytokinin oxidase/dehydrogenase attack FAD cofactor. *J Mol Biol* 380:886–899
- Kopečný D, Briozzo P, Popelková H, Šebela M, Končítíková R, Spíchal L, Nisler J, Madzak C, Frébort I, Laloue M, Houba-Hérin N (2010) Phenyl- and benzylurea cytokinins as competitive inhibitors of cytokinin oxidase/dehydrogenase: a structural study. *Biochimie* 92:1052–1062
- Kopečný D, Končítíková R, Popelka H, Briozzo P, Vigouroux A, Kopečná M, Zalabák D, Šebela M, Skopalová J, Frébort I, Moréra S (2015) Kinetic and structural investigation of the cytokinin oxidase/dehydrogenase active site. *FEBS J* 283:361–377
- Kurita K, Matsumura T, Iwakura Y (1976) Trichloromethyl chloroformate. Reaction with amines, amino acids, and amino alcohols. *J Org Chem* 41:2070–2071
- Laloue M, Fox JE (1989) Cytokinin oxidase from wheat: partial purification and general properties. *Plant Physiol* 90:899–906
- Lomin SN, Krivosheev DM, Steklov MY, Arkhipov DV, Osolodkin DI, Schmülling T, Romanov GA (2015) Plant membrane assays with cytokinin receptors underpin the unique role of free cytokinin bases as biologically active ligands. *J Exp Bot* 66:1851–1863
- Massonneau A, Houba-Hérin N, Pethe C, Madzak C, Falque M, Mercy M, Kopečný D, Majira A, Rogowsky P, Laloue M (2004) Maize cytokinin oxidase genes: differential expression and cloning of two new cDNAs. *J Exp Bot* 55:2549–2557
- Mok DW, Mok MC (2001) Cytokinin metabolism and action. *Annu Rev Plant Physiol Plant Mol Biol* 52:89–118
- Mok MC, Mok, DWS, Armstrong DJ, Shudo K, Isogai Y, Okamoto T (1982) Cytokinin activity of *N*-phenyl-*N'*-1,2,3-thiadiazol-5-ylurea (thidiazuron). *Phytochem* 21:1509–1511
- Morris RO, Bilyeu KD, Laskey JG, Cheikh NN (1999) Isolation of a gene encoding a glycosylated cytokinin oxidase from maize. *Biochem Biophys Res Commun* 255:328–333
- Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 16:2785–2791
- Nishimura C, Ohashi Y, Sato S, Kato T, Tabata S, Ueguchi C (2004) Histidine kinase homologs that act as cytokinin receptors possess overlapping functions in the regulation of shoot and root growth in *Arabidopsis*. *Plant Cell* 16:1365–1377
- Nisler J, Zatloukal M, Popa I, Doležal K, Strnad M, Spíchal L (2010) Cytokinin receptor antagonists derived from 6-benzylaminopurine. *Phytochem* 71:823–830
- Petersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612
- Romanov GA, Kieber JJ, Schmülling T (2002) A rapid cytokinin response assay in *Arabidopsis* indicates a role for phospholipase D in cytokinin signalling. *FEBS Lett* 515(:):39–43
- Romanov GA, Spíchal L, Lomin SN, Strnad M, Schmülling T (2005) A live cell hormone-binding assay on transgenic bacteria expressing a eukaryotic receptor protein. *Anal Biochem* 347:129–134
- Romanov GA, Lomin SN, Schmülling T (2006) Biochemical characteristics and ligand-binding properties of *Arabidopsis* cytokinin receptor AHK3 compared to CRE1/AHK4 as revealed by a direct binding assay. *J Exp Bot* 57:4051–4058
- Schrödinger Release 2014-3 (2014) Maestro, version 9.9. Schrödinger, LLC, New York
- Šmečilová M, Galuszka P, Bilyeu KD, Jaworek P, Kowalska M, Šebela M, Sedlářová M, English JT, Frébort I (2009) Subcellular localization and biochemical comparison of cytosolic and secreted cytokinin dehydrogenase enzymes from maize. *J Exp Bot* 60:2701–2712
- Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goetze NM, Olson BJ, Klenk DC (1985) Measurement of protein using bicinchoninic acid. *Anal Biochem* 150:76–85
- Spíchal L, Rakova NY, Riefler M, Mizuno T, Romanov GA, Strnad M, Schmülling T (2004) Two cytokinin receptors of *Arabidopsis thaliana*, CRE1/AHK4 and AHK3, differ in their ligand specificity in a bacterial assay. *Plant Cell Physiol* 45:1299–1305
- Stolz A, Riefler M, Lomin SN, Achazi K, Romanov GA, Schmülling T (2011) The specificity of cytokinin signalling in *Arabidopsis thaliana* is mediated by differing ligand affinities and expression profiles of the receptors. *Plant J* 67:157–168
- Storoni LC, McCoy AJ, Read RJ (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* 60:432–438
- Suttle JC, Mornet R (2005) Mechanism-based irreversible inhibitors of cytokinin dehydrogenase. *J Plant Physiol* 162:1189–1196
- Suzuki T, Miwa K, Ishikawa K, Yamada H, Aiba H, Mizuno T (2001) The *Arabidopsis* sensor His-kinase, AHK4, can respond to cytokinins. *Plant Cell Physiol* 42:107–113
- The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC, New York
- Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem* 31:455–461
- Vyroubalová Š, Václavíková K, Turečková V, Novák O, Šmečilová M, Hluska, T, Ohnoutková L, Frébort I, Galuszka P (2009) Characterization of new maize genes putatively involved in cytokinin

- metabolism and their expression during osmotic stress in relation to cytokinin levels. *Plant Physiol* 151:433–447
- Werner T, Motyka V, Laucou V, Smets R, Van Onckelen H, Schmülling T (2003) Cytokinin-deficient transgenic *Arabidopsis* plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. *Plant Cell* 15:2532–2550
- Whitty CD, Hall RH (1974) A cytokinin oxidase in *Zea mays*. *Can. J Biochem* 52:789–799
- Wuts PGM, Greene TW (1991) Greene's protective groups in organic synthesis, 4th edition, Wiley, New York
- Yamada H, Suzuki T, Terada K, Takei K, Ishikawa K, Miwa K, Yamashino T, Mizuno T (2001) The *Arabidopsis* AHK4 histidine kinase is a cytokinin-binding receptor that transduces cytokinin signals across the membrane. *Plant Cell Physiol* 42:1017–1023
- Zalabák D, Galuszka P, Mrízová K, Podlešáková K, Gu R, Fréborová J (2014) Biochemical characterization of the maize cytokinin dehydrogenase family and cytokinin profiling in developing maize plantlets in relation to the expression of cytokinin dehydrogenase genes. *Plant Physiol Biochem* 74:283–293
- Zatloukal M, Gemrotová M, Doležal K, Havlíček L, Spíchal L, Strnad M (2008) Novel potent inhibitors of *A. thaliana* cytokinin oxidase/dehydrogenase. *Bioorg Med Chem* 16:9268–9275

Appendix I

Nisler J, Kopečný D, Končítíková R, Zatloukal M, **Bazgier V**, Berka K, Zalabák D, Briozzo P, Strnad M, Spíchal L:

Novel thidiazuron-derived inhibitors of cytokinin oxidase/dehydrogenase.

Plant. Mol. Biol., 92(1), 235–248, 2016.

DOI: 10.1007/s11103-016-0509-0

IF = 3.356

A Novel Series of Highly Potent 2,6,9-Trisubstituted Purine Cyclin-Dependent Kinase Inhibitors

Tomáš Gucký,^{*,†} Radek Jorda,[‡] Marek Zatloukal,[†] Václav Bazgier,^{†,§} Karel Berka,[§] Eva Řezníčková,[‡] Tibor Béres,[†] Miroslav Strnad,[‡] and Vladimír Kryštof[‡]

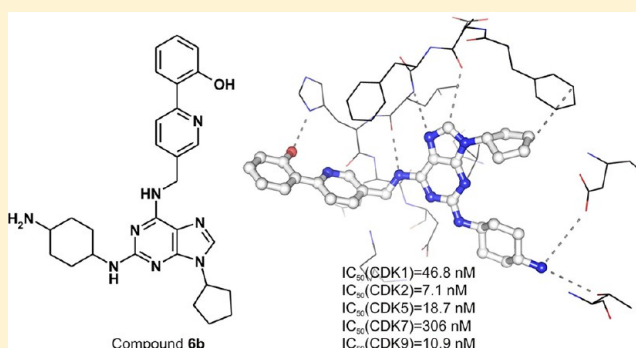
[†]Centre of the Region Haná for Biotechnological and Agricultural Research, Department of Growth Regulators, Faculty of Science, Palacký University, Šlechtitelů 11, 78371 Olomouc, Czech Republic

[‡]Laboratory of Growth Regulators, Faculty of Science, Palacký University & Institute of Experimental Botany ASCR, Šlechtitelů 11, 78371 Olomouc, Czech Republic

[§]Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacký University Olomouc, 17 Listopadu 12, 77146 Olomouc, Czech Republic

S Supporting Information

ABSTRACT: The inhibition of overactive CDKs during cancer remains an important strategy in cancer drug development. We synthesized and screened a novel series of 2-substituted-6-biarylmethylamino-9-cyclopentylpurine derivatives for improved CDK inhibitory activity and antiproliferative effects. One of the most potent compounds, **6b**, exhibited strong cytotoxicity in the human melanoma cell line G361 that correlated with robust CDK1 and CDK2 inhibition and caspase activation. In silico modeling of **6b** in the active site of CDK2 revealed a high interaction energy, which we believe is due to the 6-heterobiarylmethylamino substitution of the purine moiety.



■ INTRODUCTION

The cyclin-dependent kinases (CDKs) are pivotal regulators of the cell cycle. They are activated in a cell cycle-specific manner by cyclins and phosphorylate targets including the transcriptional regulators that in turn induce phase-specific gene expression, stimulate DNA replication, or initiate mitosis.¹ Because of their frequent deregulation in cancer cells, CDKs have been viewed as valid drug targets and, to this date, more than 20 inhibitors have entered clinical trials in cancer patients.^{2,3} The potency of CDK inhibitors designed around a purine heterocycle scaffold is largely determined by the purine's 2,6,9-substitution pattern but can also be affected by the nitrogen atom distribution in the heterocyclic core. Roscovitine (Figure 1), one of the first discovered CDK inhibitors, was identified by optimizing a 2,6,9-trisubstituted purine library.⁴ The optically pure *R*-enantiomer of roscovitine is currently being evaluated as an oncology drug candidate in patients diagnosed with nonsmall cell lung cancer and nasopharyngeal cancer⁵ or other malignancies.⁶ Using structure–activity relationship studies together with a good knowledge of ATP binding sites in various CDKs, the modification of the roscovitine molecule in its substitutable positions has given rise to purine derivatives with increased CDK inhibitory activity and cytotoxicity.^{7–13} In particular, derivatives bearing biarylmethylamino or biaryl amino substituents in the 6-position of

the purine skeleton, such as *R*-CR8 (Figure 1), are among the most active known CDK inhibitors.^{11,12}

The modifications of roscovitine-like inhibitors have not been limited to its side chains but have also touched its heterocyclic core [for review, see ref 14]. Related compounds that exhibit anti-CDK activity include trisubstituted pyrazolo[1,5-*a*][1,3,5]triazines,^{15,16} pyrazolo[4,3-*d*]pyrimidines,¹⁷ and pyrazolo[1,5-*a*]pyrimidines.^{18–20} Notably, the optimization of compounds from the latter of these classes has yielded another drug candidate, dinaciclib (Figure 1).^{21,22} Dinaciclib is at a relatively advanced stage of development: it is currently being evaluated in several phase I/II experiments in patients with various solid tumors and leukemias and also in phase III clinical trials in chronic lymphocytic leukemia patients.^{23–25}

Although many selective CDK inhibitors have been tested in clinical trials, none have yet been approved, largely due to undesired side effects arising from their unfavorable toxicological properties.^{26–29} Therefore the identification of new active compounds and pharmacophores that inhibit CDKs is still a meaningful challenge. The aim of this work was to find new pharmacophores that would increase the desired biological activities of known 2,6,9-trisubstituted purine CDK inhibitors. To this end, we synthesized and performed a SAR study on

Received: May 9, 2013

Published: July 6, 2013

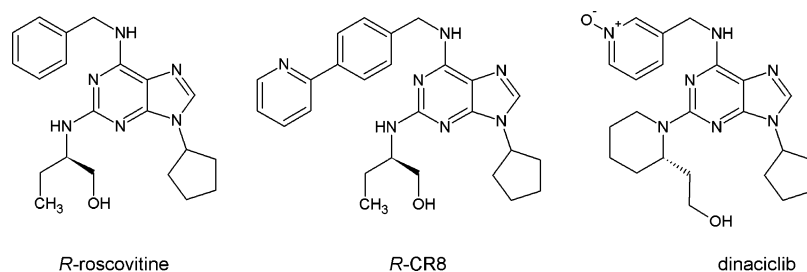
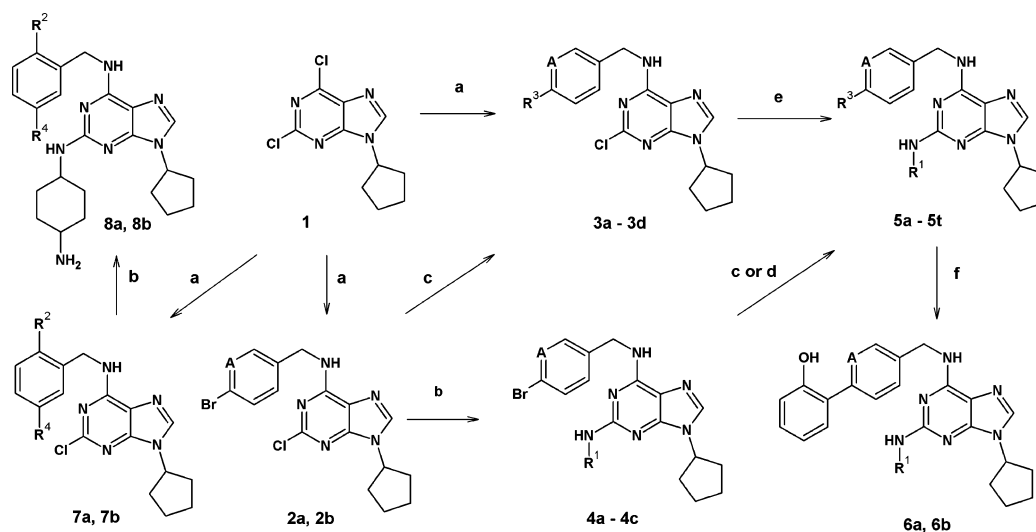


Figure 1. Structures of selected CDK inhibitors with a purine and isosteric pyrazolo[1,5-*a*]pyridine core.

Scheme 1^a



^aReagents and conditions: (a) appropriate amine, DIPEA, *n*-propanol, 80–120 °C (sealed tube); (b) appropriate amine, 160 °C (sealed tube); (c) appropriate arylboronic acid, Pd(OAc)₂, K₃PO₄, TBAB, DMF, 80–120 °C; (d) appropriate arylboronic acid, Pd(dba)₂, PPh₃, Na₂CO₃, DME, water, 80–120 °C; (e) appropriate amine, DIPEA, NMP, 160 °C, 16–36 h; (f) (1) BBr₃, DCM, rt, 18 h, (2) methanol.

novel 2-substituted-6-biarylmethylamino-9-cyclopentyl-9H-purine derivatives and found that the combination of a 9-cyclopentyl and 6-heterobiarylmethylamino substitution significantly increased CDK inhibitory activity and cytotoxic effects in cancer cell lines compared to previously described compounds of this class, including the biaryl-substituted derivative CR8.³⁰

RESULTS AND DISCUSSION

Chemistry. The target compounds **5**, **6**, and **8** were prepared by conventional chemical procedures which allowed considerable flexibility with respect to the identity of the substituents in the 2- and 6-positions of the purine moiety. The general synthetic approach is outlined in Scheme 1. Supporting Information Table S2 shows the synthesized derivatives **2a**, **2b**, **3a–3d**, **4a–4c**, **5a–5t**, and **6a**, **6b**, **7a**, **7b** and **8a**, **8b**, together with the yields in which they were obtained and the reaction conditions used.

The *C*-(6-bromopyridin-3-yl)methylamine was prepared via a multistep route from 2-bromo-5-methylpyridine, which was initially brominated with *N*-bromosuccinimide to give 2-bromo-5-bromomethylpyridine. This was then reacted with urotropine to give a salt, which was hydrolyzed in aqueous ammonia to yield the desired amine.³¹

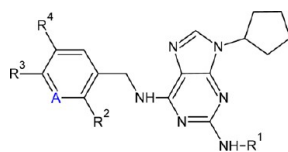
The synthesis of target compounds **5a–5t**, **6a**, **6b**, and **8a**, **8b** started from commercially available 2,6-dichloropurine, which was alkylated with cyclopentanol under Mitsunobu conditions³²

to obtain 9-cyclopentyl-2,6-dichloro-9H-purine (**1**). The alkylation proceeded with over 90% selectivity for the N9 isomer; a small amount of the N7 isomer was formed but could be removed by a single crystallization from ethanol.

The substitution of the 6-position 9-cyclopentyl-2,6-dichloro-9H-purine (**1**) with 4-bromobenzylamine or *C*-(6-bromopyridin-3-yl)methylamine gave compounds **2a** and **2b**. In addition, the reactions of (**1**) with commercially available 1-[4-(furan-2-yl)phenyl]methanamine, 1-[6-(furan-2-yl)pyridin-3-yl]methanamine, 1-[6-(thiophen-2-yl)pyridin-3-yl]methanamine, and 1-[4-(1*H*-pyrazol-1-yl)phenyl]methanamine gave the desired compounds **3a–3d** in high yield and purity. Compound **3a** has previously been synthesized by the Suzuki coupling of **2a** with 2-furanylboronic acid.

The substitution of the purine moiety at position 2 proceeded smoothly to afford compounds **4a**, **4b**, **5i**, **5o**, **5s**, and **5q** with high yields and in good purity. This was achieved by reacting compound **1** with a large excess of *trans*-1,4-diaminocyclohexane (or *trans*-4-aminocyclohexan-1-ol in the case of **5s**) at 160 °C for a few hours. However the preparation of compounds **4c** and **5j** required the use of Hunig's base and a longer reaction time. Compound **5j** was only obtained in a moderate yield (37%), and further heating led to the decomposition of the starting material.

The cross-coupling reactions of compounds **4a–4c** with the appropriate aryl or heteroaryl boronic acids for the synthesis of compounds **5a–5c**, **5e**, **5f**, **5j–5m**, **5o**, and **5p** were performed

Table 1. Inhibition of CDK1 and 2 and in Vitro Antiproliferative Activity of Selected Novel 2-Substituted-6-biarylmethylamino-9-cyclopentyl-9H-purines^a

code	A	R1	R2	R3	R4	inhibitory concentration IC ₅₀ (μM)						enzyme inhibition IC ₅₀ (nM)	
						K562	MCF-7	G361	HOS	HCT-116	HeLa	CDK2	CDK1
roscovitine	CH					45.5	12.3	22.4	24.3	14.4	28.6	170	2421
CR8	CH					0.175	0.160	0.503	0.170	0.35	0.145	40.7	90
H717	CH											48	52
4a	CH	4-aminocyclohexyl	-H	-Br	-H	1.168	0.850	1.475	0.995	0.790	1.510		
4b	N	4-aminocyclohexyl	-H	-Br	-H	1.177	0.483	0.640	0.935	0.577	0.633		
5a	CH	4-aminocyclohexyl	-H	phenyl	-H	0.530	0.583	1.020	0.290	0.745	0.240	26	232
5b	CH	4-aminocyclohexyl	-H	2-aminophenyl	-H	1.107	0.977	4.12	1.881	0.930	1.013	20.5	152
5c	CH	4-aminocyclohexyl	-H	2-methoxyphenyl	-H	0.347	0.303	0.545	0.150	0.430	0.333	33	301
5d	CH	4-aminocyclohexyl	-H	3-fluorophenyl	-H	0.290	0.293	0.815	0.211	0.551	0.445	33.5	422
5e	CH	4-aminocyclohexyl	-H	2-furanyl	-H	0.096	0.085	0.164	0.102	0.129	0.071	11.4	148
5f	CH	4-aminocyclohexyl	-H	3-furanyl	-H	0.287	0.247	0.670	0.140	1.035	0.140	13	202
5g	CH	4-aminocyclohexyl	-H	2-thienyl	-H	0.307	0.290	0.655	0.247	0.592	0.358	20	119
5h	CH	4-aminocyclohexyl	-H	3-thienyl	-H	0.150	0.145	0.570	0.120	0.315	0.220	14	183
5i	CH	4-aminocyclohexyl	-H	pyrazol-1-yl	-H	0.119	0.139	0.220	0.103	0.320	0.170	12	58
5j	CH	2-hydroxy-2-methylpropan-1-yl	-H	2-furanyl	-H	16.847	15.263	17.380	21.410	16.660	14.750	478	>999
5k	N	4-aminocyclohexyl	-H	phenyl	-H	0.563	0.483	0.895	0.206	0.563	0.114	31	184
5l	N	4-aminocyclohexyl	-H	2-aminophenyl	-H	0.200	0.117	0.730	0.398	0.635	0.535	10	100
5m	N	4-aminocyclohexyl	-H	2-methoxyphenyl	-H	0.066	0.082	0.140	0.122	0.177	0.039	34	118
5n	N	4-aminocyclohexyl	-H	3-fluorophenyl	-H	0.061	0.033	0.049	0.044	0.074	0.036	14	77
5o	N	4-aminocyclohexyl	-H	2-furanyl	-H	0.067	0.033	0.097	0.044	0.086	0.030	10	50
5p	N	4-aminocyclohexyl	-H	3-furanyl	-H	0.056	0.032	0.097	0.024	0.046	0.015	8	66
5q	N	4-aminocyclohexyl	-H	2-thienyl	-H	0.100	0.092	0.150	0.032	0.097	0.040	4	49
5r	N	4-aminocyclohexyl	-H	3-thienyl	-H	0.150	0.100	0.272	0.089	0.145	0.075	18	169
5s	N	4-hydroxycyclohexyl	-H	2-furanyl	-H	0.303	0.390	0.443	0.180	0.617	0.247	23	215
5t	N	2-hydroxy-2-methylpropan-1-yl	-H	2-furanyl	-H	1.425	1.305	2.095	0.560	1.270	0.667	120	555
6a	CH	4-aminocyclohexyl	-H	2-hydroxyphenyl	-H	2.610	2.910	4.150	1.160	4.350	4.177	68	777
6b	N	4-aminocyclohexyl	-H	2-hydroxyphenyl	-H	0.016	0.018	0.024	0.014	0.035	0.016	7.1	47
8a	CH	4-aminocyclohexyl	-OH	-H	-CL	0.69	0.43	0.39	0.49	0.35	0.75	8.0	23.0
8b	CH	4-aminocyclohexyl	-NH ₂	-H	-H	4.04	2.33	2.07	2.03	2.30	1.03	18.0	93.0

^aBold values are from the literature.

as described elsewhere.³³ We used palladium diacetate as the catalyst in the cross-coupling reactions, potassium phosphate as the base, a tetrabutylammonium bromide phase transfer catalyst, and *N,N*-dimethylformamide as the solvent. This method is rapid and simple because the catalyst is not air-sensitive, the reaction time is short, and the products are readily isolated and purified. The yields obtained in these reactions ranged from 75% to 85%, and the method is readily amenable to scaling up.

However, these ligand-free coupling reactions failed for sulfur-containing heteroaryl boronic acids and 3-fluorophenyl boronic acid. In these cases, we adopted a modified procedure that uses a catalyst system consisting of bis-(dibenzylideneacetone)palladium and triphenylphosphine, sodium carbonate as the base, and a mixture of dimethoxyethane and water as the solvent. This method was used to prepare compounds **5d**, **5g**, **5h**, **5n**, and **5r**.

Compounds **5c** and **5m** were efficiently demethylated using boron tribromide in dichloromethane under mild conditions to yield compounds **6a** and **6b**.

The structure of the newly synthesized compounds was verified using ¹H NMR spectrometry and for selected compounds also ¹³C NMR spectrometry and HPLC-MS. The structure of model compound **5m** was further verified using 2D NMR spectrometry (¹H-¹H COSY, ¹³C NMR, and ¹H-¹³C HSQC). The purity of all synthesized compounds was checked by HPLC-DAD-MS and elemental analysis. All of the synthesized and compounds were of >95% purity.

SAR of CDK Inhibition. All compounds were tested in kinase inhibition assays for their inhibitory potency toward recombinant human CDK1 and CDK2. Almost all of the derivatives' IC₅₀ values were approximately 10 times lower than those for roscovitine (Table 1). The most potent compounds (**5p**, **5q**, and **6b**) had single-digit nanomolar IC₅₀ values against CDK2, i.e., they were approximately 20 times more potent inhibitors of this kinase than roscovitine. The compounds were less potent inhibitors of CDK1, yielding IC₅₀ values greater than 50 nM against this kinase. However, some of the most potent CDK2 inhibitors in the new series (specifically, compounds **5o-q** and **6b**) also had IC₅₀ values against

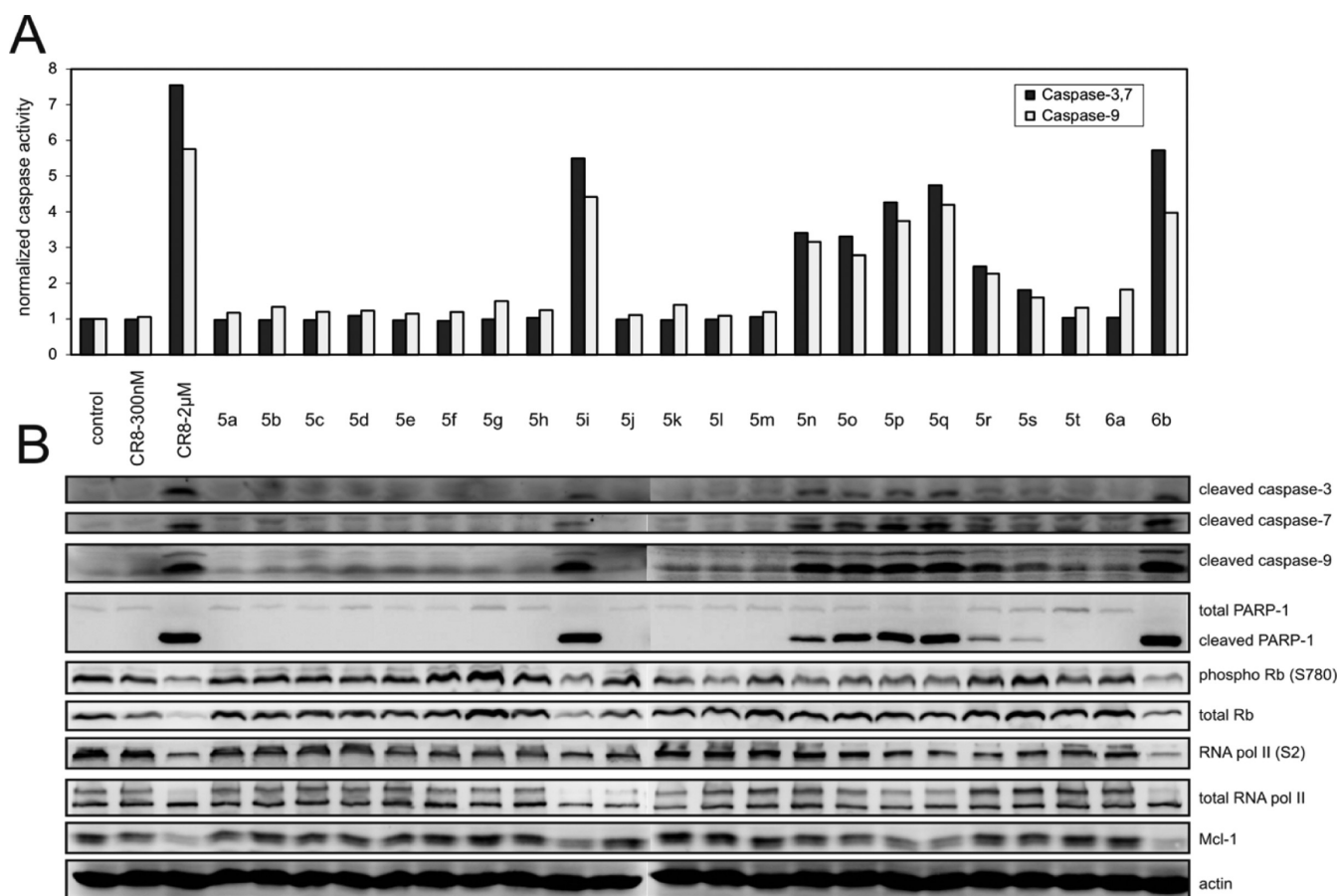


Figure 2. Induction of apoptosis by test compounds after a single 300 nM 24 h dose in G361 melanoma cells. (A) The activities of caspases-3/7/9 were measured in cell lysates using the fluorogenic substrate Ac-DEVD-AMC or Ac-DEVD-AMC (for caspase-9). (B) Immunoblotting analysis of apoptosis proteins and substrates of CDKs. Actin levels are included as controls for equal protein loading. Two concentrations of CR8 were used as an internal control.

CDK1 that were around 40 times lower than that for roscovitine. As with roscovitine, most compounds retained the same ratio between IC_{50} values for CDK1 and CDK2.

The majority of compounds bear the same side groups in positions 2- and 9- of the purine, i.e., 4-aminocyclohexylamino and cyclopentyl, respectively, and a combination of these was found to be advantageous for cancer cell proliferation inhibition. In fact, CDK affinity was reduced when the 4-aminocyclohexylamino side chain at position 2- was replaced with a 4-hydroxycyclohexylamino or hydroxyalkylamino group (compare compound **5o** to **5s** or **5t**). Unexpectedly, **5j** was even less active than roscovitine. This is probably because it adopts a suboptimal position within the active site of CDK2 due to the combined influence of the substituents at the 2- and 6- positions.

Minimizing the diversity of the substituents in the 2- and 9- positions enabled us investigate how varying the substituent at the 6-position of the purine affected activity against CDK1/CDK2. The most potent compounds in this series had substituents with a furanyl (**5e**, **5f**, **5p**, **5o**), thienyl (**5g**, **5h**, **5q**, **5r**), or pyrazolyl (**5i**) ring at the 6-position of the purine. Side chains that incorporated phenyl groups with polar substituents also conferred increased potency against CDK1 and CDK2 (cf. compounds **5l**, **5b**, **6b**). In general, the increased cytotoxicity of the 6-[4-arylpyridin-3-yl] derivatives (A = N) over 6-[4-arylphenyl] derivatives (A = CH) appears to be linked to more effective CDK1/CDK2 inhibition; this is

particularly clear when one compares the IC_{50} values for compound pairs such as **5b** and **5l**, **5d** and **5n**, **5g** and **5q**, **5j** and **5t**, and **6a** and **6b**.

The CDK selectivity of the most potent compounds, **5k**–**5t** and **6b**, was further characterized by performing kinase assays with CDK5, CDK7, and CDK9 at two doses. The data in Supporting Information Table S1 clearly show that these compounds were also potent inhibitors of these kinases. A dose–response assay conducted using the most potent derivative from this series (**6b**) with CDK5, CDK7, and CDK9 yielded IC_{50} values of 19, 306, and 11 nM, respectively. These results show that compounds from the new series can inhibit multiple kinases from the CDK family, which may be beneficial because it means they can exert anticancer effects via multiple pathways.

To further explore compound **6b**'s greater affinity for CDK2 relative to roscovitine and CR8, we used our flexible docking procedure in conjunction with the Autodock Vina molecular docking program to model its binding interactions with CDK2 (see the Apparatus and Methods section for details). Our model suggests that all of the tested compounds bind to the ATP-binding site of CDK2 in an orientation similar to that of roscovitine and other purine inhibitors; the predicted binding energies correlated well with the measured IC_{50} values for kinase inhibition ($r^2 = 0.73$; see Supporting Information Table S3 and Figure S1). Compound **6b** forms conserved hydrogen bonds with the backbone carbonyl and NH moieties of Leu83,

and an additional conserved hydrogen bond is formed with the backbone carbonyl of Glu81 (Figure 3). The 9-cyclopentyl chain is located in the small hydrophobic patch deep in the active site close to the side chain of Phe80, and the 6-biaryl chain points outward from the active site of CDK2. The model shows that the *trans*-4-amino-cyclohexylamino group in the 2-position forms two hydrogen bonds with Asp145 and Thr 14, which may be why it was the best of the tested substituents at the 2-position; alternatives such as a 4-hydroxycyclohexyl group can only form a single hydrogen bond with Asp145, while the roscovitine-like 2-hydroxy-2-methyl-propan-1-yl substituent forms no such contacts (compare **5o** to **5s** and **5t**). The 2-hydroxyphenyl ring of **6b** may interact with His84. Importantly, the basic nitrogen in the pyridine ring of **6b** helps to stabilize the rigid planar conjugated conformation of both rings by forming an H-bond with the hydroxyl group. Of the tested compounds, **6b** has the highest barrier for the rotation of the second ring, followed by **5o**, **5q**, and **5l**; a high barrier to rotation reduces the thermal motion of this fragment (Supporting Information Figure S1 and S2). Planarity is also promoted by the presence of a functional group on the six-membered ring that can interact with the pyridine nitrogen (compare compounds **5l** and **6b** to **5k** and **5m**). Overall, the interactions described above may significantly contribute to the ability of **6b** to bind more tightly and rigidly to CDK2 than roscovitine can.

Compound Antiproliferative Activity in Cancer Cells

Lines. To explore the SAR of our novel compound series, we determined the compounds' cytotoxicity (IC_{50}) in representative cancer cell lines incubated with increasing compound concentrations for 72 h. We found that individual compounds exhibited similar levels of activity across all cell lines tested (Table 1), with IC_{50} values ranging from 14 nM to 4 μ M for both series **5** and **6**. In general, the most active CDK inhibitors identified using the kinase inhibition assays (see below and Table 1) had a 4-aminocyclohexylamino chain in the 2-position. Compounds of this type generally had midnanomolar IC_{50} values. To verify the suitability of the 4-amino-cyclohexylamino substitution in position 2- of the purine core, we synthesized derivatives bearing 2-hydroxy-2-methylpropan-1-yl-amino (**5j**, **5t**) and 4-hydroxy-cyclohexyl-amino (**5s**) groups in the 2-position. All of these derivatives proved to be more antiproliferative than the original compounds. The most potent derivatives **5n–p** and **6b** have IC_{50} values of less than 100 nM. The least potent compounds from this series, **5b** and **6a**, both have similar substituents on the 6-position of the second aryl ring (a 2-aminophenyl and a 2-hydroxyphenyl group, respectively), which may be responsible for their comparatively poor activity.

Cells in our cytotoxicity assay displayed morphological signs of apoptosis (data not shown), prompting us to evaluate the ability of the prepared compounds to induce apoptosis at 300 nM in the melanoma cell line G361 (Figure 2). We analyzed caspase 3/7 and 9 activity using a fluorimetry-based assay in cell lysates (Figure 2A). In parallel, levels of known CDK substrates and apoptosis proteins were determined by immunoblotting (Figure 2B). The potent CDK inhibitors **5i**, **5n–5r**, and **6b** induced strong caspase activity (Figure 2A) and caused the cleavage of caspases and their substrate PARP, along with decreases in the abundance of the antiapoptotic protein Mcl-1 (Figure 2B). Importantly, the control compound CR8 induced caspase-dependent cell death in G361 at micromolar (2 μ M) but not nanomolar (300 nM) concentrations.

Along with the apoptosis-related markers, we also monitored the phosphorylation of retinoblastoma protein and RNA polymerase II at Ser780 and Ser2, sites known to be phosphorylated by CDK2/4 and CDK9, respectively (Figure 2B). The level of phosphorylation at both sites was reduced in the treated cells, particularly in those exposed to potent compounds such as **5i**, **5n–5q**, and **6b**. For compounds **5i** and **6b**, which strongly activate caspases at the tested doses, this could have been caused by reductions in the total abundance of Rb or RNA polymerase II.

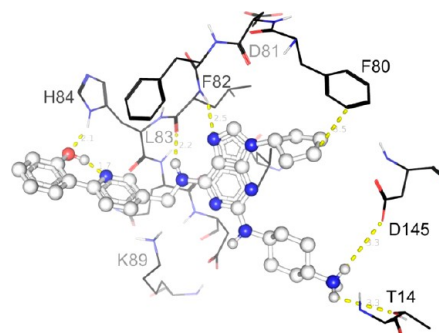


Figure 3. Structure of the binding pose of compound **6b** in CDK2 (template PDB ID: 2A4L). Lines represent important amino acid residues of CDK2. The ligand is shown in a ball and stick representation, with all heteroatoms shown in black. Interactions with K89, T14, D145, E8, and with the backbone of L83 and E81 are illustrated with gray dots.

Interestingly, the potent CDK2 inhibitors **5k** and **5l** blocked the phosphorylation of Rb but not RNA polymerase II. This correlates with their weak inhibition of CDK9 (see Supporting Information Table S1), an RNA polymerase II activator. Neither **5k** nor **5l** induces apoptosis in G361 cells, and because CDK inhibitor cytotoxicity is supposedly based on the inhibition of the RNA polymerase II activators CDK7 and CDK9, **5k** and **5l** should be considered to be cytostatic rather than cytotoxic agents.³⁴

Molecular Mechanisms of Activity for Compound 6b.

Compound **6b** emerged as one of the most potent CDK inhibitors from this series, being much more potent than the parent compounds roscovitine and CR8 at the same dose. Because of its superior cytotoxicity in the cell lines tested here, its clear cell cycle suppression (Supporting Information Figure S3a), and its significant induction of apoptosis (Supporting Information Figure S3b), we decided to characterize the cellular effects of **6b** in more detail.

We treated G361 cells with increasing doses of **6b** for 24 h and immunoblotted the lysates with antibodies against proteins involved in apoptosis (Figure 4A). Treated cells showed elevated levels of cleaved (activated) caspases 3 and 7 and a concomitant decrease in the levels of the inactive forms of caspases 3 and 7. The presence of an 89 kDa cleavage fragment of PARP, the known substrate of caspase 3, in cells treated with doses as low as 80 nM, correlates well with the activation of caspases. Moreover, treated cells exhibited dose-dependent reductions in the abundance of the antiapoptotic protein Mcl-1, which is consistent with the known mechanisms of action for other CDK inhibitors.^{17,35,36} Compound **6b** also causes a dose-dependent increase in the abundance of p53, which is known to be activated and stabilized by CDK inhibitors^{35,37,38} and whose induction results in apoptosis via the mitochondrial pathway.

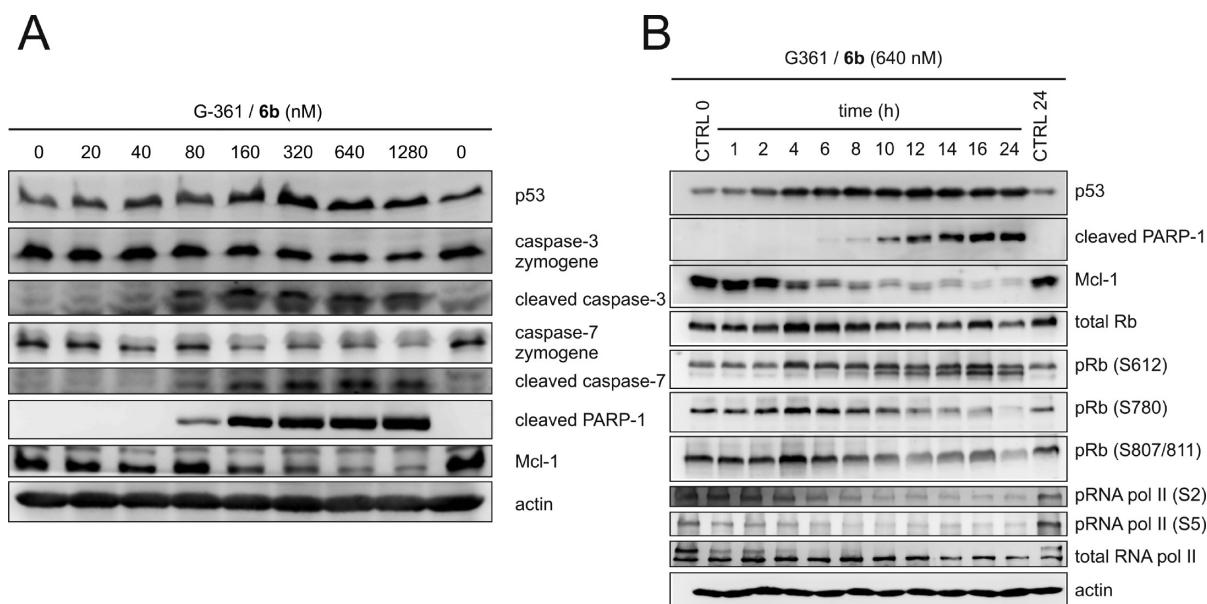


Figure 4. Compound **6b** induces apoptosis and dephosphorylation of Rb protein in G361 cells in a (A) dose and (B) time-dependent manner. (A) The cleavage of caspases, fragmentation of PARP, total p53 protein level, and down-regulation of Mcl-1 were detected by immunoblot analysis. (B) Immunoblotting analysis of selected apoptotic proteins and Rb protein in G361 cells treated with 640 nM of **6b** for the indicated times. Actin was included as a control for equal protein loading.

Our data therefore indicates that in G361 cells, **6b** induces the mitochondrial pathway of apoptosis.

To study the kinetics of its cellular effects, **6b** (640 nM) was added to G361 cells for periods of between 1 and 24 h (Figure 4B). The 640 nM dose was selected because it strongly activated apoptosis within 24 h (Figure 4A), and we therefore anticipated that it would cause pronounced changes in the levels of important cellular markers at earlier time points. Compound **6b** gradually reduced the levels of phosphorylated Rb and RNA polymerase II over 24 h, with the first signs of dephosphorylation appearing within 1 h of treatment. The dephosphorylation of Rb at Ser612 and Ser780 and RNA polymerase II at Ser2 and Ser5 suggests that at least CDK2, CDK7, and CDK9 were effectively inhibited. However, the levels of both substrates started decreasing after 16 h. This was probably due to the initiation of cell death, as evidenced by the strong cleavage of PARP at this time point (Figure 4B) and by the activation of caspase-3,7 observed in the enzymatic assay (Supporting Information Figure S2). In our time-course experiment, CDK substrate dephosphorylation preceded changes in the apoptotic markers. This implies that compound **6b** may also induce cell death by transcriptional repression: an inhibition of transcriptional CDKs would cause a reduced transcription of their target genes and a subsequent down-regulation of proteins with a rapid turnover like Mcl-1, which we note became significantly less abundant within 2 h of exposure to the **6b**.

To determine whether the antiproliferative activity of **6b** may involve mechanisms independent of CDK2 or CDK9, we investigated the phosphorylation of less explored CDK1 and CDK2 substrates: protein phosphatase-1 α (PP1 α) and nucleophosmin (NPM).^{39,40} These proteins are phosphorylated during mitosis, but any changes in their phosphorylation caused by CDK inhibitors are barely detectable in asynchronous cells, where approximately 10% of cells undergo mitosis. We therefore treated G361 cells with nocodazole to block mitosis (94%, data not shown) and visualized the phosphor-

ylation of PP1 α and NPM at threonines 320 and 199, respectively (Figure 5, second lane). When these cells were dosed with **6b** for 4 h, the abundance of the phospho-forms of both PP1 α and NPM decreased (Figure 5). After 8 h of treatment, NPM phosphorylation returned to asynchronous levels and its dephosphorylation due to **6b** exposure was no longer evident. However, the difference in PP1 α phosphorylation remained.

Additionally we observed decreases in the phosphorylation of CDK1 (T161) and CDK2 (T160).⁴¹ These sites are phosphorylated by CDK7, and their lower levels of phosphorylation are consistent with the inhibition of CDK7 by compound **6b** (Supporting Information Table S1).

DISCUSSION AND CONCLUSIONS

There is considerable controversy regarding the suitability of CDKs as targets for anticancer drugs. Experiments in mice have shown that most CDKs are dispensable for the cell cycle,⁴² confirming earlier findings that some cancer cells proliferate well even without CDK2.^{34,43} However, other studies have suggested that CDK2 might be a good target in the treatment of melanoma,^{44,45} which has recently been shown to be sensitive to dinaciclib both in vitro and in vivo.⁴⁶ Dinaciclib is a nanomolar CDK inhibitor that was developed as a bioisostere of roscovitine and is currently undergoing phase II and III clinical trials.^{19,21–25,47}

The experiments described herein demonstrate that 2-substituted-6-biarylmethylamino-9-cyclopentylpurines are strong CDK inhibitors; the prototype compound **6b** proved to be even more potent than the experimental purine-based inhibitor CR8^{30,48} and the pyrazolo[1,5-*a*]pyrimidine drug candidate dinaciclib, which is active in melanoma cell lines at micromolar concentrations.⁴⁶ Its remarkable potency was achieved by introducing a cyclopentyl group in the 9-position of the purine unit (as demonstrated by its superior performance compared to compounds **8a** and **8b**, which have 9-isopropyl units instead of the cyclopentyl group¹³), along with a 4-

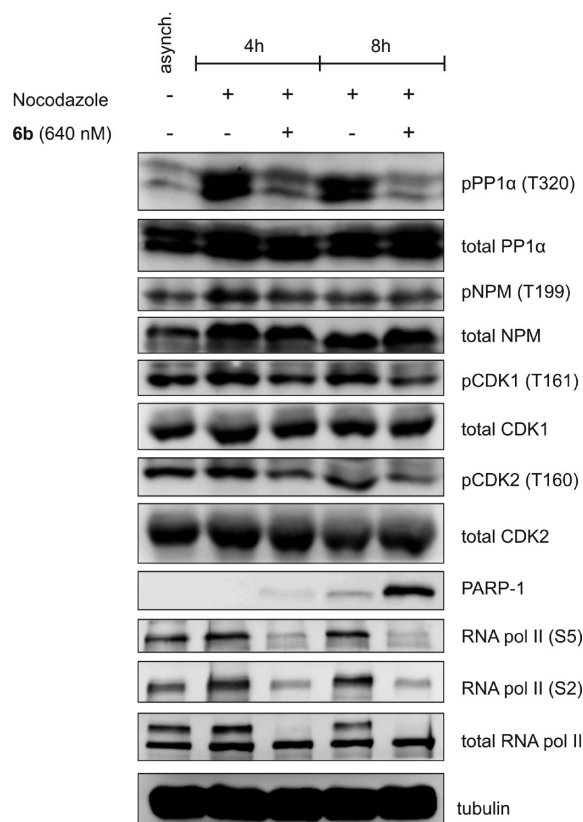


Figure 5. Compound **6b** inhibits the phosphorylation of CDK2- or CDK9-independent substrates. G361 cells were first synchronized with nocodazole (3 ng/mL for 16 h) and then treated with 640 nM **6b** for the indicated times in the presence of nocodazole. Substrates of CDKs and other cell cycle regulatory proteins were detected by immunoblotting analysis. Tubulin was included as a control for equal protein loading.

aminocyclohexyl group in the 2-position, and by modifying the 6-substituent that confers CDK selectivity.^{10–12,30,32,49}

A theoretical model of compound **6b** bound to the active site of CDK2 revealed that it has one of the highest interaction energies of the whole series with this active site. This is because its purine unit has an optimal substitution pattern arising from the presence of the (2-hydroxy-phenyl)-pyridin-3-ylmethylamino group in the 6-position, a cyclopentyl group in the 9-position, and a 4-aminocyclohexyl group in the 2-position of the purine. Its increased inhibitory activity is also partly due to the nitrogen in the pyridine ring (**5k–5r**, **6b**), which allows the rings of the substituent at position 6 to assume a more planar conformation and imposes a higher barrier to rotation than is present in the analogous phenyl derivatives (**5a–5h**), thereby restricting the thermal motion of the two rings. This enforced planarity is especially pronounced when five-membered rings are added on to the pyridine ring (**5o–5q**) or when an additional group on the second six-membered ring can interact with the pyridine nitrogen (compare **5l**, **6b** with **5k**, **5m**). Compound **6b**'s second ring has the highest barrier to rotation of the entire series, followed by those of **5o**, **5l**, and **5q**. The 2-thienyl ring (**5q**) has an advantage over the 2-furanyl ring (**5o**) due to the stronger interaction of nitrogen with its sulfur. A similar skeleton-stabilizing N–S interaction has also been found in MAPK and aurora A inhibitors.⁵⁰

The high theoretical and in vitro activity of the novel biaryl purine derivatives described herein is consistent with their in

vivo performance. In addition, more recent experiments have demonstrated that **6b** and related compounds have strong antitumorigenic effects in mouse models of hepatocarcinoma (manuscript submitted to *Mol. Cancer Ther.*). The new purine CDK inhibitors impinge on multiple cancer phenotypes: they are not only directly cytotoxic in cancer cells but also block their invasion and migration and, in addition, interfere with angiogenesis, probably via CDK-dependent noncell cycle mechanisms. For example, CDK7 limits the migration of cancer cell lines⁵¹ and CDK5 is critical for endothelial cell migration and angiogenesis.⁵² Because the novel inhibitors are pan-selective and act on all of these CDKs, their interference with multiple malignant hallmarks of cancer makes them attractive drug candidates that should be considered for clinical evaluation.

EXPERIMENTAL SECTION

Apparatus and Methods. Melting points were determined on a Boetius stage and are corrected. ¹H NMR spectra were measured in DMSO-*d*₆ or CDCl₃ at 300 K on a Bruker Avance 300 spectrometer (300 MHz); chemical shifts are reported in ppm and coupling constants in Hz. Mass spectra were recorded by using an LCQ ion trap mass spectrometer (Finnigan MAT, San Jose, CA, USA). The chromatographic purity of the compounds were determined using HPLC-DAD-MS. An Alliance 2695 separations module (Waters) linked simultaneously to a PDA 996 (Waters) and a Q-ToF micro (Waters) benchtop quadrupole orthogonal acceleration time-of-flight tandem mass spectrometer were used. Samples were dissolved in methanol and diluted to a concentration of 10 μg·mL⁻¹ in the mobile phase (initial conditions). Then 10 μL of the solution were injected on a RP-column (150 mm × 2.1 mm; 3.5 μm; Symmetry C18, Waters). The column was kept in a thermostat at 25 °C. Solvent (A) consisted of 15 mM formic acid adjusted to pH 4.0 by ammonium hydroxide. Methanol was used as the organic modifier (solvent B). At flow rate of 0.2 mL·min⁻¹, the following binary gradient was used: 0 min, 10% B; 0–24 min, a linear gradient to 90% B, followed by 10 min isocratic elution of 90% B. At the end of the gradient, the column was re-equilibrated to initial conditions for 10 min. The effluent was introduced into the DAD (scanning range 210–400 nm, with 1.2 nm resolution) and an electrospray source (source temperature 110 °C, capillary voltage +3.0 kV, cone voltage +20 V, desolvation temperature 250 °C). Nitrogen was used both as desolvation gas (500 L·h⁻¹) and as cone gas (50 L·h⁻¹). The mass spectrometer was operated in positive (ESI+) ionization mode. The data were acquired in the 50–1000 *m/z* range. Elemental analyses were performed by using an EA 1108 elemental analyzer (Fison Instruments); their values (C, H, N) agreed with the calculated ones within acceptable limits. Merck silica gel Kieselgel 60 (230–400 mesh) was used for column chromatography. The purity of biologically evaluated compounds was >95% as determined by HPLC-DAD-MS and elemental analysis.

Prepared Compounds. *Preparation of 9-Cyclopentyl-2,6-dichloro-9H-purine.* 2,6-Dichloro-9H-purine (30.0 mmol), cyclopentanol (60.0 mmol), and triphenylphosphine (36.0 mmol) were dissolved in dry tetrahydrofuran (120 mL) and cooled to 0 °C. To the stirred solution diisopropyl azodicarboxylate (36.0 mmol) was added dropwise under an argon atmosphere, and the temperature was kept between 0 and 20 °C. The reaction mixture was stirred under an argon atmosphere at 20 °C for a further 2 h. The reaction mixture was then evaporated under reduced pressure, and the residue was dissolved in boiling toluene (100 mL). After cooling to room temperature, the solution was inoculated with a small amount of triphenylphosphine oxide and the solution was kept at 5 °C for 24 h. The triphenylphosphine oxide was filtered off, and the filtrate was evaporated under reduced pressure. The residue was crystallized from ethanol to obtain pure 9-cyclopentyl-2,6-dichloro-9H-purine. Yield 56%; mp 118–120 °C. Elemental analysis Calcd for C₁₀H₁₀Cl₂N₄ (257.12): C, 46.71; H, 3.92; N, 21.79. Found: C, 46.95; H, 3.81; N, 21.70. HPLC-MS (ESI+): 288.10 (99.6%). ¹H

NMR (DMSO- d_6): 1.64–1.69 (m, 2H), 1.81–1.96 (m, 4H), 2.09–2.15 (m, 2H), 4.92 (qui, $J = 7.53$, 1H, CH), 8.82 (s, 1H, CH).

Preparation of C-(6-Bromopyridin-3-yl)methylamine. 2-Bromo-5-methylpyridine (70.0 mmol) and *N*-bromosuccinimide (80.0 mmol) were dissolved in 1,2-dichloroethane (150 mL), and to this mixture 2,2'-azobis(2-ethylpropionitrile) (1.50 mmol) was added. The reaction mixture was heated under reflux at 85 °C for 15 min and the next portion of 2,2'-azobis(2-methylpropionitrile) (1.50 mmol) was added and the reaction mixture was heated at 85 °C for a further 15 min. After cooling to room temperature, the reaction mixture was kept at 5 °C for 2 h and the precipitate was filtered off and washed with a small amount of 1,2-dichloroethane. The filtrate was evaporated under reduced pressure, and the crude product was used for further reaction steps without purification.

The crude 2-bromo-5-bromomethylpyridine was dissolved in chloroform (100 mL), and urotropine (70.0 mmol) was added. The reaction mixture was stirred at room temperature for 16 h. The precipitate was filtered off, washed with a small amount of chloroform, and dried on air. The crude urotropine salt was refluxed in a mixture of concd ammonium hydroxide (12 mL) and water (80 mL) for 90 min, and after cooling to room temperature, 40% formaldehyde (5.0 mL) was added while stirring. The precipitate was filtered off, washed with ice-cold water, and dried in a vacuum desiccator. The crude product was crystallized from ethanol.

Yield 40%; mp 105–106 °C. Elemental analysis Calcd for $C_6H_7BrN_2$ (187.04): C, 38.53; H, 3.77; N, 14.98. Found: C, 38.22; H, 3.72; N, 14.71. HPLC-MS (ESI+): 188.02 (97.2%). 1H NMR (DMSO- d_6): 4.04 (t, $J = 5.67$, 2H, CH_2), 7.71 (d, $J = 8.19$, 1H, ArH), 7.95 (dd, $J = 8.19$, $J' = 1.95$, 1H, ArH), 8.51 (d, $J = 1.95$, 1H, ArH), 8.74 (s(br), 2H, NH_2).

General Procedure A for Preparing Compounds 2a, 2b, 3a–3d, 7a and 7b. To the suspension of 9-cyclopentyl-2,6-dichloro-9H-purine (7.78 mmol) in a mixture of *n*-propanol (40 mL) and *N,N*-diisopropyl-*N*-ethylamine (23.34 mmol), the appropriate amine (8.56 mmol) was added. The suspension was heated with stirring in a sealed tube under an argon atmosphere (for detailed conditions see Table 1). After cooling to room temperature, the reaction mixture was evaporated under reduced pressure and the residue was partitioned between water (50 mL) and dichloromethane (50 mL). In addition, the water phase was extracted twice with dichloromethane. The combined organic phases were washed with water and brine and evaporated under reduced pressure.

(4-Bromobenzyl)-(2-chloro-9-cyclopentyl-9H-purin-6-yl)-amine (2a). Yield 98%; mp 152–154 °C. Elemental analysis Calcd for $C_{17}H_{17}ClBrN_5$ (406.71): C, 50.20; H, 4.21; N, 17.22. Found: C, 50.00; H, 3.99; N, 16.95. HPLC-MS (ESI+): 408 (99.9%). 1H NMR (DMSO- d_6): 1.64–1.69 (m, 2H), 1.81–1.96 (m, 4H), 2.09–2.15 (m, 2H), 4.59 (d, $J = 6.72$, 2H, CH_2), 4.77 (qui, $J = 7.05$, 1H, CH), 7.28 (d, $J = 8.22$, 2H, ArH), 7.49 (d, $J = 8.22$, 2H, ArH), 8.26 (s, 1H, CH), 8.83 (t, $J = 6.72$, 1H, NH).

(6-Bromopyridin-3-ylmethyl)-(2-chloro-9-cyclopentyl-9H-purin-6-yl)-amine (2b). Yield 71%; mp 178–179 °C. Elemental analysis Calcd for $C_{16}H_{16}ClBrN_6$ (407.70): C, 47.14; H, 3.96; N, 20.61. Found: C, 47.35; H, 3.88; N, 20.48. HPLC-MS (ESI+): 409 (98.5%). 1H NMR (DMSO- d_6): 1.64–1.69 (m, 2H), 1.81–1.96 (m, 4H), 2.09–2.15 (m, 2H), 4.61 (s(br), 2H, CH_2), 4.77 (qui, $J = 7.20$, 1H, CH), 7.59 (d, $J = 8.19$, 1H, ArH), 7.70 (d, $J = 8.19$, 1H, ArH), 8.26 (s, 1H, CH), 8.38 (s, 1H, ArH), 8.82 (s(br), 1H, NH). ^{13}C NMR (DMSO- d_6): 10.25, 23.88, 25.98, 33.00, 56.13, 64.74, 118.34, 128.16, 133.25, 138.39, 141.40, 149.87.

(2-Chloro-9-cyclopentyl-9H-purin-6-yl)-(4-furan-2-yl-benzyl)-amine (3a). Yield 85%; mp 135–137 °C. Elemental analysis Calcd for $C_{21}H_{20}ClN_5O$ (393.87): C, 64.04; H, 5.12; N, 17.78. Found: C, 64.25; H, 4.98; N, 17.67. HPLC-MS (ESI+): 394 (97.4%). 1H NMR (CDCl $_3$): 1.72–1.93 (m, 6H), 2.22–2.28 (m, 2H), 4.85–4.92 (m, 3H, CH, CH_2), 6.48 (d, $J = 3.28$, 1H, ArH), 6.55 (s(br), 1H, NH), 6.65 (d, $J = 3.33$, 1H, ArH), 7.40 (d, $J = 8.13$, 2H, ArH), 7.48 (t, $J = 3.33$, 1H, ArH), 7.64–7.69 (m, 3H, ArH, CH).

(2-Chloro-9-cyclopentyl-9H-purin-6-yl)-(6-furan-2-yl-pyridin-3-yl-methyl)-amine (3b). Yield 96%; mp 119–122 °C. Elemental analysis

Calcd for $C_{20}H_{19}ClN_6O$ (394.86): C, 60.84; H, 4.85; N, 21.28. Found: C, 60.56; H, 4.92; N, 21.48. HPLC-MS (ESI+): 396 (97.6%). 1H NMR (CDCl $_3$): 1.76–1.91 (m, 6H), 2.22–2.28 (m, 2H), 4.85–4.92 (m, 3H, CH, CH_2), 6.54 (d, $J = 3.42$, 1H, ArH), 6.59 (s(br), 1H, NH), 7.05 (d, $J = 3.42$, 1H, ArH), 7.53 (d, $J = 3.42$, 1H, ArH), 7.64–7.69 (m, 2H, ArH), 7.75 (d, $J = 6.27$, 1H, ArH) 8.61 (s, 1H, CH).

(2-Chloro-9-cyclopentyl-9H-purin-6-yl)-(6-thiophen-2-yl-pyridin-3-ylmethyl)-amine (3c). Yield 92%; mp 111–114 °C. Elemental analysis Calcd for $C_{20}H_{19}ClN_6S$ (410.92): C, 58.46; H, 4.66; N, 20.45; S, 7.80. Found: C, 58.56; H, 4.72; N, 20.37, S, 7.55. HPLC-MS (ESI+): 411.3 (97.3%). 1H NMR (DMSO- d_6): 1.61–1.70 (m, 2H), 1.78–1.96 (m, 4H), 2.09–2.16 (m, 2H), 4.64 (d, $J = 5.37$, 2H, CH_2), 4.77 (qui, $J = 7.20$, 1H, CH), 7.14 (t, $J = 4.52$, 1H, ArH), 7.59 (d, $J = 5.01$, 1H, ArH), 7.74 (d, $J = 5.01$, 1H, ArH), 7.80 (d, $J = 4.52$, 1H, ArH), 7.85 (d, $J = 4.52$, 1H, ArH), 8.27 (s, 1H, CH), 8.51 (s, 1H, ArH), 8.87 (t, $J = 5.37$, 1H, NH).

(2-Chloro-9-cyclopentyl-9H-purin-6-yl)-(4-pyrazol-1-yl-benzyl)-amine (3d). Yield 72%; mp 165–167 °C. Elemental analysis Calcd for $C_{20}H_{19}ClN_6O$ (394.86): C, 60.84; H, 4.85; N, 21.28. Found: C, 60.56; H, 4.92; N, 21.48. HPLC-MS (ESI+): 394.3 (97.6%). 1H NMR (DMSO- d_6): 1.61–1.71 (m, 2H), 1.80–1.98 (m, 4H), 2.09–2.18 (m, 2H), 4.66 (d, $J = 5.25$, 2H, CH_2), 4.77 (qui, $J = 7.05$, 1H, CH), 6.51 (t, $J = 2.16$, 1H, ArH), 7.45 (d, $J = 8.37$, 2H, ArH), 7.71 (d, $J = 2.16$, 1H, ArH), 7.77 (d, $J = 8.37$, 2H, ArH), 8.27 (s, 1H, CH), 8.43 (d, $J = 2.16$, 1H, ArH), 8.86 (t, $J = 5.25$, 1H, NH).

4-Chloro-2-[(2-chloro-9-cyclopentyl-9H-purin-6-ylamino)-methyl]-phenol (7a). Yield 96%; mp 121–123 °C. Elemental analysis Calcd for $C_{17}H_{17}Cl_2N_5O$ (378.26): C, 53.98; H, 4.53; N, 18.51. Found: C, 54.09; H, 4.59; N, 18.11. HPLC-MS (ESI+): 379.3 (98.5%). 1H NMR (DMSO- d_6): 1.67–1.71 (m, 2H), 1.75–1.98 (m, 4H), 2.09–2.13 (m, 2H), 4.56 (d, $J = 5.19$, 2H, CH_2), 4.79 (qui, $J = 7.20$, 1H, CH), 6.84 (d, $J = 8.43$, 1H, ArH), 7.05–7.12 (m, 2H, ArH), 8.29 (s, 1H, CH), 8.64 (t, $J = 5.19$, 1H, NH), 9.97 (s, 1H, OH).

(2-Aminobenzyl)-(2-chloro-9-cyclopentyl-9H-purin-6-yl)-amine (7b). Yield 97%; mp 128–130 °C. Elemental analysis Calcd for $C_{17}H_{19}ClN_6$ (342.83): C, 59.56; H, 5.59; N, 24.51. Found: C, 59.59; H, 5.32; N, 24.33. HPLC-MS (ESI+): 343.3 (96.6%). 1H NMR (DMSO- d_6): 1.67–1.71 (m, 2H), 1.75–1.98 (m, 4H), 2.09–2.14 (m, 2H), 4.44 (d, $J = 5.13$, 2H, CH_2), 4.77 (qui, $J = 7.44$, 1H, CH), 5.21 (s(br), 2H, NH_2), 6.48 (t, $J = 7.41$, 1H, ArH), 6.64 (d, $J = 7.86$, 1H, ArH), 6.93 (t, $J = 7.23$, 1H, ArH), 7.07 (d, $J = 7.23$, 1H, ArH), 8.26 (s, 1H, CH), 8.67 (t, $J = 5.13$, 1H, NH).

General Procedure B for Preparing Compounds 4a, 4b, 5i, 5o, 5q, 8a, and 8b. Well powdered bromoderivative 2 or 3 (7.36 mmol) and *trans*-1,4-diaminocyclohexane (110.0 mmol) were mixed and heated with stirring in a sealed tube under an argon atmosphere (for detailed conditions see Table 1). After cooling to 100 °C, water (50 mL) was added to the reaction mixture and the resulting suspension was extracted three times with ethyl acetate (50 mL). Combined organic phases were washed with water and brine, dried over sodium sulfate, and evaporated under reduced pressure. The residue was dissolved in ethyl acetate (10 mL) and triturated with diethyl ether to obtain a white crystalline mass, which was filtered off and dried at 80 °C for 4 h. The crude product was purified by column chromatography, mobile phase chloroform–methanol (19:1, v/v), if necessary.

N^2 -(4-Aminocyclohexyl)- N^6 -(4-bromobenzyl)-9-cyclopentyl-9H-purine-2,6-diamine (4a). Yield 91%; mp 123–124 °C. Elemental analysis Calcd for $C_{23}H_{30}BrN_7$ (484.44): C, 57.02; H, 6.24; N, 20.24. Found: C, 56.88; H, 6.19; N, 20.02. HPLC-MS (ESI+): 484.2, 486.1 (100.0%). 1H NMR (DMSO- d_6): 0.85–1.22 (m, 4H), 1.64–2.04 (m, 12H) 3.29–3.37 (m, 3H, CH, NH_2), 3.52 (sex, $J = 7.11$, 1H, CH), 4.57 (s(br), 2H, CH_2), 4.62 (qui, 1H, $J = 7.38$, CH), 6.02 (d, $J = 7.89$, 1H, NH), 7.28 (d, $J = 8.31$, 2H, ArH), 7.46 (d, $J = 8.31$, 2H, ArH), 7.73 (s, 1H, CH), 7.84 (s(br), 1H, NH).

N^2 -(4-Aminocyclohexyl)- N^6 -(6-bromo-pyridin-3-ylmethyl)-9-cyclopentyl-9H-purine-2,6-diamine (4b). Yield 33%; mp 114–116 °C. Elemental analysis Calcd for $C_{22}H_{29}BrN_8$ (485.42): C, 54.43; H, 6.02; N, 23.08. Found: C, 54.29; H, 6.15; N, 23.00. HPLC-MS (ESI+): 487.2, 488.2 (98.1%). 1H NMR (CDCl $_3$): 1.13–1.29 (m, 4H), 1.50

(s(br), 2H, NH₂), 1.71–2.22 (m, 12H), 2.75 (sep, *J* = 7.43, 1H, CH), 3.67 (sex, *J* = 7.52, 1H, CH), 4.59 (d, *J* = 7.52, 1H, NH), 4.70 (qui, *J* = 7.20, 1H, CH), 4.82 (d, *J* = 7.20, 2H, CH₂), 6.21 (t, *J* = 5.25, 1H, NH), 7.40 (d, *J* = 8.16, 1H, ArH), 7.55 (dd, *J* = 8.16, *J'* = 2.4, 1H, ArH), 8.34 (s(br), 1H, NH), 8.39 (s, 1H, CH). ¹³C NMR (CDCl₃): 24.25, 31.99, 32.66, 35.21, 41.30, 50.11, 50.27, 55.56, 114.53, 127.91, 134.77, 135.72, 137.96, 140.67, 149.49, 152.52, 154.58, 158.77.

N²-(4-Aminocyclohexyl)-9-cyclopentyl-N⁶-(4-pyrazol-1-yl-benzyl)-9H-purine-2,6-diamine (5i). Yield 78%; mp 186–187 °C. Elemental analysis Calcd for C₂₆H₃₃N₉ (471.60): C, 66.22; H, 7.05; N, 26.73. Found: C, 66.48; H, 7.24; N, 16.51. HPLC-MS (ESI+): 472.4 (99.8%). ¹H NMR (DMSO-*d*₆): 1.02–1.21 (m, 4H), 1.64–2.05 (m, 12H), 2.90–3.15 (m, 3H, CH, NH₂), 3.59 (sex, *J* = 5.05, 1H, CH), 4.58–4.67 (m, 3H, CH₂, CH), 6.05 (d, *J* = 7.29, 1H, NH), 6.51 (t, *J* = 2.28, 1H, ArH), 7.45 (d, *J* = 8.34, 2H, ArH), 7.70–7.86 (m, 4H, ArH), 7.95 (s(br), 1H, NH), 8.42 (d, *J* = 7.29, 1H, CH). ¹³C NMR (DMSO-*d*₆): 24.26, 31.93, 32.70, 34.83, 50.01, 50.30, 55.47, 107.61, 114.56, 119.35, 126.83, 128.67, 135.45, 137.78, 139.32, 141.08, 154.84, 158.90.

N²-(4-Aminocyclohexyl)-9-cyclopentyl-N⁶-(6-furan-2-yl-pyridin-3-ylmethyl)-9H-purine-2,6-diamine (5o). Yield 89%; mp 184–186 °C. Elemental analysis Calcd for C₂₆H₃₄N₈O (472.59): C, 66.08; H, 6.83; N, 23.71. Found: C, 66.32; H, 6.59; N, 23.99. HPLC-MS (ESI+): 473.5 (98.6%). ¹H NMR (CDCl₃): 1.12–1.28 (m, 4H), 1.71–2.15 (m, 12H), 2.60–2.68 (m, 3H, CH, NH₂), 3.68 (sex, *J* = 7.02, 1H, CH), 4.65–4.73 (m, 4H, CH, CH₂, NH), 6.50 (t, *J* = 3.42, 1H, ArH), 6.62 (s(br), 1H, NH), 7.00 (s, 1H, ArH), 7.41–7.69 (m, 4H, ArH), 8.57 (s, 1H, CH).

N²-(4-Aminocyclohexyl)-9-cyclopentyl-N⁶-(6-thiophen-2-yl-pyridin-3-ylmethyl)-9H-purine-2,6-diamine (5q). Yield 88%; mp 151–153 °C. Elemental analysis Calcd for C₂₆H₃₄N₈S (497.64): C, 63.64; H, 6.98; N, 22.84; S, 6.53. Found: C, 63.72; H, 7.08; N, 23.02; S, 6.28. HPLC-MS (ESI+): 489.4 (99.9%). ¹H NMR (DMSO-*d*₆): 1.04–1.17 (m, 4H), 1.64–2.05 (m, 12H), 3.25–3.38 (m, 3H, CH, NH₂), 3.54 (sex, *J* = 7.83, 1H, CH), 4.59–4.65 (m, 3H, CH₂, CH), 6.09 (d, *J* = 7.83, 1H, NH), 7.13 (t, *J* = 4.05, 1H, ArH), 7.58 (d, *J* = 4.05, 1H, ArH), 7.71–7.84 (m, 4H, ArH), 7.90 (s(br), 1H, NH), 8.51 (s, 1H, CH).

2-[[2-(4-Aminocyclohexylamino)-9-cyclopentyl-9H-purin-6-ylamino]-methyl]-4-chloro-phenol (8a). Yield 85%; mp 134–136 °C. Elemental analysis Calcd for C₂₃H₃₀ClN₇O (455.99): C, 60.58; H, 6.63; N, 21.50. Found: C, 60.69; H, 6.67; N, 21.34. HPLC-MS (ESI+): 448.4 (99.5%). ¹H NMR (DMSO-*d*₆): 1.05–1.22 (m, 4H), 1.60–2.02 (m, 12H), 3.53 (sex, *J* = 7.68, 1H, CH), 4.44–4.52 (m, 3H, CH, CH₂), 4.61 (qui, *J* = 7.17, 1H, CH), 6.07 (d, *J* = 7.68, 1H, NH), 6.77 (d, *J* = 8.70, 1H, ArH), 7.00–7.03 (m, 2H, ArH), 7.62 (s(br), 1H, NH), 7.73 (s, 1H, CH).

N⁶-(2-Aminobenzyl)-N²-(4-amino-cyclohexyl)-9-cyclopentyl-9H-purine-2,6-diamine (8b). Yield 82%; mp 135–136 °C. Elemental analysis Calcd for C₂₃H₃₂N₈ (420.57): C, 65.69; H, 7.67; N, 26.64. Found: C, 65.34; H, 7.84; N, 26.42. HPLC-MS (ESI+): 421.4 (99.6%). ¹H NMR (DMSO-*d*₆): 1.04–1.29 (m, 4H), 1.61–2.03 (m, 12H), 3.34 (sep, *J* = 7.20, 1H, CH), 3.65 (sex, *J* = 6.96, 1H, CH), 4.45 (s(br), 2H, CH₂), 4.65 (qui, *J* = 6.63, 1H, CH), 5.21 (s(br), 2H, NH₂), 6.08 (d, *J* = 7.11, 1H, NH), 6.44 (t, *J* = 7.05, 1H, ArH), 6.58 (d, *J* = 7.77, 1H, ArH), 6.91 (t, *J* = 7.77, 1H, ArH), 7.12 (d, *J* = 7.05, 1H, ArH), 7.57 (s(br), 1H, NH), 7.71 (s, 1H, CH).

General Procedure C for the Preparation of Compounds 3a, 5a–5c, 5e, 5f, 5j, 5k–5m, 5o, and 5p. To the suspension of bromoderivative 2 or 4 (0.21 mmol), the appropriate arylboronic acid (0.31 mmol), potassium phosphate trihydrate (0.80 mmol) and tetrabutylammonium bromide (0.003 mmol) in *N,N*-dimethylformamide (5.0 mL), and palladium diacetate (2.5 μmol) were added under an argon atmosphere. The suspension was heated with stirring in a sealed tube under an argon atmosphere (for detailed conditions see Table 1). After cooling to room temperature, the reaction mixture was diluted with water (20 mL) and the resulting suspension was extracted twice with ethyl acetate (25 mL). Combined organic phases were washed with brine, dried over anhydrous sodium sulfate, and evaporated under reduced pressure. The crude product was purified

by column chromatography on silica, mobile phase chloroform–methanol–concd ammonium hydroxide (9:1:0.05).

(2-Chloro-9-cyclopentyl-9H-purin-6-yl)-(4-furan-2-yl-benzyl)-amine (3a). Yield 55%; mp 135–137 °C.

N²-(4-Aminocyclohexyl)-9-cyclopentyl-N⁶-biphenyl-4-ylmethyl-9H-purine-2,6-diamine (5a). Yield 78%; mp 146–148 °C. Elemental analysis Calcd for C₂₉H₃₆N₈ (496.65): C, 72.32; H, 7.32; N, 20.36. Found: C, 72.61; H, 7.16; N, 20.19. HPLC-MS (ESI+): 482.4 (99.7%). ¹H NMR (DMSO-*d*₆): 1.06–1.20 (m, 4H), 1.64–2.05 (m, 12H), 3.17–3.30 (m, 3H, NH₂, CH), 3.60 (sex, *J* = 6.71, 1H, CH), 4.60–4.65 (m, 3H, CH, CH₂), 6.03 (d, *J* = 7.20, 1H, NH), 7.35 (t, *J* = 7.29, 1H, ArH), 7.41–7.46 (m, 4H, ArH), 7.56–7.62 (m, 4H, ArH), 7.74 (s, 1H, CH).

N⁶-(2'-Aminobiphenyl-4-ylmethyl)-N²-(4-aminocyclohexyl)-9-cyclopentyl-9H-purine-2,6-diamine (5b). Yield 85%; mp 168–170 °C. Elemental analysis Calcd for C₂₉H₃₆N₈ (496.65): C, 70.13; H, 7.31; N, 22.56. Found: C, 70.32; H, 7.28; N, 22.46. HPLC-MS (ESI+): 497.4 (99.9%). ¹H NMR (DMSO-*d*₆): 1.14–1.26 (m, 4H), 1.72–1.82 (m, 2H), 1.83–1.96 (m, 10H), 1.98–2.22 (m, 4H), 2.71 (sex, *J* = 6.72, 1H, CH), 3.69–3.80 (m, 3H, NH₂, CH), 4.66 (d, *J* = 7.71, 1H, NH), 4.74 (qui, *J* = 7.08, 1H, CH), 4.81 (d, *J* = 5.43, 2H, CH₂), 6.15 (s(br), 1H, NH), 6.76 (d, *J* = 7.41, 1H, ArH), 6.82 (t, *J* = 7.41, 1H, ArH), 7.11 (d, *J* = 7.41, 1H, ArH), 7.15 (t, *J* = 7.41, 1H, ArH), 7.38–7.46 (m, 5H, ArH, CH).

N²-(4-Aminocyclohexyl)-9-cyclopentyl-N⁶-(2'-methoxybiphenyl-4-ylmethyl)-9H-purine-2,6-diamine (5c). Yield 88%; mp 178–180 °C. Elemental analysis Calcd for C₃₀H₃₇N₇O (511.66): C, 70.42; H, 7.29; N, 19.16. Found: C, 70.58; H, 7.10; N, 19.45. HPLC-MS (ESI+): 512.4 (99.8%). ¹H NMR (CDCl₃): ¹H NMR (CDCl₃): 1.14–1.34 (m, 4H), 1.71–2.22 (m, 14H), 2.74 (sep, *J* = 6.33, 1H, CH), 3.78 (sex, *J* = 7.05, 1H, CH), 4.00 (s, 3H, CH₃), 4.59 (d, *J* = 5.87, 1H, NH), 4.71 (qui, *J* = 6.87, 1H, CH), 4.79 (d, *J* = 7.20, 2H, CH₂), 6.12 (s(br), *J* = 7.42, 1H, NH), 7.05 (d, *J* = 8.05, 1H, ArH), 7.09 (t, *J* = 8.05, 1H, ArH), 7.32 (t, *J* = 8.05, 1H, ArH), 7.49 (d, *J* = 8.05, 1H, ArH), 7.38–7.46 (m, 4H, ArH), 8.65 (s, 1H, CH).

N²-(4-Aminocyclohexyl)-9-cyclopentyl-N⁶-(4-furan-2-yl-benzyl)-9H-purine-2,6-diamine (5e). Yield 87%; mp 157–159 °C. Elemental analysis Calcd for C₂₇H₃₃N₇O (471.60): C, 68.76; H, 7.05; N, 20.79. Found: C, 68.81; H, 7.22; N, 20.51. HPLC-MS (ESI+): 472.4 (99.8%). ¹H NMR (DMSO-*d*₆): 1.04–1.17 (m, 4H), 1.64–2.05 (m, 12H), 2.65–2.72 (m, 3H, CH, NH₂), 3.58 (sex, *J* = 7.55, 1H, CH), 4.58–4.63 (m, 3H, CH, CH₂), 6.04 (d, *J* = 7.55, 1H, NH), 7.28 (d, *J* = 7.89, 2H, ArH), 7.38 (d, *J* = 5.95, 1H, ArH), 7.46 (d, *J* = 7.89, 2H, ArH), 7.61 (d, *J* = 5.95, 1H, ArH), 7.70–7.73 (m, 2H, ArH, NH), 7.95 (s, 1H, CH). ¹³C NMR (DMSO-*d*₆): 24.24, 31.93, 32.70, 34.74, 49.97, 50.29, 55.44, 104.90, 111.69, 114.52, 128.01, 129.27, 129.92, 131.60, 135.36, 138.55, 142.03, 151.63, 153.95, 154.78, 158.91.

N²-(4-Aminocyclohexyl)-9-cyclopentyl-N⁶-(6-furan-3-yl-pyridin-3-ylmethyl)-9H-purine-2,6-diamine (5f). Yield 94%; mp 154–156 °C. Elemental analysis Calcd for C₂₇H₃₃N₇O (471.60): C, 68.76; H, 7.05; N, 20.79. Found: C, 68.52; H, 7.16; N, 20.49. HPLC-MS (ESI+): 472.4 (97.8%). ¹H NMR (DMSO-*d*₆): 1.04–1.17 (m, 4H), 1.64–2.05 (m, 12H), 3.15–3.19 (m, 3H, CH, NH₂), 3.58 (sex, *J* = 7.32, 1H, CH), 4.58–4.63 (m, 3H, CH, CH₂), 6.05 (d, *J* = 7.32, 1H, NH), 6.91 (s, 1H, ArH), 7.34 (d, *J* = 7.92, 2H, ArH), 7.51 (d, *J* = 7.92, 2H, ArH), 7.70–7.73 (m, 2H, ArH), 7.78 (s(br), 1H, NH), 8.12 (s, 1H, CH). ¹³C NMR (DMSO-*d*₆): 24.39, 31.74, 32.28, 35.06, 49.13, 50.13, 50.37, 56.10, 109.24, 125.80, 126.25, 128.32, 130.66, 136.80, 138.31, 139.51, 144.70, 153.10, 158.88.

1-[9-Cyclopentyl-6-(4-furan-2-yl-benzylamino)-9H-purin-2-ylamino]-2-methyl-propan-2-ol (5j). Yield 80%; mp 121–123 °C. Elemental analysis Calcd for C₂₅H₃₀N₆O₂ (446.54): C, 67.24; H, 6.77; N, 18.82. Found: C, 67.59; H, 6.37; N, 18.62. HPLC-MS (ESI+): 447.4 (99.8%). ¹H NMR (CDCl₃): 1.27 (s, 6H, CH₃), 1.70–1.91 (m, 6H), 2.20–2.35 (m, 2H), 3.40 (d, *J* = 6.18, 2H, CH₂), 4.70 (qui, *J* = 5.01, 1H, CH), 4.79 (s(br), 2H, CH₂), 5.29 (s(br), 1H, OH), 5.62 (t, *J* = 6.21, 1H, NH), 7.21–7.63 (m, 7H, ArH), 7.68 (s(br), 1H, NH), 8.63 (s, 1H, CH).

N²-(4-Aminocyclohexyl)-9-cyclopentyl-N⁶-(6-phenyl-pyridin-3-ylmethyl)-9H-purine-2,6-diamine (5k). Yield 86%; mp 136–137 °C. Elemental analysis Calcd for C₂₈H₃₄N₈ (482.62): C, 69.68; H, 7.10; N,

23.22. Found: C, 69.55; H, 7.16; N, 23.28. HPLC-MS (ESI+): 483.5 (99.9%). ¹H NMR (DMSO-*d*₆): 1.05–1.18 (m, 4H), 1.64–2.05 (m, 12H), 2.65–2.72 (m, 3H, CH, NH₂), 3.58 (sex, *J* = 7.55, 1H, CH), 4.67–4.78 (m, 3H, CH, CH₂), 7.27–7.72 (m, 6H, ArH), 7.80 (d, *J* = 7.8, 2H, ArH), 7.70 (s(br), 1H, NH), 7.95 (s, 1H, CH).

*N*²-(4-Aminocyclohexyl)-*N*⁶-[6-(2-aminophenyl)-pyridin-3-ylmethyl]-9-cyclopentyl-9*H*-purine-2,6-diamine (**5l**). Yield 56%; mp 173–175 °C. Elemental analysis Calcd for C₂₈H₃₃N₉ (497.64): C, 67.58; H, 7.09; N, 25.33. Found: C, 67.69; H, 7.19; N, 25.02. HPLC-MS (ESI+): 498.4 (99.9%). ¹H NMR (CDCl₃): 1.14–1.34 (m, 4H), 1.71–2.05 (m, 12H), 2.10–2.23 (m, 4H, 2 × NH₂), 2.75 (sep, *J* = 7.32, 1H, CH), 3.73 (sex, *J* = 7.52, 1H, CH), 4.59 (d, *J* = 7.52, 1H, NH), 4.70 (qui, *J* = 7.20, 1H, CH), 4.82 (d, *J* = 7.20, 2H, CH₂), 5.92 (t, *J* = 7.20, 1H, NH), 6.75–6.81 (m, 2H, ArH), 7.20 (t, *J* = 7.89, 1H, ArH), 7.47–7.51 (m, 2H, ArH), 7.61 (d, *J* = 8.34, 1H, ArH), 7.79 (d, *J* = 8.34, 1H, ArH), 8.63 (s, 1H, CH).

*N*²-(4-Aminocyclohexyl)-*N*⁶-[6-(2-methoxyphenyl)-pyridin-3-ylmethyl]-9-cyclopentyl-9*H*-purine-2,6-diamine (**5m**). Yield 85%; mp 184–186 °C. Elemental analysis Calcd for C₂₉H₃₆N₈O (512.65): C, 67.94; H, 7.08; N, 21.86. Found: C, 67.78; H, 7.01; N, 21.59. HPLC-MS (ESI+): 513.5 (99.6%). ¹H NMR (CDCl₃): 1.14–1.34 (m, 4H), 1.71–2.22 (m, 12H), 2.72 (sep, *J* = 5.87, 1H, CH), 3.75 (sex, *J* = 6.25, 1H, CH), 3.85 (s, 3H, CH₃), 4.61 (d, *J* = 5.87, 1H, NH), 4.69 (qui, *J* = 6.87, 1H, CH), 4.82 (d, *J* = 7.20, 2H, CH₂), 6.00 (s(br), *J* = 7.20, 1H, NH), 7.00 (d, *J* = 8.22, 1H, ArH), 7.09 (t, *J* = 7.32, 1H, ArH), 7.35 (t, *J* = 7.32, 1H, ArH), 7.49 (s, 1H, ArH), 7.70–7.74 (m, 3H, ArH), 8.72 (s, 1H, CH). ¹³C NMR (CDCl₃): 24.21, 31.76, 32.71, 33.82, 41.84, 49.81, 55.43, 55.68, 111.42, 114.51, 121.18, 125.01, 128.89, 130.01, 131.13, 133.23, 135.14, 135.44, 148.64, 152.50, 154.71, 155.03, 156.95, 158.79.

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-(6-furan-2-yl-pyridin-3-ylmethyl)-9*H*-purine-2,6-diamine (**5o**). Yield 81%; mp 180–183 °C. Elemental analysis Calcd for C₂₆H₃₂N₈O (472.59): C, 66.08; H, 6.83; N, 23.71. Found: C, 66.18; H, 6.59; N, 23.88. HPLC-MS (ESI+): 473.5 (99.8%).

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-(6-furan-3-yl-pyridin-3-ylmethyl)-9*H*-purine-2,6-diamine (**5p**). Yield 76%; mp 165–167 °C. Calcd for C₂₆H₃₂N₈O (472.59): C, 66.08; H, 6.83; N, 23.71. Found: C, 66.01; H, 6.93; N, 23.51. HPLC-MS (ESI+): 473.26 (98.6%). ¹H NMR (CDCl₃): 1.12–1.28 (m, 4H), 1.71–2.15 (m, 12H), 2.60–2.68 (m, 3H, CH, NH₂), 3.68 (sex, *J* = 10.00, 1H, CH), 4.65–4.73 (m, 4H, CH, CH₂, NH), 6.50 (t, *J* = 3.42, 1H, ArH), 6.62 (s(br), 1H, NH), 6.91 (s, 1H, ArH), 7.00 (s, 1H, ArH), 7.61–7.73 (m, 3H, ArH), 8.57 (s, 1H, CH).

General Procedure D for the Preparation of Compounds 5d, 5g, 5h, 5n, 5r, and 5o. To the suspension of bromoderivative **4** (0.50 mmol), the appropriate arylboronic acid (1.50 mmol), triphenylphosphine (0.25 mmol), and sodium carbonate (2.0 mmol) in a mixture of 1,2-dimethoxyethane (3.0 mL) and water (2.0 mL), and bis(dibenzylideneacetone)palladium (15.0 μmol) were added under an argon atmosphere. The suspension was heated with stirring in a sealed tube at 120 °C for 65 h under an argon atmosphere. After cooling to room temperature, the reaction mixture was diluted with water (25 mL) and the suspension was extracted twice with ethyl acetate (25 mL). Combined organic phases were washed with brine, dried over anhydrous sodium sulfate, and evaporated under reduced pressure. The residue was purified by column chromatography on silica, mobile phase chloroform–methanol–concd ammonium hydroxide (9:1:0.05).

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-(3'-fluorobiphenyl-4-ylmethyl)-9*H*-purine-2,6-diamine (**5d**). Yield 75%; mp 146–148 °C. Elemental analysis Calcd for C₂₉H₃₄FN₇ (499.63): C, 69.71; H, 6.86; N, 19.62. Found: C, 69.95; H, 7.12; N, 19.45. HPLC-MS (ESI+): 500.4 (99.9%). ¹H NMR (DMSO-*d*₆): 1.02–1.21 (m, 4H), 1.61–2.06 (m, 12H), 2.65–2.72 (m, 3H, CH, NH₂), 3.59 (sex, *J* = 7.19, 1H, CH), 4.60–4.66 (m, 3H, CH₂, CH), 6.01 (d, *J* = 6.60, 1H, NH), 7.12–7.18 (s, 1H, ArH), 7.42–7.48 (m, 5H, ArH), 7.61 (d, *J* = 8.01, 2H, ArH), 7.73 (s, 1H, CH), 7.86 (s(br), 1H, NH). ¹³C NMR (DMSO-*d*₆): 24.35, 31.89, 32.28, 35.83, 43.11, 50.24, 50.53, 55.44, 113.57, 113.85, 114.27, 114.55, 123.08, 123.12, 127.05, 128.44, 129.36,

131.26, 131.38, 131.95, 132.09, 136.32, 137.46, 137.49, 141.43, 143.09, 143.19, 152.40, 155.01, 158.92, 161.62, 164.85.

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-(4-thiophen-2-yl-benzyl)-9*H*-purine-2,6-diamine (**5g**). Yield 71%; mp 225–226 °C. Elemental analysis Calcd for C₂₇H₃₃N₇S (487.66): C, 66.50; H, 6.82; N, 20.11; S, 6.58. Found: C, 66.58; H, 6.51; N, 20.35; S, 6.41. HPLC-MS (ESI+): 488.5 (99.8%). ¹H NMR (CDCl₃): 1.20–1.28 (m, 4H), 1.61–2.22 (m, 14H), 2.71 (sep, *J* = 5.52, 1H, CH), 3.72 (sex, *J* = 7.44, 1H, CH), 4.61 (d, *J* = 7.44, 1H, NH), 4.71 (qui, *J* = 6.36, 1H, CH), 4.78 (d, *J* = 5.25, 2H, CH₂), 5.93 (s(br), 1H, NH), 7.08 (t, *J* = 4.50, 1H, ArH), 7.30 (d, *J* = 4.50, 1H, ArH), 7.38 (d, *J* = 7.95, 2H, ArH), 7.47 (d, *J* = 4.50, 1H, ArH), 7.57 (d, *J* = 7.95, 2H, ArH), 8.63 (s, 1H, CH).

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-(6-thiophen-3-yl-benzyl)-9*H*-purine-2,6-diamine (**5h**). Yield 71%; mp 114–118 °C. Elemental analysis Calcd for C₂₇H₃₃N₇S (487.66): C, 66.50; H, 6.82; N, 20.11; S, 6.58. Found: C, 66.49; H, 7.06; N, 20.39; S, 6.32. HPLC-MS (ESI+): 488.4 (99.9%). ¹H NMR (DMSO-*d*₆): 1.07–1.22 (m, 4H), 1.64–2.04 (m, 12H), 2.62–2.75 (m), 3H, CH, NH₂), 3.58 (sex, *J* = 7.25, 1H, CH), 4.60–4.65 (m, 3H, CH, CH₂), 6.02 (d, *J* = 7.20, 1H, NH), 7.37 (d, *J* = 7.71, 2H, ArH), 7.50 (d, *J* = 4.83, 1H, ArH), 7.61 (d, *J* = 7.71, 2H, ArH), 7.72–7.78 (m, 3H, ArH, NH), 8.32 (s, 1H, CH). ¹³C NMR (DMSO-*d*₆): 24.37, 31.88, 32.29, 35.78, 49.14, 50.24, 50.53, 55.40, 114.22, 120.95, 126.33, 126.68, 127.47, 128.36, 133.98, 136.29, 140.31, 141.96, 152.40, 155.03, 158.93.

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-[6-(3-fluorophenyl)-pyridin-3-ylmethyl]-9*H*-purine-2,6-diamine (**5n**). Yield 92%; mp 121–122 °C. Elemental analysis Calcd for C₂₈H₃₃FN₈O (500.61): C, 67.18; H, 6.64; N, 22.38. Found: C, 67.41; H, 6.69; N, 22.09. HPLC-MS (ESI+): 501.4 (99.5%). ¹H NMR (CDCl₃): 1.12–1.42 (m, 4H), 1.71–2.21 (m, 12H), 2.81 (sex, *J* = 5.87, 1H, CH), 3.12 (s(br), 2H, NH₂), 3.73 (sex, *J* = 7.44, 1H, CH), 4.62–4.72 (m, 2H, CH, NH), 4.81 (d, *J* = 5.77, 1H, CH₂), 6.33 (t, *J* = 5.77, 1H, NH), 7.12 (t, *J* = 8.25, 1H, ArH), 7.38–7.44 (m, 2H, ArH), 7.61–7.77 (m, 4H, ArH), 8.72 (s, 1H, CH).

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-(6-furan-2-yl-pyridin-3-ylmethyl)-9*H*-purine-2,6-diamine (**5o**). Yield 78%; mp 184–186 °C.

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-(6-thiophen-3-yl-pyridin-3-ylmethyl)-9*H*-purine-2,6-diamine (**5r**). Yield 68%; mp 139–140 °C. Elemental analysis Calcd for C₂₆H₃₄N₈S (497.64): C, 63.64; H, 6.98; N, 22.84; S, 6.53. Found: C, 63.62; H, 6.78; N, 22.59; S, 6.76. HPLC-MS (ESI+): 498.4 (98.9%). ¹H NMR (DMSO-*d*₆): 1.04–1.17 (m, 4H), 1.64–2.05 (m, 12H), 3.25–3.38 (m, 3H, CH, NH₂), 3.54 (sex, *J* = 7.56, 1H, CH), 4.59–4.65 (m, 3H, CH₂, CH), 6.09 (d, *J* = 7.56, 1H, NH), 7.11 (s, *J* = 4.12, 1H, ArH), 7.62 (d, *J* = 4.05, 1H, ArH), 7.72–7.82 (m, 4H, ArH), 7.90 (s(br), 1H, NH), 8.53 (s, 1H, CH).

General Procedure E for Preparing Compounds 4c, 5s, and 5t. The mixture of compound **2a** or **3b** (4.92 mmol), 1-amino-2-methylpropan-2-ol (25.00 mmol), *N,N*-diisopropyl-*N*-ethylamine (10.83 mmol), and *N*-methylpyrrolidone (5.0 mL) was heated with stirring in a sealed tube under an argon atmosphere (for detailed conditions see Table 1). After cooling to room temperature, the mixture was partitioned between water (25 mL) and ethyl acetate (25 mL) and the water phase was extracted twice with ethyl acetate. The combined organic phases were washed with water and brine and concentrated in vacuo. The residue was treated with 1% hydrochloric acid (25 mL) and extracted twice with dichloromethane. The combined organic phases were dried with sodium sulfate and concentrated in vacuo. The crude product was used for further reactions without purification. An analytical sample was obtained after column chromatography on silica (chloroform–methanol 9:1, v/v).

1-[6-(4-Bromobenzylamino)-9-cyclopentyl-9*H*-purin-2-ylamino]-2-methylpropan-2-ol (**4c**). Yield 82%; mp 108–110 °C. Elemental analysis Calcd for C₂₁H₂₉BrN₆O (461.40): C, 54.67; H, 6.34; N, 18.21. Found: C, 54.59; H, 6.12; N, 18.07. HPLC-MS (ESI+): 482.3 (98.6%). ¹H NMR (CDCl₃): 1.28 (s, 6H, CH₃), 1.74–1.90 (m, 6H), 2.05–2.38 (m, 2H), 2.84 (d, *J* = 2.32, 2H, CH₂), 4.75–4.83 (m, 3H, CH₂, CH), 5.20 (s(br), 1H, OH), 7.28 (d, *J* = 7.75, 2H, ArH), 7.45 (d, *J* = 7.75, 2H, ArH), 7.62 (s, 1H, CH).

1-[9-Cyclopentyl-6-[(6-furan-2-yl-pyridin-3-ylmethyl)-amino]-9H-purin-2-ylamino]-2-methyl-propan-2-ol (**5t**). Yield 37%; mp 128–129 °C. Elemental analysis Calcd for C₂₄H₂₉N₇O₂ (447.53): C, 64.41; H, 6.53; N, 21.91. Found: C, 64.65; H, 6.44; N, 21.58. HPLC-MS (ESI+): 448.4 (99.5%). ¹H NMR (CDCl₃): 1.27 (s, 6H, CH₃), 1.70–1.91 (m, 6H), 2.20–2.35 (m, 2H), 3.40 (d, J = 6.21, 2H, CH₂), 4.69 (qui, J = 6.42, 1H, CH), 4.79 (s(br), 2H, CH₂), 5.24 (t, J = 6.21, 1H, NH), 5.61 (s(br), 1H, OH), 6.04 (s(br), 1H, NH), 6.54 (t, J = 3.42, 1H, ArH), 7.03 (d, J = 3.42, 1H, ArH), 7.50–7.54 (m, 2H, ArH), 7.64 (d, J = 8.25, 1H, ArH), 7.74 (dd, J = 8.25, J' = 3.42, 1H, ArH), 8.63 (s, 1H, CH).

4-[9-Cyclopentyl-6-[(6-furan-2-yl-pyridin-3-ylmethyl)-amino]-9H-purin-2-ylamino]-cyclohexanol (**5s**). *trans*-4-Aminocyclohexan-1-ol hydrochloride (9.43 mmol) was suspended in methanol (10 mL), and to the suspension sodium methoxide (9.43 mmol) was added. The reaction mixture was stirred for 10 min at room temperature, and sodium chloride was filtered off. The filtrate was evaporated under reduced pressure, and to the residue (2-chloro-9-cyclopentyl-9H-purin-6-yl)-(6-furan-2-yl-pyridin-3-ylmethyl)-amine (0.25 mmol) and *N*-methylpyrrolidone (1 mL) was added. The reaction mixture was heated at 160 °C for 16 h under an argon atmosphere. After cooling to room temperature, water (10 mL) was added and resulting suspension was extracted twice with ethyl acetate (25 mL). Combined organic phases were washed with water, and brine, dried over anhydrous sodium sulfate, and evaporated under reduced pressure. The crude product was purified by column chromatography on silica, mobile phase chloroform–methanol (9:1). Yield 33%; mp 164–166 °C. Elemental analysis Calcd for C₂₆H₃₁N₇O₂ (473.57): C, 65.94; H, 6.60; N, 20.70. Found: C, 66.08; H, 6.48; N, 20.34. HPLC-MS (ESI+): 474.4 (99.6%). ¹H NMR (CDCl₃): 1.22 (q, J = 10.2, 2H), 1.43 (q, J = 10.2, 2H), 1.72–1.81 (m, 2H), 1.90–2.02 (m, 6H), 2.11–2.25 (m, 4H), 2.86 (s(br), 1H, OH), 3.61–3.76 (m, 2H), 4.64 (d, J = 7.68, 1H, NH), 4.69 (qui, J = 7.14, 1H, CH), 4.79 (d, J = 5.43, 2H, CH₂), 6.12 (t, J = 5.43, 1H, NH), 6.52 (dd, J = 3.39, J' = 1.77, 1H, ArH), 7.02 (d, J = 3.39, 1H, ArH), 7.48 (s, 1H, ArH), 7.52 (d, J = 3.39, 1H, ArH), 7.63 (d, J = 8.13, 1H, ArH), 7.73 (dd, J = 8.13, J' = 2.07, 1H, ArH), 8.61 (s, 1H, CH).

General Procedure F for the Preparation of Compounds 6a and 6b. To the solution of methoxyderivative (0.48 mmol) in dichloromethane (10 mL), boron tribromide (2.40 mmol) solution in dichloromethane (10 mL) was slowly added with stirring at room temperature. The mixture was stirred for further 18 h, and then methanol (20 mL) was added dropwise. The mixture was evaporated under reduced pressure, and the residue was purified by column chromatography on silica and mobile phase chloroform–methanol–ammonium hydroxide (4:1:0.025).

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-(2'-hydroxybiphenyl-4-ylmethyl)-9H-purine-2,6-diamine (**6a**). Yield 95%; mp 168–170 °C. Elemental analysis Calcd for C₂₉H₃₅N₇O (497.63): C, 69.99; H, 7.09; N, 19.70. Found: C, 69.68; H, 7.23; N, 19.57. HPLC-MS (ESI+): 498.5 (99.9%). ¹H NMR (CDCl₃): 1.16–1.40 (m, 4H), 1.71–2.22 (m, 12H), 2.52 (s(br), 2H, NH₂), 2.76 (sep, J = 5.43, 1H, CH), 3.67 (sex, J = 7.41, 1H, CH), 4.59 (d, J = 7.25, 1H, NH), 4.72 (qui, J = 7.00, 1H, CH), 4.79 (d, J = 5.63, 2H, CH₂), 6.10 (s(br), 1H, NH), 7.10 (t, J = 7.43, 1H, ArH), 7.16 (d, J = 8.04, 1H, ArH), 7.24 (t, J = 7.43, 1H, ArH), 7.48–7.66 (m, 4H, ArH), 8.63 (s, 1H, CH).

*N*²-(4-Aminocyclohexyl)-9-cyclopentyl-*N*⁶-[6-(2-hydroxyphenyl)-pyridin-3-ylmethyl]-9H-purine-2,6-diamine (**6b**). Yield 86%; mp 202–203 °C. Elemental analysis Calcd for C₂₈H₃₄N₈O (498.62): C, 67.45; H, 6.87; N, 22.47. Found: C, 67.28; H, 7.11; N, 22.41. HPLC-MS (ESI+): 499.5 (97.8%). ¹H NMR (CDCl₃): 1.16–1.40 (m, 4H), 1.71–2.22 (m, 12H), 2.49 (s(br), 2H, NH₂), 2.80 (sep, J = 5.31, 1H, CH), 3.62 (sex, J = 7.25, 1H, CH), 4.61 (d, J = 7.77, 1H, NH), 4.69 (qui, J = 7.17, 1H, CH), 4.82 (d, J = 5.43, 2H, CH₂), 6.13 (s(br), 1H, NH), 6.91 (t, J = 7.38, 1H, ArH), 7.02 (d, J = 8.19, 1H, ArH), 7.29 (t, J = 7.38, 1H, ArH), 7.50 (s, 1H, ArH), 7.78 (d, J = 8.19, 1H, ArH), 7.82–7.86 (m, 2H, ArH), 8.54 (s, 1H, CH). ¹³C NMR (CDCl₃): 24.22, 29.77, 31.86, 32.69, 34.38, 49.94, 50.31, 55.52, 114.53, 118.62, 118.84, 118.88, 118.93, 126.14, 131.41, 133.34, 135.69, 137.11, 145.10, 153.25, 154.64, 156.70, 158.78, 159.87.

Cell Maintenance and Cytotoxicity Assays. The cytotoxicity of the studied compounds was determined using cell lines of different histological origin as described earlier.^{17,53} Briefly, compounds in 3-fold dilutions were added to the cells in triplicate. The treatment lasted for 72 h, after which Calcein AM solution was added, and the fluorescence of live cells at 485 nm/538 nm (excitation/emission) was measured with a Fluoroskan Ascent microplate reader (Labsystems). IC₅₀ (the drug concentration that reduced the number of viable cells to 50%) values were determined from the dose–response curves. Roscovitine and CR8, used as standard drugs, were prepared according to published procedures.^{4,10}

Immunoblotting and Antibodies. Immunoblotting analysis was performed as described earlier.^{17,53} Briefly, cellular lysates were prepared by harvesting cells in Laemmli sample buffer. Proteins were separated on SDS-polyacrylamide gels and electroblotted onto nitrocellulose membranes. After blocking, the membranes were incubated with specific primary antibodies overnight, washed, and then incubated with peroxidase-conjugated secondary antibodies. Finally, peroxidase activity was detected with ECL+ reagents (AP Biotech) using a CCD camera LAS-4000 (Fujifilm). Specific antibodies were purchased from Sigma-Aldrich (anti- α -tubulin, clone DM1A; peroxidase-labeled secondary antibodies; anti-pRb antibody phosphorylated at S612; anti-CDK2, clone PSTAIR), Santa Cruz Biotechnology (anti-Mcl-1, clone S-19; anti-PARP, clone F-2; anti- β -actin, clone C4; anti-pCdk2 (T160); anti-pCDK1 (T161); anti-CDK1, clone B-6), Roche Applied Science (anti-5-bromo-2'-deoxyuridine-fluorescein, clone BMC 9318), Bethyl Laboratories (anti-pRNA polymerase II antibodies phosphorylated at S5 and (S2)), Millipore (anti-RNA polymerase II, clone ARNA-3), Cell Signaling (anti-PP1 α (T320); anti-PP1 α ; anti-pNPM (T199); anti-NPM; anticaspase-3, clone 3G2; anticaspase-7; anticaspase-9 (D330); anti-pRb antibodies phosphorylated at S780 and S807/811; anti-Rb, clone 4H1) or were a generous gift from Dr. B. Vojtěšek from Masaryk Memorial Cancer Institute, Brno, Czech Republic (anti-p53, clone DO-1).

Kinase Inhibition Assays. CDK2/Cyclin E kinase was produced in Sf9 insect cells via baculoviral infection and purified on a NiNTA column (Qiagen). CDK5/p35, CDK7/Cyclin H/MAT1, and CDK9/Cyclin T1 were purchased from ProQinase GmbH. The kinases were assayed with 1 mg/mL histone H1 (for CDK2 and CDK5) or (YSPTSPS)2KK peptide (for CDK7 and CDK9) in the presence of 15/0.15/1.5/1.5 μ M ATP (for CDK2/CDK5/CDK7/CDK9), 0.05 μ Ci [γ -³³P]ATP, and of the test compound in a final volume of 10 μ L, all in a reaction buffer (60 mM HEPES-NaOH, pH 7.5, 3 mM MgCl₂, 3 mM MnCl₂, 3 μ M Na-orthovanadate, 1.2 mM DTT, 2.5 μ g/50 μ L PEG20.000). The reactions were stopped by adding 5 μ L of 3% aq H₃PO₄. Aliquots were spotted onto P-81 phosphocellulose (Whatman), washed 3 \times with 0.5% aq H₃PO₄, and finally air-dried. Kinase inhibition was quantified using a FLA-7000 digital image analyzer (Fujifilm). The concentration of the test compounds required to decrease the CDK activity by 50% was determined from dose–response curves and designated as IC₅₀.^{17,53}

Cell Cycle Analysis. Subconfluent cells were treated with test compounds at different concentrations for 24 h. The cultures were pulse-labeled with 10 μ M 5-bromo-2'-deoxyuridine (BrdU) for 30 min at 37 °C prior to harvesting. The cells were then washed in PBS, fixed with 70% ethanol, and denatured in 2 M HCl. Following neutralization, the cells were stained with anti-BrdU fluorescein-labeled antibodies, washed, stained with propidium iodide, and analyzed by flow cytometry using a 488 nm laser (Cell Lab Quanta SC, Beckman Coulter) as described previously.^{17,53}

Caspases Activity Assay. The cells were homogenized in an extraction buffer (10 mM KCl, 5 mM HEPES, 1 mM EDTA, 1 mM EGTA, 0.2% CHAPS, inhibitors of proteases, pH 7.4) on ice for 20 min. The homogenates were clarified by centrifugation at 10000g for 30 min at 4 °C, and then the proteins were quantified and diluted to equal concentrations. Lysates were then incubated for 5 h with 100 μ M Ac-DEVD-AMC as a substrate of caspases-3,7 in the assay buffer (25 mM PIPES, 2 mM EGTA, 2 mM MgCl₂, 5 mM DTT, pH 7.3) or for 24 h with 100 μ M Ac-LEHD-AMC as a substrate of caspase-9 in the assay buffer (100 mM HEPES, pH 7.5; 0.5 mM EDTA; 20%

glycerol, 5 mM DTT). The fluorescence of the product was measured using a Fluoroskan Ascent microplate reader (Labsystems) at 355/460 nm (excitation/emission) as described previously.¹⁷

Molecular Modeling. 3D structures of compounds were prepared with Marvin, a software used for drawing, displaying, and characterizing chemical structures, substructures, and reactions [Marvin 5.10.3, 2012, ChemAxon (<http://www.chemaxon.com>)]. 3D structures were optimized and all hydrogens were added within the MarvinSketch 5.10 program. All compounds contain mostly aromatic residues, which were kept planar as well as secondary amino groups attached to C2 and C6 positions on the purine ring. The cyclohexane ring was kept in the most stable chair conformation, and the cyclopentane ring was attached in the N9 position to the equatorial position on the endo carbon as this combination poses the smallest sterical clashes with the purine central ring. Ionization of groups was considered for the range of pH 5.5–7.5. Therefore, the primary amino groups attached to the aliphatic carbons were kept charged ($pK_{a \text{ est. Marvin}} < 5$), nitrogens attached to the aromatic rings were kept neutral ($pK_a > 10$). All nonaromatic and nonring bonds were set as rotatable within AutoDock Tools program.⁵⁴

The crystal structure for CDK2 with roscovitine (PDB ID: 2A4L) was used as the protein docking template. Glu8 and His84 were set as flexible. Polar hydrogens were added to receptor or selected for all ligands with the AutoDock Tools program⁵⁴ prior to docking with the Autodock Vina program.⁵⁵ Docking grid box was set to 14 Å around the center of the ligand in the crystal structure, which was deleted prior to docking. Vina docking parameters were extended in order to increase docking accuracy and exhaustiveness with $\text{energy_range} = 10$ (default 3), $\text{num_modes} = 500$ (default 9), $\text{exhaustiveness} = 20$ (default 8).

As the docking results with the default setting did not correlate well with the experimental IC_{50} data ($r^2 < 0.2$, data not shown), we analyzed binding motifs of CDK2 inhibitors found in crystal structures. In all cases, the binding motif interacted with conserved hydrogen bonds of the Leu83 and Glu81 backbones as in the roscovitine crystal structure (PDB ID: 2A4L). For this reason, we applied a three-step docking procedure, taking into account the conserved binding motif: (1) after the first round of docking, the pose of the best crystal-like position was selected, (2) then the ligand was constrained by the atoms forming conserved hydrogen bonds and it was treated like a flexible side chain and the dummy atom of hydrogen was docked, (3) the final pose of the ligand stemming from the previous step was then rescored to obtain binding free energies while retaining conserved binding motif.

Dihedral scans for assessing the rigidity of two rings were calculated at the DF-PBE/6-311++g(3df,3pf) level of theory with the C-PCM implicit solvent model for water. All relaxed scans were done with the Gaussian 09 program, revision A.2.⁵⁶

■ ASSOCIATED CONTENT

Supporting Information

CDK 5, CDK7, and CDK9 inhibitory activity of selected compounds, list of synthesized compounds with detailed reaction conditions, experimental and theoretical free energies calculated from IC_{50} s and Vina results and their comparison, induction of apoptosis and replication of compound **6b**, and dihedral scans between the two rings for compounds **5a–6b**. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +420585634953. Fax: +420585634870. E-mail: akrylhydrazid@seznam.cz.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Czech Science Foundation (grants P305/12/0783 and 203/09/H046), the Ministry of Education, Youth and Sports, Czech Republic (grant ED0007/01/01; Centre of the Region Haná for Biotechnological and Agricultural Research), the Operational Program Research and Development for Innovations (European Regional Development Fund (CZ.1.05/2.1.00/03.0058), and European Social Fund (CZ.1.07/2.3.00/20.0017)) and by the Student Projects PrF_2013_023 and PrF_2013_028 of Palacký University.

■ REFERENCES

- (1) Malumbres, M.; Barbacid, M. Mammalian cyclin-dependent kinases. *Trends Biochem. Sci.* **2005**, *30*, 630–641.
- (2) Krystof, V.; Uldrijan, S. Cyclin-dependent kinase inhibitors as anticancer drugs. *Curr. Drug Targets* **2010**, *11*, 291–302.
- (3) Lapenna, S.; Giordano, A. Cell cycle kinases as therapeutic targets for cancer. *Nature Rev. Drug Discovery* **2009**, *8*, 547–566.
- (4) Havlicek, L.; Hanus, J.; Vesely, J.; LeClerc, S.; Meijer, L.; Shaw, G.; Strnad, M. Cytokinin-derived cyclin-dependent kinase inhibitors: synthesis and cdc2 inhibitory activity of olomoucine and related compounds. *J. Med. Chem.* **1997**, *40*, 408–412.
- (5) Hsieh, W. S.; Soo, R.; Peh, B. K.; Loh, T.; Dong, D.; Soh, D.; Wong, L. S.; Green, S.; Chiao, J.; Cui, C. Y.; Lai, Y. F.; Lee, S. C.; Mow, B.; Soong, R.; Salto-Tellez, M.; Goh, B. C. Pharmacodynamic effects of seliciclib, an orally administered cell cycle modulator, in undifferentiated nasopharyngeal cancer. *Clin. Cancer Res.* **2009**, *15*, 1435–1442.
- (6) Le, T. C.; Faivre, S.; Laurence, V.; Delbaldo, C.; Vera, K.; Girre, V.; Chiao, J.; Armour, S.; Frame, S.; Green, S. R.; Gianella-Borradori, A.; Dieras, V.; Raymond, E. Phase I evaluation of seliciclib (R-roscovitine), a novel oral cyclin-dependent kinase inhibitor, in patients with advanced malignancies. *Eur. J. Cancer* **2010**, *46*, 3243–3250.
- (7) Chang, Y. T.; Gray, N. S.; Rosania, G. R.; Sutherlin, D. P.; Kwon, S.; Norman, T. C.; Sarohia, R.; Leost, M.; Meijer, L.; Schultz, P. G. Synthesis and application of functionally diverse 2,6,9-trisubstituted purine libraries as CDK inhibitors. *Chem. Biol.* **1999**, *6*, 361–375.
- (8) Gray, N. S.; Wodicka, L.; Thunnissen, A. M.; Norman, T. C.; Kwon, S.; Espinoza, F. H.; Morgan, D. O.; Barnes, G.; LeClerc, S.; Meijer, L.; Kim, S. H.; Lockhart, D. J.; Schultz, P. G. Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* **1998**, *281*, 533–538.
- (9) Legraverend, M.; Ludwig, O.; Bisagni, E.; LeClerc, S.; Meijer, L.; Giocanti, N.; Sadri, R.; Favaudon, V. Synthesis and in vitro evaluation of novel 2,6,9-trisubstituted purines acting as cyclin-dependent kinase inhibitors. *Bioorg. Med. Chem.* **1999**, *7*, 1281–1293.
- (10) Oumata, N.; Bettayeb, K.; Ferandin, Y.; Demange, L.; Lopez-Giral, A.; Goddard, M. L.; Myrianthopoulos, V.; Mikros, E.; Flajolet, M.; Greengard, P.; Meijer, L.; Galons, H. Roscovitine-derived, dual-specificity inhibitors of cyclin-dependent kinases and casein kinases 1. *J. Med. Chem.* **2008**, *51*, 5229–5242.
- (11) Trova, M. P.; Barnes, K. D.; Alicea, L.; Benanti, T.; Bielaska, M.; Bilotta, J.; Bliss, B.; Duong, T. N.; Haydar, S.; Herr, R. J.; Hui, Y.; Johnson, M.; Lehman, J. M.; Peace, D.; Rainka, M.; Snider, P.; Salamone, S.; Tregay, S.; Zheng, X.; Friedrich, T. D. Heterobiaryl purine derivatives as potent antiproliferative agents: inhibitors of cyclin dependent kinases. Part II. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6613–6617.
- (12) Trova, M. P.; Barnes, K. D.; Barford, C.; Benanti, T.; Bielaska, M.; Burry, L.; Lehman, J. M.; Murphy, C.; O'Grady, H.; Peace, D.; Salamone, S.; Smith, J.; Snider, P.; Toporowski, J.; Tregay, S.; Wilson, A.; Wyle, M.; Zheng, X.; Friedrich, T. D. Biaryl purine derivatives as potent antiproliferative agents: inhibitors of cyclin dependent kinases. Part I. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6608–6612.
- (13) Zatloukal, M.; Jorda, R.; Gucky, T.; Reznickova, E.; Voller, J.; Pospisil, T.; Malinkova, V.; Adamcova, H.; Krystof, V.; Strnad, M. Synthesis and in vitro biological evaluation of 2,6,9-trisubstituted

purines targeting multiple cyclin-dependent kinases. *Eur. J. Med. Chem.* **2013**, *61*, 61–72.

(14) Jorda, R.; Paruch, K.; Krystof, V. Cyclin-dependent kinase inhibitors inspired by roscovitine: purine bioisosteres. *Curr. Pharm. Des.* **2012**, *18*, 2974–2980.

(15) Bettayeb, K.; Sallam, H.; Ferandin, Y.; Popowycz, F.; Fournet, G.; Hassan, M.; Echaliier, A.; Bernard, P.; Endicott, J.; Joseph, B.; Meijer, L. N-&-N, a new class of cell death-inducing kinase inhibitors derived from the purine roscovitine. *Mol. Cancer Ther.* **2008**, *7*, 2713–2724.

(16) Popowycz, F.; Fournet, G.; Schneider, C.; Bettayeb, K.; Ferandin, Y.; Lamigeon, C.; Tirado, O. M.; Mateo-Lozano, S.; Notario, V.; Colas, P.; Bernard, P.; Meijer, L.; Joseph, B. Pyrazolo[1,5-*a*]-1,3,5-triazine as a purine bioisostere: access to potent cyclin-dependent kinase inhibitor (*R*)-roscovitine analogue. *J. Med. Chem.* **2009**, *52*, 655–663.

(17) Jorda, R.; Havlicek, L.; McNae, I. W.; Walkinshaw, M. D.; Voller, J.; Sturc, A.; Navratilova, J.; Kuzma, M.; Mistrik, M.; Bartek, J.; Strnad, M.; Krystof, V. Pyrazolo[4,3-*d*]pyrimidine bioisostere of roscovitine: evaluation of a novel selective inhibitor of cyclin-dependent kinases with antiproliferative activity. *J. Med. Chem.* **2011**, *54*, 2980–2993.

(18) Heathcote, D. A.; Patel, H.; Kroll, S. H.; Hazel, P.; Periyasamy, M.; Alikian, M.; Kanneganti, S. K.; Jogalekar, A. S.; Scheiper, B.; Barbazanges, M.; Blum, A.; Brackow, J.; Siwicka, A.; Pace, R. D.; Fuchter, M. J.; Snyder, J. P.; Liotta, D. C.; Freemont, P. S.; Aboagy, E. O.; Coombes, R. C.; Barrett, A. G.; Ali, S. A novel pyrazolo[1,5-*a*]pyrimidine is a potent inhibitor of cyclin-dependent protein kinases 1, 2, and 9, which demonstrates antitumor effects in human tumor xenografts following oral administration. *J. Med. Chem.* **2010**, *53*, 8508–8522.

(19) Paruch, K.; Dwyer, M. P.; Alvarez, C.; Brown, C.; Chan, T. Y.; Doll, R. J.; Keertikar, K.; Knutson, C.; McKittrick, B.; Rivera, J.; Rossmann, R.; Tucker, G.; Fischmann, T. O.; Hruza, A.; Madison, V.; Nomeir, A. A.; Wang, Y.; Lees, E.; Parry, D.; Sgambellone, N.; Seghezzi, W.; Schultz, L.; Shanahan, F.; Wiswell, D.; Xu, X.; Zhou, Q.; James, R. A.; Paradkar, V. M.; Park, H.; Rokosz, L. R.; Stauffer, T. M.; Guzi, T. J. Pyrazolo[1,5-*a*]pyrimidines as orally available inhibitors of cyclin-dependent kinase 2. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 6220–6223.

(20) Williamson, D. S.; Parratt, M. J.; Bower, J. F.; Moore, J. D.; Richardson, C. M.; Dokurno, P.; Cansfield, A. D.; Francis, G. L.; Hebbon, R. J.; Howes, R.; Jackson, P. S.; Lockie, A. M.; Murray, J. B.; Nunns, C. L.; Powles, J.; Robertson, A.; Surgenor, A. E.; Torrance, C. J. Structure-guided design of pyrazolo[1,5-*a*]pyrimidines as inhibitors of human cyclin-dependent kinase 2. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 863–867.

(21) Parry, D.; Guzi, T.; Shanahan, F.; Davis, N.; Prabhavalkar, D.; Wiswell, D.; Seghezzi, W.; Paruch, K.; Dwyer, M. P.; Doll, R.; Nomeir, A.; Windsor, W.; Fischmann, T.; Wang, Y.; Oft, M.; Chen, T.; Kirschmeier, P.; Lees, E. M. Dinaciclib (SCH 727965), a novel and potent cyclin-dependent kinase inhibitor. *Mol. Cancer Ther.* **2010**, *9*, 2344–2353.

(22) Paruch, K.; Dwyer, M. P.; Alvarez, C.; Brown, C.; Chan, T. Y.; Doll, R. J.; Keertikar, K.; Knutson, C.; McKittrick, B.; Rivera, J.; Rossmann, R.; Tucker, G.; Fischmann, T.; Hruza, A.; Madison, V.; Nomeir, A. A.; Wang, Y. L.; Kirschmeier, P.; Lees, E.; Parry, D.; Sgambellone, N.; Seghezzi, W.; Schultz, L.; Shanahan, F.; Wiswell, D.; Xu, X. Y.; Zhou, Q. A.; James, R. A.; Paradkar, V. M.; Park, H.; Rokosz, L. R.; Stauffer, T. M.; Guzi, T. J. Discovery of Dinaciclib (SCH 727965): A Potent and Selective Inhibitor of Cyclin-Dependent Kinases. *ACS Med. Chem. Lett.* **2010**, *1*, 204–208.

(23) Gorlick, R.; Kolb, E. A.; Houghton, P. J.; Morton, C. L.; Neale, G.; Keir, S. T.; Carol, H.; Lock, R.; Phelps, D.; Kang, M. H.; Reynolds, C. P.; Maris, J. M.; Billups, C.; Smith, M. A. Initial testing (stage 1) of the cyclin dependent kinase inhibitor SCH 727965 (dinaciclib) by the pediatric preclinical testing program. *Pediatr. Blood Cancer* **2012**, *59*, 1266–1274.

(24) Johnson, A. J.; Yeh, Y. Y.; Smith, L. L.; Wagner, A. J.; Hessler, J.; Gupta, S.; Flynn, J.; Jones, J.; Zhang, X.; Bannerji, R.; Grever, M. R.; Byrd, J. C. The novel cyclin-dependent kinase inhibitor dinaciclib (SCH727965) promotes apoptosis and abrogates microenvironmental cytokine protection in chronic lymphocytic leukemia cells. *Leukemia* **2012**, *26*, 2554–2557.

(25) Zhang, D.; Mita, M.; Shapiro, G. I.; Poon, J.; Small, K.; Tzontcheva, A.; Kantesaria, B.; Zhu, Y.; Bannerji, R.; Statkevich, P. Effect of aprepitant on the pharmacokinetics of the cyclin-dependent kinase inhibitor dinaciclib in patients with advanced malignancies. *Cancer Chemother. Pharmacol.* **2012**, *70*, 891–898.

(26) Benson, C.; White, J.; De, B. J.; O'Donnell, A.; Raynaud, F.; Cruickshank, C.; McGrath, H.; Walton, M.; Workman, P.; Kaye, S.; Cassidy, J.; Gianella-Borradori, A.; Judson, I.; Twelves, C. A phase I trial of the selective oral cyclin-dependent kinase inhibitor seliciclib (CYC202; R-Roscovitine), administered twice daily for 7 days every 21 days. *Br. J. Cancer* **2007**, *96*, 29–37.

(27) Chen, R.; Keating, M. J.; Gandhi, V.; Plunkett, W. Transcription inhibition by flavopiridol: mechanism of chronic lymphocytic leukemia cell death. *Blood* **2005**, *106*, 2513–2519.

(28) Massard, C.; Soria, J. C.; Anthony, D. A.; Proctor, A.; Scaburri, A.; Pacciarini, M. A.; Laffranchi, B.; Pellizzoni, C.; Kroemer, G.; Armand, J. P.; Balheda, R.; Twelves, C. J. A first in man, phase I dose-escalation study of PHA-793887, an inhibitor of multiple cyclin-dependent kinases (CDK2, 1 and 4) reveals unexpected hepatotoxicity in patients with solid tumors. *Cell Cycle* **2011**, *10*, 963–970.

(29) Tong, W. G.; Chen, R.; Plunkett, W.; Siegel, D.; Sinha, R.; Harvey, R. D.; Badros, A. Z.; Popplewell, L.; Coutre, S.; Fox, J. A.; Mahadocon, K.; Chen, T.; Kegley, P.; Hoch, U.; Wierda, W. G. Phase I and pharmacologic study of SNS-032, a potent and selective Cdk2, 7, and 9 inhibitor, in patients with advanced chronic lymphocytic leukemia and multiple myeloma. *J. Clin. Oncol.* **2010**, *28*, 3015–3022.

(30) Bettayeb, K.; Oumata, N.; Echaliier, A.; Ferandin, Y.; Endicott, J. A.; Galons, H.; Meijer, L. CR8, a potent and selective, roscovitine-derived inhibitor of cyclin-dependent kinases. *Oncogene* **2008**, *27*, 5797–5807.

(31) Liu, Y. X.; Wei, D. G.; Zhu, Y. R.; Liu, S. H.; Zhang, Y. L.; Zhao, Q. Q.; Cai, B. L.; Li, Y. H.; Song, H. B.; Liu, Y.; Wang, Y.; Huang, R. Q.; Wang, Q. M. Synthesis, herbicidal activities, and 3D-QSAR of 2-cyanoacrylates containing aromatic methylamine moieties. *J. Agric. Food Chem.* **2008**, *56*, 204–212.

(32) Dreyer, M. K.; Borchering, D. R.; Dumont, J. A.; Peet, N. P.; Tsay, J. T.; Wright, P. S.; Bitonti, A. J.; Shen, J.; Kim, S. H. Crystal structure of human cyclin-dependent kinase 2 in complex with the adenine-derived inhibitor H717. *J. Med. Chem.* **2001**, *44*, 524–530.

(33) Zim, D.; Monteiro, A. L.; Dupont, J. Pd(Cl)₂(SEt₂)₂ and Pd(OAc)₂: simple and efficient catalyst precursors for the Suzuki cross-coupling reaction. *Tetrahedron Lett.* **2000**, *41*, 8199–8202.

(34) Payton, M.; Chung, G.; Yakowec, P.; Wong, A.; Powers, D.; Xiong, L.; Zhang, N.; Leal, J.; Bush, T. L.; Santora, V.; Askew, B.; Tasker, A.; Radinsky, R.; Kendall, R.; Coats, S. Discovery and evaluation of dual CDK1 and CDK2 inhibitors. *Cancer Res.* **2006**, *66*, 4299–4308.

(35) MacCallum, D. E.; Melville, J.; Frame, S.; Watt, K.; Anderson, S.; Gianella-Borradori, A.; Lane, D. P.; Green, S. R. Seliciclib (CYC202, R-Roscovitine) induces cell death in multiple myeloma cells by inhibition of RNA polymerase II-dependent transcription and down-regulation of Mcl-1. *Cancer Res.* **2005**, *65*, 5399–5407.

(36) Santo, L.; Vallet, S.; Hideshima, T.; Cirstea, D.; Ikeda, H.; Pozzi, S.; Patel, K.; Okawa, Y.; Gorgun, G.; Perrone, G.; Calabrese, E.; Yule, M.; Squires, M.; Ladetto, M.; Boccadoro, M.; Richardson, P. G.; Munshi, N. C.; Anderson, K. C.; Raje, N. AT7519, A novel small molecule multi-cyclin-dependent kinase inhibitor, induces apoptosis in multiple myeloma via GSK-3beta activation and RNA polymerase II inhibition. *Oncogene* **2010**, *29*, 2325–2336.

(37) Demidenko, Z. N.; Blagosklonny, M. V. Flavopiridol induces p53 via initial inhibition of Mdm2 and p21 and, independently of p53, sensitizes apoptosis-reluctant cells to tumor necrosis factor. *Cancer Res.* **2004**, *64*, 3653–3660.

- (38) Kotala, V.; Uldrijan, S.; Horky, M.; Trbusek, M.; Strnad, M.; Vojtesek, B. Potent induction of wild-type p53-dependent transcription in tumour cells by a synthetic inhibitor of cyclin-dependent kinases. *Cell. Mol. Life Sci.* **2001**, *58*, 1333–1339.
- (39) Kwon, Y. G.; Lee, S. Y.; Choi, Y.; Greengard, P.; Nairn, A. C. Cell cycle-dependent phosphorylation of mammalian protein phosphatase 1 by cdc2 kinase. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 2168–2173.
- (40) Tokuyama, Y.; Horn, H. F.; Kawamura, K.; Tarapore, P.; Fukasawa, K. Specific phosphorylation of nucleophosmin on Thr(199) by cyclin-dependent kinase 2-cyclin E and its role in centrosome duplication. *J. Biol. Chem.* **2001**, *276*, 21529–21537.
- (41) Laroche, S.; Merrick, K. A.; Terret, M. E.; Wohlbold, L.; Barboza, N. M.; Zhang, C.; Shokat, K. M.; Jallepalli, P. V.; Fisher, R. P. Requirements for Cdk7 in the assembly of Cdk1/cyclin B and activation of Cdk2 revealed by chemical genetics in human cells. *Mol. Cell* **2007**, *25*, 839–850.
- (42) Malumbres, M.; Barbacid, M. Cell cycle, CDKs and cancer: a changing paradigm. *Nature Rev. Cancer* **2009**, *9*, 153–166.
- (43) Tetsu, O.; McCormick, F. Proliferation of cancer cells despite CDK2 inhibition. *Cancer Cell* **2003**, *3*, 233–245.
- (44) Caporali, S.; Alvino, E.; Starace, G.; Ciomei, M.; Brasca, M. G.; Levati, L.; Garbin, A.; Castiglia, D.; Covaciu, C.; Bonmassar, E.; D'Atri, S. The cyclin-dependent kinase inhibitor PHA-848125 suppresses the in vitro growth of human melanomas sensitive or resistant to Temozolomide, and shows synergistic effects in combination with this triazene compound. *Pharmacol. Res.* **2010**, *61*, 437–448.
- (45) Du, J.; Widlund, H. R.; Horstmann, M. A.; Ramaswamy, S.; Ross, K.; Huber, W. E.; Nishimura, E. K.; Golub, T. R.; Fisher, D. E. Critical role of CDK2 for melanoma growth linked to its melanocyte-specific transcriptional regulation by MITF. *Cancer Cell* **2004**, *6*, 565–576.
- (46) Abdullah, C.; Wang, X.; Becker, D. Expression analysis and molecular targeting of cyclin-dependent kinases in advanced melanoma. *Cell Cycle* **2011**, *10*, 977–988.
- (47) Blachly, J. S.; Byrd, J. C. Emerging Drug Profile: Cyclin-Dependent Kinase Inhibitors. *Leuk. Lymphoma* **2013**.
- (48) Bettayeb, K.; Baunbaek, D.; Delehouze, C.; Loaec, N.; Hole, A. J.; Baumli, S.; Endicott, J. A.; Douc-Rasy, S.; Benard, J.; Oumata, N.; Galons, H.; Meijer, L. CDK Inhibitors Roscovitine and CR8 Trigger Mcl-1 Down-Regulation and Apoptotic Cell Death in Neuroblastoma Cells. *Genes Cancer* **2010**, *1*, 369–380.
- (49) Wilson, S. C.; Atrash, B.; Barlow, C.; Eccles, S.; Fischer, P. M.; Hayes, A.; Kelland, L.; Jackson, W.; Jarman, M.; Mirza, A.; Moreno, J.; Nutley, B. P.; Raynaud, F. I.; Sheldrake, P.; Walton, M.; Westwood, R.; Whittaker, S.; Workman, P.; McDonald, E. Design, synthesis and biological evaluation of 6-pyridylmethylaminopurines as CDK inhibitors. *Bioorg. Med. Chem.* **2011**, *19*, 6949–6965.
- (50) Meanwell, N. A. Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. *Chem. Res. Toxicol.* **2011**, *24*, 1420–1456.
- (51) Collins, C. S.; Hong, J.; Sapinoso, L.; Zhou, Y.; Liu, Z.; Micklash, K.; Schultz, P. G.; Hampton, G. M. A small interfering RNA screen for modulators of tumor cell motility identifies MAP4K4 as a promigratory kinase. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 3775–3780.
- (52) Liebl, J.; Weitensteiner, S. B.; Vereb, G.; Takacs, L.; Furst, R.; Vollmar, A. M.; Zahler, S. Cyclin-dependent kinase 5 regulates endothelial cell migration and angiogenesis. *J. Biol. Chem.* **2010**, *285*, 35932–35943.
- (53) Krystof, V.; Cankar, P.; Frysova, I.; Slouka, J.; Kontopidis, G.; Dzubak, P.; Hajdich, M.; Srovnal, J.; de Azevedo, W. F. J.; Orsag, M.; Paprskarova, M.; Rolcik, J.; Latr, A.; Fischer, P. M.; Strnad, M. 4-Arylazo-3,5-diamino-1H-pyrazole CDK inhibitors: SAR study, crystal structure in complex with CDK2, selectivity, and cellular effects. *J. Med. Chem.* **2006**, *49*, 6500–6509.
- (54) Sanner, M. F. Python: a programming language for software integration and development. *J. Mol. Graphics Modell.* **1999**, *17*, 57–61.
- (55) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.
- (56) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E. Jr.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 2009*, revision A.02; Gaussian, Inc.: Wallingford, CT 2009.

Appendix J

Pozharitskaya O, Shikov A, Makarova M, Ivanova S, Kosman V, Makarov V, **Bazgier V**,
Berka K, Otyepka M, Ulrichová J:

Antiallergic Effects of Pigments Isolated from Green Sea Urchin (*Strongylocentrotus*
droebachiensis) Shells.

Planta Med., 79(18), 1698-1704, 2013.

DOI: 10.1055/s-0033-1351098

IF=2.342

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/259003610>

Antiallergic Effects of Pigments Isolated from Green Sea Urchin (*Strongylocentrotus droebachiensis*) Shells

Article in *Planta Medica* · November 2013

DOI: 10.1055/s-0033-1351098 · Source: PubMed

CITATIONS

14

READS

285

10 authors, including:



Alexander N Shikov

Saint Petersburg Institute of Pharmacy

131 PUBLICATIONS 913 CITATIONS

SEE PROFILE



Vaclav Bazgier

Palacký University Olomouc

17 PUBLICATIONS 66 CITATIONS

SEE PROFILE



Karel Berka

Palacký University Olomouc

81 PUBLICATIONS 1,160 CITATIONS

SEE PROFILE



Michal Otyepka

Palacký University Olomouc

258 PUBLICATIONS 7,146 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Bio.Tools registry [View project](#)



Force field parameterization [View project](#)

Personal pdf file for

Olga N. Pozharitskaya, Alexander N. Shikov, Marina N. Makarova, Svetlana A. Ivanova, Vera M. Kosman, Valery G. Makarov, Václav Bazgier, Karel Berka, Michal Otyepka, Jitka Ulrichová

With compliments of Georg Thieme Verlag

www.thieme.de

Antiallergic Effects of Pigments
Isolated from Green Sea Urchin
(*Strongylocentrotus droebachiensis*)
Shells

DOI 10.1055/s-0033-1351098
Planta Med 2013; 79: 1698–1704

For personal use only.
No commercial use, no depositing in repositories.

Publisher and Copyright:

© 2013 by
Georg Thieme Verlag KG
Rüdigerstraße 14
70469 Stuttgart
ISSN 0032-0943

Reprint with the
permission by
the publisher only

 **Thieme**

Antiallergic Effects of Pigments Isolated from Green Sea Urchin (*Strongylocentrotus droebachiensis*) Shells

Authors

Olga N. Pozharitskaya¹, Alexander N. Shikov¹, Marina N. Makarova¹, Svetlana A. Ivanova¹, Vera M. Kosman¹, Valery G. Makarov¹, Václav Bazgier², Karel Berka³, Michal Otyepka³, Jitka Ulrichová⁴

Affiliations

The affiliations are listed at the end of the article

Key words

- ◉ *Strongylocentrotus droebachiensis*
- ◉ Strongylocentrotidae
- ◉ spinochrome
- ◉ histamine H₁-receptor
- ◉ guinea pig ileum
- ◉ allergic conjunctivitis
- ◉ molecular docking

Abstract

▼ This study was undertaken to evaluate possible antiallergic effects of an extract of pigments from green sea urchin (*Strongylocentrotus droebachiensis*) shells. Effects were studied on animal models – guinea pig ileum contraction, rabbit eyes allergic conjunctivitis, and rabbit local skin irritation. The extract significantly reduced, in a dose-dependent manner, the histamine-induced contractions of the isolated guinea pig ileum with ID₅₀ = 1.2 µg/mL (in equivalents of spinochrome B), had an inhibitory effect on the model of ocular allergic inflammation surpassing the reference drug olo-

patadine, and did not show any irritating effect in rabbits. The extract predominantly contained polyhydroxy-1,4-naphthoquinone which would be responsible for the pharmacological activity. The active compounds of the extract were evaluated *in silico* with molecular docking. Molecular docking into H₁R receptor structures obtained from molecular dynamic simulations showed that all spinochrome derivatives bind to the receptor active site, but spinochrome monomers fit better to it. The results of the present study suggest possibilities for the development of new agents for treating allergic diseases on the base of pigments from sea urchins shells.

Introduction

▼ Among biologically active substances of natural origin, the secondary metabolites of marine organisms are of special interest. Sea urchins are major consumers in the shallow waters of the world oceans. They are commonly found in northern waters all around the world including both the Pacific and Atlantic Oceans to a northerly latitude of 81 degrees. Various species of sea urchins are widely distributed in the world's oceans and some of them have been caught for their gonads, used as high-priced Sushi foodstuff known as “Uni” in Japanese traditional food. After removal of the edible gonads, the residual shells with spines, which comprise 40% of the sea urchin's weight, are generally discarded as a waste. The characteristic color of spines and shells of sea urchins is due to calcium salts of naphthoquinone derivatives. The total content of polyhydroxylated 1,4-naphthoquinone pigments (spinochromes) in the shells is 121–163 mg/g [1]. The presence of several phenolic OH groups in these compounds' structure suggests that they could have several activities. In particular, pigment echinochrome A and its synthetic analogues are known today as

biologically active compounds that present antimicrobial, antialgal, cardioprotective, and antioxidant activities [2,3]. Spinochrome monomers B and D and their dimers showed potent antiradical properties [4]. Significant antipruritic activity was reported for 2-hydroxy-3-(2-hydroxyethyl)-1,4-naphthoquinone, 2,2'-methylenebis(3-hydroxy-1,4-naphthoquinone) and 2,2'-ethylidenebis(3-hydroxy-1,4-naphthoquinone) isolated from *Impatiens balsamina* L. [5].

Allergic conjunctivitis is a term used to describe a highly prevalent group of heterogeneous diseases of the ocular surface affecting 20% of the population worldwide [6]. It is usually associated with type I hypersensitivity reaction, and the spectrum of clinical expression varies according to individual cases [6,7]. The incidence of allergies in developed countries has been increasing in recent years [8–11]. Although a single cause of this increase cannot be pinpointed, experts are considering the contribution of numerous factors, including genetics, air pollution in urban areas, pets, and early childhood exposure to infections. As an increasing number of people require treatment for allergies, the associated costs have increased substantially [12]. Ocular allergy, which

received June 24, 2013
revised October 10, 2013
accepted October 23, 2013

Bibliography

DOI <http://dx.doi.org/10.1055/s-0033-1351098>
Published online November 28, 2013
Planta Med 2013; 79:
1698–1704 © Georg Thieme
Verlag KG Stuttgart · New York ·
ISSN 0032-0943

Correspondence

Prof. Dr. Alexander N. Shikov
Saint Petersburg Institute of
Pharmacy
47/5 Piskarevskiy pr.
195067 St. Petersburg
Russia
Phone: + 7 81 23 22 56 05
Fax: + 7 81 25 43 13 74
alex79@mail.ru

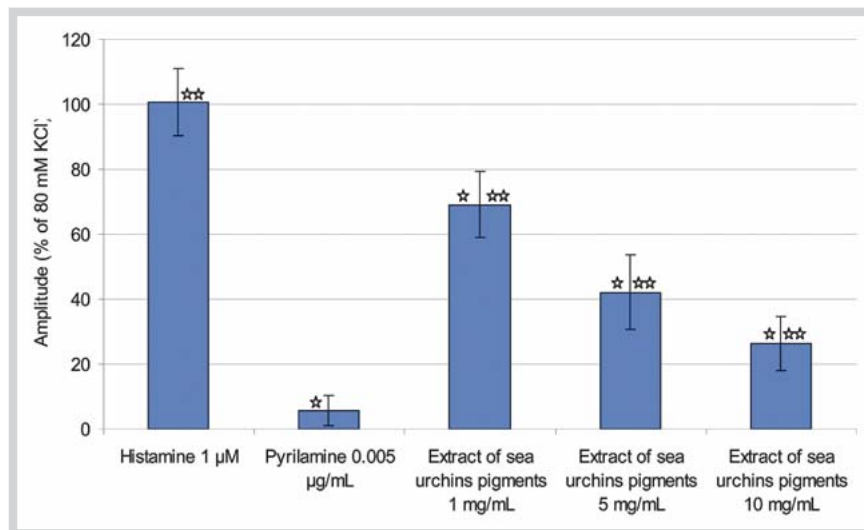


Fig. 1 Effect of extract of sea urchin pigments on histamine-induced contraction of the isolated guinea pig ileum; * $p < 0.05$, vs. control group; ** $p < 0.05$, vs. pyrilamine-treated group. (Color figure available online only.)

Table 1 Means of cumulative clinical scores (\pm SEM) recorded 5 minutes, 15 minutes, 1 hour, 2 hours, and 24 hours after instillation of compound 48/80 in group II (untreated allergic conjunctivitis model), group III (allergic conjunctivitis model pretreated with 0.1% olopatadine eye drops), and group IV (allergic conjunctivitis model pretreated with 1% pigments solution).

Time	Clinical score		
	Group II	Group III	Group IV
After 5 minutes (n = 6)	8.66 \pm 2.52	5.65 \pm 1.04	2.82 \pm 0.86
After 15 minutes (n = 6)	9.67 \pm 1.76	5.32 \pm 1.84	4.66 \pm 2.21
After 1 hour (n = 6)	9.50 \pm 2.64	4.50 \pm 1.78	4.50 \pm 0.63
After 2 hours (n = 6)	10.16 \pm 1.64	3.49 \pm 1.23	1.83 \pm 0.67
After 24 hours (n = 6)	5.53 \pm 2.31	0.50 \pm 0.21	0.33 \pm 0.18

is often overlooked in the presence of asthma and nasal symptoms, can itself produce irritating symptoms, and severe forms such as atopic keratoconjunctivitis which could lead to visual loss.

For the treatment of allergic conjunctival diseases, several local ocular drugs, such as antihistamines, antiallergic agents, and corticosteroids, have been developed as commercially available eye drops recently. Antiallergic eye drop products which are the basic therapy for allergic conjunctival diseases are classified into 2 types based on their pharmacological characteristics: drugs that suppress the release of a mediator (e.g., disodium cromoglycate) and histamine H_1 antagonists. Especially histamine H_1 antagonists are recommended as the first choice of treatment for cases with severe nasal symptoms and itching because prompt symptom relief is expected.

In the present paper, the analysis of possible antiallergic effects of sea urchins' extract is provided on animal models – guinea pig ileum contraction, rabbit eyes allergic conjunctivitis, and rabbit local skin irritation. The most active compounds were evaluated *in silico* with docking experiments to propose a molecular mode of action.

Results

Histamine is an important chemical mediator in allergic reactions and inflammation. Guinea pig ileum is commonly used to define whether a potential drug has an antihistaminic effect [13]. In the present study, we found an antihistaminic action of extract of sea urchins' pigments on H_1 -receptor. Histamine-induced contraction was dose-dependently suppressed by pyr-

amine, an H_1 -receptor antagonist treatment (► Fig. 1). As compared with vehicle treatment, the histamine-induced contraction of the isolated guinea pig ileum was largely inhibited by the extract of sea urchin pigments. Fifty percent inhibitory dose (ID_{50}) for the total extract from green sea urchin shells in equivalents of spinochrome B was 1.2 μ g/mL.

All rabbit eyes in the control group (group I) had a clinical score of zero when examined at the previously mentioned time intervals, so they are not included in the result table (► Table 1). In group II, instillation of compound 48/80 on the ocular surface produced almost immediate manifestation in the form of lid edema, conjunctival edema, discharge, and conjunctival redness. The latter three scoring criteria were evident, and the scores ranged from 1–4 for each one (data not shown). However, lid edema was not severe in these animals and was scored from 0–2 in all other eyes examined (data not shown). The manifestations increased gradually and were most severe 2 hours after instillation of the compound with a cumulative clinical score of 10.16 \pm 1.64, then decreased gradually to give a cumulative clinical score of 5.53 \pm 2.31 at the end of the experiment (24 hours after compound instillation).

Instillation of pigment solution from sea urchins shells (group IV) exhibited clinical scores that were less than those seen with the untreated model at all time intervals of examination. Each criterion had a score from 1–4 except lid edema as all the eyes had a score from 1–2 (data not shown). The decrease in clinical scores was significant after 5 minutes and 2 and 24 hours after compound application. However, this decrease was insignificant from 15 minutes to 1 hour of application of the compound. In group III, olopatadine treatment caused a uniform decrease in the clinical scores. All scoring criteria ranged from 1–3. The clinical scores

were significantly lower than those for untreated animals at all examinations performed at different time intervals. When the cumulative clinical scores of pigments from sea urchin shells were compared to those of olopatadine, there was an insignificant difference after 15 minutes and 60 minutes and a significant one after 5 minutes and 2 hours of compound application as the efficacy of pigments surpassed that of olopatadine.

Macroscopically, no congestion, swelling, or blue color in the subcutaneous tissues and peripheral vessels of the injection sites was observed after treatment in both the control and treatment groups.

To elucidate the mode of action, the binding of spinochromes to the H₁R receptor was analyzed by molecular docking. As the protein structures are not rigid and undergo conformational fluctuations, we generated a set of H₁R structures from classical molecular dynamic simulation of H₁R immersed into the membrane with ligand bound to account for flexibility of the receptor structure. The most flexible parts of the H₁R were the loops covering the active site, while transmembrane helices stayed almost rigid. Afterwards, a molecular docking experiment with spinochrome derivatives to the set of H₁R structures was carried out. The most favorable binding energies were found for spinochrome dimers (Table 2), especially for the more planar anhydroethylidene-6,6'-bis(2,3,7-trihydroxynaphthazarin). It should be noted that docked dimers' structures did not fully fit the binding site and remained partially interacting also with the loops shielding the binding side of the protein because they are too bulky (Fig. 2C, D). On the other hand, the poses of all monomeric ligands were similar to the position of doxepin in the crystal structure (Fig. 2A, B). The binding energies of spinochrome dimers (Table 2) are higher than the binding energies of monomers by ~2 kcal/mol, however, the dimers do not exactly complement the ligand binding site and bind further from the site up to the flexible loops region. As monomers fit better to the active site (resembling the binding mode of H₁R antagonist doxepin [14]), they may compete with histamine more effectively than dimers and therefore may be more effective inhibitors than the dimers.

Discussion

In this study we have shown the ability of an extract of pigments from sea urchin shells to act as an antiallergic agent in reducing the contraction evoked by histamine on the guinea pig isolated ileum. It should be noted that an action on intestinal motility has not been reported previously for sea urchin pigments. The extract had a significant inhibitory effect on a model of ocular allergic inflammation. In the local skin irritation test of pigments from sea urchin shells in rabbits, no toxic signs were observed.

The HPLC profile of the pigments extract is presented in Fig. 3, while chemical structures of identified pigments are presented in Fig. 4. Based on our previous HPLC-DAD-MS analysis of the pigments extract from sea urchin shells [15] and literature data [1, 16, 17], spinochrome monomers B and D as well as spinochrome dimers anhydroethylidene-6,6'-bis(2,3,7-trihydroxynaphthazarin) with its isomer and ethylidene-6,6'-bis(2,3,7-trihydroxynaphthazarin) were identified. Spinochromes are polyhydroxylated 1–4 naphthoquinone pigments.

Quinones constitute a structurally diverse class of phenolic compounds with a wide range of pharmacological properties which are the basis for different applications in the broad fields of pharmacy and medicine. In traditional medicine all over the world,

Table 2 Autodock Vina results from docking of spinochromes into the set of H₁R structures obtained from molecular dynamics. Average values of the most favorable poses as well as the respective standard deviations with maximal and minimal binding energies are reported.

Ligand	Binding energy [kcal/mol]
Spinochrome B	-8.7 ± 0.3 -9.4 – 7.9
Spinochrome D	-8.8 ± 0.4 -9.6 – 8.1
Ethylidene-6,6'-bis(2,3,7-trihydroxynaphthazarin)	-10.6 ± 0.9 -12.4 – 9.2
Anhydroethylidene-6,6'-bis(2,3,7-trihydroxynaphthazarin)	-10.8 ± 1.0 -13.0 – 8.9

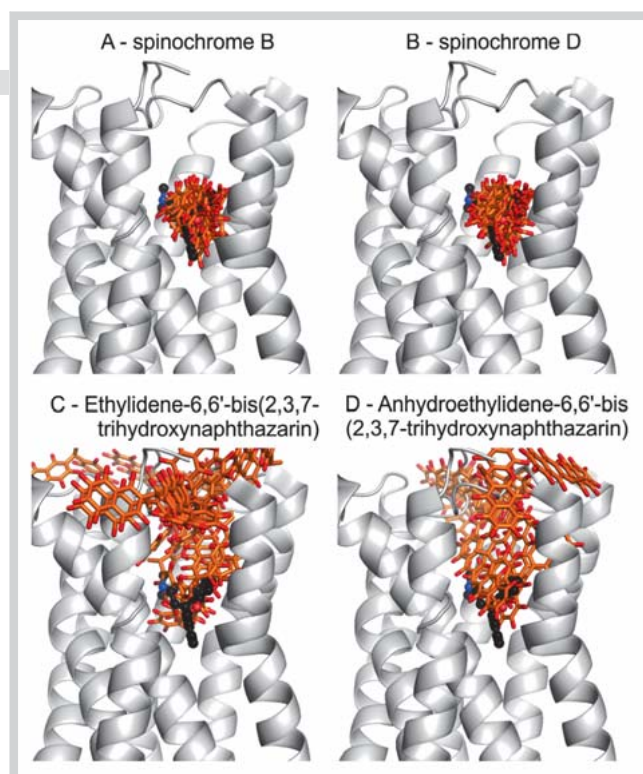


Fig. 2 Docking poses of individual spinochrome ligands (shown in orange sticks) superimposed onto H₁R crystal structure (3RZE) shown in grey. While monomeric spinochrome ligands bind preferentially to the similar position as doxepin (black sticks with balls), dimeric spinochrome ligands bound mostly to the loops in the access of the active site. (Color figure available online only.)

plants which are rich in quinones are used for treatment of a variety of diseases [14]. But there is very little data in the literature describing the antiallergic effect of natural quinones. Several investigations reported the ability of some natural quinones to be therapeutically useful in the treatment of asthma. The natural anthraquinone chrysophanol-8-O-β-D-glucopyranoside isolated from the rhizomes of *Rheum undulatum* (Polygonaceae) has potent inhibitory activity on the histamine release from mast cells, and its effect was similar to that of disodium cromoglycate, one of the commercial antiallergic drugs [18]. Cheng et al. [19] described the inhibitory effect of 9,10-anthraquinone 2-carboxylic acid on IgE-mediated passive cutaneous anaphylaxis reaction.

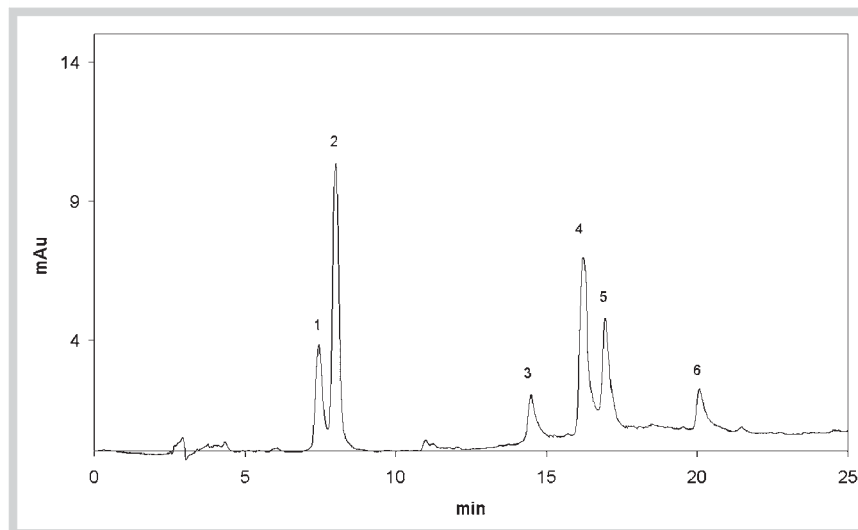


Fig. 3 Representative HPLC profile of pigments identified in the extract from sea urchin shells at 330 nm: spinochrome D (1), spinochrome B (2), non-identified spinochrome dimmer (3), anhydroethylidene-6,6'-bis(2,3,7-trihydroxynaphthazarin) (4), ethylidene-6,6'-bis(2,3,7-trihydroxynaphthazarin) (5), anhydroethylidene-6,6'-bis(2,3,7-trihydroxynaphthazarin) isomer (6).

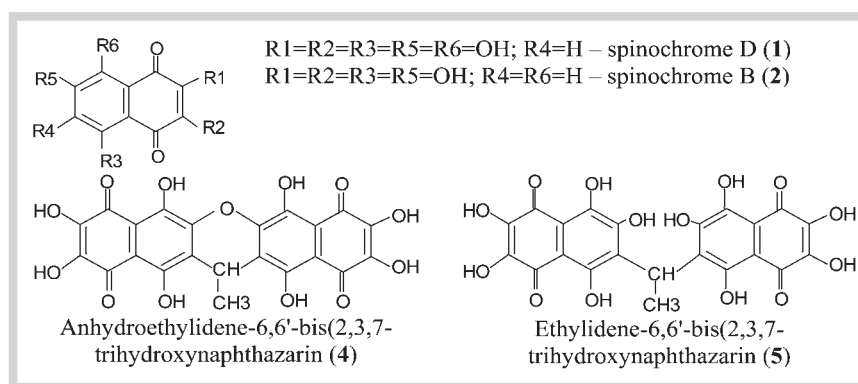


Fig. 4 Chemical structures of polyhydroxy-1,4-naphthoquinone pigments identified in the extract from sea urchin shells.

The antipruritic effects of orally administered 1,4-naphthoquinone derivatives on compound 48/80-induced scratching behavior in mice were studied. It was shown that 1,4-naphthoquinone derivatives exhibited significant antipruritic activity [5]. 2-Propoxy-1,4-naphthoquinone and 2-butoxy-1,4-naphthoquinone exhibited a potent inhibitory effect on neutrophil superoxide anion formation [20]. Three naphthoquinone derivatives isolated from the leaves of *Rhinacanthus nasutus* were tested for an antiallergic effect. The results indicated that all three compounds have potent antiallergic activity against antigen-induced β -hexosaminidase release as a marker of degranulation in the DCL-23 cells, without cytotoxic effects on non-sensitized cells [21]. Plumbagin (derivative 2-methyl-1,4-naphthoquinone, vitamin K3 analogue) modulated the levels of intracellular reactive oxygen species and glutathione and suppressed PHA-induced activation of NF- κ B in human peripheral blood mononuclear cells *in vitro* [22].

Quite a few 1,4-quinones and 1,4-naphthoquinones cause toxic side effects in cells that are related to cyclic redox reactions of the quinoid structure and formation of semiquinone and reactive oxygen species as intermediates [23]. The number and position of hydroxyl groups are important features for different activities of quinoids. β -Hydroxyl groups in a naphthol ring change quinoids metabolism and sufficiently restrict their cytotoxicity [23,24]. The position of OH groups in the quinoid fragment is important for the antiallergic action of flavonoids on hexosaminidase release from RBL-2H3 cells. Among the compounds tested, luteolin, apigenin (with OH groups at R3 and R5), quercetin (with OH groups at R2, R3, and R5), and fisetin (with OH groups at R2 and R5) were found to be most active with IC_{50} values 1.8–4.5 μ M

[25]. Hydroxyl groups in polyhydroxylated 1,4-naphthoquinones in the R1, R2, and R5 positions play key roles in both iron-ion complexing and free radical scavenging [26]. In this study, we hypothesized that antiallergic properties of sea urchin pigments could be associated with hydroxyl groups in their structure.

In summary, this is the first report on the antiallergic effects of the extract of pigments from sea urchin shells. The extract significantly reduced, in a dose-dependent manner, the histamine contractions of the isolated guinea pig ileum, had an inhibitory effect on the model of ocular allergic inflammation and did not show any irritating effect in rabbits. The extract predominantly contained polyhydroxy-1,4-naphthoquinone which would be responsible for its pharmacological activity. Molecular docking into H_1R receptor structures obtained from molecular dynamics simulations shows that all spinochrome derivatives bind to the receptor active site, but spinochrome monomers fit better to the active site. The binding of spinochrome derivatives to the H_1R receptor may rationalize the antiallergic activity of sea urchin pigments. The results of the present study suggest possibilities for the development of new agents for treating allergic diseases on the base of pigments from sea urchin shells.

Materials and Methods

Sea urchin material

Green sea urchins, *Strongylocentrotus droebachiensis* (O. F. Müller, 1776; fam. Strongylocentrotidae), were harvested in Barents Sea, a marginal sea of the Arctic Ocean, located north of Norway and

Russia, in August 2012. The organisms were identified by Dr. Svetlana Ivanova, and a voucher specimen (SDB55) was deposited in the St. Petersburg Institute of Pharmacy (St. Petersburg, Russia).

Extraction of pigments from sea urchins shells

Pigments were isolated from the shells with spines by the method of Amarowicz et al. [1] with modifications. After removal of the internal organs, like gonads, the shells were washed with a stream of cold water and air-dried at 4 °C for 2 days in the dark. The dried shells were ground and dissolved by gradually adding lactic acid, and then the pigments were extracted with diethyl ether. The ether layer was washed with 5% NaCl until the acid was almost removed. The ether solution including the pigments was dried over anhydrous sodium sulfate, and the solvent was evaporated under reduced pressure. The extract including the pigments was stored at -20 °C in the dark.

Chromatographic fingerprint analysis

The liquid chromatographic apparatus Shimadzu comprised two LC20AD pumps, a DGU-20 A3 degasser, and an SPD-M20 A diode-array detector. Separation was achieved on a 4.6 mm i.d. × 150 mm, 5 µm particle, Luna C₁₈ column (Phenomenex) with a SecurityGuard pre-column (2 mm) containing the same adsorbent, operating at room temperature and with a flow rate of 1.0 mL/min. The binary mobile phase consisted of 0.1% formic acid (A) and (B) acetonitrile-methanol (5:9, v/v). The elution profile was as follows: 0–25 min 30–80% B in A (linear gradient). Injection volume was 20 µL. Detection was carried out between the wavelengths of 200 and 800 nm. The data analysis system consisted of the LC Solution (Shimadzu). The extract was dissolved in ethanol in a concentration of 3 mg/mL, filtered (0.45 µm, Phenex; Phenomenex) and injected into the HPLC system.

Chemicals

The following chemicals were used in the present experiments: formic acid, pyrilamine maleate (purity >97%), histamine dihydrochloride, and compound 48/80, all were obtained from Sigma-Aldrich Co. HPLC-grade acetonitrile was purchased from Cryochrom, HPLC-grade methanol from J. T. Baker. Water was purified using a Simplicity UV system (Millipore).

Animals

Male guinea pigs weighing 350–400 g, male grey chinchilla and white rabbits weighting 2.7–3.1 kg (Rappolovo) were used in this study. Guinea pigs were housed in standard cages and maintained under standard laboratory conditions (temperature 20–26 °C, relative humidity 50–70%, 12-h light/12-h dark cycle) with free access to a solid pellet (Volosovo) diet and water *ad libitum* throughout the study. Rabbits free of any signs of ocular inflammation or gross abnormality were housed individually in separate cages under standard laboratory conditions with free access to a solid pellet diet and water *ad libitum*. All procedures used in the present study were approved by the Institutional Ethics Committee on the Use of Animals (approved on 10.10.2012, protocol No. 32/12.3), complied with the National Institutes of Health Guide for Care and Use of Laboratory Animals (Publication No. 85–23, revised 1985) and followed the Principles of Good Laboratory Practice (GOST R 53434–2009, identical to OECD GMP). The minimum number of animals required to obtain consistent experimental data was used.

Histamine-induced contraction of the isolated guinea pig ileum

Histamine-induced contraction of the isolated guinea pig ileum was carried out following the method described in a previous report by Matsumoto et al. [27]. The guinea pigs were anesthetized with diethyl ether and then euthanized by decapitation. The ileum was rapidly removed and immersed in ice-cold, oxygenated, modified Krebs-Henseleit solution (KHS) consisting of 118.0 mM NaCl, 4.7 mM KCl, 25.0 mM NaHCO₃, 1.8 mM CaCl₂, 1.2 mM NaH₂PO₄, 1.2 mM MgSO₄, and 11.0 mM dextrose. The ileal segments (intact ileum), approximately 1.0 cm in length, were placed longitudinally in a well-oxygenated (95% O₂, 5% CO₂) bath containing 10 mL of KHS at 30 °C. The segments were equilibrated for 60 min before the experiments were initiated. Changes in the tone of the preparation with a 1.0-g initial resting tension were measured. The contractions were recorded using computer software (PhysExp Mass; Kardioprotect). For the contraction studies, the ileal segments were first contracted using 80 mM KCl, these responses being taken as 100%. After washing and equilibrating for 1 h, the segments were treated with extract of sea urchin pigments (1, 5, 10 mg/mL), pyrilamine (0.005 µg/mL), or vehicle for 30 min. After this incubation period, histamine (10⁻⁶ M) was cumulatively applied. Contractile responses were expressed as a percentage of the contraction previously induced by 80 mM KCl.

Induction of allergic inflammation in the conjunctiva

Compound 48/80 was prepared in phosphate-buffered saline at a 50 mg/mL concentration, and the pH was adjusted to 7.4 with 0.1 N NaOH [28]. The solution was instilled once (20 µL) into the conjunctival sac of grey chinchilla rabbit eyes. The control group received 50 µL saline.

Pigments from sea urchin shells (1%) were dissolved in distilled water. The pigment solution in a dose of 50 µL was instilled once daily for three days in both rabbit eyes which were observed for any signs of irritation.

Twenty-four hours after the instillation of compound 48/80, the animals were killed. Multiple parameters of clinical symptoms evident during the acute-phase reaction and the cellular components of the late-phase reaction were evaluated in all groups of animals.

Animal groups (six animals per group)

A – Control group:

▶ group I – intact, grey chinchilla rabbits received saline.

B – Allergic conjunctivitis model groups:

▶ group II – grey chinchilla rabbits received saline in a dose of 50 µL;

▶ group III – grey chinchilla rabbits received 0.1% olopatadine hydrochloride eye drops (Alcon-Couvreur) in a dose of 50 µL;

▶ group IV – grey chinchilla rabbits received 1% pigment solution in a dose of 50 µL.

Tested solutions were instilled three times: once daily two days before instillation, then one hour before instillation of compound 48/80 in the allergic conjunctivitis model groups or of saline in the control groups.

Ophthalmological examination

Ophthalmological examination was performed using direct ophthalmoscope Beta 100 (Heine) in the mode of refraction 40 diopters. The clinical responses were graded using a modified version of the published method 5, 15, 60, 120 min, and 24 h after the al-

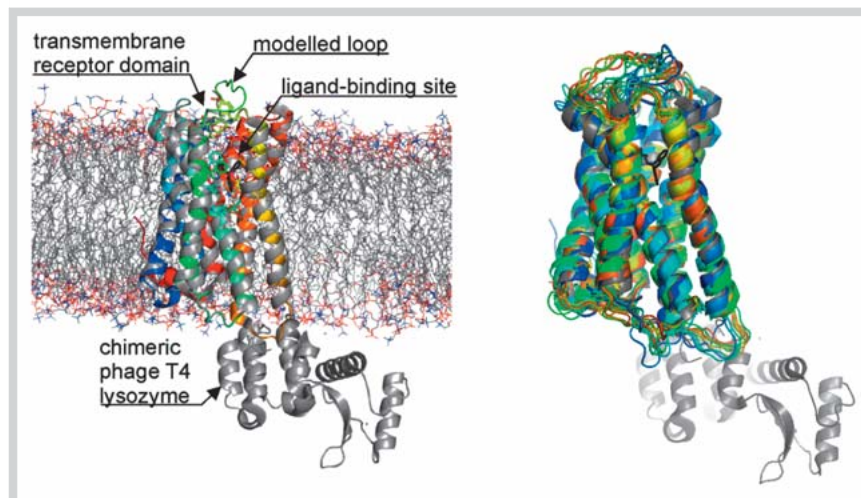


Fig. 5 Membrane bound model of histamine H₁-receptor (H₁R) in left panel. The crystal structure 3RZE containing transmembrane receptor domain and lysozyme domain is shown in grey (Shimamura, 2011). The lysozyme part was deleted from the structure, and a short loop not resolved in the X-ray was added from homology modeling. Doxepin ligand is shown in black sticks inside the ligand-binding domain (water molecules are not shown for clarity). Right panel shows flexible movements of the loops outside of the membrane during simulation while retaining the transmembrane helical region. Snapshots (taken each 10 ns) are shown; membrane is not shown for clarity. Figures were generated with Pymol 0.99rc6 (www.pymol.org). (Color figure available online only.)

lergen challenge. Conjunctival edema, lid edema, tear/discharge, and conjunctival redness were graded from 0 to 4 [29]. The cumulative clinical score was calculated as the sum of the scores of each of the four parameters with a range from 0 to 16. Before the experiment, all rabbit eyes had a clinical score of zero.

Local skin irritation test after subcutaneous administration of the extract of pigments from sea urchin shells in rabbits

Local skin irritation test followed the method described by Wang et al. [30] with some changes. White rabbits were used as the animal model, and 8 rabbits were selected. After the chest wall skin on both sides of each rabbit was disinfected with povidone iodine, 0.1 mL of the pigment extract was injected subcutaneously into the left chest wall at a concentration of 10 mg/mL, and the right chest wall was injected with an equal volume of blank solution in the same way. Thirty minutes after the subcutaneous injection, each rabbit was intravenously injected with 2 mL of 1% solution of Evans blue.

All animals were sacrificed by air injection into the ear vein 30 min after treatment. Changes in skin, subcutaneous tissues, and vessels at the injection site were observed macroscopically. Immediate-type allergic reactions are accompanied by an increase in capillaries due to the release of histamine and therefore by the appearance of subcutaneous blue spots. If the size of the subcutaneous blue spot observed was greater than 6 mm, the reaction was considered positive.

Molecular modeling

Histamine H₁-receptor (H₁R) structure has been recently cocrystallized with the first-generation antagonist drug doxepin (PDBID: 3RZE) at 3.1 Å resolution [31]. The crystal structure is chimeric and consists of two domains: human ligand-binding transmembrane receptor domain with doxepin bound and phage T4 lysozyme domain (● Fig. 5). Missing loops on top of the ligand-binding pocket of the transmembrane receptor domain (using corresponding parts of the UniprotID: P35367 sequence) were modelled by homology modeling using Modeller 9.9 [32]. As the structure of the receptor had relatively low resolution, molecular dynamic simulation in lipid bilayer was employed in order to generate receptor structures for the molecular docking procedure. For molecular dynamic simulation, the doxepin molecule and phosphate ion were kept in the structure. Topologies of

both ligands were obtained from PRODRG2Beta server [33]. The transmembrane protein with bound ligands was immersed into the dioleoylphosphatidylcholine (DOPC) bilayer with helices perpendicular to the membrane plane in the GROMACS 4.5.4 program suite [34]. Lipid bilayer containing 532 molecules of DOPC was solvated with 30,120 molecules of SPC/E water. To the system, 79 sodium and 95 chlorine ions were added to mimic their physiological concentrations and to neutralize the system prepared for simulation. Molecular dynamics were run in the Gromacs 4.5.4 program suite with a Gromos 53a6 force field [35] for protein and ligands combined with a Berger force field for DOPC [36]. After energy minimization using the steepest descent algorithm, the lipid bilayer and water were equilibrated with 1-ns-long conditioning with the positional constraint imposed on the protein structure at 310 K with a V-rescale thermostat and Berendsen anisotropic pressure coupling to 1 bar. Electrostatic interactions were calculated with particle mesh Ewald method with a cutoff of 1.0 nm, and van der Waals interactions were calculated with a cutoff of 1.1 nm. After the conditioning, ten consecutive 1-ns-long simulations with randomized starting velocities were used to equilibrate the protein structure without any positional restraint. Finally, a 100-ns-long production simulation was carried out (310 K and 1 atm) with Nose-Hoover pressure coupling to collect structures of H₁R, each 1 ns-long.

Molecular docking

Polyhydroxy-1,4-naphthoquinone pigments were modeled with Marvin (<http://www.chemaxon.com>). After the geometry of molecules was optimized with Marvin the compounds were prepared for docking in the AutoDockTools program suite. Autodock Vina program [37] was used for docking of seven ligands (● Fig. 2) into the set of H₁R structures obtained from the molecular dynamics. For docking, only the transmembrane protein without ligands, lipids, and water molecules was considered as a receptor. A 15 Å box with the center on the doxepin-binding position was used. Exhaustiveness was set to 20. Only poses with the lowest energy were selected for the analysis. Average values as well as the standard deviations with maximal and minimal binding energies are reported for all ligands.

Statistical analysis of experimental data

Data are presented as the mean ± SEM. Statistical analyses were performed by analysis of variance (ANOVA).

Acknowledgements

Part of this work has been supported by the operational programs Education for Competitiveness – European Social Fund (CZ.1.07/2.3.00/20.0017) and Research and Development for Innovations – European Regional Development Fund (CZ.1.05/2.1.00/03.0058 and/01.0030). The authors acknowledge support by the Grant Agency of the Czech Republic grant P208/12/G016 and funding by the Centre of the Region Haná for Biotechnological Agricultural Research (grant No. ED0007/01/01).

Conflict of Interest

The authors declare no conflicts of interest.

Affiliations

- ¹ Saint Petersburg Institute of Pharmacy, St. Petersburg, Russia
- ² Department of Physical Chemistry, Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacký University, Olomouc, Czech Republic
- ³ Department of Physical Chemistry, Regional Centre of Advanced Technologies and Materials, Faculty of Science, Palacký University, Olomouc, Czech Republic
- ⁴ Department of Medical Chemistry and Biochemistry, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University, Olomouc, Czech Republic

References

- ¹ Amarowicz R, Synowiecki J, Shahidi S. Chemical composition of shells from red (*Strongylocentrotus franciscanus*) and green (*Strongylocentrotus droebachiensis*) sea urchin. *Food Chem* 2012; 133: 822–826
- ² Mishchenko NP, Fedoreev SA, Bagirova VL. Histochochrome: a new original domestic drug. *Pharm Chem J* 2003; 37: 48–52
- ³ Anufriev VPh, Novikov VL, Maximov OB, Elyakov GB, Levitsky DO, Lebedev AV, Sadretidinov SM, Shvilkin AV, Afonskaya NI, Ruda MY, Cherpachenko NM. Synthesis of some hydroxynaphthazarins and their cardioprotective effects under ischemia-reperfusion *in vivo*. *Bioorg Med Chem Lett* 1998; 8: 587–592
- ⁴ Pozharitskaya ON, Ivanova SA, Shikov AN, Makarov VG. Evaluation of free radical-scavenging activity of sea urchin pigments using HPTLC with post-chromatographic derivatization. *Chromatographia* 2013; 76: 1353–1358
- ⁵ Oku H, Kato T, Ishiguro K. Antipruritic effects of 1,4-naphthoquinones and related compounds. *Biol Pharm Bull* 2002; 25: 137–139
- ⁶ Trocme SD, Karan K. Spectrum of ocular allergy. *Curr Opin Allergy Clin Immunol* 2002; 29: 423–427
- ⁷ Ono SJ, Abelson MB. Allergic conjunctivitis: updates on pathophysiology and prospects for future treatment. *J Allergy Clin Immunol* 2005; 115: 118–122
- ⁸ Barbee RA, Kaltenborn W, Lebowitz MD, Burrows B. Longitudinal changes in allergen skin test reactivity in a community population sample. *J Allergy Clin Immunol* 1987; 79: 16–24
- ⁹ Aberg N. Asthma and allergic rhinitis in Swedish conscripts. *Clin Exp Allergy* 1989; 19: 59–63
- ¹⁰ Maziak W, Behrens T, Brasky TM, Duhme H, Rzehak P, Weiland SK, Keilet U. Are asthma and allergies in children and adolescents increasing? Results from ISAAC phase I and phase III surveys in Munster, Germany. *Allergy* 2003; 58: 572–579
- ¹¹ Verlato G, Corsico A, Villani S, Cerveri I, Migliore E, Accordini S, Carolei A, Piccioni P, Bugiani M, Cascio VL, Marinoni A, Poli A, de Marco R. Is the prevalence of adult asthma and allergic rhinitis still increasing. Results of an Italian study. *J Allergy Clin Immunol* 2003; 111: 1232–1238
- ¹² Bielory L. Update on ocular allergy treatment. *Expert Opin Pharmacother* 2002; 3: 541–553
- ¹³ Larsson AK, Fumagalli F, DiGennaro A, Andersson M, Lundberg J, Edenius C, Govoni M, Monopoli A, Sala A, Dahlén SE, Folco GC. A new class of nitric oxide-releasing derivatives of cetirizine; pharmacological profile in vascular and airway smooth muscle preparations. *Br J Pharmacol* 2007; 151: 35–44
- ¹⁴ Martínez MJA, Benito PB. Biological activity of quinones. *Stud Nat Prod Chem* 2005; 30: 303–366
- ¹⁵ Shikov AN, Ossipov VI, Martiskainen O, Pozharitskaya ON, Ivanova SA, Makarov VG. The offline combination of thin-layer chromatography and high-performance liquid chromatography with diode array detection and micrOTOF-Q mass spectrometry for the separation and identification of spinochromes from sea urchin (*Strongylocentrotus droebachiensis*) shells. *J Chromatogr A* 2011; 1218: 9111–9114
- ¹⁶ Kuwahara R, Hatate H, Yuki T, Murata H, Tanaka R, Hama Y. Antioxidant property of polyhydroxylated naphthoquinone pigments from shells of purple sea urchin *Anthocidaris crassispina*. *LWT – Food Sci Technol* 2009; 42: 1296–1300
- ¹⁷ Chang CWJ, Moore JC. Pigments from some marine specimens. The isolation and spectral characterization of spinochromes. *J Chem Educ* 1971; 48: 408–409
- ¹⁸ Kim DH, Park EK, Bae EA, Han MJ. Metabolism of rhaponticin and chrysophanol 8-o- β -D-glucopyranoside from the rhizome of *Rheum undulatum* by human intestinal bacteria and their anti-allergic actions. *Biol Pharm Bull* 2000; 23: 830–833
- ¹⁹ Cheng HH, Kuo SC, Lin WC. Pharmacodynamic and pharmacokinetic studies of anthraquinone 2-carboxylic acid on passive cutaneous anaphylaxis in rats. *Res Commun Mol Pathol Pharmacol* 1999; 105: 97–103
- ²⁰ Lien JC, Huang LJ, Teng CM, Wang JP, Kuo SC. Synthesis of 2-alkoxy 1,4-naphthoquinone derivatives as antiplatelet, antiinflammatory, and anti-allergic agents. *Chem Pharm Bull* 2002; 50: 672–674
- ²¹ Tewtrakul S, Tansakul P, Panichayupakaranant P. Anti-allergic principles of *Rhinacanthus nasutus* leaves. *Phytomedicine* 2009; 16: 929–934
- ²² Kohli V, Sharma D, Sandur SK, Suryavanshi S, Sainis KB. Immune responses to novel allergens and modulation of inflammation by vitamin K3 analogue: a ROS dependent mechanism. *Int Immunopharmacol* 2011; 11: 233–243
- ²³ O'Brien PJ. Molecular mechanisms of quinone cytotoxicity. *Chem Biol Interact* 1991; 80: 1–41
- ²⁴ Ollinger K, Brunmark A. Effect of hydroxy substituent position on 1,4-naphthoquinone toxicity to rat hepatocytes. *J Biol Chem* 1991; 266: 21496–21503
- ²⁵ Cheong H, Ryu SY, Oak MH, Cheon SH, Yoo GS, Kim KM. Studies of structure-activity relationship of flavonoids for the anti-allergic actions. *Arch Pharm Res* 1998; 21: 478–480
- ²⁶ Lebedev AV, Ivanova MV, Levitsky DO. Iron chelators and free radical scavengers in naturally occurring polyhydroxylated 1,4-naphthoquinones. *Hemoglobin* 2008; 32: 165–179
- ²⁷ Matsumoto T, Horiuchi M, Kamata K, Seyama Y. Effects of *Bidens pilosa* L. var. *radiata* SCHERFF treated with enzyme on histamine-induced contraction of guinea pig ileum and on histamine release from mast cells. *J Smooth Muscle Res* 2009; 45: 75–86
- ²⁸ Udell IJ, Abelson MB. Animal and human ocular surface response to a topical nonimmune mast-cell degranulating agent (compound 48/80). *Am J Ophthalmol* 1981; 91: 226–230
- ²⁹ Miyazaki D, Tominaga T, Yakura K, Kuo C, Komatsu N, Inoue Y, Ono SJ. Conjunctival mast cell as a mediator of eosinophilic response in ocular allergy. *Mol Vis* 2008; 14: 1525–1532
- ³⁰ Wang S, Wu M, Li D, Jiao M, Wang L, Zhang H, Liu H, Wang D, Han B. Preparation, characterization and related *in vivo* release, safety and toxicity studies of long acting lanreotide microspheres. *Biol Pharm Bull* 2012; 35: 1898–1906
- ³¹ Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han GW, Kobayashi T, Stevens RC, Iwata S. Structure of the human histamine H1 receptor complex with doxepin. *Nature* 2011; 475: 65–70
- ³² Yang Z, Lasker K, Schneidman-Duhovny D, Webb B, Huang CC, Pettersen EF, Goddard TD, Meng EC, Sali A, Ferrin TE. UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *J Struct Biol* 2012; 179: 269–278
- ³³ Schüttelkopf AW, van Aalten DMF. PRODRG – a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D* 2004; 60: 1355–1363
- ³⁴ Hess B, Kutzner C, Van der Spoel D, Lindahl E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 2008; 4: 435–447
- ³⁵ Oostenbrink C, Villa A, Mark AE, Van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 2004; 25: 1656–1676
- ³⁶ Berger O, Edholm O, Jähnig F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys J* 1997; 72: 2002–2013
- ³⁷ Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multi-threading. *J Comput Chem* 2010; 31: 455–461

Appendix K

Havlíková M, Huličiak M, **Bazgier V**, Berka K, Kubala M:
Fluorone dyes have binding sites on both cytoplasmic and extracellular domains
of Na,K-ATPase.

Biochim. Biophys. Acta, Biomembr., 1828(2), 568-576, 2013.

DOI: 10.1016/j.bbamem.2012.10.029

IF=3.498



Fluorone dyes have binding sites on both cytoplasmic and extracellular domains of Na,K-ATPase

Marika Havlíková^{a,*}, Miroslav Huličiak^a, Václav Bazgier^b, Karel Berka^c, Martin Kubala^a

^a Department of Biophysics, Faculty of Science, Palacký University in Olomouc, tř. 17. listopadu 12, 771 46 Olomouc, Czech Republic

^b Department of Physical Chemistry, Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacký University in Olomouc, tř. 17. listopadu 12, 771 46 Olomouc, Czech Republic

^c Department of Physical Chemistry, Regional Center of Advanced Technologies and Materials, Faculty of Science, Palacký University in Olomouc, tř. 17. listopadu 12, 771 46 Olomouc, Czech Republic

ARTICLE INFO

Article history:

Received 28 June 2012

Received in revised form 24 October 2012

Accepted 31 October 2012

Available online 8 November 2012

Keywords:

Na,K-ATPase

C45 loop

Binding sites

Fluorone dyes

Time-resolved fluorescence

Fluorescence quenching

ABSTRACT

Combination of fluorescence techniques and molecular docking was used to monitor interaction of Na,K-ATPase and its large cytoplasmic loop connecting fourth and fifth transmembrane helices (C45) with fluorone dyes (i.e. eosin Y, 5(6)-carboxyeosin, rose bengal, fluorescein, and erythrosine B). Our data suggested that there are at least two binding sites for all used fluorone dyes, except of 5(6)-carboxyeosin. The first binding site is located on C45 loop, and it is sensitive to the presence of nucleotide. The other site is located on the extracellular part of the enzyme, and it is sensitive to the presence of Na⁺ or K⁺ ions. The molecular docking revealed that in the open conformation of C45 loop (which is obtained in the presence of ATP) all used fluorone dyes occupy position directly inside the ATP-binding pocket, while in the closed conformation (i.e. in the absence of any ligand) they are located only near the ATP-binding site depending on their different sizes. On the extracellular part of the protein, the molecular docking predicts two possible binding sites with similar binding energy near Asp897(α) or Gln69(β). The former was identified as a part of interaction site between α- and β-subunits, the latter is in contact with conserved FXVD sequence of the γ-subunit. Our findings provide structural explanation for numerous older studies, which were performed with fluorone dyes before the high-resolution structures were known. Further, fluorone dyes seem to be good probes for monitoring of intersubunit interactions influenced by Na⁺ and K⁺ binding.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Na,K-ATPase (NKA, sodium pump, EC 3.6.3.9) is one of the most important enzymes in animal cell metabolism. It is a member of the P-type ATPase superfamily of active cation transport proteins. Enzymes belonging to this protein family are found in all kingdoms of life and pump specific ions through the biological membranes at the expense of one ATP molecule per cycle [1–4].

The overall stoichiometry of NKA reaction is three Na⁺ ions transported out of the cell and two K⁺ ions into the cell for each ATP hydrolyzed [5]. Electrochemical gradient formed by this active

transport is essential for the excitability of muscle and nerve tissues. Sodium ion gradient is essential for the maintenance and accumulation of important nutrients, ions and neurotransmitters, reduction of intracellular level of calcium, and regulation of cell volume and intracellular pH [6,7]. Due to important NKA role in cell metabolism, this pump has been linked to disorders displaying various degrees of severity. Neurodegenerative disorders are related to mutations in brain NKA, such as the familial hemiplegic migraine type 2 [8] and the movement disorder familial rapid-onset dystonia parkinsonism [9].

NKA reaction cycle is conveniently described by a model proposed by Albers and Post [10,11], which describes two major conformation states, E1 state with high affinity for sodium and ATP, and E2 state with high affinity for potassium and low affinity for ATP. It was described that roughly 30% of ATP generated in cells metabolism is used for functioning of NKA [12].

Recently, high-resolution crystal structure of NKA was published [13,14]. NKA is characterized by the presence of two subunits. The catalytic α-subunit contains the conserved sequences characteristic for the P-type ATPase superfamily [3] and it is responsible for the cation translocation as well as ATP hydrolysis. The α-subunit has ten transmembrane helices and two longer cytoplasmic segments that are organized into three well-separated domains; they form the

Abbreviations: NKA, Na,K-ATPase, sodium pump; SERCA, sarco(endo)plasmic reticulum; C23, cytoplasmic segment connecting the second and third transmembrane helices, other cytoplasmic segments analogously; ATP, adenosine 5'-triphosphate; RB, Rose Bengal; ErB, Erythrosine B; EY, Eosin Y; CE, 5(6)-carboxyeosin; Fl, Fluorescein; PMSF, Phenylmethylsulfonyl Fluoride; IPTG, isopropyl β-D-thiogalactoside; DTT, dithiothreitol; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TCSPC, time-correlated single photon counting method; IRF, instrument response function; K_{SV}, Stern–Volmer quenching constant; K_D, dissociation constants; FITC, fluorescein isothiocyanate

* Corresponding author.

E-mail address: marika.havlikova@upol.cz (M. Havlíková).

N or nucleotide-binding domain, the P or phosphorylation domain, and the A or actuator domain [14]. The N and P domains are formed by the C45 loop, the A domain is formed by the N-terminal tail and the C23 loop [15].

The β -subunit is transmembrane glycoprotein consisting of a cytoplasmic N-terminal domain, a single membrane spanning segment, and a larger extracellular carboxyl-terminal domain [16]. The presence of the β -subunit is of fundamental importance as it acts as a specific chaperone assisting the correct membrane insertion of the newly synthesized Na,K-ATPase and H,K-ATPase α -subunits and, hence, in their structural and functional maturation [17]. In addition, association of the β -subunit determines some of the intrinsic transport properties of Na,K- and H,K-ATPases, it modulates cation sensitivity of the pump [18].

The γ -subunit of Na,K-ATPase is a protein belonging to the FXYP family. It is often associated with the $\alpha\beta$ -complex as a third subunit and regulates pumping activity. Transmembrane part of γ -subunit has approximately 30 amino acids with mostly α -helical structure. It was shown that extracellular part of γ -subunit containing conserved FXYP sequence moves in between α - and β -subunits where it may contact β -subunit [13].

Large cytoplasmic loop of NKA (C45 loop) is located between fourth and fifth transmembrane helices of its α -subunit and it forms two of the three cytoplasmic domains. It represents approximately 40% of α -subunit mass. C45 loop contains four of the most highly conserved P-type ATPase sequences, including the sequence D-K-T-G, which contains an invariant aspartate residue (Asp369) that is phosphorylated by ATP as part of the catalytic mechanism of these proteins. It was shown that the C45 loop can be isolated from the rest of the enzyme [19–21], retaining its structure [22], dynamic properties [21,23] and ability to bind nucleotides [24–26], and it can be expressed with (His)₆-tag in bacterial cells [19]. Moreover, fluorescence experiments, which were performed in our lab, show that changes of isolated C45 loop conformation occurring due to the ATP- and Mg²⁺-binding are the same as expected changes for this part of the entire enzyme during the catalytic cycle [21,23].

Fluorone dyes are hydroxy-xanthene dyes (Fig. 1). The xanthene dyes are among the oldest synthetic dyes and have many important applications. Xanthene (fluorone) dyes are often used as drug components (virostatic drugs), laser dyes, in food industry for food coloring, as photosensitizers in the photodynamic therapy of cancer or in medicine as a tool for diagnostic. These compounds, which include fluorescein, rose bengal, eosin Y, 5(6)-carboxyeosin, and erythrosine B, have the same ring skeleton and absorb, in general, around 500 nm depending on the different ring substituent [27].

Fluorescein and reactive derivatives of fluorescein have been the most widely used class of organic fluorophores for labeling and biomolecular sensing due to their high absorption cross-sections, high fluorescence quantum yields, and their ability to attach to biomolecules [28]. In medicine, fluorescein is used as an active drug component in fluorescein angiography.

Rose bengal (RB) is a water-soluble xanthene dye that is currently used in ophthalmology for the diagnosis of dry eyes. It is a potential photosensitizer for photodynamic therapy of tumors due its high absorption coefficient in the visible region of the spectrum and a tendency to transfer electrons from its excited triplet state, producing long-lived radicals [29].

Food colorant erythrosine B is a well-known carcinogen. It was found that it is a non-specific inhibitor of a number of protein–protein interactions within the tumor necrosis factor superfamily [30].

Eosin Y is considered as a reversible inhibitor of NKA, which is competitive with ATP in the absence of K⁺ ions. Its fluorescence behavior is very sensitive to conformational changes of this enzyme [31]. It was described that the fluorescence of eosin Y in the presence of NKA is enhanced by Mg²⁺ [32], which is an important cofactor for the enzyme activity.

	R ₁	R ₂	R ₃	R ₄
Fluorescein	H	H	H	Na
Eosin Y	H	Br	Br	Na
Erythrosine B	H			Na
Rose Bengal	Cl			Na

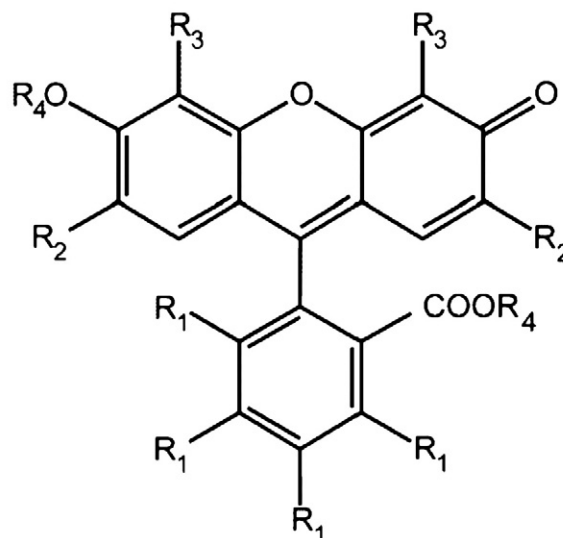


Fig 1. Chemical structure of fluorone dyes.

Based on the structure similarity with eosin Y, NKA is a possible target for all of these described dyes. Moreover, NKA belongs to the family of membrane proteins, which are in general extremely potent drug targets. This role together with the fact, that NKA has prominent position in the cell, renders it of most importance in cellular physiology.

In this work, intrinsic fluorescence of these dyes was used in combination of steady-state and time-resolved fluorescence methods to observe interactions with NKA. In order to get more detailed structural information, we tested also the interaction with its isolated large cytoplasmic loop (C45 loop).

2. Materials and methods

2.1. Materials

All solutions were made using deionized water and all reagents were of the highest purity commercially available: adenosine 5'-triphosphate (ATP, Tris salt), magnesium chloride (MgCl₂), rose bengal sodium salt, erythrosine B, fluorescein sodium salt, eosin Y, disodium salt, disodium hydrogen orthophosphate (Na₂HPO₄), leupeptin, pepstatin, phenylmethylsulfonyl (PMSF) and adenosine 5'-triphosphatase from cerebral cortex (NKA) were from Sigma. 5(6)-Carboxyeosin was from MGT Inc. Tris[hydroxymethyl]aminomethane (Tris) was from BioRad. Sodium chloride (NaCl) and potassium chloride (KCl) were from Fluka. Isopropyl β -D-thiogalactoside (IPTG) was from Roth and dithiothreitol (DTT) was from Serva.

100 μ M stock solutions of rose bengal, erythrosine B, fluorescein, 5(6)-carboxyeosin, and eosin Y were stored in the dark and 4 °C.

2.2. Expression and purification of C45 loop of Na,K-ATPase α -subunit

The DNA coding C45 loop of the Na,K-ATPase transformed into Rosetta *Escherichia coli* bacteria strain and were cultured at 37 °C to OD₆₀₀ of 0.8. Induction was carried out with 0.2 mM IPTG overnight at 17 °C. Cells were centrifuged, resuspended in 20 mM Tris-HCl, 140 mM NaCl, pH 8.8 containing protease inhibitors (2 μ g/ml leupeptin, 2 μ g/ml pepstatin and 1 mM PMSF), disrupted by sonication and the

homogenate was again centrifuged. All constructs were expressed as a (His)₆-tag fusion protein; the (His)₆-tag was attached to the N-terminus. Purification affinity chromatography was performed according to standard TALON Metal Affinity Resin (Clontech) manufacturer protocol. The protein samples were eluted with 0.5 M imidazol and were dialyzed against 1 l of 50 mM Tris-HCl, 140 mM NaCl, pH 7.5 overnight at 4 °C. The purity of protein samples was verified using 12% SDS-PAGE. Concentrations were estimated using the Bradford assay [33].

2.3. Interaction of fluorone dyes with NKA

Interactions of fluorone dyes with NKA and its large cytoplasmic loop (C45) were measured. Stock solutions of C₁₂E₈-solubilized NKA and C45 loop were dissolved to the final 2 μM concentration into 10 mM Tris-HCl buffer solution, pH 7.5, containing different concentrations of NaCl and KCl. The buffer referred to as NaCl-buffer contained 140 mM NaCl and 20 mM KCl, in turn, buffer referred as KCl-buffer contained 20 mM NaCl and 140 mM KCl.

Titration curves were made as follows. Fresh stock solutions of fluorone dyes (50 μM) were always titrated into 120 μl of buffer solution containing protein in 1 μl step, gently mixed and fluorescence emission intensity was measured. Final concentrations of fluorone dyes varied from 0.4 μM to 3.8 μM. Control curves with free fluorone dyes in solution were made in the same way, but in the absence of protein.

2.4. Steady-state fluorescence measurement

The steady-state fluorescence excitation and emission spectra were monitored on spectrofluorimeter F4500 (Hitachi, Japan). Data were collected using an excitation wavelength of 516 nm for eosin Y, 520 nm for 5(6)-carboxyeosin, 527 nm for erythrosine B, 548 nm for rose bengal, and 490 nm for fluorescein. Slits were set to 5 nm for both the excitation and emission channels, respectively, scan speed was 240 nm/min and the measurements were performed at 22 °C (bath controlled).

2.5. Time-resolved fluorescence measurement

Time-resolved fluorescence experiments were measured using time-correlated single-photon counting (TCSPC) method on the spectrometer PicoHarp300 (PicoQuant, Germany), using the pulsed diode centered at 445 nm (PicoQuant, Germany). The pulse frequency was 10 MHz. The data were collected into the histogram spanning the 100 ns time-range with the 32 ps/channel resolution. Emission was measured for eosin Y at 536 nm, for 5(6)-carboxyeosin at 540 nm, for erythrosine B at 546 nm, and for rose bengal at 575 nm. All experiments were carried out at 22 °C (bath controlled).

The experiment proceeded until 10,000 counts in the peak channel were achieved. The instrument response function (IRF) was obtained using water as a scatterer. The software FluoFit 4.2.1 (PicoQuant) was used for fitting. The model used was sum of exponentials with instrument-response-function (IRF) deconvolution:

$$I(t) = \text{IRF} \otimes \sum_i a_i \times e^{-t/\tau_i},$$

and the intensity-weighted mean fluorescence lifetime was calculated as

$$\tau_M = \frac{\sum_i \alpha_i \tau_i^2}{\sum_i \alpha_i \tau_i}.$$

For all samples we were able to get a fit with χ^2_R close to 1.00 and a random distribution of residuals.

2.6. Determination of dissociation constants from the steady-state fluorescence experiments

Increasing amounts of various fluorone dyes were added to the NaCl- or KCl-buffer in the absence or presence of 2 μM NKA or C45 loop. Data were collected on the spectrofluorometer F4500 (Hitachi, Japan) in the mode of photometry. Slits were set to 5 nm for both excitation and emission channels, integration time was 3 s and measurements were performed at 22 °C. Excitation and emission wavelengths were fixed and differed according to the used dye. Data were collected using excitation/emission wavelengths 525/560 nm for eosin Y, 520/540 nm for 5(6)-carboxyeosin, 538/575 nm for erythrosine B, 564/590 nm for rose bengal, and 490/512 nm for fluorescein. The plot of the normalized fluorescence intensity F at the selected wavelengths on the dye concentration in the presence of C45 loop was fitted to the equation [25]:

$$F^* = [L]_T + \frac{\gamma - 1}{2} \left(([L]_T + [P]_T + K_D - \sqrt{([L]_T + [P]_T + K_D)^2 - 4[L]_T[P]_T}) \right)$$

where F^* is normalized fluorescence intensity of probe, $[L]_T$ and $[P]_T$ are concentrations of ligand and protein, respectively, and γ is the ratio of the fluorescence responses of bound and free ligand. This formula represents a fitting equation for experimental data from titration experiments and thus for K_D determination of single binding site, via non-linear regression.

For the entire NKA, the plot of the normalized fluorescence intensity F at the selected wavelengths on the dye concentration in the presence of NKA was fitted to the equation [34]:

$$F = F_L[L] + F_{M_1L}K_1[M][L] + F_{M_2L}K_2[M][L] + 2F_{ML_2}K_1K_2[M][L]^2$$

where $[M]$ and $[L]$ are concentrations of free enzyme and ligand, respectively, F_{M_iL} denotes the molar fluorescence values for each bound ligand and F_{ML_2} is an average value for the two bound ligands, and where

$$[M] = \frac{[M_0]}{1 + K_1[L] + K_2[L] + K_1K_2[L]^2}$$

and $[L]$ is determined by solving a cubic equation:

$$[L]^3 K_1 K_2 + [L]^2 (2K_1 K_2 [M_0] - K_1 K_2 [L_0] + K_1 + K_2) + [L] \{1 + (K_1 + K_2)([M_0] - [L_0])\} - [L_0] = 0$$

where $[M_0]$ and $[L_0]$ are the total concentrations of the enzyme and ligand, respectively. This formula represents a fitting equation for experimental data from titration experiments and thus for the determination of association constants K_1 and K_2 for two non-identical and non-interacting binding sites, via non-linear regression.

The correction of fluorescence intensity on inner filter effect was done before fitting.

Table 1

The maxima of excitation- and emission spectra of various fluorone dyes free in solution (λ_{emF}) and in the presence of Na,K-ATPase or C45 loop (λ_{emB}). The emission wavelength λ_{obs} reflects the wavelength with the maximal ratio of fluorescence intensities of the bound- and free probes, and it was used in the titration experiments aimed at the determination of the dissociation constants. Na,K-ATPase, C45 loop: 5 μM; dyes: 0.8 μM.

Dye	λ_{ex} (nm)	λ_{emF} (nm)	λ_{emB} (nm)	λ_{obs} (nm)
Rose bengal	545	565	575	590
Erythrosine B	527	546	555	575
Eosin Y	515	536	544	560
Carboxyeosin	520	536	540	560
Fluorescein	490	512	512	512

Table 2

Dissociation constants (K_D ; nM) for rose bengal, erythrosine B, 5(6)-carboxyeosin, and eosin Y binding to Na,K-ATPase and C45 loop are expressed as mean \pm S.D. from five independent measurements. Values were obtained by fitting the experimental data as described in the **Materials and methods**. Final concentrations: Na,K-ATPase: 2 μ M; and C45 loop: 2 μ M. For experiments with C45 loop NaCl-buffer was used. Excitation/emission wavelengths: RB: 564/590; EY: 525/560; CE: 520/560; and ErB: 538/575.

	Na,K-ATPase		C45 loop
	NaCl-buffer	KCl-buffer	
Rose bengal	190 \pm 50	130 \pm 10	50 \pm 10
Erythrosine B	310 \pm 40	330 \pm 70	80 \pm 10
Eosin Y	160 \pm 20	250 \pm 20	50 \pm 10
Carboxyeosin	1000 \pm 200	793 \pm 3	120 \pm 2

2.7. Time-resolved fluorescence quenching and determination of quenching constants

The time-resolved fluorescein, rose bengal, erythrosine B, 5(6)-carboxyeosin, and eosin Y fluorescence emission was monitored using the pulsed diode centered at 445 nm and emission wavelengths of 512, 575, 546, 540, and 536 nm, respectively, in the absence or presence of 2 μ M NKA in NaCl-, KCl-, NaI-, and KI-buffers, where the iodide anion served as a fluorescence quencher. The buffer referred as NaI-buffer contained 140 mM NaI and 20 mM KCl, in opposite, buffer referred as KI-buffer contained 140 mM KI and 20 mM NaCl. Final concentration of the dyes was 1 μ M. For monitoring changes related to individual sub-states of enzyme pumping cycle [10,11] 2.5 mM MgCl₂, 2.5 mM ATP, or 1.7 mM P_i was added. Quenching efficiency was determined using the Stern–Volmer formula:

$$K_{SV} = \frac{\tau_0 - \tau}{\tau [Q]}$$

where τ_0 and τ are the mean lifetime of fluorescence in the absence (NaCl- or KCl-buffers) and presence (NaI- or KI-buffers) of quencher, respectively, $[Q]$ is the concentration of quencher and K_{SV} is the apparent Stern–Volmer quenching constant.

Accessibility of fluorophore is proportional to the apparent bimolecular quenching constant k_q , which was determined using formula:

$$k_q = K_{SV}/\tau_0.$$

2.8. Statistics

Presented values are mean \pm S.D. from 3 to 5 replicates. Statistical significance of differences was evaluated using the unpaired two-tailed Student's *t*-test.

2.9. Homology modeling

Structure of complete human NKA was prepared with use of Modeller 9.10 [35] with template structure of spiny dogfish NKA (PDB ID: 2ZXE). Model structure of C45 loop was taken from Ref. [21]. Hydrogen atoms were added with AutoDock Tools program suite [36] and Kollman charges were assigned.

2.10. Docking

AutoDock Tools program was used for the preparation of ligand molecules. Ligands (eosin Y, erythrosine B, rose bengal, fluorescein, and 5(6)-carboxyeosin) were docked to both protein structures using AutoDock Vina [37] with grid box covering the whole protein in order to find ligand binding poses without prior knowledge. To do so, we have adjusted docking parameters defining exhaustiveness of the search to the maximum values (num_modes 9999 and exhaustiveness

100). The blind docking resulted in several poses. Afterwards, the obtained the positions have been redocked with smaller grid (cube with size of 3 nm in each direction) around detected binding sites to reassess the position and energetics (Table 4).

3. Results

First, we monitored changes in emission spectra of all dyes upon their interaction with NKA (Fig. 2, and Table 1). It enabled selection of the emission wavelength of maximal fluorescence-intensity difference between the free and bound probes, and this wavelength was used for subsequent titration experiments (Fig. 3). Notably, titration curves in experiments with isolated C45 loop could have been satisfactorily fitted by the single-site model, while in experiments with the entire NKA, the two-site model was necessary, except for 5(6)-carboxyeosin, where a single-site model adequately described the data.

The molecular docking identified on C45 loop a binding site near the ATP-binding site for all dyes. Moreover, when the docking was performed into the structure that C45 loop adopts in the presence of ATP [21], the fluorone dyes were bound directly into the ATP-binding pocket. Another possible binding site was identified on the P-domain in the pocket formed by segment Cys656–Ala684, however, the binding energy was slightly lower (by \sim 3 kJ/mol) than for the former site. On the entire NKA, we found additional two binding sites (with similar energies) on the extracellular part of the protein near Asp897(α) and Gln69(β). These extracellular parts were not occupied by 5(6)-carboxyeosin. The binding sites in the transmembrane domain near Asp770(α) and Glu320(α) identified by *in vacuo* docking are unlikely to be accessible in the real experiment due to the presence of lipid or detergent molecules, and were not further considered.

Furthermore, the time-resolved experiments and quenching experiments in the presence of various combinations of transmembrane domain ligands (Na⁺ and K⁺, inducing so called E1 and E2 conformations, respectively) and cytoplasmic ligands (ATP, P_i and Mg²⁺) were performed in an attempt to distinguish subtle NKA conformational changes induced by binding of these ligands (Table 3 and Tables S1 and S2 in Suppl. Mat.). The results from spectroscopic experiments for each dye are commented below.

3.1. Eosin Y

We observed an increase of fluorescence intensity together with spectral shift upon binding of eosin Y to NKA or C45 loop (Fig. 2, and Table 1). The largest change in fluorescence intensity was observed for emission at 560 nm, and this emission wavelength was used in the titration experiments. As mentioned above, the eosin Y binding to isolated C45 loop could have been fitted by a single site model yielding the $K_D = 50 \pm 10$ nM (Table 2). In the case of binding to the entire NKA, the second binding site was needed and the K_D value for binding to this site slightly varied between E1 and E2 states (160 \pm 20 nM in the NaCl-buffer, and 250 \pm 20 nM in the KCl-buffer).

Two exponentials were necessary to fit the fluorescence decay of free eosin Y with mean fluorescence lifetime of 1.38 \pm 0.02 ns. Addition of NKA or C45 loop resulted in the necessity to use a three-exponential model to fit decays and significantly increased the eosin Y mean fluorescence lifetime to \sim 2.1 ns and \sim 1.44 ns, respectively, with little variations depending on the present ligands (Table 3 and Table S1 in Suppl. Mat.). For NKA in the presence of ATP of P_i, the eosin Y mean fluorescence lifetime significantly differed between E1 and E2 states.

Measurement of fluorescence decay in the presence of iodide anions was a further attempt to distinguish individual conformational sub-states. The fluorescence lifetime decreased in all cases, confirming that iodide serves as a dynamic fluorescence quencher. Calculated apparent bimolecular quenching constant is proportional to the steric accessibility

Table 3
Mean fluorescence lifetimes τ_M (ns) of various fluorone dyes in the presence of Na,K-ATPase, in various buffer solutions and in the presence of different substrate molecules are expressed as mean \pm S.D. from three to five independent measurements. Final concentrations: RB, ErB, and EY: 1 μ M; Na,K-ATPase: 2 μ M; Mg^{2+} : 2.5 mM; ATP: 2.5 mM; and P_i : 1.7 mM. λ_{ex} : 445 nm; λ_{em} : RB: 575 nm, ErB: 546 nm, and CE, EY: 536 nm. Significant differences of τ_M obtained in the presence of Na,K-ATPase compared with τ_M of the free probe are shown as * for $p < 0.05$ or ** for $p < 0.01$, states of the protein-bound probes that were mutually significantly different in the NaCl- and KCl buffers are denoted by # ($p < 0.05$).

Buffer	Substrate	RB	ErB	EY	Fl	CE
NaCl	No substrate	0.81 \pm 0.06**	0.81 \pm 0.01**	2.05 \pm 0.01**	4.02 \pm 0.04	1.49 \pm 0.10*
	+Mg ²⁺	0.84 \pm 0.04**	0.80 \pm 0.02**	2.04 \pm 0.06**	4.09 \pm 0.14	1.48 \pm 0.13
	+ATP	0.82 \pm 0.01**	0.80 \pm 0.02**	2.05 \pm 0.02**#	4.01 \pm 0.01	1.30 \pm 0.10
	+P _i	0.83 \pm 0.01*	0.79 \pm 0.01*	2.05 \pm 0.02**#	4.03 \pm 0.02	1.42 \pm 0.04*
	+Mg ²⁺ + ATP	0.83 \pm 0.04**	0.81 \pm 0.02**	2.04 \pm 0.02**	4.04 \pm 0.06	1.26 \pm 0.08
KCl	+Mg ²⁺ + P _i	0.84 \pm 0.03**	0.82 \pm 0.04*	2.07 \pm 0.03**	4.04 \pm 0.05	1.44 \pm 0.05*
	No substrate	0.81 \pm 0.04**	0.79 \pm 0.07**	2.13 \pm 0.07**	4.01 \pm 0.02	1.49 \pm 0.01**
	+Mg ²⁺	0.84 \pm 0.03**	0.81 \pm 0.07**	2.13 \pm 0.07**	4.00 \pm 0.02	1.61 \pm 0.14
	+ATP	0.84 \pm 0.02**	0.80 \pm 0.02*	2.10 \pm 0.02**#	4.00 \pm 0.03	1.48 \pm 0.01**
	+P _i	0.84 \pm 0.02**	0.80 \pm 0.06	2.09 \pm 0.03**#	3.99 \pm 0.02	1.50 \pm 0.01**
	+Mg ²⁺ + ATP	0.85 \pm 0.02**	0.83 \pm 0.04**	2.11 \pm 0.05**	3.99 \pm 0.02	1.50 \pm 0.00**
	+Mg ²⁺ + P _i	0.85 \pm 0.03**	0.80 \pm 0.05*	2.14 \pm 0.07**	4.01 \pm 0.03	1.51 \pm 0.01**

of the fluorophor (Table S2 in Suppl. Mat.). In the presence of protein (with all variants of ligands), the bimolecular quenching constant was significantly reduced compared to the case of free eosin Y, thus, confirming that the binding occurs. However, the sensitivity to conformational changes was rather little and we observed significant difference only between E1 and E2 states in the absence of any cytoplasmic ligand (Mg^{2+} , ATP or P_i).

3.2. 5(6)-Carboxyeosin

We observed a slight increase of fluorescence intensity together with spectral shift upon binding of 5(6)-carboxyeosin to NKA or C45 loop (not shown), and, again, we observed the largest change in fluorescence intensity at 560 nm. 5(6)-Carboxyeosin was the only dye, for which the binding to the entire NKA could have been satisfactorily fitted by a single site model and K_D value varied only slightly between E1 and E2 states (1000 \pm 200 nM in the NaCl-buffer, and 793 \pm 3 nM in the KCl-buffer).

Again, two exponentials were necessary to fit the fluorescence decay of free 5(6)-carboxyeosin with mean fluorescence lifetime of 1.22 \pm 0.05 ns. Presence of NKA or C45 loop increased the number of exponentials necessary to fit the decays to three with mean fluorescence lifetime of 1.49 ns (Table 3 and Table S1 in Suppl. Mat.). In E1 state, the mean fluorescence lifetime of 5(6)-carboxyeosin significantly differed in the presence of ATP (both in the presence or absence of Mg^{2+}) from all the other states. The quenching experiments provided no additional information in this case.

3.3. Erythrosine B

As in the case of eosin Y, we observed an increase of erythrosine B emission intensity together with red spectral shifts upon binding to NKA or C45 loop (Fig. 2, and Table 1), and we found that the largest change in fluorescence intensity occurs at the emission wavelength of 575 nm. The K_D value for binding to C45 loop was 80 \pm 10 nM,

and for the second binding site on NKA in E1 and E2 states 310 \pm 40 nM and 330 \pm 70 nM, respectively.

Again, two exponentials were necessary to fit the fluorescence decays of free erythrosine B with a mean fluorescence lifetime of 0.4 \pm 0.1 ns, while in the presence of protein, three components were necessary. In the case of NKA, the mean fluorescence lifetime increased to ~0.8 ns, and was significantly different from free probe under all conditions. The τ_M increase in the presence of C45 was only to values ~0.5 ns, and statistically significant difference from free probe was observed only in the presence of Mg^{2+} or in NaCl buffer in the presence of ATP (Table S1 in Suppl. Mat.). For the entire NKA, we found that erythrosine B fluorescence was not sensitive to difference between E1 and E2 states, nor to the presence of any cytoplasmic ligand.

In quenching experiments, we observed a significant difference in the E1-state between k_q values for conformation in the presence of $Mg^{2+} + P_i$ and conformation in the presence of Mg^{2+} only, and in the E2-state between conformations without any cytoplasmic ligand and conformation with Mg^{2+} (not shown).

3.4. Rose bengal

In the case of rose bengal, we observed the largest increase of fluorescence intensity and red spectral shift caused by the presence of NKA or C45 loop (Fig. 2, and Table 1). The K_D value for binding to C45 loop was 50 \pm 10 nM, and for the second binding site binding on NKA in E1 and E2 states was 190 \pm 50 nM and 130 \pm 10 nM, respectively.

Three exponentials were needed to fit the fluorescence decay kinetic for both free- and protein-bound rose bengal. The mean fluorescence lifetime increased from ~0.4 ns for the free probe to ~0.8 ns after binding to NKA (Fig. 4) or to ~0.7 ns after binding to C45 loop, the difference being statistically significant in all cases (Table 3 and Table S1 in Suppl. Mat.). Rose bengal was insensitive to the differences between E1 and E2 conformations, as well as to the presence of cytoplasmic ligands, and neither quenching experiments brought any new information.

Table 4
Binding energies (kJ/mol) calculated from molecular docking simulations for binding of fluorone dyes to Na,K-ATPase and isolated C45 loop in closed- or open conformation. N.D. is acronym for "none detected".

Protein	Na,K-ATPase		C45 loop closed/no ATP-like		C45 loop open/ATP-binding-like	
	Extracellular D897(Å)	Extracellular Q69(Å)	ATP-binding site	P domain	ATP-binding site	P domain
Eosin Y	-38	-38	-34	-33	-32	N.D.
Carboxyeosin	N.D.	N.D.	N.D.	-24	-31	N.D.
Erythrosine B	-35	-36	-30	-26	-31	N.D.
Rose bengal	-37	-36	-30	-28	-31	N.D.
Fluorescein	-38	-38	-38	-35	-36	N.D.

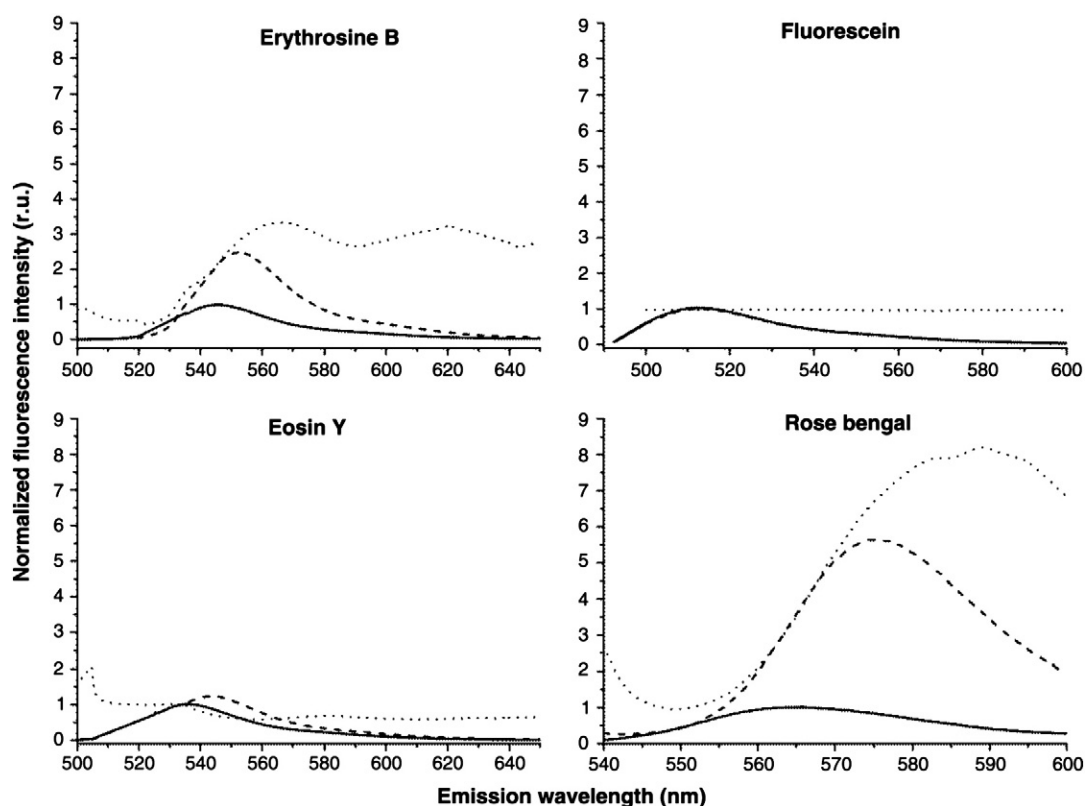


Fig. 2. Normalized fluorescence emission spectra of various fluorone dyes free in NaCl-buffer solution (solid line; 5 μM) and in the presence of C45 loop (dash line; Dye: 5 μM ; C45 loop: 5 μM); λ_{ex} : RB: 545 nm, ErB: 527 nm, EY: 515 nm, and Fl: 490 nm. Dot line represents ratio of fluorescence intensity of the dye in the presence of protein to fluorescence intensity of the dye free in solution, maximum was selected as wavelength λ_{obs} . The same results were observed in the presence of NKA.

3.5. Fluorescein

The case of fluorescein was very interesting. The molecular docking predicted that it should be able to bind to both NKA and C45 loop. However, none of the fundamental fluorescence characteristics (i.e., fluorescence intensity, shape of spectra or fluorescence lifetime) turned out to be sensitive to the interaction with protein and it hindered us to

determine the dissociation constants. The only experimental evidence that the binding does occur, we got from the quenching experiments, where we observed a significant decrease of the apparent bimolecular quenching constant in the presence of protein. In the presence of quencher, we observed also a significant difference between E1 and E2 states (Table S2 in Suppl. Mat.).

4. Discussion

NKA plays a central role in the metabolism of all animal cells, maintaining the membrane potential and steep gradient of sodium within the plasma membrane. Examined fluorone dyes are involved

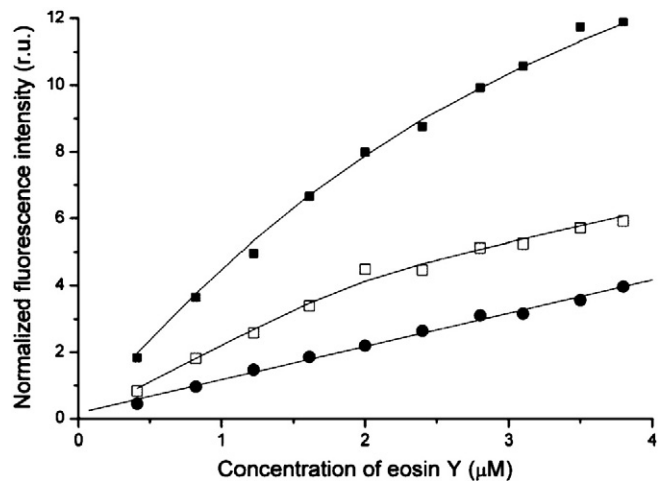


Fig. 3. Changes in normalized fluorescence intensity of eosin Y upon binding to NKA or C45 loop. Solid circles represent eosin Y free in NaCl-buffer solution, open squares represent eosin Y in the presence of C45 loop, and solid squares represent eosin Y in the presence of Na,K-ATPase. Solid lines represent the best fit to experimental data using the single-site model for C45, the two-site model for NKA, and linear fit to the free probe titration. Final concentrations: NKA: 2 μM ; C45 loop: 2 μM . λ_{ex} : 525 nm, and λ_{em} : 560 nm.

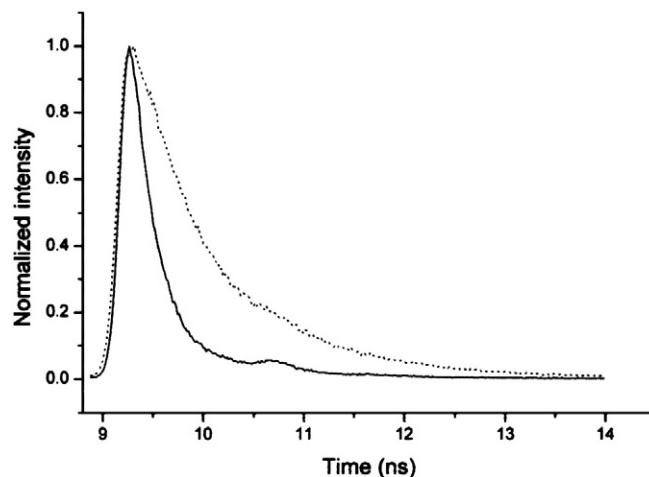


Fig. 4. Fluorescence decay of 1 μM rose bengal free in NaCl-buffer solution (black line) and in the presence of 2 μM NKA (dotted line). λ_{ex} : 445 nm; λ_{em} : 575 nm.

in food as food-dyes, in medicaments as active species or markers in diagnostic methods. They are also successfully used as photosensitizers in photodynamic therapy of cancer. However, the observed interaction with NKA with sub-micromolar dissociation constants presents a potential health-risk, and it was already demonstrated that eosin Y can influence its activity [38].

Our study demonstrated that the fluorone dyes can be also useful probes for the examination of the protein structure in solution. Steady-state fluorescence experiments showed that there is an increase of emission intensity together with red-shift in the fluorescence spectra of eosin Y, 5(6)-carboxyeosin, erythrosine B and rose bengal in the presence of NKA or C45 loop. As suggested by the early paper of Martin [39], such effect can be caused by intermolecular hydrogen bonding of these dyes. The examination of fluorescein interactions was obstructed by the fact that none of the fundamental fluorescence characteristics of fluorescein altered after binding to the enzyme. However, “sensitization” of binding by addition of quencher unambiguously revealed that also fluorescein does bind to the NKA.

4.1. Two binding sites

We found that there are at least two binding sites for eosin Y, erythrosine B, rose Bengal, and fluorescein. The first binding site is located on C45 loop; the other is located on the extracellular part of the enzyme (Fig. 5). In the case of 5(6)-carboxyeosin, we found evidence only for one binding site, and experiments with isolated C45 loop revealed that it is the one located in the cytoplasmic part of the enzyme.

Existence of two binding sites provides a plausible structural explanation for early observations of Skou and Esmann [31] who concluded that eosin Y behaves like ATP-analog and has the same binding site, which is located on N-domain, and also the finding of Ogan et al. [38] that eosin Y is a mixed-type inhibitor. In general, mixed inhibition means that both substrate (ATP) and eosin can be bound to the pump and their binding is not mutually exclusive. Binding site on the extracellular part of the enzyme could still enable eosin binding while nucleotide binding site is occupied by ATP, and this view could explain eosin action as a mixed-type inhibition. In other studies on P-type ATPases, it was shown that eosin inhibits the plasma membrane Ca^{2+} -pump (PMCA) from red cells [42], smooth muscle [43] and seedlings of *Arabidopsis thaliana* [44] and that eosin is not competitive with ATP [42,43].

4.2. Cytoplasmic binding site

Recent works demonstrated that C45 loop adopts a closed conformation in the absence of any ligand (E2-like), while in the presence of ATP it has an open conformation (E1-like) [21,23]. The molecular docking revealed that the fluorone dyes occupy position directly inside the ATP-binding pocket in the open conformation, while in the closed conformation they are located only near the ATP-binding site (Fig. 6). This presents a strong support for the results of Esmann and Fedosova [45] who showed that in the presence of Na^+ and the absence of K^+ (i.e. in E1 conformation), binding of eosin and ATP to NKA is mutually exclusive. In our time-resolved experiments, we found that mean fluorescence lifetime of eosin significantly differed in the presence of ATP, from the other states induced by the presence of other ligands.

As mentioned above, in the case of docking into closed conformation of C45 loop, the dyes were not bound directly inside the ATP-binding pocket, but only in its proximity. Position of all dyes was not identical due to their different sizes, and the smallest fluorescein was located deeper toward the ATP-binding pocket than the other analogs. Our results also shed new light on the experiments of Karlish [40] and Tsuda et al. [41], who labeled enzyme with fluorescein isothiocyanate (FITC). It was found that FITC preferentially bound to Lys501, which is located

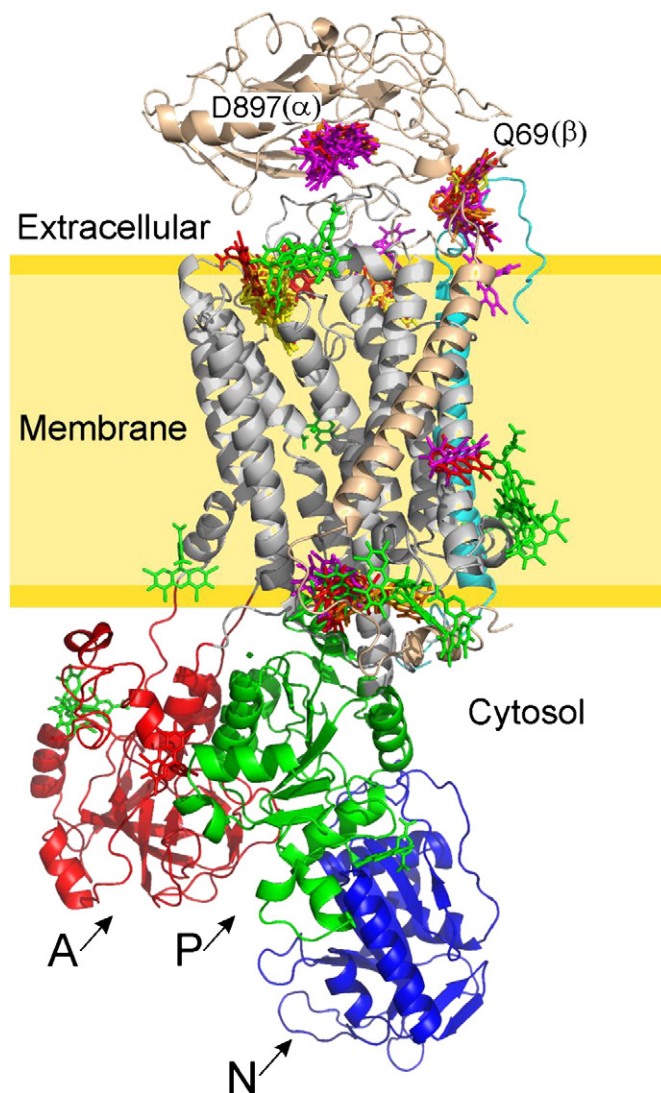


Fig. 5. Molecular docking to NKA. Transmembrane region of alpha unit is shown in gray, cytosolic A domain is shown in red, P domain in green, N domain in blue, gamma subunit is shown in cyan and beta subunit is shown in light orange. Ligands are colored as follows: fluorescein in yellow, eosin Y in orange, erythrosine B in red, rose bengal in magenta, and 5(6)-carboxyeosin in green. Five most strongly bound poses for all ligands are shown in thicker sticks. Several possible binding poses have been detected mostly in the extracellular part near residues D897 on alpha unit and Q69 on beta unit and in the transmembrane region.

in the depth of the nucleotide-binding pocket, and it was proposed that this selectivity is a consequence of extremely high reactivity of Lys501 with isothiocyanate groups. It was also shown that several non-fluorescein two-sites labels (crosslinkers) as dihydro-4,4'-diisothiocyanostilbene-2,2'-disulfonate (H_2DIDS) and 4-acetamido-4'-isothiocyanatostilbene-2,2'-disulfonic acid (SITS) label Lys501 and inhibit the NKA [46,47]. In addition to these observations, our data suggest that the affinity of fluorone dye itself toward the ATP-binding pocket is a significant factor that increases probability of reaction with Lys501 over other lysine residues. Electrostatic interactions between positively charged ATP-binding pocket and negatively charged fluorone dyes probably play a key role.

Monitoring of another fundamental parameter, the kinetic of fluorescence decay, enabled the monitoring of subtle conformational changes of the enzyme. Recent experiments [21,23] with isolated C45 loop suggested that cytoplasmic ligands ATP and Mg^{2+} can bind to the enzyme during two steps of the catalytic cycle, ATP-binding causes opening of the

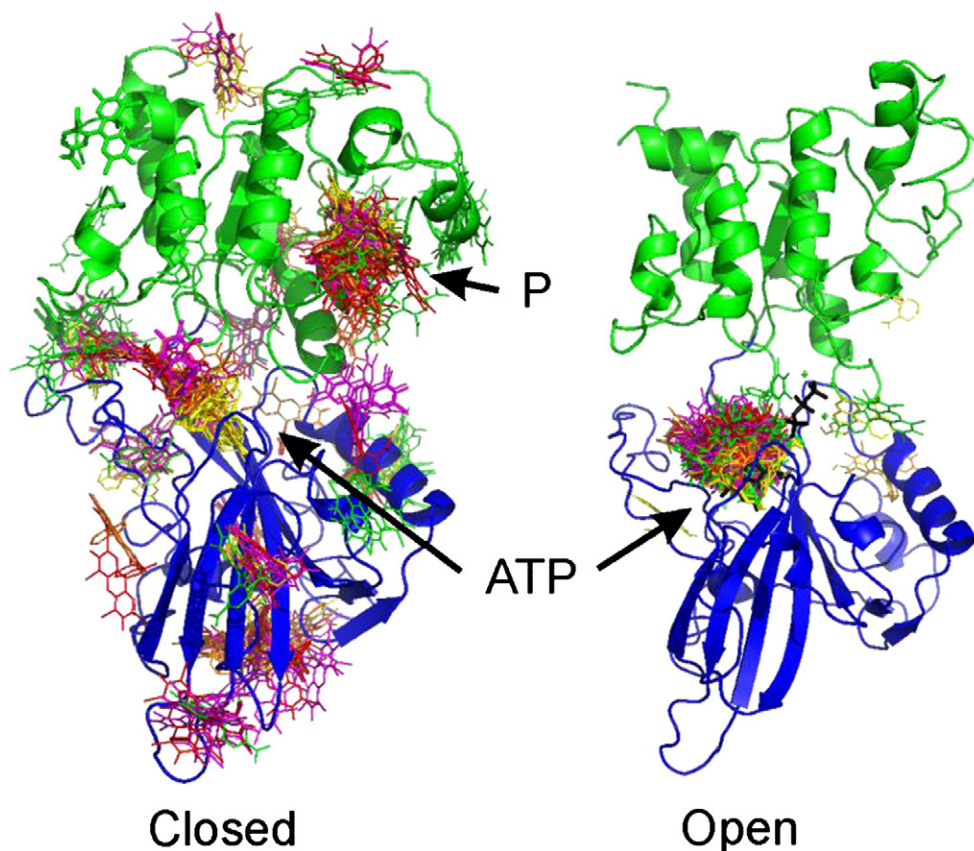


Fig. 6. Molecular docking to C45 loop in closed and open conformations. P domain is shown in green, and N domain in blue. Ligands are colored as follows: fluorescein in yellow, eosin Y in orange, erythrosine B in red, rose bengal in magenta, and 5(6)-carboxyeosin in green. Most strongly bound poses for all ligands are shown in thicker sticks. ATP is shown in black. Two possible binding poses have been detected in closed structure – near ATP binding site and on P domain, while only one predominant can be found in the case of open structure in the ATP-binding pocket.

cytoplasmic headpiece and Mg^{2+} brings it to the closed conformation again, and obviously these changes are likely to be reflected in the transmembrane domain. However, the abovementioned studies did not verify, whether conformational changes observed for isolated C45 loop could be observed also with the entire enzyme. Binding of eosin Y to the entire enzyme was sensitive to differences between E1- and E2 conformations just in the presence of ATP or P_i (without Mg^{2+}). It revealed that the enzyme conformation in the magnesium-free states is different and provided a strong support for the abovementioned hypothesis.

4.3. Extracellular binding site(s)

Moreover, we found that fluorone dyes can be a good marker of Na^+ or K^+ binding to the binding sites in the transmembrane domain, traditionally assigned to forcing the enzyme to adopt E1 or E2 conformation, respectively.

Molecular docking predicted two possible extracellular binding sites with roughly the same binding energy. One of them was located near the Asp897 of the α -subunit, and this site was shown to be part of the α/β -subunit interface [14]. The other possible site is located near Gln69 of the β -subunit, which is in direct contact with the conserved FXVD motif of the γ -subunit. At the moment, our data do not allow us to distinguish, which of these two sites is preferentially occupied by fluorone dyes, but in any case, both of these sites are likely to be very sensitive to the mutual movement of NKA subunits during the E1 \rightarrow E2 conformational transitions.

Fluorescein turned out to be the best fluorone dye to sense the E1 \rightarrow E2 conformational transitions. The differences of fluorescence decays between E1 and E2 were significant for all combinations of cytoplasmic ligands, particularly, when the fluorescence quencher iodide

was present in the buffer. Another dye enabling sensing of differences between E1 and E2 states, independently on the presence of cytoplasmic ligands, was eosin Y. Although, it was proposed that eosin Y is not optimal to monitor conformational changes in detergent-solubilized Na,K-ATPase [48], monitoring of the kinetic of fluorescence decay enabled us to observe these changes even in the presence of detergent. However, simultaneous sensitivity of eosin Y to changes in both the cytoplasmic and extracellular parts is expected to be a complicating factor in the interpretation of the spectroscopic data. Therefore, we consider the fluorescein to be the best fluorone dye to probe the E1–E2 transitions.

Note that 5(6)-carboxyeosin, which does not bind to the extracellular site, was insensitive to the presence of NaCl- or KCl-buffer.

5. Conclusion

We found that all examined fluorone dyes can bind to NKA with sub-micromolar affinity, and hence, their use in the medicine or food industry presents potential health risk. We identified two binding sites, one of them being localized in the cytoplasmic part at the ATP-binding pocket, the other on the extracellular domain. Existence of two binding sites presents a plausible structural explanation for previously observed mixed-type inhibition of eosin Y. Concerning their use as the NKA conformation sensitive probes, erythrosine B and rose bengal did not substantially reflect the presence of various NKA ligands, and hence, they seem to be useless in this respect. In agreement with Esmann's work, we conclude that the best probe for examination of the ATP-binding site is 5(6)-carboxyeosin, because it is the only probe that does not bind to the extracellular domain. The predicted two candidates for the possible extracellular binding site are located on the sites of contact between α/β - and β/γ -subunits and they can reflect the E1/E2

conformational change. The most sensitive probe for this transition turned out to be fluorescein in the time-resolved experiment, particularly, when the sensitivity was enhanced by the presence of collisional fluorescence quencher iodide. Eosin Y seems to be sensitive to both nucleotide binding and E1/E2 transition, which however, may complicate interpretation of the experimental data.

Acknowledgments

This work was supported by the grants GACR P205/11/P319 and P301/10/0883 from the Czech Science Foundation and by the grant no. CZ.1.07/2.3.00/20.0057 Operational Programme Education for Competitiveness from the Ministry of Education, Youth and Sports, Czech Republic.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bbamem.2012.10.029>.

References

- [1] K.B. Axelsen, M.G. Palmgren, Evolution of substrate specificities in the P-type ATPase superfamily, *J. Mol. Evol.* 46 (1998) 84–101.
- [2] W. Kühlbrandt, Biology, structure and mechanism of P-type ATPases, *Nat. Rev. Mol. Cell Biol.* 5 (2004) 282–295.
- [3] J.V. Møller, B. Juul, M. le Maire, Structural organization, ion transport, and energy transduction of P-type ATPases, *Biochim. Biophys. Acta* 186 (1996) 1–51.
- [4] H.-J. Apell, How do P-type ATPases transport ions, *Bioelectrochemistry* 63 (2004) 149–156.
- [5] K.Y. Xu, Activation of (Na⁺ + K⁻)-ATPase, *Biochem. Biophys. Res. Commun.* 338 (2005) 1669–1677.
- [6] A.A. Jaitovich, A.M. Bertorello, Na⁺, K⁺-ATPase: an indispensable ion pumping signaling mechanism across mammalian cell membranes, *Semin. Nephrol.* 26 (2006) 386–392.
- [7] G. Blanco, R.W. Mercer, Isozymes of the Na-K-ATPase: heterogeneity in structure, diversity in function, *Am. J. Physiol. Renal Physiol.* 275 (1998) 633–650.
- [8] M. De Fusco, R. Marconi, L. Silvestri, L. Atorino, L. Rampoldi, L. Morgante, A. Ballabio, P. Aridon, G. Casari, Haploinsufficiency of ATP1A2 encoding the Na⁺/K⁺ pump alpha 2 subunit associated with familial hemiplegic migraine type 2, *Nat. Genet.* 33 (2003) 192–196.
- [9] P. de Carvalho, K.J. Sweadner, J.T. Penniston, J. Zaremba, L. Liu, M. Caton, G. Linazasoro, M. Borg, M.A. Tijssen, S.B. Bressman, W.B. Dobyns, A. Brashear, L.J. Ozelius, Mutations in the Na⁺/K⁺-ATPase alpha3 gene ATP1A3 are associated with rapid-onset dystonia parkinsonism, *Neuron* 43 (2004) 169–175.
- [10] R.W. Albers, Biochemical aspects of active transport, *Annu. Rev. Biochem.* 36 (1967) 727–756.
- [11] R.L. Post, S. Kume, T. Tobin, B. Orcutt, A.K. Sen, Flexibility of an active center in sodium-plus-potassium adenosine triphosphatase, *J. Gen. Physiol.* 54 (1969) 306–326.
- [12] J.H. Kaplan, Biochemistry of Na, K-ATPase, *Annu. Rev. Biochem.* 71 (2002) 511–535.
- [13] J.P. Morth, B.P. Pedersen, M.S. Toustrup-Jensen, T.L.-M. Sørensen, J. Petersen, J.P. Andersen, B. Vilsen, P. Nissen, Crystal structure of the sodium-potassium pump, *Nature* 450 (2007) 1043–1050.
- [14] T. Shinoda, H. Ogawa, F. Cornelius, C. Toyoshima, Crystal structure of the sodium-potassium pump at 2.4 angstrom resolution, *Nature* 459 (2009) 446–U167.
- [15] C. Toyoshima, M. Nakasako, H. Nomura, H. Ogawa, Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution, *Nature* 405 (2000) 647–655.
- [16] M.D. Laughery, M.L. Todd, J.H. Kaplan, Mutational analysis of α-β subunit interactions in the delivery of Na, K-ATPase heterodimers to the plasma membrane, *J. Biol. Chem.* 278 (37) (2003) 34794–34803.
- [17] C. Gatto, S.M. McCloud, J.H. Kaplan, Heterologous expression of Na⁺-K⁺-ATPase in insect cells: intracellular distribution of pump subunits, *Am. J. Physiol. Cell Physiol.* 281 (2001) C982–C992.
- [18] R.E. Dempski, T. Friedrich, E. Bamberg, The β-subunit of the Na⁺/K⁺-ATPase follows the conformational state of the holoenzyme, *J. Gen. Physiol.* 125 (2005) 505–520.
- [19] C. Gatto, A.X. Wang, J.H. Kaplan, The M4M5 cytoplasmic loop of the Na, K-ATPase, overexpressed in *Escherichia coli*, binds nucleoside triphosphates with the same selectivity as the intact native protein, *J. Biol. Chem.* 273 (1998) 10578–10585.
- [20] C.M. Tran, R.A. Farley, Catalytic activity of an isolated domain of Na, K-ATPase expressed in *Escherichia coli*, *Biophys. J.* 77 (1999) 258–266.
- [21] L. Grycova, P. Sklenovsky, Z. Lansky, M. Janovska, M. Otyepka, E. Amler, ATP and magnesium drive conformational changes of the Na(+)/K(+)-ATPase cytoplasmic headpiece, *Biochim. Biophys. Acta (BBA)-Biomembr.* 1788 (2009) 1081–1091.
- [22] M. Kubala, T. Obsil, V. Obsilova, Z. Lansky, E. Amler, Protein modeling combined with spectroscopic techniques: an attractive quick alternative to obtain structural information, *Physiol. Res.* 53 (2004) S187–S197.
- [23] M. Kubala, L. Grycova, Z. Lansky, P. Sklenovsky, M. Janovska, M. Otyepka, J. Teisinger, Changes in electrostatic surface potential of Na(+)/K(+)-ATPase cytoplasmic headpiece induced by cytoplasmic ligand(s) binding, *Biophys. J.* 97 (2009) 1756–1764.
- [24] M. Kubala, K. Hofbauerova, R. Ettrich, V. Kopecky, R. Krumscheid, J. Plasek, J. Teisinger, W. Shoner, E. Amler, Phe(475) and Glu(446) but not Ser(445) participate in ATP-binding to the alphasubunit of Na⁺/K⁺-ATPase, *Biochem. Biophys. Res. Commun.* 297 (2002) 154–159.
- [25] M. Kubala, J. Plasek, E. Amler, Limitations in linearized analyses of binding equilibria: binding of TNP-ATP to the H-4–H-5 loop of Na/K-ATPase, *Eur. Biophys. J. Biophys.* 32 (2003) 363–369.
- [26] M. Kubala, J. Teisinger, R. Ettrich, K. Hofbauerova, V. Kopecky, V. Baumruk, R. Krumscheid, J. Plasek, W. Schoner, E. Amler, Eight amino acids form the ATP recognition site of Na⁺/K⁺-ATPase, *Biochemistry* 42 (2003) 6446–6452.
- [27] J. Shi, X.-P. Zhang, D.C. Neckers, Xanthenes: fluorone derivatives II, *Tetrahedron Lett.* 34 (1993) 6013–6016.
- [28] J.R. Lakowicz, Principles of Fluorescence Spectroscopy, third ed. Springer, New York, 2006.
- [29] C. Giulivi, M. Sarcansky, E. Rosenfeld, A. Boveris, The photodynamic effect of rose bengal on proteins of the mitochondrial inner membrane, *Photochem. Photobiol.* 52 (4) (1990) 745–751.
- [30] L. Ganesan, E. Margolles-Clark, Y. Song, P. Buchwald, The food colorant erythrosine is a promiscuous protein-protein interaction inhibitor, *Biochem. Pharmacol.* 81 (2011) 810–818.
- [31] J.C. Skou, M. Esmann, Effect of magnesium ions on the high-affinity binding of eosin to the Na⁺ + K⁺-ATPase, *Biochim. Biophys. Acta (BBA)-Biomembr.* 1 (1981) 101–107.
- [32] J.C. Skou, M. Esmann, Effect of magnesium-ions on the high-affinity binding of eosin to the (Na⁺ + K⁺)-ATPase, *Biochim. Biophys. Acta (BBA)* 727 (1) (1983) 101–107.
- [33] M.M. Bradford, Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding, *Anal. Biochem.* 72 (1976) 248–254.
- [34] M.R. Eftink, Fluorescence methods for studying equilibrium macromolecule-ligand interactions, *Meth. Enzymol.* 278 (1997) 221–257.
- [35] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234 (1993) 779–815.
- [36] M.F. Sanner, Python: a programming language for software integration and development, *J. Mol. Graph. Model.* 17 (1999) 57–61.
- [37] O. Trott, A.J. Olson, AutoDock Vina: improving to the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading, *J. Comput. Chem.* 31 (2010) 455–461.
- [38] J.T. Ogan, M.S. Reifenger, M.A. Milanick, C. Gatto, Kinetic characterization of Na, K-ATPase inhibition by Eosin, *Blood Cells Mol. Dis.* 38 (2007) 229–237.
- [39] M.M. Martin, Hydrogen bond effects on radiationless electronic transitions in xanthene dyes, *Chem. Phys. Lett.* 35 (1) (1975) 105–111.
- [40] S.J.D. Karlish, Characterization of conformational-changes in (Na, K)ATPase labeled with fluorescein at the active-site, *J. Bioenerg. Biomembr.* 12 (3–4) (1980) 111–136.
- [41] T. Tsuda, S. Kaya, H. Funatsu, Y. Hayashi, K. Taniguchi, Fluorescein 5'-isothiocyanate-modified Na⁺, K⁺-ATPase, at Lys-501 of α-chain, accepts ATP independent of pyridoxal 5'-diphospho-5'-adenosine modification at Lys-480, *J. Biochem.* 123 (1998) 169–174.
- [42] C. Gatto, M.A. Milanick, Inhibition of the red blood cell calcium pump by eosin and other fluorescein analogues, *Am. J. Physiol.* 264 (1993) 1577–1586.
- [43] N.N. Slinchenko, N.F. Bratkova, S.A. Kosterin, V.P. Zimina, I.G. Chernysh, Effects of eosin Y on the catalytic and functional activities of Mg²⁺, ATP-dependent calcium pump of smooth muscle cell plasma membrane, *Biochemistry* 63 (6) (1998) 685–690.
- [44] M.I. De Michelis, A. Carnelli, F. Rasi-Caldogno, The Ca²⁺ pump of the plasma membrane of *Arabidopsis thaliana*: characteristics and sensitivity to fluorescein derivatives, *Bot. Acta* 106 (1993) 20–25.
- [45] M. Esmann, N.U. Fedosova, Anion interactions with Na, K-ATPase: simultaneous binding of nitrate and eosin, *Eur. Biophys. J.* 33 (2004) 683–690.
- [46] C.H. Pedemonte, T.L. Kirley, M.J. Treuheit, J.H. Kaplan, Inactivation of th Na, K-ATPase by modification of Lys-501 with 4-acetamido-4'-isothiocyanatostilbene-2,2'-disulfonic acid (SITS), *FEBS Lett.* 314 (1) (1992) 97–100.
- [47] C. Gatto, S. Lutsenko, J.H. Kaplan, Chemical modification with dihydro-4,4'-diisothiocyanatobenzen-2,2'-disulfonate reveals the distance between K480 and K501 in the ATP-binding domain of the Na, K-ATPase, *Arch. Biochem. Biophys.* 340 (1) (1996) 90–100.
- [48] M. Esmann, Conformational transitions of detergent-solubilized Na, K-ATPase are conveniently monitored by the fluorescent probe 6-carboxy-eosin, *Biochem. Biophys. Res. Commun.* 174 (1991) 63–69.

Appendix L

Bazgier V, Berka K, Otyepka M, Banáš P:
Exponential repulsion improves structural predictability of molecular docking.
J. Comput. Chem., 37(28), 2485–2494, 2016.
DOI: 10.1002/jcc.24473

IF=3.229

Exponential Repulsion Improves Structural Predictability of Molecular Docking

Václav Bazgier,^[a,b] Karel Berka,^[b] Michal Otyepka,^[b] and Pavel Banáš^{✉*}^[b]

Molecular docking is a powerful tool for theoretical prediction of the preferred conformation and orientation of small molecules within protein active sites. The obtained poses can be used for estimation of binding energies, which indicate the inhibition effect of designed inhibitors, and therefore might be used for *in silico* drug design. However, the evaluation of ligand binding affinity critically depends on successful prediction of the native binding mode. Contemporary docking methods are often based on scoring functions derived from molecular mechanical potentials. In such potentials, nonbonded interactions are typically represented by electrostatic

interactions between atom-centered partial charges and standard 6–12 Lennard–Jones potential. Here, we present implementation and testing of a scoring function based on more physically justified exponential repulsion instead of the standard Lennard–Jones potential. We found that this scoring function significantly improved prediction of the native binding modes in proteins bearing narrow active sites such as serine proteases and kinases. © 2016 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.24473

Introduction

Molecular docking is a well-established and widely used method for *in silico* drug design.^[1,2] The predictive power of molecular docking studies heavily depends on the exhaustiveness of the searching algorithm and quality of the underlying scoring function.^[3] Molecular docking experiments are often used for high-throughput *in silico* screening,^[1] which imposes severe limitations on the duration of a single docking experiment. Preferably, a docking experiment should be finished in seconds or minutes. This dictates the use of empirical scoring functions based on molecular mechanics, which can be quickly enumerated by current computers.

Molecular docking experiments generally consist of two crucial mutually interrelated steps: (i) the search algorithm, which is aimed at identifying the energetically preferred conformation and orientation of a ligand (ligand pose) within the active site, and (ii) the scoring function, which is used to evaluate the energy of the ligand pose in the active site. The estimated binding energy may be subsequently used for qualitative prediction of the ligand affinity to the active site of a given target (usually a protein). However, the reliability of this prediction critically depends on the first step, that is, the ability of the docking algorithm to identify (and subsequently score) the native ligand pose correctly. If the docking method is not able to predict the correct binding pose of the ligand within the active site, any correlation between the ligand score and the experimental value such as K_i is meaningless. In such a case, the correlation relates scores calculated from the theoretically predicted but artificial ligand binding modes with experimentally measured quantities corresponding to the native binding poses of the ligands. In other words, a docking experiment is required to determine correct native ligand poses inside the active site.

A typical scoring function^[4] consists of intramolecular and intermolecular terms, where the latter ones represent the most computationally demanding part in the scoring function enumeration. The intermolecular terms account for electrostatic interactions, which are typically described in terms of interacting atomic centered partial charges via the Coulomb law, and van der Waals interactions covering London dispersion and Pauli repulsion forces, which are usually computed simultaneously by the Lennard–Jones potential^[5] [eq. (1)]:

$$E_{\text{vdw}} = \sum_i^{\text{ligand}} \sum_j^{\text{protein}} \varepsilon_{ij} \left[\left(\frac{R_{ij}}{r_{ij}} \right)^{12} - 2 * \left(\frac{R_{ij}}{r_{ij}} \right)^6 \right] \quad (1)$$

Where ε_{ij} and R_{ij} stand for the potential well-depth and radius of van der Waals pair parameters, respectively, r_{ij} is the distance between atoms of a given pair and the sum is over all unique protein–ligand atom pairs. The empirical potentials typically derives atom-pair parameters ε_{ij} and R_{ij} from the atomic ones using so called mixing rules. Two different mixing rules are commonly used in different force fields: arithmetic (Lorentz–Berthelot) mixing rules [eq. (2)]

[a] Václav Bazgier
Laboratory of Growth Regulators, Centre of the Region Haná for Biotechnological and Agricultural Research, Institute of Experimental Botany ASCR & Palacký University, Šlechtitelů 11, Olomouc 783 71, Czech Republic

[b] Václav Bazgier, K. Berka, M. Otyepka, P. Banáš
Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacký University Olomouc, 17. listopadu 12, Olomouc 77146, Czech Republic
E-mail: pavel.banas@upol.cz

Contract grant sponsor: Ministry of Education, Youth and Sports (Project LO1305); Contract grant number: IGA_PrF_2016_028 (V.B.) of Palacký University

© 2016 Wiley Periodicals, Inc.

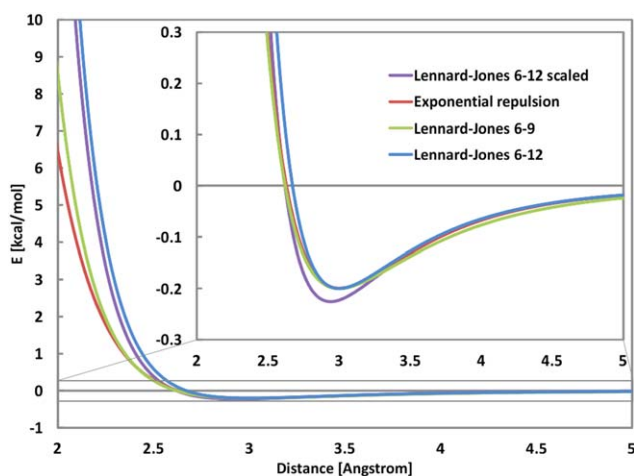


Figure 1. Different representations of van der Waals potentials with pair parameters ε_{ij} and R_{ij} equal to 0.2 kcal/mol and 3 Å, respectively: standard 6–12 Lennard–Jones potential following eq. (1) (blue), soften 6–9 Lennard–Jones potential (green), 6–12 Lennard–Jones potential with R_{ij} of repulsion term scaled by 0.99 (violet), and exponential repulsion with damped dispersion according eqs. (2) and (5) (red).

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \quad R_{ij} = R_i + R_j, \quad (2)$$

and geometric mixing rules [eq. (3)]

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j} \quad R_{ij} = 2\sqrt{R_i R_j} \quad (3)$$

where ε_i and R_i in eqs. (2) and (3) denote the atomic van der Waals parameters of the i th atom, namely the well-depth and atomic radius, and ε_{ij} and R_{ij} stand for the corresponding pair parameters used in the Lennard–Jones potential [see eq. (1)].

The Lennard–Jones potential describes London dispersion through a $k \cdot r^{-6}$ term, which is physically correct as it represents an approximation for the interaction between instantaneous dipoles.^[6] Conversely, Pauli repulsion is computed by a $k \cdot r^{-12}$ function, which although unphysical is still used because it can readily be enumerated by computer as the square of the dispersion term. The exponential function $k \cdot e^{-\alpha r}$ represents a physically more justified mathematical representation of the Pauli repulsion^[6] [note that k and α parameters used in the functional forms reported in this paragraph are typically derived from van der Waals atomic parameters ε_{ij} and R_{ij} , cf. eqs. (1) and (5)]. It should be noted that usage of the $k \cdot r^{-12}$ term in the Lennard–Jones potential causes overestimation of the intermolecular repulsion at van der Waals contact and shorter distances (Fig. 1).^[4,7] As the repulsion counterbalances the electrostatic interaction, which is generally overestimated by the molecular mechanics at the contact and shorter distances,^[7] problems may be encountered in scoring ligands docked into narrow and polar active site clefts. In such cases, the imbalance in intermolecular terms is significantly pronounced, such as in kinases and serine proteases.

For sake of clarity, henceforth the term “narrow” active site stands for an attribute of the protein active site with respect to a given ligand, so that any repulsive clashes at close distance contacts caused, for example, by inaccurate parameterization of

nonbonded van der Waals interaction, cannot be relaxed by small replacement of the ligand as other clashes immediately arise elsewhere in the active site (see Fig. 2). In addition, the polar nature of protein–ligand interaction would inevitably lead into even stronger effect of van der Waals repulsion parameterization as the typical interaction distances in polar contacts are shifted to the repulsive region of the van der Waals potential due to attractive electrostatics. Therefore, the expected effect of clashes caused by inaccurate parameterization of nonbonded van der Waals interactions in the narrow and polar active site pockets should be even more apparent. It should be noted that the “narrowness” of the active site with respect to the ligand might be easily monitored by portion of ligand surface area buried in the protein active site pocket, so that the highly buried ligands are likely tightly embedded within the protein active site and therefore such active site should be considered as narrow in terms of our definition.

A representative example of the proteins with narrow and polar active site cleft is cyclin-dependent kinase 2 (CDK2), which is a member of the protein kinase family involved in cell cycle regulation.^[8] As cell cycle deregulation has been shown to be connected with many serious human diseases, including cancer, diabetes II type, etc., CDK2 (and other cyclin-dependent kinases) has been subject of extensive research aimed at discovering potent inhibitors.^[9] CDK2 has a narrow cleft-like active site to which the native ligand—ATP binds.^[10] Thus, the active site might be the target for any of the numerous CDK2 inhibitors discovered so far.^[11–14] Until now, numerous docking experiments were performed to predict native poses and strength of the inhibitor binding to the CDK2 active site pocket.^[15–25] These studies used different scoring functions and algorithms and reported various level of correlation between estimated interaction energy of obtained pose and experimental inhibition constant with r^2 ranging from 0.15 to 0.87. The most successful agreement was achieved with scoring functions based on multiscale QM/MM^[26] or semi-empirical^[25,27] methods that benefit from fundamentally better description of the intramolecular attraction and repulsion. However, these advanced methods are computationally demanding, often requiring several hours for evaluation of one ligand pose, making their usage in docking experiments virtually impractical. In addition, to achieve high correlation between score values and thermodynamics data, other terms



Figure 2. Scheme showing definition of the “narrow” active site used in this study. The left panel shows an active site with ligand that has enough space for relaxation of any clashes, while the right panel demonstrates narrow active site, in which ligand is fully embedded within the narrow active site and its surface area is highly buried. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

such as entropic contribution, implicit and explicit solvation effects, or flexibility changes should be taken into account, which brings additional computational demands. Hence, the development of a reliable and quickly enumerable scoring function that would be able to find correct ligand binding mode represents an appealing challenge for molecular docking.

The cleft-like active site of many enzymes such as CDK2, poor performance of classical docking schemes based on empirical scoring functions and better performance of QM/MM and semi-empirical calculations motivated us to modify the empirical scoring function, which seems to be too repulsive. We replaced the $k \cdot r^{-12}$ term in the Lennard–Jones potential by the exponential function $k \cdot e^{-\alpha r}$ and implemented this function in the popular DOCK 6.6 docking software.^[28] In contrast to other ways of the softening of the repulsion term, for example, usage of modified Lennard–Jones 6–9 potential or scaled atomic radius in the repulsion term, the exponential repulsion represents physically justified description of the van der Waals repulsion.

We carried out test docking experiments with fifteen CDK2 inhibitors for which crystal structures and inhibition constants (K_i) are known.^[25] In addition, we tested the performance of the present modified scoring function using the SB2012 docking validation set of 1043 protein–ligand complexes introduced by Rizzo et al.,^[29] using docking dataset introduced by Wang et al.,^[4] and using the directory of decoys (DUD) dataset comprising 40 target proteins, 2950 active ligands, and 124,413 decoys.^[30] We mainly focused on the ability of this method to correctly reproduce the crystal-like position of the ligand as the pose with the best score and ability to correctly distinguish between ligands and decoys. In case of CDK2, we also examined the correlation between the scoring and experimental data.

Methods

We used the DOCK 6.6 software package^[28] to dock ligands into the active sites of proteins. We tested the performance of the present scoring function with exponential repulsion on three case studies. First, we carried out redocking experiments of 15 CDK2-inhibitors with the experimentally determined inhibition activities into the CDK2 structures, that is, we took crystal structures of the given CDK2-inhibitor complexes from the PDB database (PDB IDs: 1AQ1,^[31] 1E1X,^[32] 1PKD,^[33] 1PXJ,^[34] 1PXL,^[34] 1PXM,^[34] 1PXN,^[34] 1PXP,^[34] 2FVD,^[35] 1H1P,^[36] 1H1S,^[36] 1OGU,^[37] 2A4L,^[11] 2EXM,^[38] and 2X1N^[39] with resolution range from 1.80 Å to 2.53 Å), removed a ligand from the CDK2 active site and redocked ligand to CDK2 structures. Additionally, we also performed redocking experiments for the statistically robust docking test set SB2012 (<http://ringo.ams.sunysb.edu/index.php/SB2012>) containing 1043 protein–ligand complexes,^[29] and for docking dataset introduced by Wang et al.^[4] Finally, we tested the performance with enrichment docking experiment with DUD dataset of 40 protein targets with annotated ligands and decoy molecules (<http://dud.docking.org>).

All ligands and proteins were prepared with following protocol. Hydrogens were added to all ligands and protein structures using the Chimera program 1.10.1.^[40] Gasteiger partial charges were assigned to the ligand atoms (for comparison, we also tested docking to CDK2 with RESP charges of the ligand, results are shown in Supporting Information, Table S5). The partial charges of protein atoms, as well as van der Waals parameters of both ligand and protein, were taken from the AMBER ff99/ff99SB^[41,42] force field. We thus used a scoring function based on the AMBER ff99 force field (note that this covered all more recent variants of the AMBER protein force field, such as ff99SB or ff14SB, because they differ from generic ff99 only in dihedral terms, which do not contribute to the docking due to the fixed protein structure).

DOCK uses a two-step procedure to speed up docking experiments. In the first step, interaction grids containing the interaction potential derived from the given protein structure are precomputed, then in the second step, the molecule is docked and the scoring function evaluated using the precomputed data on grids. For docking using the exponential repulsion term, we used a modification of the DOCK 6.6 software package (for details see the implementation details paragraph of the Results and Discussion). The intramolecular deformation energy of the ligand was not included in the scoring function.

The success of a docking experiment (i.e., ability to identify the correct native ligand pose as the best score pose) relies on both the quality of the scoring function and effectiveness (or exhaustiveness) of the docking searching algorithm. In our study, we focused on testing the quality of the scoring function, and therefore set the parameters of the searching algorithm to be as exhaustive as possible (even at the expense of using a computational time that would be rather impractical for a docking experiment). Namely, we set the maximum orientation parameter, the maximum number of anchor orientations and conformers, and the number of conformations scored and ranked using the primary score to 100,000. The grid computation is prepared using sphgen program^[28] and used a 0.3 Å grid spacing. The distance cutoff was set to a sufficiently large value of 9,999 Å to account for all pair-wise interactions. A 4r-distance dependent dielectric constant and atom model was applied to all atoms.

During redocking experiments for either CDK2, or SB2012 datasets, docked ligand with root-mean-square deviation (RMSD) to the crystal native ligand pose below 2.0 Å were assigned as the native ones. All ligand poses above this value were considered as alternative poses. We have analyzed only the poses with the lowest score (aka best poses) as those would be the ones selected by the blind docking. We also analyzed the energy differences between the scores of the native pose and the non-native pose with the lowest energy, which measures the fidelity of the docking to distinguish the native pose over all other non-native ones. In addition, for CDK2 dataset we calculated the coefficient of determination (square of the correlation coefficient r^2) relating the experimentally measured inhibition constants K_i with the score of the best pose.

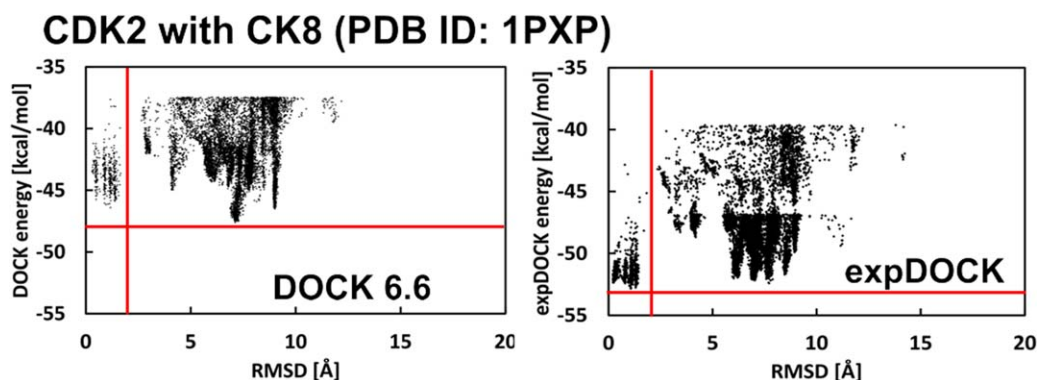


Figure 3. DOCK binding energy versus RMSD (from the crystal-structure pose) of all docked poses of the staurosporine to CDK2 (PDB ID: 1PXP) evaluated by original DOCK (left) and modified DOCK with exponential repulsion (see Fig. S2 in the Supporting Information for equivalent plots of other CDK2 containing systems). Vertical red line highlights 2.0 Å cutoff used for a definition of native and non-native poses, while horizontal red line denotes the most favorable binding energy obtained in docking. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

In the case of DUD dataset, we have docked for each protein target a set of annotated ligands together with larger number of decoy molecules. Subsequently, we have evaluated enrichment factor (EF) of annotated ligands defined as follows:

$$EF_{x\%} = \frac{p_{\text{lig},x\%}}{p_{\text{lig,tot}}} \quad (4)$$

where $p_{\text{lig},n\%}$ and $p_{\text{lig,tot}}$ represent relative occurrences of the active ligands in the subset of the $n\%$ best scoring molecules and in complete set of ligands and decoys, respectively. The quality docking, that is, better scoring of active ligands than the decoys, was evaluated by $EF_{20\%}$ and EF_{max} , which corresponds to maximal value of $EF_{x\%}$ over the percentils $x\%$.

Results and Discussions

In this study, we focused on the implementation and testing of molecular docking using scoring function with exponential repulsion (further as expDOCK) instead of classical r^{-12} function for repulsion used in the standard 12–6 Lennard–Jones term in Amber scoring function in DOCK 6.6 program (further as DOCK 6.6). As a hallmark of successful docking, we evaluated the ability to find the correct binding mode of a ligand (i.e., according to the X-ray structure) as the best score pose and/or distinguishing the active ligands from the decoys. Although this might also have been limited by the exhaustiveness and efficiency of the searching algorithm, we set the searching algorithm to be as exhaustive as possible so that we could directly monitor the predictability and quality only of the introduced scoring function. When possible, we checked that the search algorithm explored poses with (RMSDs) ranging from values below 1 Å from the known native pose (with minimal RMSD typically ~ 0.3 Å) up to values above 15 Å in all systems. Thus, the search algorithm thoroughly explored the conformational space, even though the global minima did not in all cases correspond to the crystal-like poses with small RMSD values (see Fig. 3 and Fig. S2 in Supporting Information).

Scoring function based on AMBER force field with exponential repulsion

The accuracy of the correct binding mode prediction could potentially be systematically improved by employing scoring functions based on empirical force fields that are carefully parameterized to capture nonbonded interactions, which thus represent the most promising class of scoring functions. Among these scoring functions, we decided to focus on a function based on the AMBER force field because of the broad success of the AMBER family force fields.^[42]

However, the AMBER-based scoring function implemented in DOCK does not exactly match the interaction energy calculated by AMBER owing to different mixing rules. The AMBER force field was designed to use arithmetic (Lorentz–Berthelot) mixing rules^[41] [eq. (2)], while to simplify the docking calculation and allow the use of a precomputed grid potential, DOCK uses geometric mixing rules [eq. (3)].

Hence, we decided to implement arithmetic mixing rules and test the effect of using more physically justified exponential repulsion instead of the Lennard–Jones r^{-12} term. Switching to the geometric mixing rules makes van der Waals interactions less repulsive as the final van der Waals (atom-pair) radius derived by eq. (3) is always lower or equal to the radius calculated by eq. (2), and thus the repulsive part of the interaction is apparent mainly at smaller interatomic distances. This applies particularly for atoms with significantly different van der Waals (atomic) radii, for example, hydrogen and heavy atoms in hydrogen bond interactions [Fig. 4, see also eqs. (2) and (3)].

There are several ways to derive parameters for the exponential repulsion term from classical Lennard–Jones atomic parameters. Here, we used the exponential form from the DREIDING force field^[43] with parameter $\xi = 12$ to retain AMBER-like long range dispersion interactions. In addition, we used a damped dispersion term to prevent divergence of the energy at very short interatomic distances so that the dispersion term was damped by only $\sim 0.1\%$ in the van der Waals

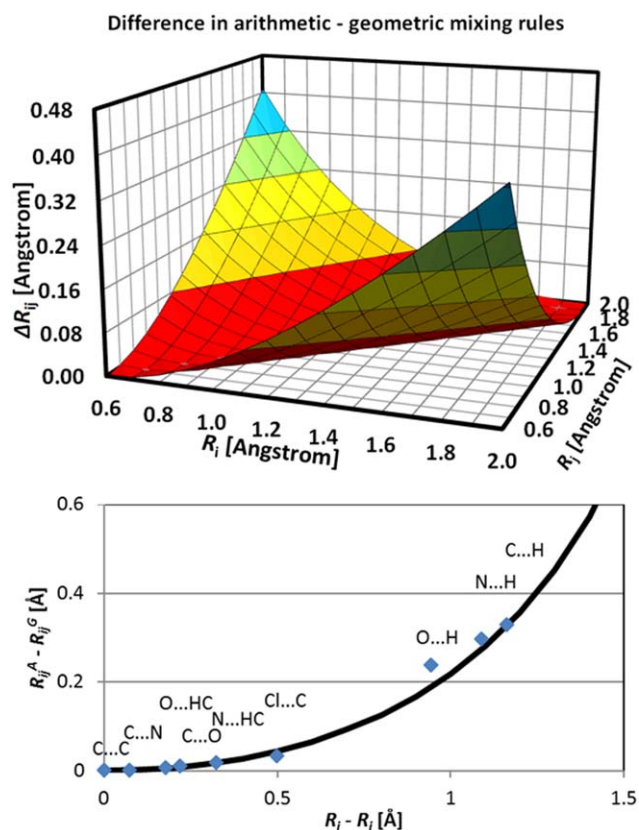


Figure 4. Difference $\Delta R_{ij} = R_{ij}^A - R_{ij}^G$ between van der Waals atom-pair radii calculated with arithmetic (Lorentz–Berthelot) mixing rules [R_{ij}^A , eq. (2)] and geometric mixing rules [R_{ij}^G , eq. (3)]. (Upper) 3D plot of this difference as a function of the corresponding atomic parameters R_i and R_j . (Lower) 2D representation of the same difference as a function of the absolute value of the difference between corresponding atomic parameters R_i and R_j . Blue diamonds show typical examples of AMBER atom pairs parameters.

minimum, which had negligible effect on the interaction energies. Such a term is commonly used in soft-core potentials and has been shown to work well in combination with exponential repulsion and AMBER parameters.^[44] Thus, the van der Waals potential with arithmetic Lorentz–Berthelot mixing rules, exponential term and damped dispersion was defined as follows:

$$E_{\text{vdw}} = \sqrt{\varepsilon_i \varepsilon_j} \left[e^{12 \left(1 - \frac{r_{ij}}{R_i + R_j} \right)} - 2 \left(\frac{(R_i + R_j)^6}{r_{ij}^6 + ((R_i + R_j)/3)^6} \right) \right] \quad (5)$$

where ε_i , ε_j , R_i , and R_j stand for the AMBER atomic van der Waals parameters, that is, the well-depth and atomic radii for the i th and j th atom. The scaling parameter of exponential repulsion and damping parameter of damped dispersion were set to 12 and 3, respectively (see eq. (1) in Ref. [44]).

The exponential repulsion is not the only way to correct the artificially overestimated repulsion of the r^{-12} Lennard–Jones term. Two other options are available in the DOCK 6.6 software for this purpose, that is, scaling of the atomic radii for the repulsive part of the van der Waals potential or use of a

modified Lennard–Jones potential, typically the softened 6–9 Lennard–Jones potential, whereby the repulsion exponent of 12 is replaced by 9. However, neither of these options is physically justified nor fully corrects the artificially strong repulsion at close contact distances (even the relatively promising 6–9 Lennard–Jones potential was found to be significantly more repulsive than the exponential repulsion potential at close contact distances; see Fig. 1). In addition, both these potentials failed to fully retain the shape of the van der Waals potential around its minimum and/or in the long-range region compared to the original 6–12 Lennard–Jones potential for which the parameters were designed. In particular, scaling of the van der Waals distance in the repulsion term shifted both the position and depth of the van der Waals minimum, whereas usage of the modified 6–9 Lennard–Jones potential slightly increased dispersion in the long-range region (Fig. 1).

Exponential repulsion and arithmetic mixing rules: Implementation details

We modified the source code of the DOCK 6.6 program to implement the exponential repulsion term and arithmetic Lorentz–Berthelot mixing rules in the docking algorithm. The main obstacle to such an implementation was both the exponential repulsion term and arithmetic mixing rules precluded the multiplicative formation of the pair van der Waals term from the corresponding atomic contributions. Consequently, it was no longer possible to express the van der Waals interaction of a ligand atom with the entire protein as a product of the protein van der Waals potential and contribution of the ligand atoms. Thus, the precomputed grid used to enhance the speed of the docking calculation could not be used directly (in its original implementation) together with these modifications.

To solve this problem, that is, implement exponential repulsion and arithmetic mixing rules but keep the benefits of grid calculation; we modified how the precomputed grid was calculated. Instead of calculating a single protein's van der Waals potential on a grid, we calculated van der Waals interaction energies for all (ligand) van der Waals atom types in all grid points. Thus, we obtained a van der Waals grid for each atom type. For the AMBER ff99 force field, this generated 38 van der Waals grids in total. Note that the electrostatic interaction could still be decomposed in terms of the electrostatic potential and charge of the ligand atom; thus, the calculation of the electrostatic grid remained unchanged.

Importantly, these modifications in expDOCK decreased the performance of the precomputed grid calculation but not the docking itself, which remained as effective as in the original implementation. A docking calculation using a precomputed grid typically consists of two steps: grid calculation on a given protein, followed by searching and scoring likely ligand poses. If other ligands need to be docked to the same protein, the precomputed grid can be reused for all ligands, that is, it is calculated only once for each protein. Therefore, the decline in performance of the grid calculation was a relatively minor issue.

A heuristic estimate of the performance reduction of the grid calculation suggested that it was ~ 38 times slower in comparison with the original docking implementation. However, we identified several ways to rationalize the algorithm for calculation of such a grid. Namely, we redesigned and sorted the nested-loops looping over protein atoms and grid points to maximize the benefits of the loop vectorizations. Thus, the final performance of the grid calculation was only ~ 4 times slower (in the case of the CDK2 protein, this corresponded to an increase in computational time from ~ 2 to ~ 8 h). It should be noted that the docking calculation with expDOCK scoring function was not affected by these modifications in comparison with original DOCK 6.6 implementation.

Case study I: Pilot docking to CDK2

The improved performance expDOCK scoring function compared to the standard DOCK 6.6 function was first tested on fifteen inhibitors of CDK2 with known inhibition constants K_i . We selected this first pilot testing system from three reasons: (i) as the exponential repulsion is expected to have the most significant effect for narrow and polar active site clefts, which is the case in CDK2, (ii) due to availability of the experimental inhibition constants, and (iii) as the docking to CDK2 still remains a challenge. We mainly focused on the ability of docking to correctly identify the native position and conformation of the ligand within the active site of CDK2 as the best pose.

We found that in the case of the scoring function using DOCK 6.6 function, the mean RMSD of the best poses to the corresponding crystal positions averaged over all 15 inhibitors was 2.8 Å, whereas usage of expDOCK function decreased mean RMSD value to 2.1 Å. However, we would like to stress that the averaged value of RMSD should be interpreted with care. While small values of RMSD (below 2 Å) represent good metric for evaluation of the accuracy of the native pose prediction, the same is not valid for high RMSD values (above 4 Å). For instance, the ligand pose being 5 Å in RMSD from the native conformation should not be considered as better docked than the pose having RMSD of 10 Å as both of them are simply incorrect. However, the high values of RMSD might typically have significant influence on the averaged value, so the corresponding difference in averaged RMSD would be misleading in this particular case. Thus we focused also on the number of correctly docked ligands. In five of the fifteen systems, DOCK 6.6 function identified as the best pose a non-native ones, with RMSD of the best pose to the corresponding crystalline position mostly above 6 Å (note that all other poses had a RMSD below cutoff value 2 Å used to define correctly docked pose, but most of them were typically less than 1 Å, see Fig. S1 in the Supporting Information). Conversely, in the case of expDOCK function, incorrect docked poses of the inhibitor were obtained only in four systems (see Fig. S1 in Supporting Information). Notably, expDOCK significantly improved docking in the systems with PDB IDs 1H1S, 1PXM, and 1XPX but worsened it in the case of 1PXN and 2A4L (Table 1, Fig. S1 Supporting Information). We found that expDOCK function favored the fidelity of native poses over

Table 1. Overview of the studied systems of CDK2 together with their inhibitors.

PDBID	$\ln(K_i)$	DOCK 6.6		expDOCK	
		RMSD [Å]	Score [kcal/mol]	RMSD [Å]	Score [kcal/mol]
1AQ1	-19.66	0.66	-53.0	0.66	-61.2
1E1X	-13.55	1.83	-38.8	1.99	-43.0
1H1P	-11.33	2.97	-45.5	2.98	-48.8
1H1S	-18.93	6.62	-50.3	0.74	-54.3
1OGU	-17.55	0.86	-61.8	1.49	-66.5
1PKD	-17.32	0.54	-60.0	0.55	-67.7
1PXJ	-11.94	0.25	-34.0	0.20	-37.8
1PXL	-15.05	8.54	-45.7	8.54	-49.1
1PXM	-16.63	7.04	-46.6	0.33	-51.8
1PXN	-16.47	0.79	-46.6	9.31	-54.7
1XPX	-15.33	7.10	-47.6	1.07	-52.9
2A4L	-13.63	1.89	-54.4	2.23	-61.3
2EXM	-9.46	1.50	-35.2	1.65	-39.8
2FVD	-19.62	0.29	-60.5	0.31	-68.7
2X1N	-17.59	0.88	-48.2	0.80	-52.0

Data are shown for the experimentally observed inhibition constants and root-mean-square deviation (RMSD) of the highest scored pose and its score for scoring functions using either standard 6–12 Lennard–Jones potential (DOCK 6.6) or exponential repulsion with a damped dispersion term (expDOCK). The RMSD values of incorrectly docked poses are shown in red bold.

non-native ones by ~ 0.4 kcal/mol on average (Table S1 in Supporting Information). Taken together, the scoring function with exponential repulsion improved the predictability of the native pose of CDK2 inhibitors, although the criteria of blind docking were still not met.

To check the stability of our docking results, we repeated the aforementioned docking experiments using different random seeds. We did not observe any qualitative effect on our results (see Table S6 in the Supporting Information). In addition, we tested the effect of molecular symmetry on the calculation of RMSD. We reanalyzed the obtained poses using Hungarian symmetry corrected RMSD metric,^[45,46] which revealed that even the standard (symmetry uncorrected) RMSD is sufficient to correctly predict docking success.

Beside standard back-docking calculations, we tested also the performance of both studied scoring functions for the crossdocking calculations. In particular, we tested docking of all ligands from CDK2 dataset to each CDK2 structure used in our dataset. We found that while standard 6–12 repulsion succeeded in 53 out of total 225 protein–ligand combinations (including those corresponding to back-docking), the exponential repulsion was able to find correct native pose 60 combinations. More specifically, the exponential repulsion improved results of crossdocking in 19 cases, while worsen them in another 12 cases. (see Tables S8 and S9 in the Supporting Information)

Finally, we examined also the effect of exponential repulsion on the correlation between experimentally measured inhibition constants, K_i , and the score of the best pose. We found a rather negligible effect of the proposed scoring with expDOCK function. Namely, docking with DOCK 6.6 function yielded a coefficient of determination $r^2 = 0.62$, whereas usage of

expDOCK function increased it to $r^2 = 0.64$ with exclusion of incorrectly docked poses for both functions (see Fig. S3 in the Supporting Information for regression graphs). In other words, there is still room for improvement of the scoring function, particularly in the approximate description of the electrostatics by atom centered (conformational independent) partial charges and consideration of polarization effects. However, as we argued above, the improvement of the predictability of the correct native pose is promising.

Case study II: SB2012 and Wang et al. docking validation datasets

Subsequently, we aimed to test the studied scoring function using a robust validation datasets. For this purpose, we selected SB2012 docking validation dataset introduced by Rizzo et al., which represents an updated version of the older SB2010 dataset containing 1043 protein–ligand complexes in total,^[29] and docking validation dataset introduced by Wang et al.,^[4] which was used for comparison of 11 scoring functions.

Similarly to the CDK2 case study discussed above, we found that the mean RMSD of the best poses from the corresponding crystal positions averaged over entire SB2012 dataset was decreased from 2.6 Å with DOCK 6.6 function to 2.4 Å with expDOCK function (see Table S2 in Supporting Information for complete data). Although the averaged RMSD values likely document the modest improvement obtained by exponential repulsion, we would like to reiterate that the averaged values of RMSD should be interpreted with care (see previous section). More importantly, we found that expDOCK function qualitatively improved docking in nine systems of 1043 in total (i.e., in ~1% of entire dataset, assuming RMSD cutoff for docking success of 2.0 Å). Detailed analysis, however, showed that expDOCK function worsens the docking in 49 while improving it in 58 cases leading to the overall nine improved systems. Therefore, although the usage of expDOCK function slightly improves the docking results, its effect on a particular enzyme–ligand system is ambiguous. Such behavior is however typical for systems that rely on a cancelation of errors and is rather typical for force field development. In other words, the overestimated repulsion in DOCK 6.6 function is most likely not the only problem of the contemporary force fields, so its correction by more physically justified exponential repulsion might be crucial in the systems, where such artifact significantly affect the ligand binding (e.g., in case of the ligand tightly bound within the narrow and polar cleft-like active site), but insignificant in other systems.

Finally, we analyzed the docking success among the SB2012 docking validation dataset as a function of RMSD cutoff defining the successful docking (Fig. 5). We observed that the docking success of expDOCK function was slightly increased compared to data obtained by DOCK 6.6 function for RMSD cutoffs below 1 Å. This corresponds to the fact that expDOCK function increased number of systems docked to the native pose with RMSD below 1 Å by additional 20 systems (i.e., by ~2% of entire dataset) compared to DOCK 6.6. The expDOCK function thus qualitatively improves docking in ~1% of the

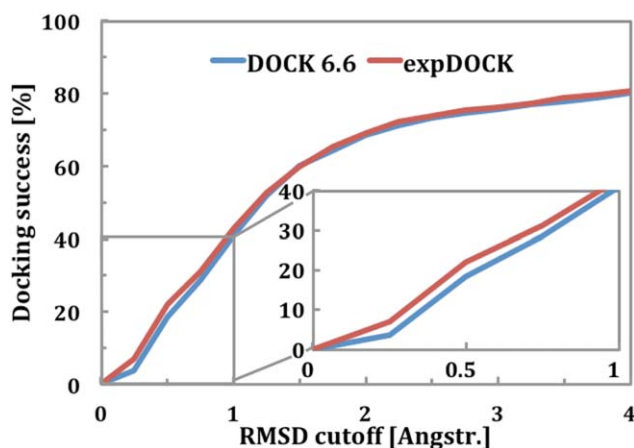


Figure 5. The docking success in SB2012 docking validation dataset as a function of RMSD cutoff defining successful docking. The scoring function with expDOCK function reveals slightly increased docking success compared to DOCK 6.6 function for RMSD cutoff below 1 Å, while its effect at higher RMSD cutoffs is insignificant. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

dataset (with RMSD cutoff of 2.0 Å), but also fine-tunes the ligand positioning in additional ~1% of the dataset, i.e., dock the ligand even below RMSD of 1 Å from crystalline position. Taken together, while exponential repulsion generally improved docking on SB2012 dataset, the overall effect is rather modest.

The Wang et al. docking validation dataset containing 100 protein–ligand complexes was used to compare results obtained by both studied scoring functions with the set of 11 scoring function tested by Wang et al.^[4] We utilized the same docking protocol as described in both aforementioned cases and again we observed that the scoring function with exponential term perform slightly better than scoring function based on standard Lennard–Jones 6–12 potential (see Table 2). When comparing the performance of both studied scoring functions with other scoring functions tested by Wang et al. we found that the exponential repulsion did not alter the ranking of AMBER scoring function, that is, both studied scoring functions performed better than ChemScore^[47] but worse than Autodock.^[48]

Case study III: DUD docking validation dataset

Finally, we tested expDOCK function on large DUD dataset (Directory of useful decoys)^[30] containing 40 protein targets divided to six groups: Nuclear Hormone Receptors (NHR), Kinases (KIN), Serine Proteases (SEP), Metalloenzymes (MET),

Table 2. Number of correctly docked systems out of 100 protein–ligand complexes of Wang et al. docking dataset as a function of RMSD criteria for native pose.

	RMSD ≤1.0 Å	RMSD ≤1.5 Å	RMSD ≤2.0 Å	RMSD ≤2.5 Å	RMSD ≤3.0 Å
DOCK 6.6	23	32	45	50	61
expDOCK	23	39	55	64	67

See Table 2 in Ref. [4] for comparison with other scoring functions.

Table 3. Enrichment factors $EF_{20\%}$ for each target protein with standard DOCK 6.6 and expDOCK scoring function.

Group	Protein family	DOCK 6.6	expDOCK	
		$EF_{20\%}$	$EF_{20\%}$	
Nuclear Hormone Receptors	AR - Androgen receptor	8.6	2.2	
	ER_agonist - Estrogen receptor agonist	3.9	4.7	
	ER_antagonist - Estrogen receptor antagonist	1.0	2.9	
	GR - Glucocorticoid receptor	4.7	2.1	
	MR - Mineralcorticoid receptor	2.5	1.0	
	PPARg - Peroxisome proliferator activated receptor gamma	0.0	4.8	
	PR - Progesterone receptor	2.6	0.5	
	RXRa - Retinoic X receptor alpha	1.6	2.2	
	Kinases	CDK2 - Cyclin dependent kinase 2	2.3	3.6
		EGFr - Epidermal growth factor receptor kinase	4.0	17.8
FGFr1 - Fibroblast growth factor receptor kinase		0.0	3.7	
HSP90 - Human heat shock protein 90 kinase		1.6	0.2	
P38 - P38 mitogen activated protein kinase		0.8	3.2	
PDGFrB - Platelet derived growth factor receptor kinase		0.0	1.9	
SRC - Tyrosine kinase SRC		0.5	7.8	
TK - Thymidine kinase		2.8	2.5	
VEGFr2 - Vascular endothelial growth factor receptor kinase		1.9	3.3	
Serine Proteases		FXa - Factor Xa	3.9	5.6
	Thrombin	3.8	8.1	
	Trypsin	2.1	4.4	
Metalloenzymes	ACE - Angiotensin-converting enzyme	3.4	2.3	
	ADA - Adenosine deaminase	2.9	2.9	
	COMT - Catechol O-methyltransferase	2.0	2.7	
	PDE5 - Phosphodiesterase V	3.9	1.5	
Folate Enzymes	DHFR - Dihydrofolate reductase	14.0	9.7	
	GART - glycinamide ribonucleotide transformylase	3.4	0.5	
other enzymes	ACHE - Acetylcholine esterase	2.2	2.4	
	ALR2 - Aldose reductase	2.6	2.1	
	AmpC - AmpC beta lactamase	0.9	1.5	
	COX-1 - Cyclooxygenase 1	1.9	2.5	
	COX-2 - Cyclooxygenase 2	3.8	2.7	
	GPB - Glycogen phosphorylase beta	5.5	2.2	
	HIVPR - HIV protease	0.9	3.3	
	HIVRT - HIV reverse transcriptase	1.6	0.7	
	HMGR - Hydroxymethylglutaryl-CoA reductase	4.2	3.1	
	InhA - Enoyl ACP reductase	0.0	6.4	
	NA - Neuraminidase	5.8	7.1	
	PARP - Poly(ADP-ribose) polymerase	5.4	3.6	
	PNP - Purine nucleoside phosphorylase	2.2	1.1	
	SAHH - S-adenosyl-homocysteine hydrolase	5.5	4.8	

EF results were classified by values in the range: very good (green; $EF_{\max} \geq 30$ and $EF_{20\%} \geq 3$), good (yellow; $30 > EF_{\max} \geq 20$ and $3 > EF_{20\%} \geq 2.5$), medium (orange; $20 > EF_{\max} \geq 10$ and $2.5 > EF_{20\%} \geq 2$) and poor (red; $EF_{\max} < 10$ and $EF_{20\%} < 2$). In case of disagreement between classification based on EF_{\max} and $EF_{20\%}$ we selected the worse one. EF_{\max} values and the corresponding percentiles of EF maximum are for simplicity documented in Table S4 in Supporting Information.

Folate Enzymes (FOE), and others (OTH). Each target has a list of annotated ligands that are experimentally proven to bind to the protein target, but also larger list of similar molecules not recognized by the target protein—decoys. We aimed to simulate a typical drug discovery scenario for identification of unknown binding ligands within some selection from larger chemical library. For this reason, we analyzed EF (see Methods for its definition) of the annotated ligands within 20%-percentil of ligands sorted by scoring ($EF_{20\%}$) for each protein target (Table 3). In addition, we analyzed the effect of expDOCK function in all protein families separately to reveal influence of different types of active site arrangement.

Although the expDOCK function showed only modest improvement in all previous cases (CDK2, SB2012, and Wang

et al. datasets), the DUD dataset reveals clear and significant improvement of the selectivity for native ligands among the other decoys reflected in high EF values. Particularly, the usage of DOCK 6.6 function lead to very good or good level of EF (i.e., EF_{\max} and $EF_{20\%}$ above 20 and 2.5, respectively, see Table 3) in 11 cases out of 40 protein targets (27.5% of the dataset). The usage of expDOCK represented significant improvement as, at the same level of EF, total of 22 cases were identified (i.e., in 55% of the dataset, Table 4). ExpDOCK also showed improvement of ~ 0.6 in $EF_{20\%}$ and ~ 13.5 in EF_{\max} on average. It should be however noted that the improvement in terms of EF values was not distributed among all protein families equally. Almost all improvements associated with expDOCK function were observed in KIN and SER protein families. Conversely,

Table 4. Overview of the results obtained by docking in DUD dataset.

Protein family	Number of proteins	Proteins with (very) good EF	
		DOCK 6.6	expDOCK
NHR	8	3 (38%)	3 (38%)
KIN	9	0 (0%)	7 (78%)
SEP	3	1 (33%)	3 (100%)
MET	4	2 (50%)	2 (50%)
FOE	2	1 (50%)	1 (50%)
OTH	14	4 (29%)	6 (43%)
Total	40	11 (28%)	22 (55%)

Number of proteins, for which good or very good EF was observed together with portion of the total number of proteins in given family is shown in both DOCK 6.6 and expDOCK. Total numbers of proteins in the families are also shown for the sake of completeness. Data representing significant improvement obtained by usage of expDOCK function are highlighted in bold. NHR – Nuclear Hormone Receptors, KIN – Kinases, SEP – Serine Proteases, MET – Metalloenzymes, FOE – Folate Enzymes, OTH – other enzymes.

NHR, MET, and FOE families were not improved at all. This observation is not surprising as the active sites of proteins in MET and FOE families contain metal ions, which are typically poorly described by empirical force field. In the case of NHR family, the active site pocket is rather large, so inaccuracies in description of repulsion is likely not crucial. Both KIN and SER families revealing most significant effect of expDOCK function typically contain proteins with narrow cleft-like active sites and docking to such sites is expected to be extremely sensitive to accuracy of the scoring function repulsion terms.

Conclusions

Here, we reported the implementation and testing of a new scoring function based on the AMBER ff99 force field employing an exponential repulsion term with damped dispersion instead of the commonly used r^{-12} term of the 6–12 Lennard–Jones potential. We argued that the use of softer exponential repulsion is important particularly in proteins with narrow active sites such as kinases and serine proteases, and leads to a more physically justified description of van der Waals interactions. In addition, we showed that beside exponential repulsion, some attention should be paid to the choice of mixing rules for derivation of pair van der Waals parameters from atomic ones. This was shown to be particularly important for scoring functions based on the AMBER force field that were parameterized for arithmetic Lorentz–Berthelot mixing rules. In contrast, the original implementation of AMBER force field in DOCK 6.6 uses geometric mixing rules to decrease the computational cost. We succeeded in implementing both exponential repulsion and arithmetic mixing rules with relatively low computational demands on the calculation of the precomputed grid and particularly with no computational overhead on docking itself.

We tested the performance of the docking with exponential repulsion and arithmetic mixing rules on three test cases: (i) pilot test case on redocking of fifteen inhibitors into CDK2 and corresponding crossdocking calculations, (ii) docking validation

datasets SB2012 and Wang et al., and (iii) DUD validation dataset. The pilot test case on CDK2 showed that exponential repulsion improved prediction of ligand native binding position, but had no effect on the correlation between inhibition constant and the score of the highest scored pose. However, it should be noted that this pilot test case has rather limited statistical significance. Some improvement of the docking results was also obtained on more robust SB2012 docking validation dataset and also on Wang et al. dataset. We hypothesize that the small improvement observed in SB2012 validation dataset might be explained by the fact that the exponential repulsion is expected to have significant effect only on systems, where inaccurate description of the repulsion is crucial, e.g., proteins with narrow and polar active site clefts. The effect of exponential repulsion on the systems suffering from the other force field inaccuracy might be ambiguous, which limits the overall performance. This was subsequently confirmed by DUD validation set, which clearly revealed significant improvement by 27% in correct recognition of the native active ligands among the set of challenging decoys. In addition, this improvement was almost solely observed for kinases and serine proteases, that is, enzyme families having typically narrow cleft-like actives. The fact that the reparametrization has significant effect only on a specific set of the systems is however typical for force field development. While the improvement might be expected in the systems, where inaccuracy of the repulsion term is the dominant problem, rather insignificant effect is expectable in other systems suffering from inaccuracy of other force field terms. For instance, docking to the metalloproteins using force-field-based scoring function likely suffer from lack of the polarization in the empirical potentials that cannot be compensated by improvement of other force field terms.


We can conclude that usage of the exponential repulsion represents modest improvement of the force field-based scoring functions. Its effect is expected to be significant in proteins with narrow and polar active sites such as kinases and serine proteases, while in other cases it remain hidden behind other force field deficiencies.

Acknowledgment

We would like to thank Professor Robert C. Rizzo for his valuable comments.

Keywords: molecular docking · DOCK 6.6 · drug design · cyclin-dependent kinase 2 · directory of decoys

How to cite this article: V. Bazgier, K. Berka, M. Otyepka, P. Banáš. *J. Comput. Chem.* **2016**, *37*, 2485–2494. DOI: 10.1002/jcc.24473

 Additional Supporting Information may be found in the online version of this article.

- [1] D. B. Kitchen, H. Decornez, J. R. Furr, J. Bajorath, *Nat. Rev. Drug Discov.* **2004**, *3*, 935.
- [2] X. Y. Meng, H. X. Zhang, M. Mezei, M. Cui, *Curr. Comput. Aided Drug Des.* **2011**, *7*, 146.

- [3] T. J. A. Ewing, I. D. Kuntz, *J. Comput. Chem.* **1997**, *18*, 1175.
- [4] R. X. Wang, Y. P. Lu, S. M. Wang, *J. Med. Chem.* **2003**, *46*, 2287.
- [5] J. E. Jones, *R. Soc. Phys. Sci. Res. J.* **1924**, *106*, 14.
- [6] A. J. Stone, *Int. Ser. Monographs Chem.* **1996**, 264, xi.
- [7] M. Zgarbova, M. Otyepka, J. Sponer, P. Hobza, P. Jurecka, *Phys. Chem. Chem. Phys.* **2010**, *12*, 10476.
- [8] D. O. Morgan, *Nature* **1995**, *374*, 131.
- [9] J. Cicenias, K. Kalyan, A. Sorokinas, A. Jatulyte, D. Valiunas, A. Kaupinis, M. Valius, *Cancers (Basel)* **2014**, *6*, 2224.
- [10] K. Bettayeb, N. Oumata, A. Echaliier, Y. Ferandin, J. A. Endicott, H. Galons, L. Meijer, *Oncogene* **2008**, *27*, 5797.
- [11] W. F. DeAzevedo, S. Leclerc, L. Meijer, L. Havlicek, M. Strnad, S. H. Kim, *Eur. J. Biochem.* **1997**, *243*, 518.
- [12] M. Knockaert, P. Greengard, L. Meijer, *Trends Pharmacol. Sci.* **2002**, *23*, 417.
- [13] H. Hirai, N. Kawanishi, Y. Iwasawa, *Curr. Top. Med. Chem.* **2005**, *5*, 167.
- [14] C. McInnes, P. M. Fischer, *Curr. Pharm. Des.* **2005**, *11*, 1845.
- [15] M. Otyepka, V. Krystof, L. Havlicek, V. Siglerova, M. Strnad, J. Koca, *J. Med. Chem.* **2000**, *43*, 2506.
- [16] H. N. Bramson, J. Corona, S. T. Davis, S. H. Dickerson, M. Edelstein, S. V. Frye, R. T. Gampe, P. A. Harris, A. Hassell, W. D. Holmes, R. N. Hunter, K. E. Lackey, B. Lovejoy, M. J. Luzzio, V. Montana, W. J. Rocque, D. Rusnak, L. Shewchuk, J. M. Veal, D. H. Walker, L. F. Kuyper, *J. Med. Chem.* **2001**, *44*, 4339.
- [17] D. A. Ibrahim, A. M. El-Metwally, *Eur. J. Med. Chem.* **2010**, *45*, 1158.
- [18] J. S. Duca, V. S. Madison, J. H. Voigt, *J. Chem. Inf. Model.* **2008**, *48*, 659.
- [19] H. Sato, L. M. Shewchuk, J. Tang, *J. Chem. Inf. Model.* **2006**, *46*, 2552.
- [20] P. Ferrara, A. Curioni, E. Vangrevelinghe, T. Meyer, T. Mordasini, W. Andreoni, P. Acklin, E. Jacoby, *J. Chem. Inf. Model.* **2006**, *46*, 254.
- [21] C. R. W. Guimaraes, M. Cardozo, *J. Chem. Inf. Model.* **2008**, *48*, 958.
- [22] B. C. Pearce, D. R. Langley, J. Kang, H. W. Huang, A. Kulkarni, *J. Chem. Inf. Model.* **2009**, *49*, 1797.
- [23] T. Gucky, R. Jorda, M. Zatloukal, V. Bazgier, K. Berka, E. Reznickova, T. Beres, M. Strnad, V. Krystof, *J. Med. Chem.* **2013**, *56*, 6234.
- [24] M. Mojzycz, V. Subertova, A. Bielawska, K. Bielawski, V. Bazgier, K. Berka, T. Gucky, E. Fornal, V. Krystof, *Eur. J. Med. Chem.* **2014**, *78*, 217.
- [25] P. Dobes, J. Fanfrlik, J. Rezac, M. Otyepka, P. Hobza, *J. Comput. Aided Mol. Des.* **2011**, *25*, 223.
- [26] J. H. Alzate-Morales, R. Contreras, A. Soriano, I. Tunon, E. Silla, *Biophys. J.* **2007**, *92*, 430.
- [27] P. Dobes, J. Rezac, J. Fanfrlik, M. Otyepka, P. Hobza, *J. Phys. Chem. B* **2011**, *115*, 8581.
- [28] I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, T. E. Ferrin, *J. Mol. Biol.* **1982**, *161*, 269.
- [29] S. Mukherjee, T. E. Balius, R. C. Rizzo, *J. Chem. Inf. Model.* **2010**, *50*, 1986.
- [30] N. Huang, B. K. Shoichet, J. J. Irwin, *J. Med. Chem.* **2006**, *49*, 6789.
- [31] A. M. Lawrie, M. E. M. Noble, P. Tunnah, N. R. Brown, L. N. Johnson, J. A. Endicott, *Nat. Struct. Biol.* **1997**, *4*, 796.
- [32] C. E. Arris, F. T. Boyle, A. H. Calvert, N. J. Curtin, J. A. Endicott, E. F. Garman, A. E. Gibson, B. T. Golding, S. Grant, R. J. Griffin, P. Jewsbury, L. N. Johnson, A. M. Lawrie, D. R. Newell, M. E. M. Noble, E. A. Sausville, R. Schultz, W. Yu, *J. Med. Chem.* **2000**, *43*, 2797.
- [33] L. N. Johnson, E. De Moliner, N. R. Brown, H. W. Song, D. Barford, J. A. Endicott, M. E. M. Noble, *Pharmacol. Therapeut.* **2002**, *93*, 113.
- [34] S. Y. Wu, I. McNae, G. Kontopidis, S. J. McClue, C. McInnes, K. J. Stewart, S. D. Wang, D. I. Zheleva, H. Marriage, D. P. Lane, P. Taylor, P. M. Fischer, M. D. Walkinshaw, *Structure* **2003**, *11*, 399.
- [35] X. J. Chu, W. DePinto, D. Bartkovitz, S. S. So, B. T. Vu, K. Packman, C. Lukacs, Q. J. Ding, N. Jiang, K. Wang, P. Goelzer, X. F. Yin, M. A. Smith, B. X. Higgins, Y. S. Chen, Q. Xiang, J. Moliterni, G. Kaplan, B. Graves, A. Lovey, N. Fotouhi, *J. Med. Chem.* **2006**, *49*, 6549.
- [36] T. G. Davies, J. Bentley, C. E. Arris, F. T. Boyle, N. J. Curtin, J. A. Endicott, A. E. Gibson, B. T. Golding, R. J. Griffin, I. R. Hardcastle, P. Jewsbury, L. N. Johnson, V. Mesguiche, D. R. Newell, M. E. M. Noble, J. A. Tucker, L. Wang, H. J. Whitfield, *Nat. Struct. Biol.* **2002**, *9*, 745.
- [37] K. L. Sayle, J. Bentley, F. T. Boyle, A. H. Calvert, Y. Z. Cheng, N. J. Curtin, J. A. Endicott, B. T. Golding, I. R. Hardcastle, P. Jewsbury, V. Mesguiche, D. R. Newell, M. E. M. Noble, R. J. Parsons, D. J. Pratt, L. Z. Wang, R. J. Griffin, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3079.
- [38] U. Schulzegahmen, J. Brandsen, H. D. Jones, D. O. Morgan, L. Meijer, J. Vesely, S. H. Kim, *Proteins* **1995**, *22*, 378.
- [39] N. A. McIntyre, C. McInnes, G. Griffiths, A. L. Barnett, G. Kontopidis, A. M. Z. Slawin, W. Jackson, M. Thomas, D. I. Zheleva, S. D. Wang, D. G. Blake, N. J. Westwood, P. M. Fischer, *J. Med. Chem.* **2010**, *53*, 2136.
- [40] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605.
- [41] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- [42] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins* **2006**, *65*, 712.
- [43] S. L. Mayo, B. D. Olafson, W. A. Goddard, *J. Phys. Chem.* **1990**, *94*, 8897.
- [44] P. Banas, A. Mladek, M. Otyepka, M. Zgarbova, P. Jurecka, D. Svozil, F. Lankas, J. Sponer, *J. Chem. Theory Comput.* **2012**, *8*, 2448.
- [45] H. W. Kuhn, *Nav. Res. Logist. Q.* **1955**, *2*, 83.
- [46] J. Munkres, *J. Soc. Ind. Appl. Math.* **1957**, *5*, 32.
- [47] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, R. P. Mee, *J. Comput. Aided Mol. Des.* **1997**, *11*, 425.
- [48] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639.

Received: 7 March 2016

Revised: 21 June 2016

Accepted: 16 July 2016

Published online on 13 September 2016