UNIVERZITA PALACKÉHO V OLOMOUCI

PŘÍRODOVĚDECKÁ FAKULTA

KATEDRA BIOFYZIKY

DIPLOMOVÁ PRÁCA

Structure analysis and characterization of Organic cation transporters (OCTs) in the Solute carrier family 22 (SLC22)

Vypracovala: Bc. Nina Kadášová

Študijný obor: Molekulární biofyzika

Vedúci práce: doc. RNDr. Berka Karel, Ph.D.

Olomouc 2023

Bibliografická identifikácia:

Meno a priezvisko autora: Bc. Nina Kadášová

Názov práce: Štrukturálna analýza a charakteristika transportérov organických katiónov (OCTs) zo Solute carrier family 22 (SLC22)

Typ práce: Teoretická

Pracovisko: Katedra Fyzikální chemie Přírodovědecké fakulty Univerzity Palackého v Olomouci

Vedúci práce: doc. RNDr. Berka Karel, Ph.D.

Rok obhajoby práce: 2023

Abstrakt:

Transportéry prenášajúce rozpustené látky (SLC) tvoria rôznorodú skupinu membránových proteínov s viac ako 500 zástupcami rozdelenými do 65 rodín na základe sekvenčnej homológie a transportnej funkcie. SLC zohrávajú kľúčovú úlohu pri transporte širokého spektra rozpustných látok cez biologické membrány. Dysfunkcie týchto transportérov prispievajú k rozvoju rôznych ochorení. Preto je skúmanie týchto proteínov dôležité pre pochopenie princípov metabolizmu v ľudskom tele a mechanizmov ochorení. Táto štúdia sa konkrétne zameriava na rodinu SLC22, presnejšie na podskupinu transportérov organických katiónov (OCT). Na skúmanie väzbových vlastností ligandov OCT1 a OCT2 sa použili AI-predikované modely vytvorené pomocou AlphaFold a SWISS-MODEL spolu s dostupnou kryo-EM štruktúrami OCTs. Na identifikáciu vhodných ligandov transportovaných týmito proteínmi sa vykonalo molekulové dokovanie pomocou programov AutoDock Vina. Cieľom výsledkov je určiť účinnosť predikčných metód a preferovaných štrukturálnych modelov na pochopenie väzby ligandov v OCTs.

Kľúčové slová: OCT, SLC, štruktúrna biológia, molekulové dokovanie, membránový transport, ligand, AlphaFold, SWISS-MODEL

Počet strán: 66 strán

Počet príloh: 10 príloh

Jazyk: Angličtina

Bibliographical identification:

Author´s first name and surname: Bc. Nina Kadášová

Title of thesis: Structure analysis and characterization of Organic cation transporters (OCTs) in the Solute carrier family 22 (SLC22)

Department: Department of Physical Chemistry, Faculty of Science, Palacký University in Olomouc

Supervisor: doc. RNDr. Berka Karel, Ph.D.

The year of defence: 2023

Abstract:

Solute carrier transporters (SLC) comprise a diverse group of membrane-bound proteins, with over 500 representatives classified into 65 families based on sequence homology and transport function. SLCs are crucial for transporting a wide range of solutes across biological membranes. Dysfunctions in these transporters contribute to the development of various diseases. Therefore, investigating these proteins is essential for understanding metabolic principles in the human body and disease mechanisms. This study focuses on the SLC22 family, specifically, organic cation transporters (OCTs) subclades. To investigate the ligand binding properties of OCT1 and OCT2, AI-predicted models generated by AlphaFold and SWISS-MODEL were utilized, along with the available cryo-EM structures of OCTs. Molecular docking using AutoDock Vina was performed to identify suitable ligands transported by these proteins. The results aim to determine the efficacy of the prediction methods and the preferred structural models for understanding the ligand binding of OCTs.

Keywords: OCT, SLC, structural biology, molecular docking, membrane transport, ligand, AlphaFold, SWISS-MODEL

Number of pages: 66 pages

Number of attachments: 10 attachments

Language: English language

I declare that I have prepared my diploma thesis independently, under the supervision of doc. RNDr. Karel Berka, Ph.D. and using the literature listed in the conclusion of the diploma thesis.

I want to thank my supervisor, doc. RNDr. Karel Berka, Ph.D., for his guidance and helpfulness in thesis compilation, and to Mgr. Ing. Václav Bazgier, Ph.D., for docking in AutoDock Vina 1.2.0 and his help with datasets. In the end, I want to thank my family for their support.

In Olomouc, day ....

Content

1. Introduction 1
2. Overview of the issue 3
2.1. Solute carrier transporters 3
2.1.1. SLC structures fold MFS (major facilitator superfamily) type 4
2.1.1.1. Binding sites and coupling in MFS folding 7
2.1.2. Structures folding: LeuT type 9
2.1.2.1. Features of binding sites for substrate and ion binding and coupling in LeuT folding 10
2.1.2.2. Disease relationship of MFS and LeuT folds 12
2.2. Solute carrier transporters, family 22 13
2.2.1. OCT subgroup (SLC22A1, SLC22A2 and SLC22A3) 16
2.2.1.1. SLC22A1 (OCT1) 17
2.2.1.2. SLC22A2 (OCT2) 18
2.2.1.3. SLC22A3 (OCT3) 18
2.2.2. OCTN/OCTN-Related group (SLC22A4, SLC22A5, SLC22A15, SLC22A16 and SLC22A21) 19
2.2.3. OATS1 group (SLC22A6, SLC22A8, and SLC22A20) 20
2.2.4. OATS2 group (SLC22A7) 21
2.2.5. OATS3 group (SLC22A11, SLC22A12, and Slc22a22) 21
2.2.6. OATS4 group (SLC22A9, SLC22A10, SLC2A24, and SLC22A25) 22
2.2.7. OAT-Like (SLC22A13 and SLC22A14) 22
2.2.8. OAT-Related (SLC22A17, SLC22A18, SLC22A23, and SLC22A31) 23
2.2.9. Recognition of substrates and inhibitors of OAT and OCT groups and the transport mechanism 23
2.2.9.1 Substrates and Inhibitors 26
3. Aim of the work 28
4. Materials and methods 29
4.1. Methods 29
4.1.1. Structure search 29
4.1.1.1. AlphaFold: structure predicting AI software 29
4.1.1.2. SWISS-MODEL: Protein homology modelling server 30
4.1.1.3. RCSB PDB 32

4.1.2. Ligand search 33
4.1.2.1. Molecule on Membranes Database: MolMeDB 33
4.1.2.2. Ligands from in vitro studies 34
4.1.2.3 ChEMBL Database 34
4.1.2.4 UniProt Database 35
4.1.3. Method of molecular docking 35
4.13.1. AutoDock Vina 35
4.1.3.2. AutoDock Vina 1.2.0 36
4.1.4. Description of the binding site of the OCTs proteins 37
4.1.4.1. Schrödinger Maestro 13.6 37
4.2. Software for results processing 37
5. Results 38
5.1. MolMeDB ligands docking 38
5.1.1. Bigger dataset of ligands – second ligand docking 42
5.1.2. AutoDock Vina, protein-ligand docking 42
5.1.3. AutoDock Vina 1.2.0, protein-ligand docking 45
5.2. Analysis of binding sides 48
5.2.1 Binding pocket of OCT1 48
5.2.2. Binding pocket of OCT2 49
5.2.2.1. Comparison with recent findings on OCT1 and OCT2 structures 52
5.2.2.2. Possible mechanism of inhibition/activation 54
5.2.2.3. Comparison with OCT3 experimental structure 55
6. Discussion 57
7. Conclusion 60
8. List of references 62

List of abbreviations

3D – three-dimensional

ADP – adenosine diphosphate

AI – Artificial intelligence

Ala – alanine

ApcT – amino acid/H+ symporter

Arg – arginine

Asn – asparagine

Asp – aspartic acid

ASP – 4-[4-(dimethylamino)-styryl]-N-methylpyridinium

Asx – asparagine or aspartic acid

ATP – adenosine triphosphate

Blue, dark – positive amino acids

Blue, light – polar amino acids

Ca – calcium

CORT – corticosterone

Cys – cysteine

cryo-EM – cryogenic electron microscopy

D22 – decynium-22

ΔG – Gibbs free energy

$dG°$ – standard Gibbs binding free energy

EMBL – European Molecular Biology Laboratory

EmrD – multidrug transporter

GlcP – Glucose/H+ symporter

GLUT – glucose uniporter

Glu – glutamic acid

Gln – glutamine

Glx – glutamine or glutamic acid

Gly – glycine

GlpT – glycerol-3-phosphate transporter

HEK293T – Human embryonic kidney 293 cells

Green – hydrophobic amino acids

His – histidine

H – hydrogen

IC50 – concentration of drug required for 50% inhibition

Ile – isoleucine

K – binding constant

K_m – Michaelis-Menten constant

K_i – constant of inhibition

LacY – lactose/H+ symporter

Leu – leucine

LeuT – leucine transporters

Lys – lysine

MATE – multidrug and toxin extrusion family

Met – methionine

MFS – Major facilitator superfamily

MolMeDB – Molecules on Membranes Database

MPP – 1-methyl-4-phenylpyridinium

MSAs – multiple sequence alignments

MhsT – dopamine transporter/Na+ symporter

Na – natrium

NMDA – N-methyl-D-aspartate

NMN – N-methylnicotinamide

NSS – Neurotransmitter Sodium Symporter

NRT1.1 – nitrate transporter

OAT – Organic anion transporters

Oat-pg – prostaglandin-specific organic anion transporter

OCT – Organic cation transporter

OCTN – Organic cation transporters novel

pLDDT – per-residue confidence score

PDB – Protein data bank

PDB ID – Protein data bank identification

PiPT – phosphate/H+ symporter

Phe – phenylalanine

POT – proton-dependent oligopeptide transporter

Pro – proline

Purple – aromatic amino acids

QMEAN – Qualitative Model Energy Analysis

QMEANDisCo – model quality estimation method in SWISS-MODEL

RCSB PDB – US Research Collaboratory for Structural Bioinformatics Protein Data Bank

R – universal gas constant (8,314 $J \cdot K^{-1} \cdot mol^{-1}$)

Red – negative amino acids

Rst – renal-specific transporter

S1, S2 – binding sites 1 and 2

Ser – serine

SERT – serotonin/Na+ symporter

SLC – Solute carrier transporters

SMTL – SWISS-MODEL Template Library

TEA – tetraethyl-ammonium

Thr – threonine

TM – transmembrane segments

Trp – tryptophan

Tyr – tyrosine

URAT1 – Human urate transporter 1

Val – valine

XylE – D-xylose/H+ symporter

Zn – zinc

1. Introduction

Solute carrier (SLC) family transporters transport various substrates across biological membranes using an ion gradient generated by primary active transporters. Therefore they are categorized as secondary active transporters (X. Liu 2019). Over 500 SLCs sorted in 65 families were described according to the latest review (Schumann et al. 2020).

SLC transporters relate to drug transport via mainly SLC15A, SLC21, SLC22A and SLC47A gene subfamilies. This work is focused on the SLC22A gene subfamily, more specifically on the organic cation transporters (OCT) group, which consists of three representative proteins, OCT1 (SLC22A1), OCT2 (SLC22A2) and OCT3 (SLC22A3) (Engelhart et al. 2020).

The function of OCTs is extremely broad, with the major sites of expression being the basolateral membrane of tissue cells of the liver (predominantly OCT1), kidney (predominantly OCT2), intestine, brain, eye, skin, placenta, and plexus choroudeus. The range of ligands they carry is highly overlapping. They are extensively involved in intestinal absorption, uptake of various substances by the liver, and excretion of hydrophilic drugs by the kidneys. Further, their functions may include modulation of the distribution of endogenous compounds, transport of organic cations, zwitterions, and some uncharged compounds. They significantly impact energy metabolism, pharmacokinetics, and toxicity of drugs and drug interactions (Koepsell 2020), and their study is needed for future innovations in pharmacokinetics and better drug targeting.

The structural studies of OCTs are inherently challenging due to their status as membrane proteins, making determination of their structures difficult. Their hydrophobic regions make them hard to purify and stabilize for analysis. Moreover, their large size and conformational flexibility of loops add complexity to obtaining stable crystals or cryogenic electron microscopy (cryo-EM) samples. Their structure also depends on the native lipid environment, making controlled experimental conditions essential (Marconnet et al. 2020; Schmidpeter et al. 2020). OCTs also show low expression in transiently transfected HEK293T (Suo et al., 2023), there were any structures of OCT transporters captured by Roentgen crystallography or cryo-EM, which changed during composing of this work, and all three representatives were captured by cryo-EM. When this work began to emerge, the

AlphaFold software was introduced (Jumper et al. 2021). Due to this innovation, predicted OCT models from AlphaFold and SWISS-MODEL were used (Jumper et al. 2021; Waterhouse et al. 2018).

This work is therefore focused on bioinformatic analysis of available structures of OCT transporters and their binding sites using molecular docking techniques as a possible method for the determination of substrates of OCT and, if it is successful, could be used in further studies of other proteins and their ligands. Also, the other aim of the work was to analyse the binding site of OCTs, the similarities and differences. Thanks to recent studies describing captured structures of OCTs (Khanppnavar et al. 2022; Suo et al. 2023), this work can compare the results from predicted protein structures to the real ones and describe the accuracy of the predicted models.

2. Overview of the issue

2.1. Solute carrier transporters

Solute carrier transporters (SLC) are a group of membrane-bound proteins with over 500 representatives, sorted into 65 families according to their sequence homology and transport function. In the case of SLCs, as transport proteins, they transport diverse solutes across biological membranes (Bai et al. 2017). Their key role is mediating the influx and efflux of substances as ions, nucleotides, sugars and other small molecules through biological membranes (Schumann et al. 2020). SLCs are ubiquitous in prokaryotic biological membranes as well as in eukaryotic biological membranes (Bai et al., 2017). They play an important role in physiological processes such as cellular uptake of nutrients, absorption of drugs, or xenobiotics metabolism. Dysfunction of these proteins causes a great variety of diverse diseases. From anaemias (SLC4) to defects in zinc transport (SLC30) (Bai et al., 2017). The most well-known SLCs are targets for many classes of marketed drugs. More recently characterized SLC transporters may play a role in developing rare or common diseases. They seem to have a big potential for new drug development or therapeutic opportunities (Lin et al. 2015).

According to Cesar-Razquin et al. (2015), SLC family studies are hindered by the lack of available crystal structures, which gives important functional insight. New protein structure prediction software Alphafold 2 changed the situation in 2020 (Jumper et al. 2021). This work compares known experimental SLC structures (Bai et al. 2017) (Figure 1) with AlphaFold prediction (Figure 2).

The SLC transporters undergo conformational changes to translocate the substrate through the cell membrane. Transporter activity relates to substrate concentration. The part of the transport cycle of SLC proteins, which involves the binding of the solute on one side of the membrane, conformational changes, and subsequent release of the solute on the other side, consists of three distinct states, outward-facing state, occluded state and inward-facing state (Perez et al. 2012; Srikant & Gaudet 2019). The outward-facing state is a conformation of a protein that faces the extracellular space or the lumen of an organelle, and the substrate-binding site is exposed to the outer side of the cell membrane when the transporter is ready to bind substrate and transport it across the membrane (Fan et al. 2021). The occluded state refers to the intermediate state, when the bind substrate is transported through the membrane,

neither described as outward-facing nor inward-facing state. The occluded state is enclosed with transmembrane domains and shielded from the outer environment. In the outward-facing state, the solute is released from the transporter, and the whole transport cycle can begin (Perez et al. 2012; Srikant & Gaudet 2019).

When substrate concentration is high on both sides, SLCs transform to the occluded state, equivalent to the transition state, when the binding site for a substrate is blocked off. This conformation state is critical in maintaining the cell's electrochemical gradients (Drew et al. 2021; Henderson et al. 2019). Metabolites, or ligands, can be translocated uphill thanks to the coupling of the translocation itself with the ion's movement down the transporter's electrochemical gradients. This process, called secondary active transport, is further classified as either symport or antiport depending on if ion and substrate motion – the same way (symport), opposite direction (antiport) (Drew et al. 2021).

Structures of solute carrier proteins were sorted and grouped, according to their protein folds, into four categories, MFS fold, LeuT fold, other antiparallel folds and others, focusing mainly on the MFS category and LeuT category (Bai et al., 2017).

2.1.1. SLC structures fold MFS (major facilitator superfamily) type

The largest evolutionarily related super-family of secondary transporters, which facilitates ligands transport across membranes, are major facilitator superfamily (MFS) (Pao et al. 1998). A typical example of MFS fold is the protein PepTso. PepTso comprises 12 transmembrane (TM) segments, often separated into two separately and individually folded bundles, amino-(TM1-6) and carboxyl-(TM7-12) terminal domains (Figure 3). The domains contain "3+3" inverted repeats, and importantly, transmembrane helices in the amino and carboxyl-terminal domains do not interlace (Bai et al. 2017). An intracellular loop between TM6 and TM7 typically connects two or more separate domains. In SLC family, MFS fold has been found in following proteins: glucose uniporter (GLUT), D-xylose/H^+ symporter (XylE), Glucose/H^+ symporter (GlcP), lactose/H+ symporter (LacY), phosphate/H^+ symporter (PiPT), peptide/H^+ symporter (POTs), nitrate transporter (NRT1.1), multidrug transporter (EmrD) and glycerol-3-phosphate transporter (GlpT).

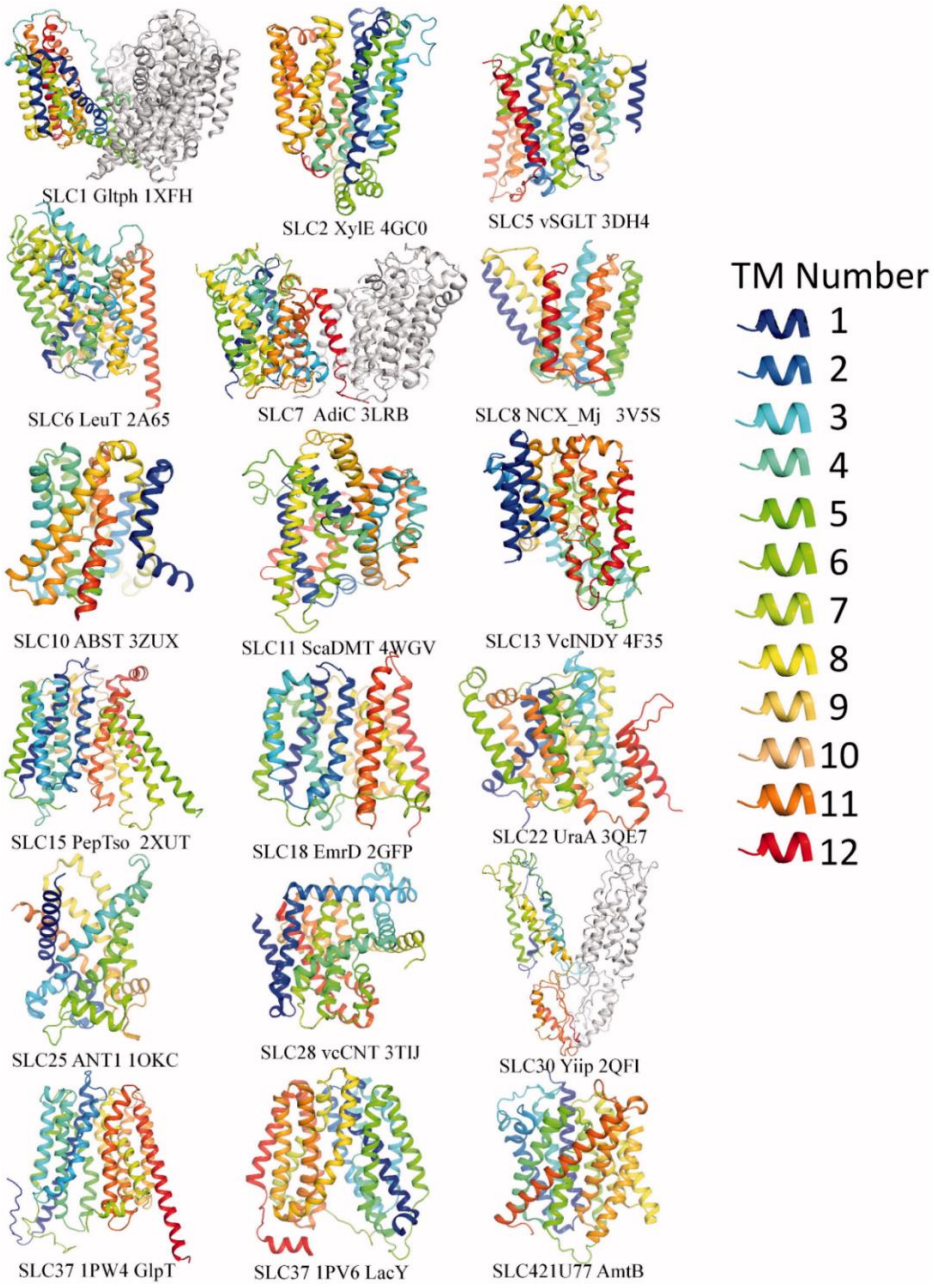


Figure 1: Examples of SLCs structures with representatives from different families. Most structures are from prokaryotic species and were obtained using roentgen crystallography. Nine structures are from mammalian species; another six were from other species. Helices were coloured using rainbow colours from N- to C-transmembrane segment (TM). Shades from blue to cyan are closer to the N-terminal TM, while red to yellow is closer to the C-terminal TM. Structures are originally obtained from PDB (www.pdb.org) and displayed using PyMOL (www.pymol.org) (taken from Bai et al. 2017)

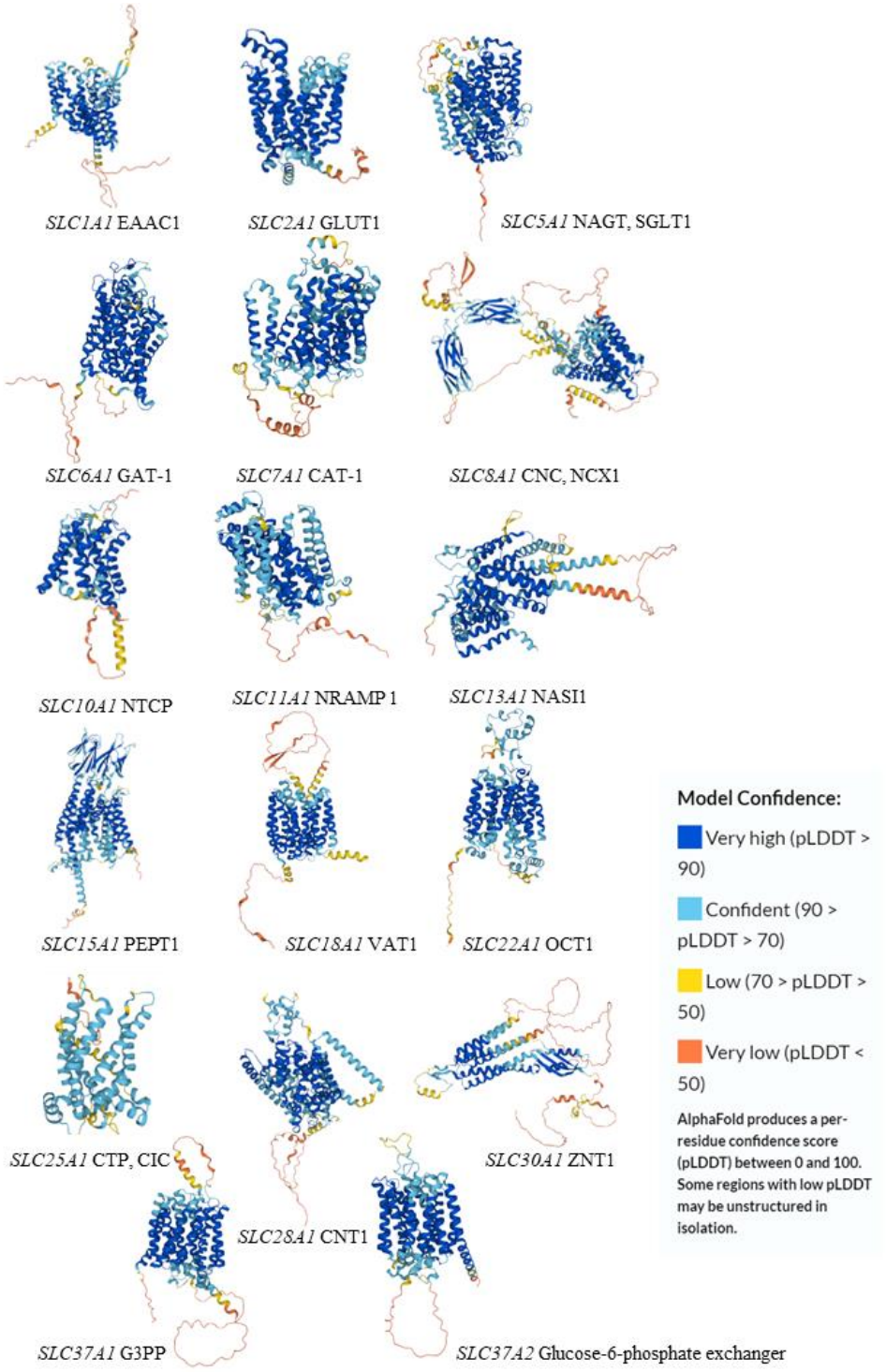


Figure 2: Examples of SLCs structures with representatives from different families. Structures were predicted by AlphaFold 2 program. All structures are obtained from Homo sapiens. The colouring of bundles is according to model confidence. Structures are downloaded from AlphaFold Protein Structure Database (AlphaFold 2023; https://alphafold.ebi.ac.uk/).

Many MFS fold-type structures were captured in an inward-facing conformation (Figure 3). The outward-facing state seems to represent a transient state in the transporting cycle as it is less likely to be captured for most of this family of transporters. It also indicates that the conformational change that brings the empty carrier to the outward-facing state could be the rate-limiting step in the transporting cycle (Bai et al. 2017). Sugar transporters represent an exception to that rule. They have been crystallised in an outward-facing state with antibodies or inhibitors (Sun et al. 2012).

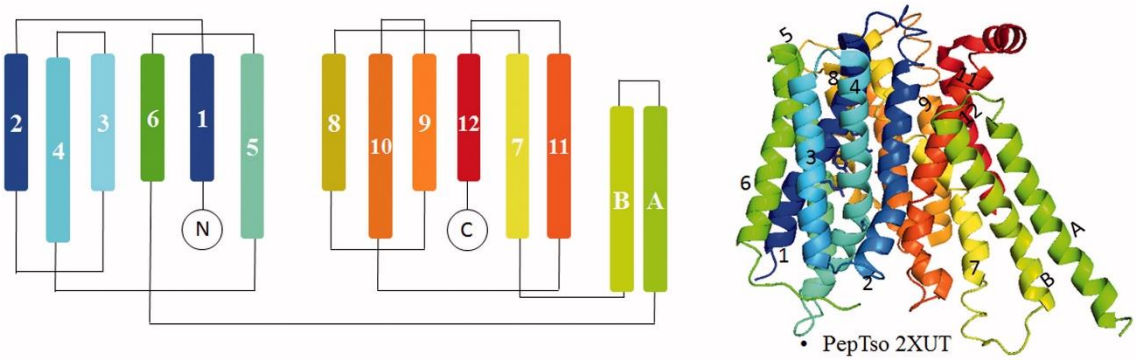


Figure 3: Examples of MFS fold using topology diagram on left and structure itself on the right of model transporter PepTso 2XUT. Helices were coloured using rainbow colours from N to C transmembrane segment (TM). Shades from blue to cyan are closer to the N TM, while red to yellow is closer to the C TM. The structure on the right is in inward-facing conformation (taken from Bai et al. 2017).

2.1.1.1. Binding sites and coupling in MFS folding

The binding sites' molecular features in SLC families vary depending on the transmitted substrate's properties. For example, the usually used model SLC protein PepTso has conserved positively charged residues **Arg**25, **Arg**32 and **Lys**127 on its N-terminal, which are supposed to coordinate the substrate or ligand to the binding site. Two conserved tyrosines, **Tyr**29 and Ty68, are close to these amino acids. On the C-terminal side, the substrate binding site consists of strictly conserved **Glu**419 and **Ser**423 (Newstead et al. 2011).

Another example from the POT family is that the substrate's phosphonate group is recognised by the guanidium group of **Arg**43, the side-chain amide group of **Gln**310, and the hydroxyl groups of **Tyr**40 and **Tyr**78. The N-terminal amino groups are recognised by the conserved **Glu**413 and **Asn**342 (Doki et al. 2013).

According to Bai et al. (2017), this type of binding arrangement is important in the recognition and orientation of the peptides. Different types of peptides with various features could be adapted appropriately within the cavity by the interactions among hydrogens bond donors and acceptors.

Another example of the dependence of substrate properties according to the substrate binding site is the phosphate substrate of PipT, coordinated by **Tyr**150 on TM4, **Gln**177 and **Phe**174 on TM5, **Trp**320, **Asp**324, **Tyr**328 on M7 and **Asn**431 on TM10 by a combination of possible hydrogen bonds, also electrostatic interactions (Pedersen et al. 2013). Before engaging the phosphate, **Asp**324 was proposed to be protonated. The positive charge of **Lys**459 should also be noted, which is conserved at the corresponding position in organic anion transporters, as an OAT3 (**Arg**454), in organic cation transporters, as an OCT1 (**Asp**474). The charge is reversed (**Arg** changed to **Asp**), confirming the residue's function at this site in the transporter cycle (Bai et al. 2017).

Most substrates are bound between the N and C terminal domains on the symmetry axis. The main location of substrate binding sites in MSF fold transporters is at the interface between terminals. PepTso can be used as an example in this case. Residues from C-terminal TM10 and TM11 and N-terminal TM4 and TM5 mainly contribute to the central cavity of the protein (Newstead et al. 2011). Other SLC proteins also have this feature (Bai et al. 2017). In the other proteins with MFS type of folding, the substrate binding site shows a similar location, indicating the relative movements between C- and N-terminals. It can be speaking of a rocket-switch type of motion. This provides alternating access of the substrate binding site to the outside and inside during the transport cycle, but the movement between the domains is considered not rigid, mainly because the binding sites are distributed unevenly on either the C-terminal or N-terminal. According to this, one domain provides most of the binding sites for substrates (Bai et al. 2017).

Regarding ion binding and coupling, binding and substrate binding are usually coupled in symporters, but in antiporters, substrates compete for the same binding site. For substrate binding, mainly in H^+-coupled symporters, protonation was proposed to lower the energy barrier, which has been proved in many examples (Bai et al. 2017). A very detailed example is the protonation mimicking mutant structure of GtPOT **Glu**310Gln (Doki et al. 2013) when substrate peptide cannot

bind into the central pocket because of the electrostatic repulsion between the substrate carboxylate group and the negative charge of deprotonated **Glu**310. The occurrence of an outward and inward conformation transition was only detectable when both protonation and peptide binding happened sequentially. The weakened interaction between **Arg**43 and protonated **Glu**310 cannot be the driving force to bring together the C-terminal close to the N-terminal when the proton is bound. A similar case can be found in other works and studies (Zhao et al. 2014).

Antiporters follow different mechanisms. Antiporters undergo a conformational change from inward- to outward-facing states in the presence of substrate. Without the substrate's presence, they cannot undergo conformal change.

On the contrary, symporters undergo an outward to an inward conformational change in the substrate presence. Also, they can return empty to an outward-facing state without substrate presence. So, in symporters, the energy barrier for the movement of the empty carrier to the outward-facing state is not prohibitively high, in contrast to the energy barrier in antiporters for empty carriers, which is high. Binding the substrate in antiporters lowers the energy barrier (Bai et al. 2017).

The local conformational changes of transporters in response to the binding of ligands can be closing the gate on one side, opening a gate on the other, and forming a transient occluded state when both ends are closing. Ligand is thus not accessible to both sides of the membrane simultaneously during transport. The known bacterial transporters showed thick periplasmatic barriers formed by different bond clusters. Barriers form as strong prevention of leakage or escape of ions and substrates in the state where the transporter is fully opened to the cytoplasmatic side (Bai et al. 2017).

2.1.2. Structures folding: LeuT type

A similar LeuT (leucine transporters) fold to the MFS fold was observed in several unrelated transporters from the SLC family. Including the neurotransmitter/Na^+ symporter LeuT, the amino acid/H^+ symporter ApcT, the dopamine transporter/Na^+ symporter MhsT and serotonin/Na^+ symporter SERT, and others (Bai et al. 2017).

This type of folding usually consists of a “5+5” reversed repeat with an anti-parallel orientation, an in addition, with a pseudo-symmetry axis in the membrane

(Figure 4). Four-helix domains (core domain, TMs 1, 2, 6 and 7) are formed from repeats of the first two helices. The second domain surrounds this domain, consisting of TMs 3, 4, 8 and 9. In addition, the first TM of each repeat contains an unwound region important in ligand binding and transport itself. However, despite the similarities in the main fold, individual SLC transporters usually have specific structural modifications.

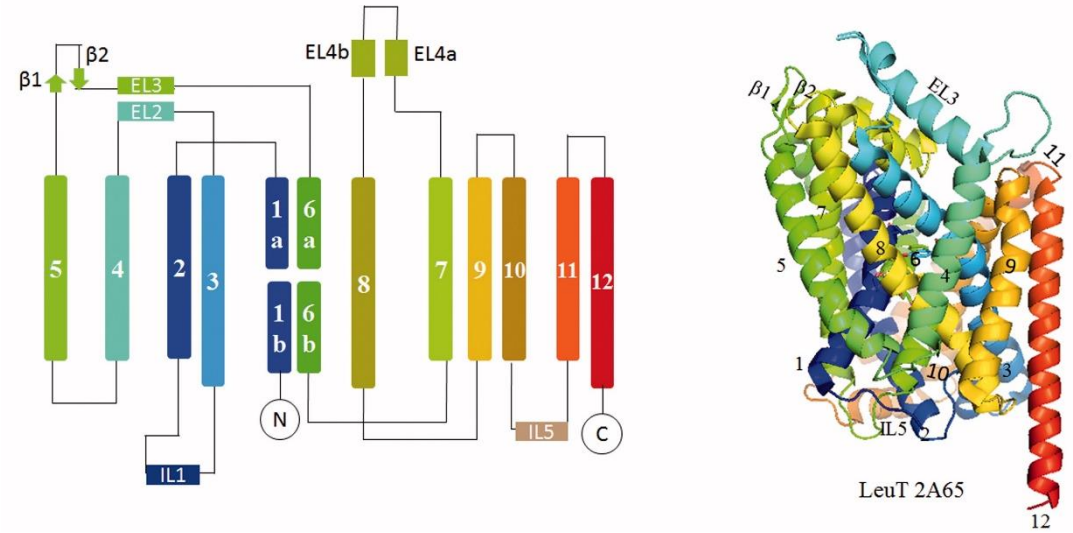


Figure 4: Examples of LeuT fold using topology diagram on the left and structure itself on the right of model transporter LeuT 2A65. Helices were coloured using rainbow colours from N to C transmembrane segment (TM). Shades from blue to cyan are closer to the N-terminal, while red to yellow is closer to the C-terminal. The structure on the right is in outward-facing, partially occluded conformation. (taken from Bai et al. 2017).

2.1.2.1. Features of binding sites for substrate and ion binding and coupling in LeuT folding

According to Yamashida et al. (2005), binding sites for substrates in LeuT folding are formed by residues, mainly regarding TM1, 6 and TM3, 8 (Figure 4). Moreover, according to Bai et al. (2017), binding sites are located within non-α-helical regions of TMs. They are usually connected with ion binding sites. Speaking about polar residues in TM1 and 6, it seems that they make most of the contacts to the bound leucine, which interacts with the main-chain carbonyl oxygens of **Ala**22 (TM1), **Phe**253 (TM6), **Thr**254 (TM6) and side- hain hydroxyl from **Ser**256 (TM6). Directly to the Na+, the amide nitrogen's **Leu**25 and **Gly**26

of TM1 and hydroxyl from **Tyr**108 of TM3 are connected to carboxyl from leucine. The role also plays a hydrophobic pocket consisting of **Val**104 and **Tyr**108 from TM3, **Phe**253, **Ser**256, **Phe**259 from TM6, **Ser**355 and **Ile**359 from TM8, which holds the side chain of leucine. In conclusion, many interactions on the side-chain and back-bone contribute to the specificity of substrate recognition by LeuT (Bai et al. 2017; Yamashita et al. 2005).

In addition, there is still controversy about two binding sites in LeuT, the primary Na^+ binding site (S1) and the second binding site (S2). The S1 site is closer by 10 Å than the S2 site towards the extracellular environment (Shi et al., 2008). The S2 site existence was questioned due to the interference of the detergent β-OG and other studies performed in detergents that do not support its existence (Khelashvili et al. 2013; Quick et al. 2009).

Bai et al. (2017) describe substrate binding sites' position in LeuT as conserved. In other experiments, it was described that the putative binding site for substrate in inward-occluded AcpT is similar to that of LeuT, which was formed by residues from non-α-helical regions of TM1, TM6 and TM3, TM8 (Shaffer et al. 2009). The similarity in the outward-facing stage was found in outward-facing DAT, where the pockets were equivalent (Penmatsa et al. 2013). The residues responsible for sugar coordination differs from MFS folding-type transporters.

An example of unique sugar coordinates can be vSGLT protein, where though lactose is bound about halfway across the membrane similarly to substrates in the transporters mentioned before, there were found interactions between lactose and **Gln**428 (TM11), **Gln**69 (TM2), **Lys**294 (TM8), **Ser**79 (TM3), **Tyr**263 (TM7), **Asn**260 (TM7) and **Tyr**87 (TM10) through hydrogen bonds. An interesting case is Lys294, also observed in LacY from the MFS folding type group of transporters and other sugar transporters. This indicates a potential conserved substrate binding mechanism across MFS and LeuT fold types (Bai et al. 2017).

Two sodium ions are important in stabilization of LeuT core region when sodium ion binding site 1 (Na1) is linking directly to the leucine carboxyl oxygen and other carbonyl oxygens of **Ala**22 (TM1) and **Thr**254 (TM6), the hydroxyl oxygen of **Thr**254, the side-chain carboxyl oxygens of **Asn**27 (TM1) and **Asn**286 (TM7). Sodium binding site 2 (Na2) is located between TM8 and TM1 unwound region and is less conserved. Na2 is coordinated by side chain hydroxyl oxygens of **Ala**351 (TM8), **Thr**354 and **Ser**355, the carbonyl oxygens of **Gly**20 and **Val**23 (TM1).

Ions and substrates can be found in stoichiometry ranging from 1:1 to 1:3. The Na2 site in LeuT superimposes with the ion binding sites of several other SLCs. Ion binding and ion releasing directly connect with substrate binding and releasing (Bai et al. 2017).

The transition from outward open to the occluded semi-open state can involve the movement inward of TM1b and TM6a towards TM3 and TM10 (Claxton et al. 2010). Afterwards, upset from occluded to inward-facing state seems to be contributed by TM1a outward movement with TM6b (Bai et al. 2017; Zhao et al. 2011). Newest studies also suggest a structural, maybe regulatory role of cholesterol (Penmatsa et al. 2013), moreover indicated flexing of the intracellular half of TM1 at 13° away in the inward-open state without substrate compared to the inward-occluded state (Watanabe et al. 2010). This effect shows that the TM1 movement can be a major event through the transition to an inward open conformation of LeuT, but it is clear that conformational changes connected with translocating of substrate involve more (Bai et al. 2017).

2.1.2.2. Disease relationship of MFS and LeuT folds

Many available SLC homologs or paralogs in the MFS fold have pharmacological or medical relevance. The best example mentioned in many studies last couple of years is the POT transporter located in the intestinal tract providing a major route for the uptake of orally administered drugs, like β-lactam antibiotics. Conversely, the LeuT fold group, the second largest group of fold clusters in SLCs, are proteins from the Neurotransmitter Sodium Symporter (NSS) family. NSS correlate with many diseases, such as orthostatic intolerance, depression, attention deficit hyperactivity disorder, Parkinson's disease, or infantile parkinsonism dystonia (Bai et al. 2017). Good to mention that this group of transporters has been targeted by antidepressants, although numerous reports of side effects (Anderson 2000).

Further studies of SLCs could show the effect of disease-related mutations on these proteins' folding and transport activity. For this reason, is a structural investigation of MFS, either LeuT folds essential for understanding the molecular basis of substrate recognition and transport, such as the transport investigation. This knowledge would help with drug improvement and quicker drug implementation.

2.2. Solute carrier transporters, family 22

This work mainly focuses on ligand transport by SLCs from family 22, which falls under MFS folding. The main aim of family 22 is moving small molecule endogenous metabolites between tissues and interfacing body fluids. Nevertheless, not only them, they also play a role in the transport of drugs and toxins, which can be exogenous or endogenous (Nigam 2018). An example of an SLC22 transporter in tissue is shown (Figure 5). The focus of this family is mainly on their pharmaceutical importance, which makes them one of the best-studied SLC families. According to recent studies, it has been shown that SLC22 transporters have a bigger role in whole organisms than we thought before, including interorgan and inter-organism communication (Nigam 2018). These transporters are expressed in the kidneys, liver, heart and brain but can also be found in erythrocytes, monocytes and macrophages (Lamhonwah et al. 2008; Minuesa et al. 2008; Sager et al. 2018).

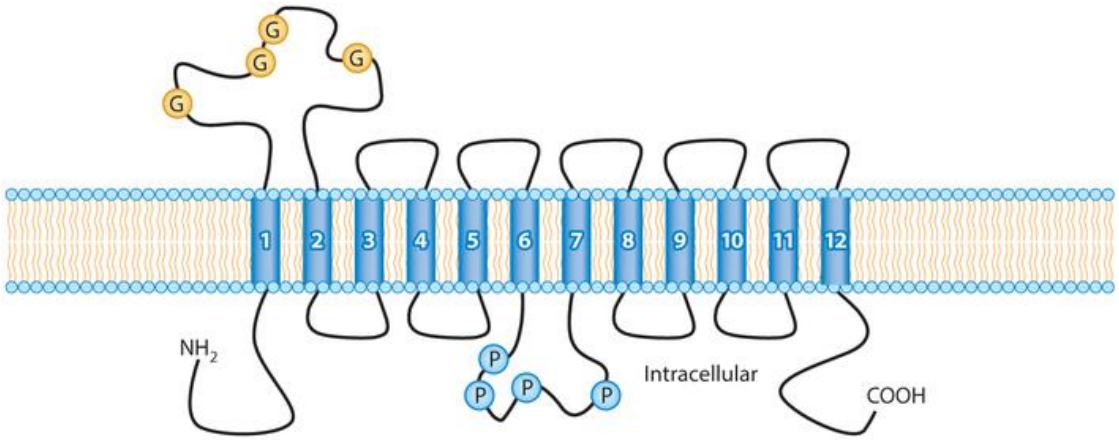


Figure 5: The deployment of 12 transmembrane domains (numbered blue cylinders) representative from the SLC22 family. The representative is placed in epithelial tissue as the main place of expression of SLC22 members. The placement is according to the transporter expression turn on the basolateral side of the cell membrane (taken from Eraly et al. 2004; Nigam 2018).

There are many discussions on how to order individual transporters of family 22. There are two possible ways mostly used in reviews and studies whenever transporters are discussed in order of their numeral designation as an SLC22, but a more extended approach is a separation of transporters into three main groups (OAT, OCT, OCTN) and other subfamilies (Engelhart et al. 2020). This approach is older and based on transporters' structural or functional relatedness. In another

review, authors divided the SLC22 family into at least six groups, which are OAT, OAT-like, OAT-related, OCT, OCTN and OCT/OCTN-related based on evolutionarily distinct groups of SLC22 groups (Nigam 2018). In another, more current study, authors corrected this division to eight SLC22 transporter subgroups, which are OATS1, OATS2, OATS3, OATS4, OAT-like, OAT-related, OCT and OCT/OCTN-related (Engelhart et al. 2020). This reanalysis and following correction was done because of new data of structures from the family, which cannot be truly included in any of the subclades before. Many of the transporters of SLC22 were ill-defined or remain unknown (Engelhart et al., 2020). Each subclade focused on OAT, like transporters, which were this work's main subject of interest. The work will follow the recognized subclades from 2020 (Engelhart et al. 2020).

The old division into subfamilies or subclades is demonstrated in Figure 6. The new division can be found in Figure 7. Despite all similarities, there are doubts about at least one particular gene, SLC22A18, which may not be a legitimate family member (Nigam 2018; Zhu et al. 2015).

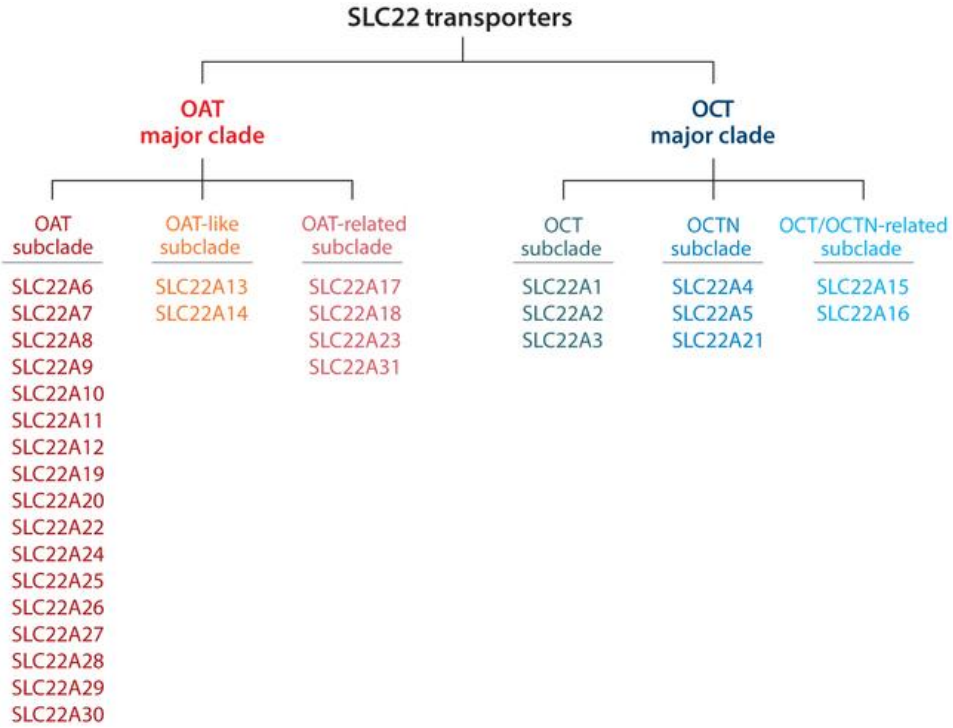


Figure 6: The old division of SLC22 members into six subgroups, two major subclades, OAT (organic anion transporter) major clade and OCT (organic cation transporter) major clade. Both are next divided into three subcategories/subgroups. The rest transporters with related features are included in OAT-related or OCT/OCTN-related subclades (taken from Nigam 2018).

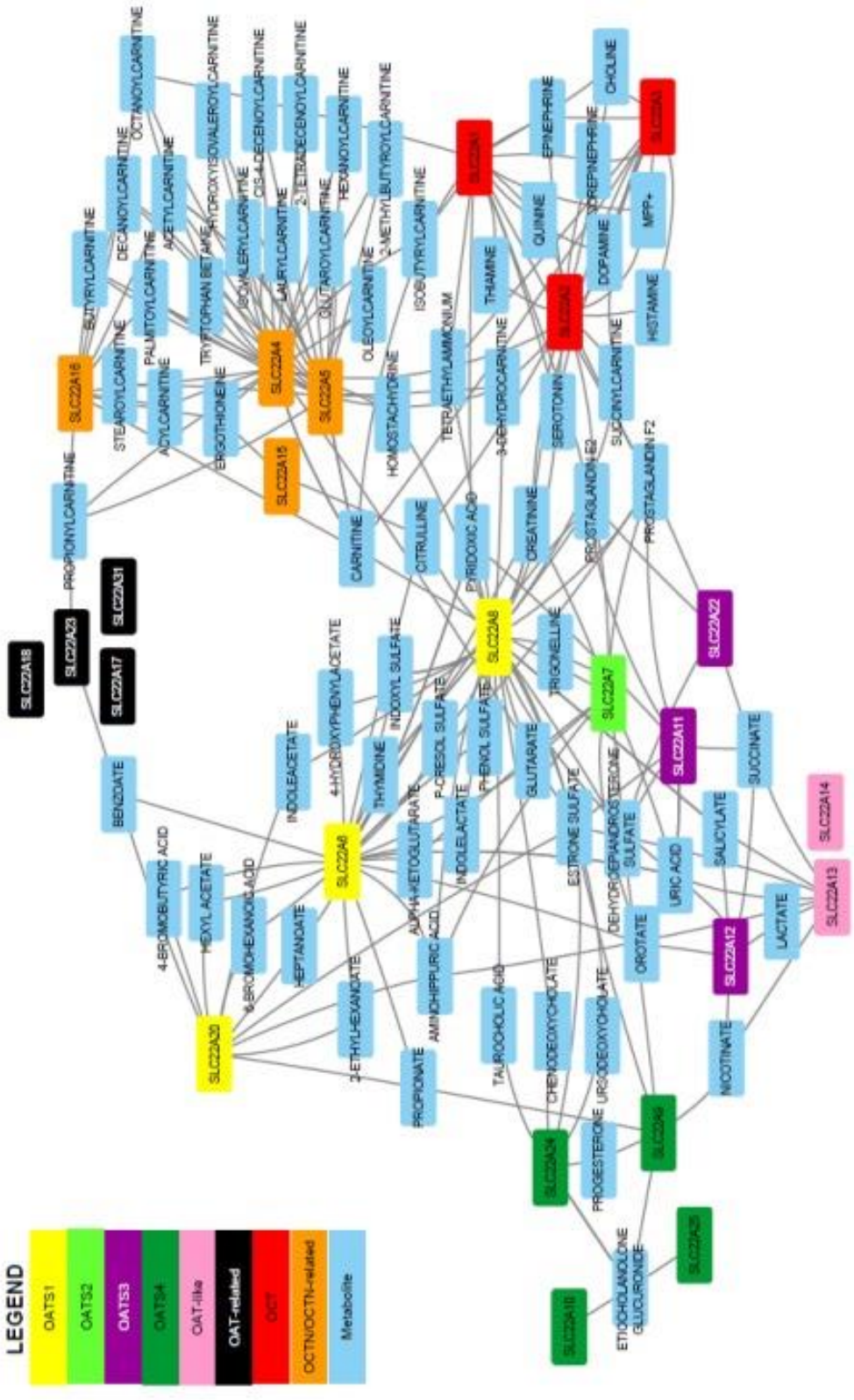


Figure 7: The new division, according to the study done in 2020, is displayed as a pruned network with the inclusion of all functional data about SLC22 members. Each edge shows a noteworthy transporter-metabolite association (taken from Engelhart et al. 2020).

2.2.1. OCT subgroup (SLC22A1, SLC22A2 and SLC22A3)

The OCT means a short version of organic cation transporters, which is also the main characterization of this type. They have a high affinity for monoamine neurotransmitters, signalling molecules and other biologically important metabolites. In this subgroup can be found 3 transporters, SLC22A1 (OCT1), SLC22A2 (OCT2) and SLC22A3 (OCT3) (Figure 8). They transport monoamine transporters and carnitine derivates, creatinine, MPP+, TEA, and others, as shown in Figure 8 (Engelhart et al. 2020). They are expressed in the brain, liver, and kidneys.

Moreover, there are some considerations about their role in inter-organ communication between the brain and the kidney-liver axis, together with neurotransmitter transport. The systematic level of neurotransmitters can be regulated by OCT transporter enzymatically. This can affect neurotransmitter availability for brain tissue, also the expression of metabolites in the liver and kidneys, which represents the excretory route (Engelhart et al. 2020; Uhlén et al. 2015).

The distribution of OCT in organs is different. Major liver OCT is OCT1 and for kidneys OCT2. OCT3 is also expressed in the liver and kidneys but is important in neurotransmitter uptake in the brain (Nigam 2018). OCT1 and OCT2 have similar substrate specificities, with any known predisposition to transport by OCT1 versus OCT2 (Liu et al. 2016). If one speaks of an endogenous function, the transport of polyamines, thiamine and carnitine is done by OCT1 and the transport of dopamine, creatinine and acetylcholine is done by OCT2 (Nigam 2018).

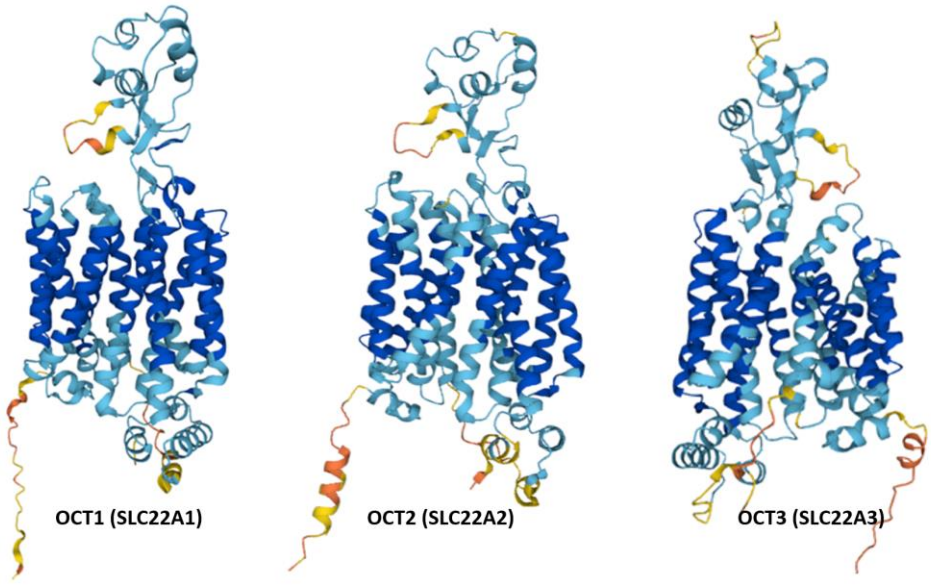


Figure 8: Structures predicted by AlphaFold. From left to right, OCT1, OCT2 ł OCT3 are all in inward-facing conformation. The colouring of bundles is according to model confidence. (taken from AlphaFold 2023).

2.2.1.1. SLC22A1 (OCT1)

As mentioned in the previous part, OCT1 (Figure 8) is mainly expressed in the liver, strongly in the sinusoidal membrane of hepatocytes. Weaker expression of hOCT1 can be found in various tissues, such as the small intestine, lungs, heart, kidneys, skeletal muscle, brain, placenta, eyes, adrenal glands, immune cells, and skin. According to transported substrates, in the last part were some of them mentioned. OCT1 also transported monoamine neurotransmitters, the polyamine putrescine, tetraethyl-ammonium (TEA), 1-methyl-4-phenylpyridinium (MPP+), N-methyl nicotinamide (NMN), and 4-[4-(dimethylamino)-styryl]-N-methylpyridinium (ASP). From various drugs, OCT1 transport the antidiabetic metformin, antiviral drugs such as acyclovir, and antineoplastics such as oxaliplatin (Volk 2014). Mediating the uptake of organic cations in hepatocytes seems to be the major function of OCT1 to start biliary secretion. OCT1 was also detected in the luminal membrane of proximal and distal tubules. This can suggest the contribution to the reabsorption of organic cations from the primary urine (Koepsell 2013). Besides these functions, OCT1 can be important in organic cations absorption in the small intestine and in the substrate transport through the blood-brain barrier (Koepsell 2013; Volk 2014).

2.2.1.2. SLC22A2 (OCT2)

SLC22A2, known as OCT2, is mainly expressed in kidneys, typically in the basolateral membrane of proximal tubules (Volk 2014). OCT2 has from other OCTs from the family 22 specific structure, where the last three TMs are lacking in splice variant (Koepsell et al. 2007). As mentioned, OCT2 relates to kidneys, which are responsible for the first step of tubular secretion with cellular uptake of organic cations. Another family does the luminal release of cations - SLC47 (Koepsell 2013). Connections have been found in neurotransmitter uptake and some parts of the brain, such as the hippocampus. Here OCT2 controls neurotransmitter concentration and facilitates transport across the brain-blood barrier. Several drugs are transported this way. OCT2 may also do the epithelial release of acetylcholine in the lungs. This process occurs during non-neuronal cholinergic regulation (Lips et al. 2005; Volk 2014; Wessler & Kirkpatrick 2008).

The substrates broadly overlap with OCT1, but some significant differences were described from the neurotransmitter group (acetylcholine, epinephrine, norepinephrine, dopamine, histamine, and serotonin) and the endogenous group (choline, putrescine, and guanidine). Also, similar to OCT1, this transporter facilitates model compounds (TEA, MPP, ASP and NMN) and drugs (cisplatin, oxaliplatin, picoplatin, amantadine, memantine, cimetidine, famotidine, ramitidine, and metformin) and some toxic compounds (ethidium bromide). Moreover, some studies described the transport of inorganic cations (Cs^+ and Cd^{2+}) (Koepsell 2013; Nies et al. 2011; Volk 2014).

2.2.1.3. SLC22A3 (OCT3)

SLC22A3 is expressed in many tissues, including skeletal muscle, liver, brain, heart, lungs, small intestine, kidneys, skin, placenta and mammary glands (Koepsell 2013), but the substrate spectrum is significantly smaller compared to OCT1 and OCT2. SLC22A3 is also known as OCT3 (Figure 10). Between facilitating substrates can be found neurotransmitters, like epinephrine, norepinephrine, agmatine and histamine. The drugs category includes metformin, oxaliplatin, lamivudine, lidocaine, quinidine, and etilefrine. Only MPP is known to have facilitated transport with OCT3 from model compounds. The function of OCT3 is different according to the tissue. In the liver, also with OCT1, it is expressed in the sinusoidal membrane of hepatocytes and mediates the initial

step in the biliary secretion of organic cations. In brain tissue, OCT3 can be found in the substantia nigra, cortex and hippocampus, contributing to the regulation of aminergic neurotransmission (Volk 2014). The study on mice's brains leads to the statement that the lack of this transporter in the brain causes anxiety and stress (Vialou et al. 2008). In other organs, OCT3 may release acetylcholine during non-neuronal cholinergic regulations. However, its function has not been clarified in all tissues where OCT3 is expressed. Interesting ones are heart and skeletal muscle tissue, where OCT3 is highly expressed, but its function is still unknown (Volk 2014).

2.2.2. OCTN/OCTN-Related group (SLC22A4, SLC22A5, SLC22A15, SLC22A16 and SLC22A21)

This group is a combination of OCTN and OCTN-related. According to a new division from 2020 (Engelhart et al. 2020), this group is joining previous groups together. Firstly, writing something little about the old division for OCTN and OCTN-related is good.

OCTN had underneath three transporters, SLC22A4 (OCTN1), SLC22A5 (OCTN2) and SLC22A21. Like the previous group, OCTN1 is highly expressed in kidneys, intestine, liver, and other tissues; OCTN2 is mostly expressed in kidneys, but there is some expression in other tissues like the intestine, placenta and others. Both are likely carnitine transporters, but OCTN1 seems to be also ergothioneine transporter (Nigam 2018).

The OCTN-related group contains two transporters, SLC22A15 (FLIPT1) and SLC22A16 (FLIPT2). Their names "FLIPT" are from fly-like putative transporter. They are also carnitine transporters, expressed in kidneys and other tissues, but have not been well studied yet (Nigam 2018).

These two groups have been joined because of new data from the last few years. Also, the transported ligands have been more clarified. There was a mistake when the transporters were first named. SLC22A15 was wrongly named CT1, and like a carnitine transporter, this name now belongs to SLC22A5 (OCTN2). According to a new study, the transporters with high affinity to carnitine and its derivates are, according to a new study, SLC22A4, SLC22A5 and SLC22A16. However, lower affinity seems to have SLC22A4 but has a high affinity for ergothioneine. This feature shares with SLC22A15 and SLC22A16 (Engelhart et al. 2020;

Gründemann et al. 1997; Shin et al. 2014). FLIPT1 also has one specificity unique in the SLC22 family: its affinity to complex lipids. This unique characteristic has not been found in any other known SLC22 member. Although data are still very limited, it can be said that this subgroup is expressed at least in five types of tissues and in circulating immune cells. All the known information about this subgroup suggests their main purpose is transporting carnitine and carnitine derivates. Carnitine metabolism is an energy-producing mechanism in every cell. Another described substrate, ergothioneine, is regulated by transporters from this subgroup, which is unique for this subgroup (Engelhart et al. 2020).

2.2.3. OATS1 group (SLC22A6, SLC22A8, and SLC22A20)

The best-studied group OAT had underneath around seventeen transporters. According to the new data, the group was divided into four unique subgroups when transporters were divided for their ligand affinity, structure, and other specialities. Starting with OATS1, which handles a wide variety of metabolites. There can be found signalling molecules, odorants and uremic toxins. Best studied transporters SLC22A6 (OAT1) and SLC22A8 (OAT3) also belong to this group. As substrates for these two transporters have been identified several metabolites, many of which are unique but with notable overlap. SLC22A6 and SLC22A8 interact with gut microbiome-derived products, uremic toxins and other general SLC22 metabolites, for example, prostaglandin E2, prostaglandin F2, creatinine and uric acid (Engelhart et al. 2020). OAT6 (SLC22A20) has not been studied as well as the previous two transporters, but according to a study from 2015 (Wu et al. 2015), OAT6 has an affinity to several odorants and short-chain fatty acids, and it appears like oligo-specified, compared to the previous OAT1 and OAT3, which are multispecific. The tissue localization of these transporters is notably different. The first two are highly expressed in kidneys, but OAT6 were found in mice's olfactory mucosa, reflecting its affinity to odorants (Monte et al. 2004; Wu et al. 2015). OAT1 with OAT3 in kidneys helps regulate urine levels of many metabolites and signalling molecules with other transporters from the SLC22 family. It has been described as potential inter-organism communication as well, which should go by mechanism when OAT6 notices some volatile substance in olfactory mucosa but has been before excreted from another organism with urine through

OAT1. This hypothesis was published in 2009 but was not studied well afterwards (Ahn & Nigam 2009; Engelhart et al. 2020).

2.2.4. OATS2 group (SLC22A7)

In this group, only one transporter, SLC22A7 (OAT2). It is associated with prototypical substrates, such as carnitine, creatinine, prostaglandins, and uric acid, all characteristic of the SLC22 family. OAT2 seems to have a different evolutionary pattern with a single common ancestor, according to the research done in previous years and has connections with cyclic nucleotides and dicarboxylic acids. This gives it unique characteristics, which have not been noticed in any previous transporters from SLC22 (Sun et al. 1993). Besides that, compared with other transporters from SLC22, OAT2 is not expressed in the liver or kidneys but in circulating red blood cells, which may play a role in cycling nucleotide transport. This feature makes it act as an avenue for signalling (Engelhart et al. 2020; Sager et al. 2018).

2.2.5. OATS3 group (SLC22A11, SLC22A12, and Slc22a22)

OATS3 group consist of three transporters. Starting with SLC22A11 (OAT4) and SLC22A12 (URAT1) have just two substrates, succinate and uric acid. URAT1 have a main role in uric acid reabsorption in the proximal tubule of kidneys. High levels of uric acid in the blood are harmful. Uric acid is responsible for more than half of human antioxidant activity. Reabsorption of uric acid makes URAT1 relatively mono-specific. OAT4 relates to prostaglandins and conjugated sex hormones in addition to uric acid. According to this, OAT4 is oligo specific. Speaking about tissue localization, for URAT1, it is just kidney tissue, but for OAT4, the places of expression are kidneys, placenta, and epididymis, which is logical according to its facilitating substrates. In rodents, the subgroup differs for Slc22a12 and Slc22a22. Slc22a12 is known as the renal-specific transporter (Rst) in mice, and Slc22a22 is known as the prostaglandin-specific organic anion transporter (Oat-pg) when they do not share substrate specificity but together handle the role of URAT1 and OAT4 in mice (Engelhart et al. 2020; Shiraya et al. 2010).

2.2.6. OATS4 group (SLC22A9, SLC22A10, SLC2A24, and SLC22A25)

OATS4 group is characterized by handling common metabolite, etiocholanolone glucuronide, a conjugated sex hormone. All the members from the OATS4 group are specifically associated with conjugated sex hormones (Long et al. 2017). SLC22A24 and SLC22A9 seem to be more oligo-specific transporters. SLC22A24 is linked with bile acids, and SLC22A9 has a bigger affinity to short-chain fatty acids. Another one, SLC22A25 and SLC22A10, are linked just to the conjugated sex hormones, so that makes them mono-specific ones. There is quite a similarity in tissue expression amongst OATS4 members of Homo sapiens, also shared function. A study done by 2020 (Engelhart et al. 2020) spoke about the OATS4 group as a conjugated sex steroid transporter, showing high expression in the liver. This tissue relates to conjugating glucuronides and sulphates to androgens and other gonadal steroids. SLC22A24 have connections with the reabsorbance of conjugated steroids in the proximal tubule of the kidney, where it is also highly expressed (Engelhart et al. 2020; Yee et al. 2019).

2.2.7. OAT-Like (SLC22A13 and SLC22A14)

This subgroup remains unknown, and very little about them is still available. SLC22A13 (OAT10/ORCTL3) has been described as a nicotine and urate transporter (Bahn et al. 2008), but there is no information about substrates for SLC22A14. There have been smoking cessation studies when SLC22A14 seemed connected with N'-methyl nicotinate, but it is not directly related to nicotine (Uhl et al. 2008, 2009). The expression specificity is also blurry when SLC22A14 have no human protein expression data and transcribed genes are found at low levels in the kidneys, the main tissue for SLC22A13 expression. Interestingly, high levels of the transcribed gene of SLC22A14 can be found in the testis. This is in concordance with the critical role of Slc22a14 in mice's sperm motility and mice male fertility (Maruyama et al. 2016). According to the sequence-based analysis, SLC22A13 and SLC2214 belonged in their subgroup. As was said, this group require more studies and functional classification (Engelhart et al. 2020).

2.2.8. OAT-Related (SLC22A17, SLC22A18, SLC22A23, and SLC22A31)

The OAT-related groups orphan transporters SLC22A17, SLC22A18, SLC22A31, and SLC22A23. Same features shares SLC22A23 (BOTC2) with SLC2217 (BOTC1) and share more than 30 % of amino acid identity. Both have high expression in brain tissue. It also showed a non-conserved amino terminus, which could negate SLC22's prototypical function (Bennett et al. 2021). SLC22A17 (LCN2-R), also known as lipocalin receptor 2, was defined as a mediator of iron homeostasis thanks to the binding and endocytosis of iron-bound lipocalin. There was also described its high affinity to proteins like calbindin (Wan-Jie et al. 2016). There is not enough information to confirm SLC22A23 substrates, but it has mutated gene forms with medically relevant phenotypic associations, including inflammatory bowel disease, endometriosis and clearance of antipsychotic drugs (Engelhart et al. 2020). Not very well-studied SLC22A31 has been associated with right-side colon cancer (Bien et al. 2019). SLC22A18 membership in SLC22 remains highly arguable, mostly because of its connections with the DHA H^+ antiporter family through sequence similarity (Zhu et al. 2015). The substrate specificity of the OAT-related group needs further study. This group shows high sequence diversity and many deviations from canonical physical transporters from the SLC22 family (Engelhart et al., 2020).

2.2.9. Recognition of substrates and inhibitors of OAT and OCT groups and the transport mechanism

The substrate binding site is exposed to one side of the cell membrane at a time, which describes the foundation of the transport mechanism of SLCs lacking the source of energy. This is called the "alternating-access model" (Jardetzky 1966), which can be described by the following models: the "rock-switch", the "rock-bundle", or the "elevator" mechanism (Figure 9) (Drew et al. 2021).

The typical structure of transporters from the SLC22 family, which comes under MFS folding, indicates transport type "alternating access". The mentioned type of transport conditions the conformational change of the transporter through an outward-facing state to an occluded state when the metabolite bind and separates from both sites of the cellular membrane (Figure 10). When the substrate relates to the intracellular environment and is released, the final conformation is called the inward-facing state. However, it is assumed that the SLC22 family shares

a transport mechanism with other members from MFS folding. The described general mechanism has not been proven in the case of OATs and OCTs (Matsson & Karlgren 2021).

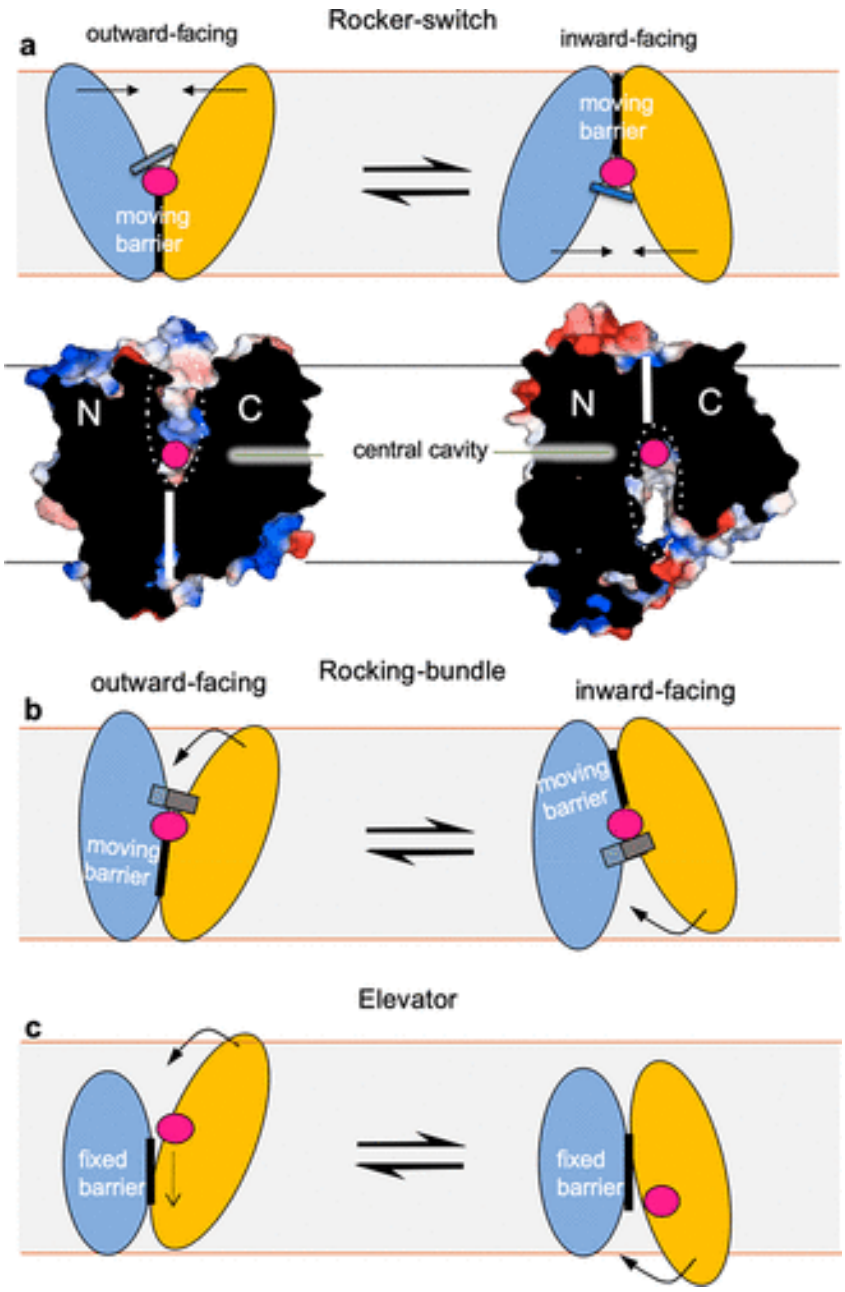


Figure 9: The representations of three "alternating-access model" types as a transport mechanism for SLCs. There can be seen the substrate translocation itself, (a) rock-switcher, (b) rocking-bundle and (c) elevator. Substrate translocation through the membrane from extracellular space to the intracellular environment is shown (taken from Drew et al. 2021).

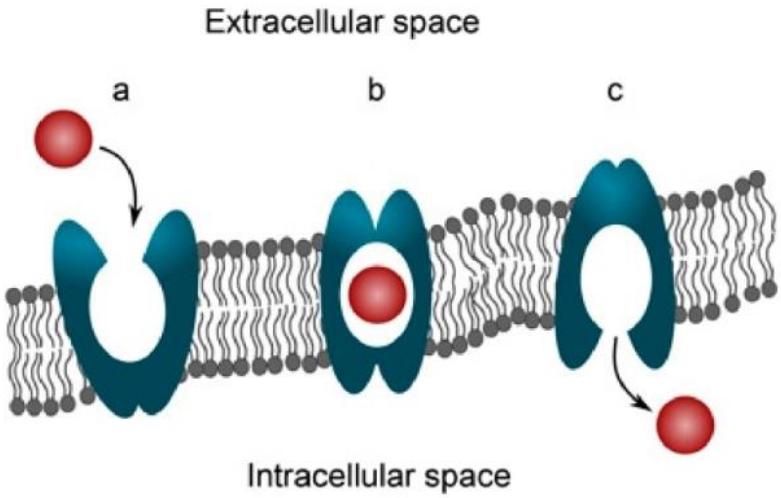


Figure 10: The draft of the transport mechanism for the SLC22 family based on the alternating-access model, a typical system of substrate translocation of MFS fold type. All three conformational states are shown: (a) outward-facing state, (b) occluded state and (c) inward-facing state. The substrate is translocated through the cell membrane from the extracellular space to the intracellular space of the cell. The reverse direction is possible depending on the cell's needs (taken from Matsson & Karlgren 2021).

The known and described substrates for both subclades, OAT and OCT, have been described in a previous chapter, but the number of drugs with clinical effect, clear and approved, is to date relatively small (Matsson & Karlgren 2021). Metformin is the most prominent example in the case of OCT1 and OCT2. Other examples, like dofetilide, pilsicainide or ranitidine, showed drug-drug interactions involving OCT substrates. The suggested problem in the case of these drugs is that the main effect of these drugs for renal excretion is related to the multidrug and toxin extrusion transporter's inhibition. The group of multidrug and toxin extrusion transporters have similar substrate selectivity with OCT2 but are expressed on the apical membrane of the renal epithelium when the SLC22s are expressed on the basal membrane (Matsson & Karlgren 2021; Wittwer et al. 2013).

In the case of the OAT subclade, the reported clinical interaction has been reported with the established OAT inhibitor probenecid, which can be used as a comedication to increase systemic exposure to antibiotics, as well as a nephroprotective agent for protection against renal toxins associated with cidofovir

treatment (Lalezari et al. 1997). This proposed use is based on the probenecid effect on drug uptake in the renal epithelium and secretion to urine. Similar properties have the structurally related drug tenofovir (Fernandez-Fernandez et al. 2011).

Research focused on finding new substrates or designing them is important to know these facts. Firstly, the evidence that the compound has been successfully bound to the transporter is not necessarily proof that the compound itself is the substrate that is transported. The other fact is that the previous studies drew information on drug interactions with transporters through anecdotal observation of substrates. Another way was large-scale screening assays of inhibition of the transporter. Lastly, the high concentration of substrates triggers inhibition of the transporter, and some substrates can competitively inhibit other substrates (Matsson & Karlgren 2021).

2.2.9.1 Substrates and Inhibitors

The molecular charge is the defining characteristic of the facilitated substrates and a smaller fraction of inhibitors. Typically, OCT at physiological pH reaches ligands with a positive charge, but on the other side, in the same conditions, OAT carries ligands with a negative charge. This suggests the alternation of substrate and inhibitor affinity by modification of charged functional groups.

Moving to the molecular size of substrates, both OAT and OCT seem relatively small, with molecular weight rarely reaching 400 Da, which is abhorrent with the overall chemical space of the registered drugs (Fernandez-Fernandez et al. 2011). Some studies described the negative effect of bulky parts near charged groups on substrate translocation. In the case of OCT1 and OCT2, a decrease in transporter-mediated uptake was also observed when the molecular size of the substrate increased, which was more significant for OCT1 than OCT2 (Hendrickx et al. 2013). These described properties are quite discriminating for inhibitors because bigger molecular volumes have a higher likelihood of inhibition. Nevertheless, the differences between inhibitors and non-inhibitors are small (Matsson & Karlgren 2021).

It was observed that the poor substrates for OCTs and OATs are lipophilic drugs, but there needs to be said that the increase of lipophilicity increased membrane permeability. Because of this phenomenon cannot be said clearly, if there is a lower transport activity or the transmembrane diffusion is higher, which masks

the transporter effect. Only for the transporter, OCT2 was determined that the increase of polar surface correlated with the lower uptake mediated by the transporter (Hendrickx et al. 2013), accompanied by a suchlike mechanism of higher uptake at pH 7,4. However, compound lipophilicity is a common feature of SLC22 transporter inhibition when the higher octanol-water partition, according to the structure-activity relationships (SARs) studies, increases the likelihood of inhibition for OCT1, OCT2, OAT1 and OAT3. This statement used datasets of a few hundred or even thousands of compounds in accordance with Matsson and Karlgren (2021). To distinguish non-inhibitor from real inhibitors was described other molecular features, like localized charge distribution (OCT1) or a lower number of hydrogen bond donors and acceptors for OCT1, OCT2, OAT1 and OAT3 (Matsson & Karlgren 2021).

3. Aim of the work

1. Study of the literature about SLCs.

2. Bioinformatic analysis of available structures of OCT transporters and their binding sites.

3. Molecular docking of substrates into OCT transporters through AutoDock Vina and AutoDock Vina 1.2.0.

4. Analysis of docking results and determination of potential ligand substrates for OCT1 and OCT2 transporters.

5. Determination of binding sites. Identifying the specific regions or pockets within the protein structures where ligands bind and form stable complexes.

6. Comparison of the results with a pre-prepared in vitro structure of OCT1, OCT2 and OCT3 with a bound ligand.

4. Materials and methods

4.1. Methods

4.1.1. Structure search

Solute carrier transporters from family 22 are diverse and differ by the carried ligand for two major clades Organic anoint transporters (OAT) and Organic cation transporters (OCT) (Nigam 2018). As well as the first clade being well described (Engelhart et al. 2020), this work mainly focused on the second group OCT. According to the aims of the work, there was a need to find appropriate protein models for molecular docking and research the binding sides and ligand affinity to the transporters.

Determining OCTs protein structures as a membrane protein is quite difficult, which was also why till 2022 (Khanppnavar et al. 2022), there were no possible structures of OCTs in the Protein Data Bank (PDB). Their hydrophobic regions make them hard to purify and stabilize for analysis, they have low expression levels in model cell lines (Suo et al., 2023), and their size and conformational variability complicate the formation of stable crystals or cryo-EM samples. In addition, they also depend on the native lipid environment (Schmidpeter et al., 2020).

For purposes of this work were used two new structure predicting software and one structure searched by electron microscopy from Protein Data Bank (Khanppnavar et al. 2022). Predicting software based on Artificial intelligence (AI) was used due to Röntgen crystallography's lack of searched structures in the Protein Data Bank. In the “Discussion”, will also be used the recent article from 2023 (Suo et al. 2023), which describes the OCT1 and OCT2 with and also without ligands by cryogenic electron microscopy (cryo-EM). This article was published at the end of the finalizing results so that the structures will be used just for comparison with the result of this work.

4.1.1.1. <u>AlphaFold: structure predicting AI software</u>

AlphaFold is a software developed by DeepMind company, which has a novel machine learning approach that combines bioinformatics and physical approaches with incorporating biological knowledge about the structure of the protein. Also, leveraging multi-sequence alignments into the design of the deep learning algorithm (Jumper et al. 2021). This software improves structure prediction accuracy

by incorporating novel neural networks and training procedures based on the structures' evolutionary, geometric, and physical constraints. This novel architecture also incorporates multiple sequence alignments (MSAs) and pairwise features.

AlphaFold uses physical and geometric inductive bias to build components according to the structure knowledge gained from Protein Data Bank (PDB) data and then represents an effective tool for accelerating the prediction of 3D protein structures. AlphaFold, thanks to the described features, can do accurate endo-to-end structure prediction with self-estimates of accuracy. It uses the primary amino acid sequence and aligned sequences of homologues as inputs (Jumper et al. 2021). As a result, AlphaFold obtains a new network using the open data in the PDB for very efficient learning. Moreover, it can cope with the wide variety and complexity of structural data and also handle missing the physical context for producing accurate 3D models in challenging cases (Jumper et al. 2021). In our case, we could find structures of all transporters from the OCT clade (OCT1, OCT2 and OCT3) (Figure 13).

4.1.1.2. SWISS-MODEL: Protein homology modelling server

The first fully automated protein homology modelling server was the SWISS-MODEL (https://SWISS-MODEL.expasy.org), which currently generates over 3000 models daily. Modelling functionality has been recently extended with the inclusion of the modelling of homo- and heteromeric complexes. Other features include the development of the new modelling engine, ProMod3, increasing accuracy of the models produced by SWISS-MODEL, and the improved local model quality estimation method (QMEANDisCo) based on the QMEAN, the Qualitative Model Energy Analysis (Waterhouse et al. 2018). SWISS-MODEL extended the scope of automated homology modelling based on the modelling of protein assemblies by efficiently using the information on quaternary structures from PDB. The template for the prediction served as an evolutionary-related protein sequence, which also provides extrapolating experimental information used to generate a 3D model of the targeting sequence. The whole workflow of SWISS-MODEL consists of 5 steps, Input data, Template search, Template selection, Model building and Model quality estimation. The server also provides a curated template library, SWISS-MODEL Template Library (SMTL), which is updated on a weekly basis according to the new

PDB release, available at https://SWISS-MODEL.org/templates. The database of automatically generated homology models is the SWISS-MODEL Repository which provides models for experimental structure and relevant organisms for all sequences in UniProtKB (Waterhouse et al., 2018). We obtained 6 structures through SWISS-MODEL: protein structures in outward-facing, occluded and inward-facing forms for the first two OCT proteins, OCT1 and OCT2 (Figure 11). Structures may seem similar, mostly because they were predicted on the same template proteins.

The human OCT3 (SLC22A3) structure was used as the first template. It was resolved in lipid nanodisc from cryo-EM with the resolution of 3,20 Å in an outward open conformation. The PDB identifier (PDB ID) is 7zh0 (Khanppnavar et al. 2022). According to the SWISS-MODEL, this protein structure with OCT1 at 49.81 % and OCT2 has OCT3 sequence identity at 50.28 %. The experimental structure of OCT3 was used as a template for OCT1 and OCT2 in an outward-facing state and further will be discussed as a 7zh0.

The second template protein is Di- or tripeptide H^+ symporter, the crystal structure of a POT family peptide transporter in an inward open conformation obtained by x-ray diffraction with the resolution of 3,30 Å. The PDB ID of this experimental structure is 4APS (Solcan et al., 2012). This protein structure with OCT1 9,66 % and with OCT2 has sequence identity 10,73 % according to the SWISS-MODEL. The experimental structure of Di-or tripeptide H^+ symporter was used as a template for OCT1 and OCT2 occluded state and further will be discussed as a 4APS. The appearance of being partially occluded selected this structure, but the next research shows the same abilities as the inward-facing structure, which was simulated according to the next template protein.

The third template structure is Di- or tripeptide H^+ symporter, the inward open PepTSt from Streptococcus thermophilus crystallized in space group P3121 obtained by x-ray diffraction with the resolution of 3,40 Å. The PDB ID of this experimental structure is 5MMT (Quistgaard et al. 2017). This protein structure with OCT1 9,97 % and with OCT2 has sequence identity 10,90 % according to the SWISS-MODEL. The experimental structure of Di- or tripeptide H^+ symporter was used as a template for OCT1 and OCT2 inward-facing state and further will be discussed as a 5MMT.

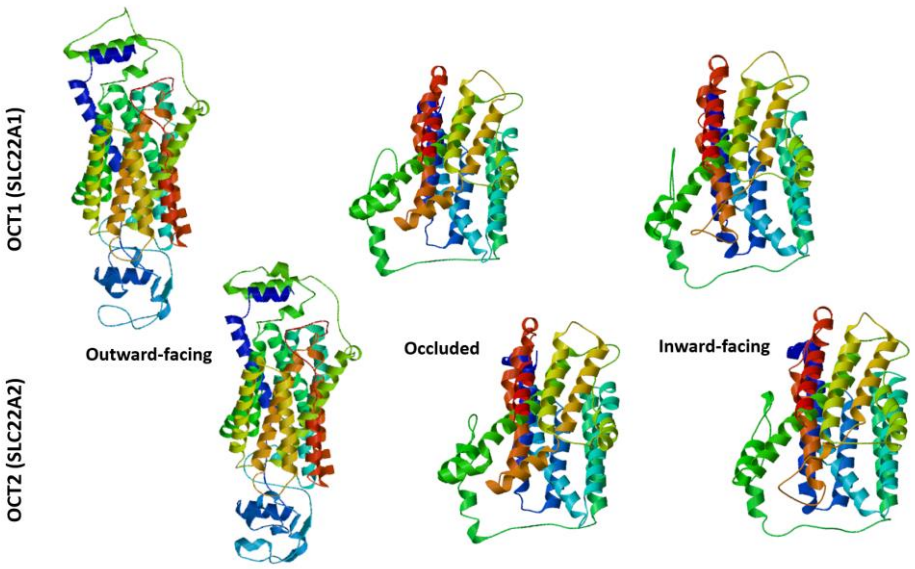


Figure 11: Structures of OCT1 (on the top) and OCT2 (on the bottom) done from automated protein homology modelling from SWISS-MODEL. On the right side can be seen structures in the outward-facing conformation, in the middle structures in occluded conformation and on the left structures in the outward-facing conformation. Helices were coloured using rainbow colours from N- to C-transmembrane segment TM. Shades from blue to cyan are closer to the N-TM, while red to yellow is closer to the C-TM (taken from SWISS-MODEL 2023).

4.1.1.3. RCSB PDB

The US Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) is a database many researchers worldwide use to deliver experimentally-determined 3D structures of biomolecules (Rose et al. 2021). This database found one experimental structure of OCT3 (SLC22A3) from the article published in 2022 (Khanppnavar et al. 2022). This structure was described by cryo-EM at 3.2 Å resolution. More specifically, it is speaking about OCT3 from *Homo sapiens.* The structure is available in three forms: one empty outward-facing structure in a lipid nanodisc, the second one with a ligand in a binding pocket: corticosterone and the third one with a ligand in a binding pocket: decynium-22 (Khanppnavar et al. 2022). The PDB ID is 7zh0 (in lipid nanodiscs), 7zh6 (with inhibitor corticosterone) and 7zha (with inhibitor Decynium-22). All three available structures are used in this work (Figure 12).

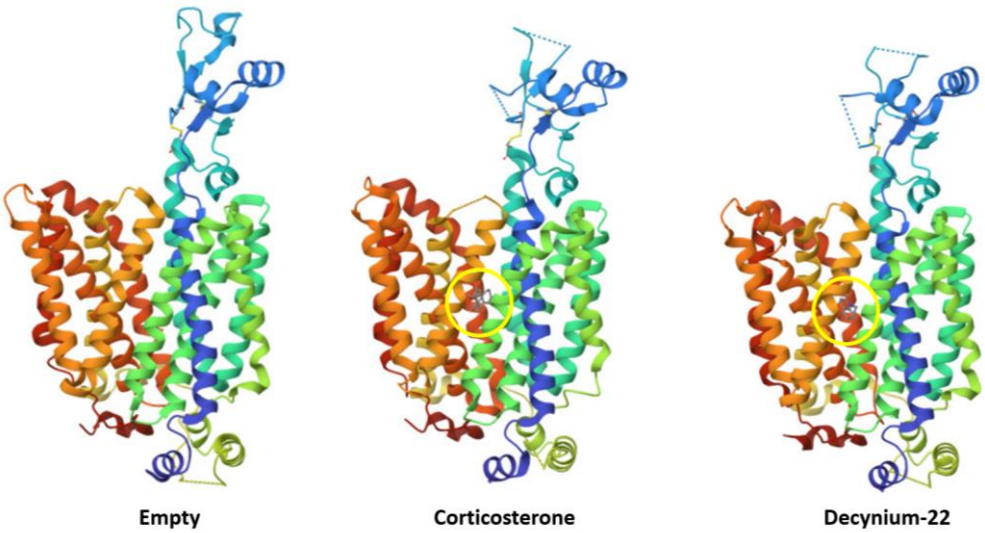


Figure 12: OCT3 structure from cryo-EM in outward-facing conformation. Three structures were obtained from Protein Data Bank; the first one from the left is displayed in lipid nanodisc with an empty binding pocket, the middle structure has bounded corticosterone, and the structure on the right has bounded decynium-22. Helices were coloured using rainbow colours from N- to C- transmembrane segment TM. Shades from blue to cyan are closer to the N- TM, while red to yellow is closer to the C- TM (taken from PDB 2023).

4.1.2. Ligand search

Search for suitable ligands for proteins from the OCT clade was done in two ways. Firstly, we used the Molecules on Membranes Database (MolMeDB) (Juračka et al. 2019) and then the list of described ligands from *in vitro* studies (Koepsell et al. 2007b).

4.1.2.1. Molecule on Membranes Database: MolMeDB

Molecule on Membranes Database (MolMeDB) gathers data about compound-membrane interactions (Juračka et al. 2019). In 2019 the number of interactions was over 3600; nowadays, it is an order of magnitude more - these interactions including partitioning, penetration and positioning. MolMeDB collects data from scientific articles and is complemented by in-house calculations from a high-throughput COSMOmic approach (Juračka et al. 2019). This database found for OCT1 172 interaction, from which seven suitable ligands and possible substrates or inhibitors were picked. They are tetrylammonium (TEA), azidoprocainamide methiodide, ganciclovir, acyclovir, ipratropium, dinoprostone and dinoprost tromethamine. For

the second protein, OCT2, it was found 78 interactions. In the case of OCT2, we picked 12 ligands, also possible substrates or inhibitors. They are nicotinyl methylamide, memantine, TEA, cimetidine, choline, norepinephrine, agmatine, serotonin, dopamine, quinine, dinoprostone and dinoprost tromethamine.

4.1.2.2. Ligands from in vitro studies

The possible ligands from *in vitro* studies were collected in an expert review from 2007 (Koepsell et al. 2007b). This review describes ligands from nine categories, namely: Metabolites, Neurotransmitters, Hormones, Miscellaneous, Receptor antagonists, Receptor agonists, Ion channel and transporter blockers, Drugs and xenobiotics and other compounds. Also is good to mention that OCT1, OCT2 and OCT3 translocate a variety of organic cations with widely differing molecular structures. Moreover, they are inhibited by many additional compounds, but these cannot be transported by OCTs (Koepsell et al. 2007b). Together were determined from all group 98 ligands which were used for purposes of this work. All the tables with the compounds can be found in the attachments (Attachment 1- 4).

4.1.2.3 ChEMBL Database

ChEMBL (https://www.ebi.ac.uk/chembl) is an open database, first described in the article from 2012 (Gaulton et al. 2012), when ChEMBL contains 5200 protein targets and more than 1 million compounds, 5,4 million bioactivity measurements (Gaulton et al. 2012). According to the article from the 2019 database contains 1,8 million distinct compounds over 15 million bioactivity measurements (Mendez et al. 2019). Nowadays, it can be even more. In the last couple of years, the ChEMBL database has undergone many changes and prompted many innovations, like enhanced search and filtering capabilities and a new data deposition system, which allows updating data sets and others (Mendez et al. 2019). Through this database can be found ligands with described constants like the Michaelis-Menten constant (K_m) and constant of inhibition (K_i), which was used for the aims of this work for comparison and possible correction of previous dataset constants. A list of ChEMBL ligands for OCT1 and OCT2 can be found in the attachments (Attachment 5, Attachment 6)

4.1.2.4 UniProt Database

UniProt is an open database containing protein sequences annotated with functional information (UniProt Consortium 2021), which are helpful in biological and biomedical research. UniProt is also extracting detail annotations about protein sequences from the literature. Automated systems provide annotations. This database found ligands described in articles, such as substrates for the proteins OCT1 and OCT2. Thanks to this information, it was possible to describe amino acid residues of the binding pocket of the substrates.

4.1.3. Method of molecular docking

The molecular docking method is highly used in virtual screening and provides information about the interaction of two or more molecular structures, such as a drug with protein, and how they fit together (Kirkpatrick 2004; Roy et al. 2015). Molecular docking predicts the behaviour of ligands in the protein's binding pocket when it aims to identify the correct poses of ligands in the binding pocket. The main output from molecular docking is the affinity prediction between the protein and the ligand.

This work will focus on protein-ligand docking, a simpler end of the complexity spectrum. Many available programs perform molecular docking. In this work was used AutoDock Vina and AutoDock Vina 1.2.0.

4.1.3.1. AutoDock Vina

AutoDock Vina is a program used for molecular docking and virtual screening, which uses a sophisticated gradient optimization method in its local optimization procedure (Trott & Olson 2010). Also, it is the fastest and most widely used open-source program for molecular docking (Eberhardt et al. 2021). The software is available from http://vina.scripps.edu. It achieves approximately two orders of magnitude speed-up compared to the older molecular docking software. It also has improved the accuracy of the binding mode predictions (Trott & Olson 2010). For purposes of molecular docking, this work used docking protocol from open source: https://github.com/DweipayanG/Multiple_Ligand_Docking_Vinna. The code is attached underneath.

```
#!/usr/bin/perl
print"Ligand_file:\t";
$ligfile=<STDIN>;
chomp $ligfile;
open (FH,$ligfile)||die "Cannot open file\n";
@arr_file=<FH>;
for($i=0;$i<@arr_file;$i++)
{
 print"@arr_file[$i]\n";
 @name=split(/\./,@arr_file[$i]);
}
for($i=0;$i<@arr_file;$i++)
{
 chomp @arr_file[$i];
 print"@arr_file[$i]\n";
 system("vina --config conf.txt --ligand @arr_file[$i] --
log @arr_file[$i]_log.log");
}
```

Condition of the first docking done by AutoDock Vina was the following: Cube shape grid box of the side of 30 Å, number of poses to writing (num_modes) was 10 and maximum energy difference from best pose (energy_range) was 4 kcal/mol. The grid box was also tested in sizes 20 Å and 15 Å.

4.1.3.2. AutoDock Vina 1.2.0

AutoDock Vina 1.2.0 is a new upgraded open-source program for molecular docking under the AutoDock Suite package. Version 1.2.0 facilitates the design and execution of simple and complex docking simulations and provides Python binding, enabling scripting for virtual screening and other advanced applications (Eberhardt et al. 2021). This work compares in "Results" output of molecular docking done by AutoDock Vina and AutoDock Vina 1.2.0. The molecular docking by 1.2.0 version was done in the Department of Physical Chemistry, Faculty of Science, Palacky University Olomouc by Mgr. Ing. Václav Bazgier, Ph.D.

4.1.4. Description of the binding site of the OCTs proteins

4.1.4.1. Schrödinger Maestro 13.6

Program Maestro (Schrödinger, Inc., USA) is a molecular modelling and structural activity relationship research tool. Maestro is used for interpreting, managing, and sharing the results of computational experiments, able to manage obtained data in many ways when computed results are returned automatically and incorporated into projects for the next study (Jamkhande et al. 2017). Maestro offers insight into molecular properties and studies intermolecular interaction in more detail. Besides providing tools for building and visualising molecular models, Maestro offers a "Ligand interaction diagram" (Jamkhande et al. 2017). This tool was used to study binding pockets and similarities in amino acid residues, which coordinate ligands in binding pockets. The amino acid residues were studied at distances up to 5 Å near the ligands.

4.2. Software for results processing

PyMOL (Schrödinger, Inc., USA)
Microsoft Excel (MS Office, USA)
ChatGPT (OpenAI, USA)
DeepL (DeepL SE, Germany)
Grammarly (Grammarly, Inc., USA)

5. Results

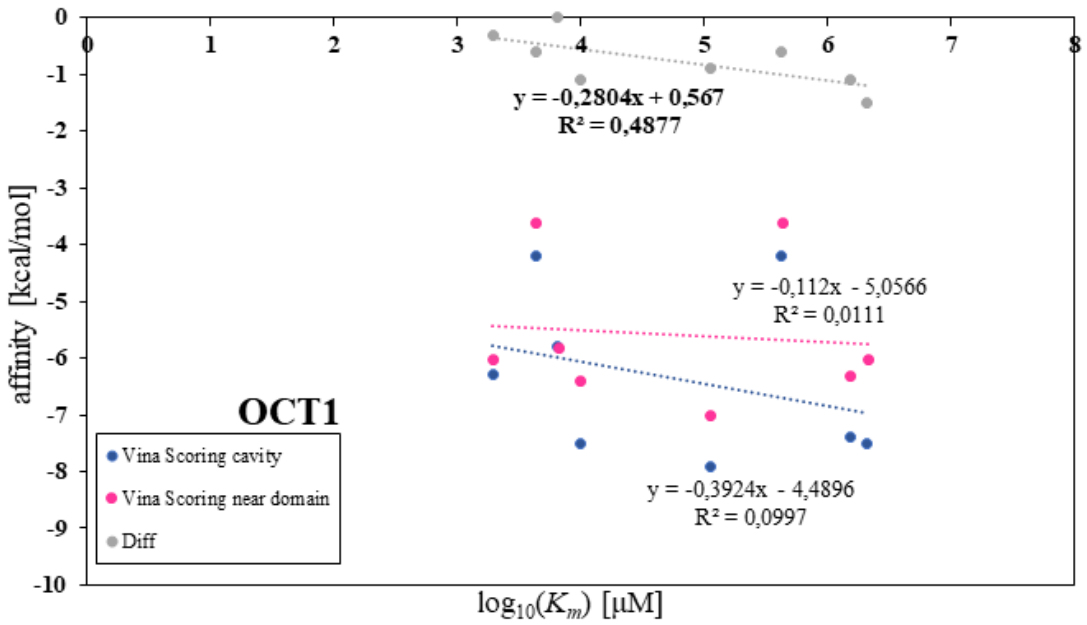
5.1. MolMeDB ligands docking

The first ligand docking through AutoDock Vina was done with protein structures from AlphaFold, specifically structures of OCT1 and OCT2. In the case of the first docking, there were exact constants of Michaelis-Menten (K_m) or constants of inhibition (K_i) from various articles collected by MolMeDB. Because of the lack of structures in inward, occluded, and outward states, the first docking was done in a potential binding cavity and outside of it. The docking score, which represents the estimated stability of the complex, can be related to Gibbs free energy (ΔG) of binding, where a more negative docking score implies a more stable and energetically favourable binding pose (Du et al. 2016). By investigating ligand docking outside the binding cavity, potential alternative binding sites or off-target interactions can be discovered (Owoloye et al., 2022). The difference between binding energies in a cavity and outside of the cavity correlates with K_m or K_i through the equation:

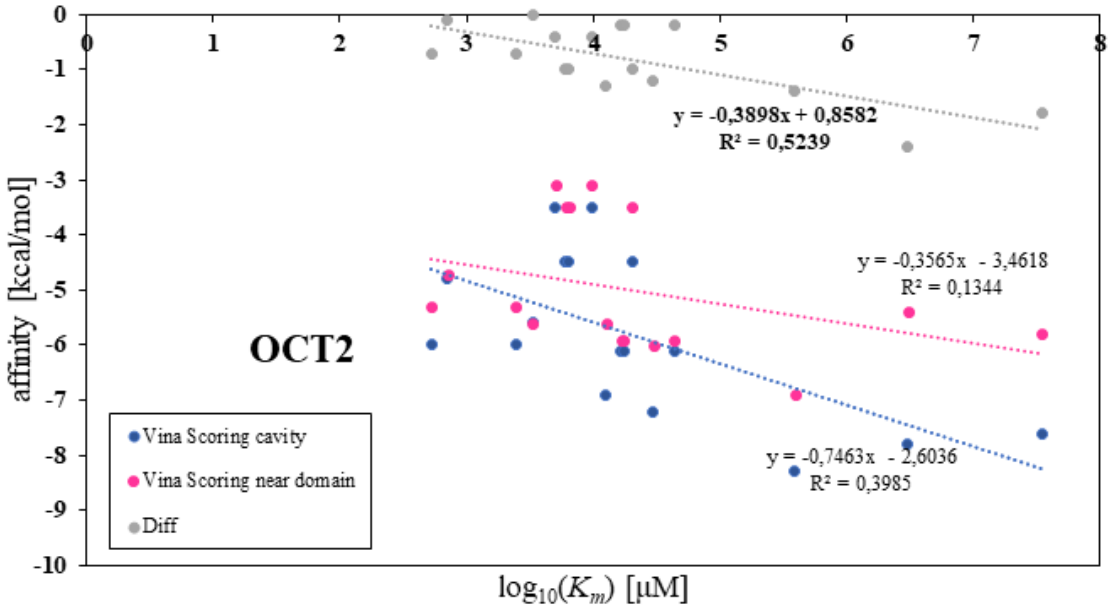
$$dG_0 = -RTlnK \quad (1)$$

Here $dG°$ is a standard binding free energy, R is the universal gas constant (1,987 cal·K−1·mol−1), K is a binding constant, and T is the temperature (Du et al. 2016). From equation (1), it becomes evident that a higher binding constant K corresponds to a more negative standard free energy of binding. In this case, K stands for K_m and K_i. Michaelis-Menten constant describes the state when the chemical reaction rate reaches half its maximum value (V_{max}). It indicates an enzyme's efficiency when a smaller Km value suggests higher enzyme activity (Srinivasan, 2022). K_i stands for the constant of inhibition, which describes the strength of an inhibitor in enzyme inhibition, when a smaller K_i value indicates a stronger inhibitor, as it requires a lower inhibitor concentration to effectively inhibit the enzyme's activity and vice versa (Yung-Chi 1973). The article from 2013 (Koepsell 2013) also used IC50 for ligand description and is defined as a concentration of drug required for 50% inhibition (Swinney 2011).

Due to the equation (1), the K_m (or K_i, or $IC50$) values were logarithmized by the decadic logarithm. Thus, their values will be reported as pK_m or log10(Km/IC50) in the results. Results are summarized in attachments (Attachments 7 and 8) and Graphs 1 and 2 for proteins OCT1 and OCT2.



Graph 1: Results of docking into OCT1 protein from AlphaFold. Blue dots represent the binding energy (affinity) of ligand in the protein binding cavity; magenta dots represent binding energy (affinity) near potential extra or intracellular domain, and grey dots show the difference (Diff) of these two values.



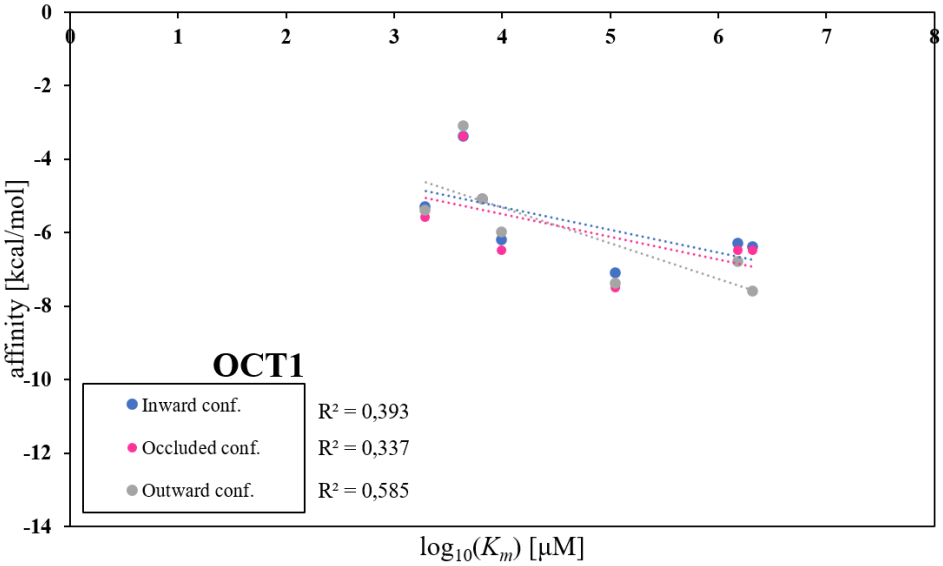
Graph 2: Results of docking into OCT2 protein from AlphaFold. Blue dots represent the binding energy (affinity) of ligand in the protein binding cavity, and magenta dots represent binding energy (affinity) near potential extra or intra-cellular domain and grey dots difference (Diff) of these two values.

Since the ligands are widely different in size, we use normalization. Since larger compounds, in general, binds stronger than small ligands due to larger interaction areas, we normalize their binding affinity by the comparison to their binding to the surface of the protein to mimic how such compound would generally bind to any protein surface. Hence, we also report Vina binding energy differences to establish how the compound would bind specifically to the OCTs' cavity.

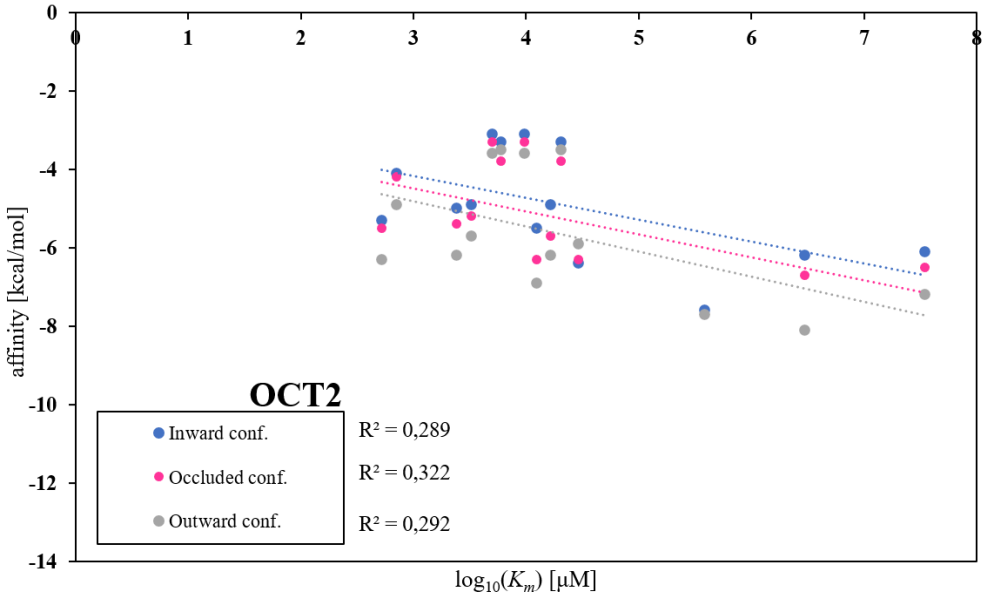
The results from the first docking show the dependence of the difference of binding affinities on pK_m, according to the theory described in equation (1). Increasing the logarithmic value of K_m is the same as lowering binding energy. According to Chicco et al. (Chicco et al. 2021), the coefficient of determination (R^2) is interpreted as the proportion of the variance in the dependent variable that is predictable from the independent variables. Values range from 0 to 1 when values closer to 1 indicate a better fit, while values closer to 0 suggest a weaker fit in the case of linear regression (Chicco et al. 2021). The R^2 of difference of docking values in and out cavity is around 0,5, which is interesting as it is higher than each of the individual docking runs showing promise for selectivity filtering.

According to the comparison of binding energies and pK_m or pK_i, the frontier between ligands which are substrates and those which are not substrates has been established on value ≥ -1,5 kcal/mol in case of binding energy (affinity) and on the $pK_m \geq 5$ in case of pK_m/pK_i. This observation was tested on a bigger dataset, so it was picked from an article from 2013 (Koepsell, 2013). Ligands from MolMeDB were also used in other OCT1 and OCT2 protein models, specifically from SWISS-MODEL. Here were available to dock into all three predicted stages of proteins, in the inward, occluded, and outward state. The results are described in graphs 3 and 4 (Graph 3 and Graph 4) and attachments (Attachment 9 and 10).

The results from docking into the proteins models of OCT1 and OCT2, from SWISS-MODEL, show a weak fit to linear regression with the coefficient of determination in general under 0,4. Only in the case of outward-facing conformation of OCT1 was higher R2 detected, around 0,6. Possible reasons for this phenomenon will be discussed in part "Discussion".



Graph 3: Results of docking into OCT1 protein from SWISS-MODEL. Blue dots represent the binding energy (affinity) of ligands in protein in the inward state according to template 4ASP, and magenta dots represent the binding energy of ligands in protein in the occluded state according to template 5MMT and grey dots difference (Diff) of ligand in protein in the outward state according to the template 7ZH0.



Graph 4: Results of docking into OCT2 protein from SWISS-MODEL. Blue dots represent the binding energy (affinity) of ligands in protein in the inward state according to template 4ASP, and magenta dots represent the binding energy (affinity) of ligands in protein in the occluded state according to template 5MMT and grey dots difference (Diff) of ligand in protein in the outward state according to the template 7ZH0.

5.1.1. Bigger dataset of ligands – second ligand docking

Second docking was done with a larger dataset consisting of ~ one hundred ligands, which can be divided into categories of "Ion channel and transporter blockers" (7 ligands), "Drugs and Xenobiotics" (46 ligands), "Metabolites" (8 ligands), "Neurotransmitters" (13 ligands), "Miscellaneous" (1 ligand), "Receptor antagonists" (14 ligands) and "Receptor agonists" (5 ligands). The article described these ligands from 2013 (Koepsell 2013) with known *Km* or *IC50,* which were also logarithmized. Docking of these ligands can be divided into docking done by AutoDock Vina with the usage of protein models predicted by AlphaFold and SWISS-MODEL and docking done by AutoDock 1.2.0 with the usage of only protein model predicted by AlphaFold done by Mgr. Ing. Václav Bazgier, Ph.D.

5.1.2. AutoDock Vina, protein-ligand docking

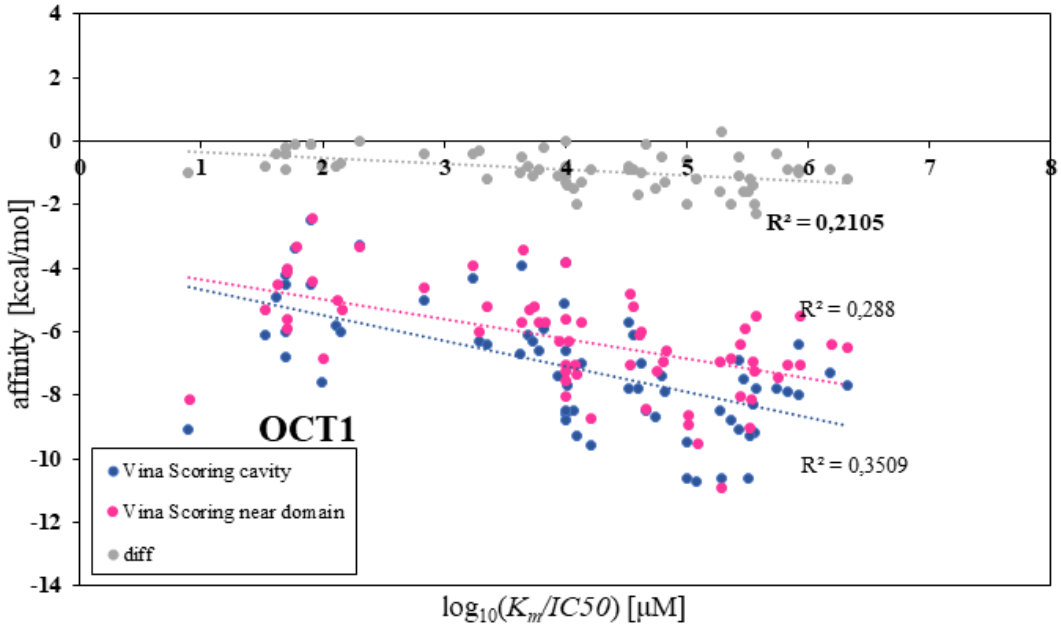
First, protein-ligand docking was done with protein models predicted by AlphaFold, with the default configuration the same as in docking using ligands found by MolMeDB. Results are visible on graphs 5 and 6 (Graph 5, Graph 6).

According to this result, the model does not explain any variability of the dependent variable around its mean value, with R^2 values under 0,3, in the case of OCT2 around 0 for the difference of binding affinities in and outside the cavity. Because of this data divergence, we tried several further approaches.

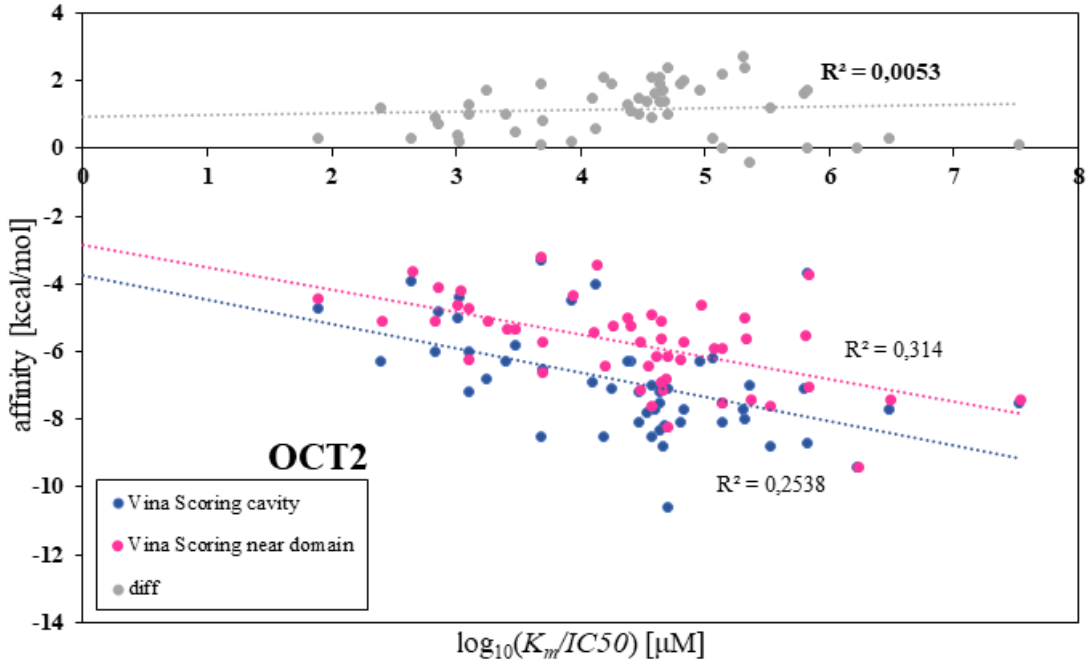
We tried to shrink the grid box, mainly because of the different sizes of the ligands. The grid box was shrunk into a cube with a side length of 20 Å and then to the 15 Å. No significant change was detected, and R^2 was still under 0,3. Subsequently, docking of ligands into the protein models created using SWISS-MODEL was performed. The results of this docking can be seen in Graphs 7 and 8.

The results from docking ligands from a bigger dataset into the protein models predicted by SWISS-MODEL show the same weak fit in both proteins with R2 under 30 % (0,3) in every state of proteins.

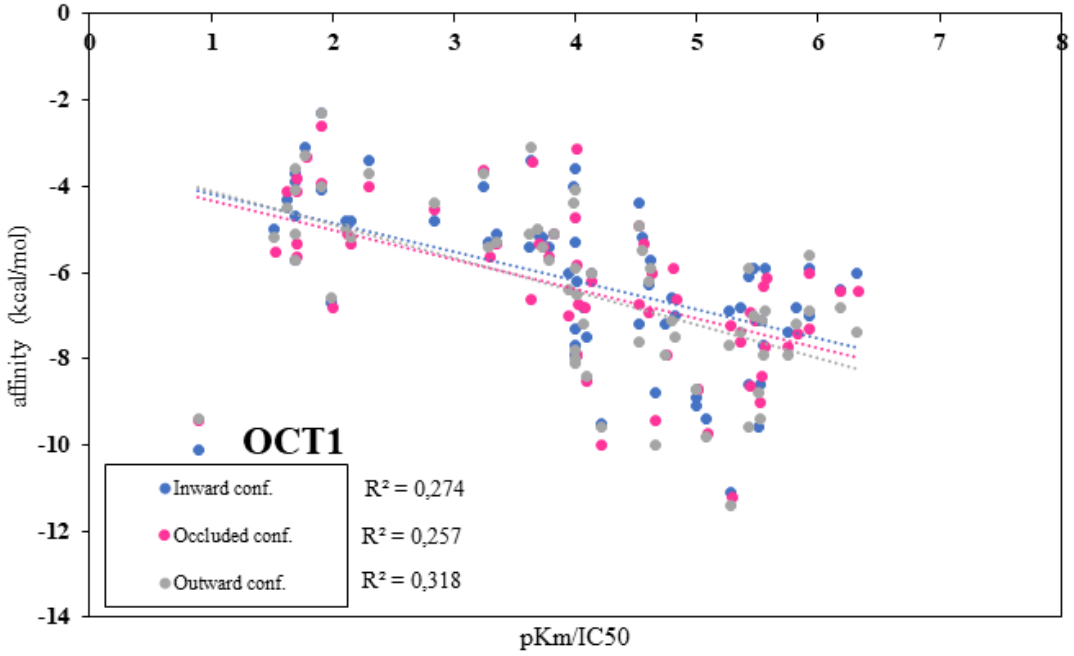
A possible explanation for why the original positive results from a smaller dataset could not be verified on a larger dataset will be discussed in the "Discussion" section.



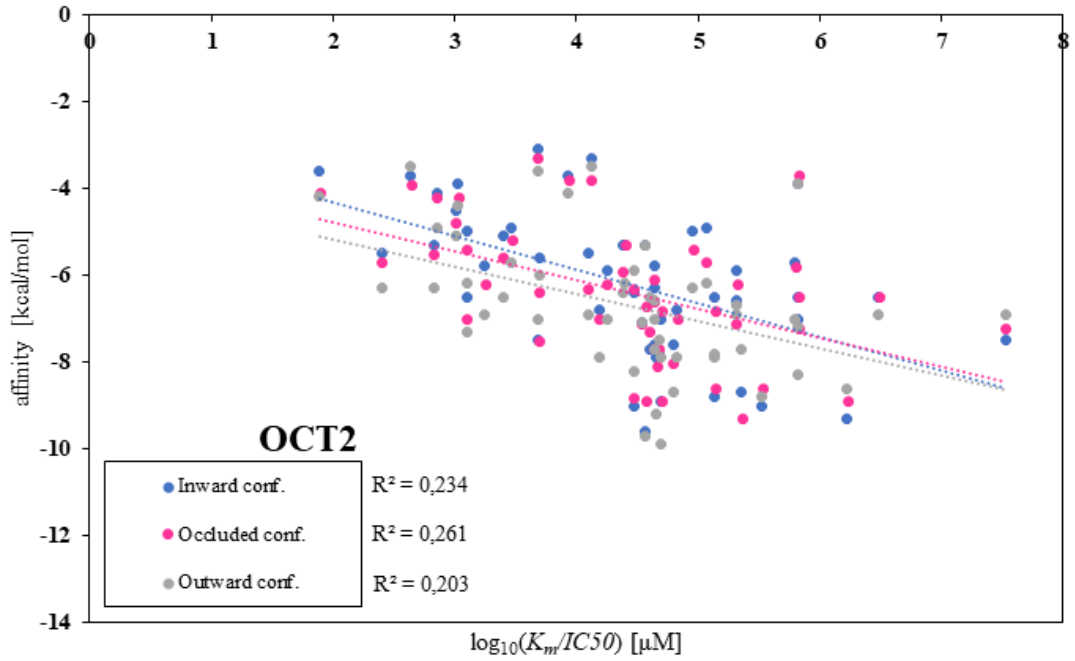
Graph 5: Results of docking ligand from article Koepsell (2013) into OCT1 protein from AlphaFold. Blue dots represent the binding energy (affinity) of ligands in the protein binding cavity, and magenta dots represent binding energy near potential extra or intra-cellular domain and grey dots difference (Diff) of these two values.



Graph 6: Results of docking ligand from article Koepsell (2013) into OCT2 protein from AlphaFold. Blue dots represent the binding energy (affinity) of ligands in the protein binding cavity, and magenta dots represent binding energy near potential extra or intra-cellular domain and grey dots difference (Diff) of these two values.



Graph 7: Results of docking ligands from a bigger dataset into OCT1 protein from SWISS-MODEL. Blue dots represent the binding energy (affinity) of ligands in protein in the inward state, magenta dots in the occluded state, and grey dots in the outward state.



Graph 8: Results of docking ligands from a bigger dataset into OCT2 protein from SWISS-MODEL. Blue dots represent the binding energy (affinity) of ligands in protein in the inward-facing state, magenta dots in the occluded state and grey dots in the outward-facing state.

5.1.3. AutoDock Vina 1.2.0, protein-ligand docking

This work was performed docking using the new AutoDock Vina version with the same default configuration as in the previous cases. This type of docking was performed only with protein models predicted by AlphaFold. This type of docking was not performed also with the SWISS-MODEL protein models was first to try a new version of AutoDock Vina, if there would be any significant improvement in the fit of linear regression. The result shows little difference; however, binding near the domain shows almost zero correlation with the binding energy; hence, the difference normalisation approach is unnecessary (Graph 9, Graph 10).

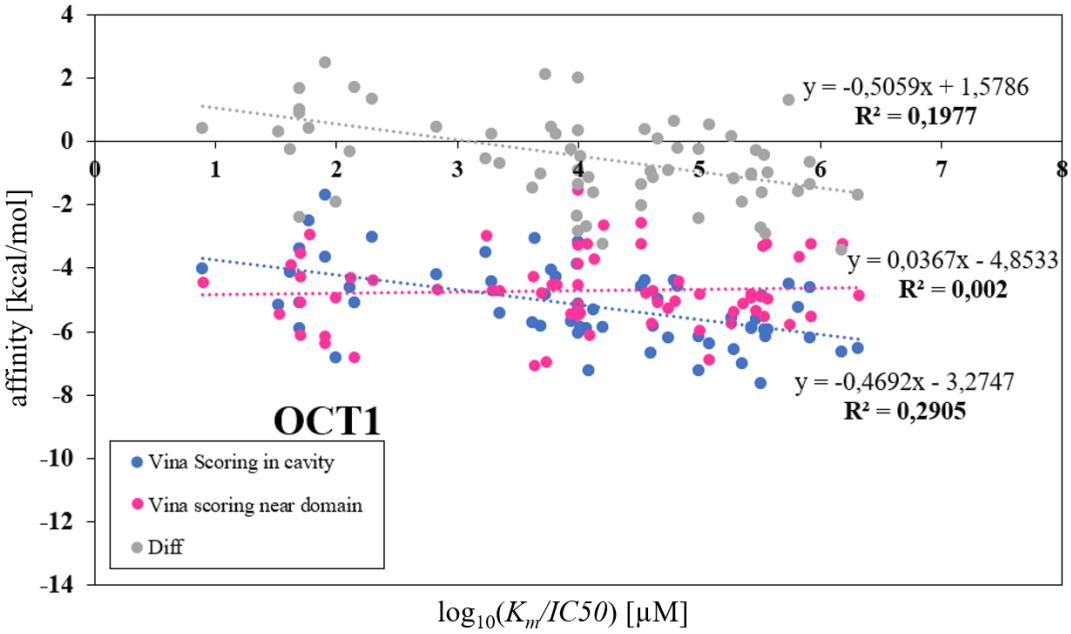
Another approach was to sort the dataset into individual parts according to the ligand type and grid boxes (30 Å, 20 Å,15 Å). This sorting and its results can be found in the following tables (Table 1, Table 2).

Table 1: Groups of ligands with the best R^2 at different grid box sizes for OCT1 ligands. Ligands came from the article (Koepsell 2013). Docking was performed by AutoDock Vina.

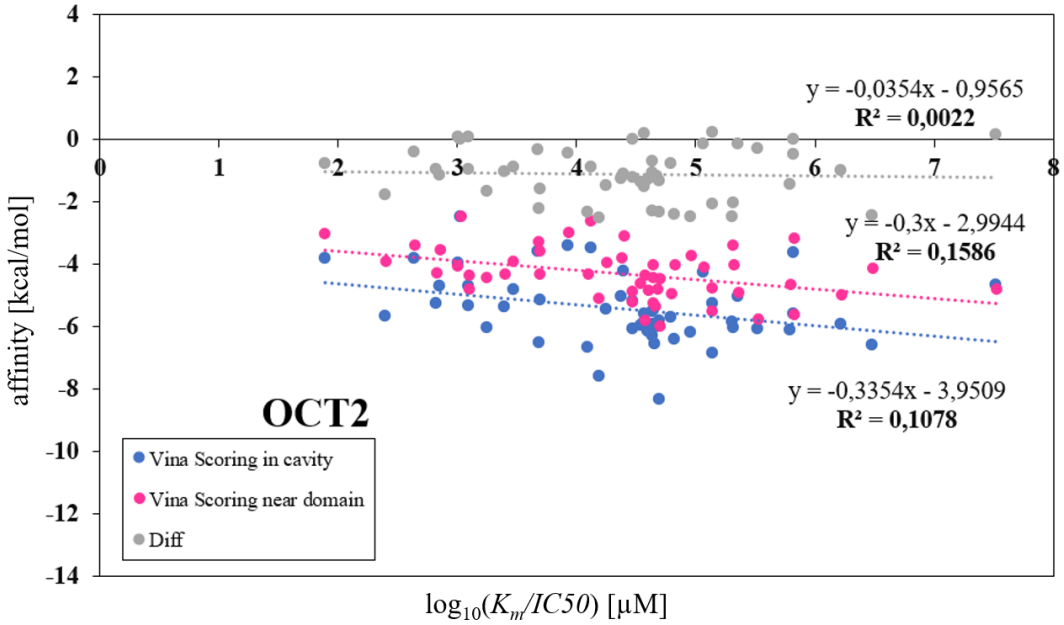
OCT1				
Ligand type	#	R^2 of Diff in 30 Å	R^2 of Diff in 20 Å	R^2 of Diff in 15 Å
Ion channel and transporter blockers	7	0,22	**0,45**	0,30
Metabolites	4	0,56	**0,96**	**0,96**
Neurotransmitters	5	0,05	0,01	**0,93**

Table 2: Groups of ligands with the best R^2 at different grid box sizes for OCT2 ligands. Ligands came from the article (Koepsell 2013). Docking was performed by AutoDock Vina.

OCT2				
Ligand type	#	R^2 of Diff in 30 Å	R^2 of Diff in 20 Å	R^2 of Diff in 15 Å
Ion channel and transporter blockers	3	0,67	**0,99**	0,78
Metabolites	3	**0,82**	**0,82**	0,40
Compounds	5	**0,86**	0,01	0,39
Neurotransmitters	5	**0,71**	0,65	0,67



Graph 9: Results of docking ligand from article Koepsell (2013) into OCT1 protein from AlphaFold using AutoDock Vina 1.2.0. Blue dots represent the binding energy (affinity) of ligands in the protein binding cavity, and magenta dots represent binding energy (affinity) near potential extra or intra-cellular domain and grey dots difference (Diff) of these two values.



Graph 10: Results of docking ligand from article Koepsell (2013) into OCT2 protein from AlphaFold using AutoDock Vina 1.2.0. Blue dots represent the binding energy (affinity) of ligands in the protein binding cavity, and magenta dots represent binding energy (affinity) near potential extra or intra-cellular domain and grey dots difference (Diff) of these two values.

According to the results after ligand partitioning, 3 groups of compounds came out for OCT1 protein and 4 groups for OCT2 protein of compounds in the same category with similar size and approximately similar molecular weight. However, their number is small; even ten ligands represent neither category. Therefore, no definitive statement can be drawn from this alignment, and it would be necessary to augment the categories with additional ligands. The reason for such a small number, even though there were more ligands in the original set for a given category, is that not all the ligands had known *Km* or *IC50* for a given protein, so the categories were greatly reduced. At first, one can notice differences between the two proteins regarding the R^2 values for each grid box size and between the groups themselves. For the "Ion channel and transporter blockers" group, the best R^2 values are offered by a grid box with a side size of 20 Å for both proteins; for the "Metabolites" group, the same R^2 values can be observed for grid boxes with a side length of 20 Å and 15 Å for OCT1, for OCT2 the grid boxes are 30 Å and 20 Å. For the "Neurotransmitters" group, which is represented by 5 ligands for both proteins, the best R^2 is for grid box 15 Å (OCT1) and grid box 30 Å (OCT2). The "Compounds" group shows the best R^2 values for grid box 30 Å. Thus, it can be concluded that for the following ligands from the given groups, the most favourable grid box sizes are 20 Å to 30 Å, but this statement cannot be taken as definitive.

5.2. Analysis of binding sides

Another aim of the work was to analyse the binding sites of OCT1 and OCT2 proteins. For this analysis, the natural ligands of the proteins were selected using the ChEMBL database. Thus, 30 natural ligands were found for the OCT1 protein and 22 for OCT2 with known Km and Ki. Using the Maestro Schrodinger 13.6 program, the binding amino acid residues forming the binding pocket were analysed. For the analysis, protein models were predicted by SWISS-MODEL because the protein model was inward-facing and outward-facing for better comparison, and amino acid residues play a role in the OCT1 and OCT2 transport mechanism.

5.2.1 Binding pocket of OCT1

The OCT1 protein had 30 natural ligands found by ChEMBL (Attachment 5 and 6). Mostly repeating amino acids residues (AMK) between these ligands were for outward-facing conformation **Ile**446 (present in all ligands binding pockets) and **Tyr**361 (in 29 cases from 30) (example of binding pocket on at Figure 16), another AMK followed are **Cys**450 (21 cases from 30), **Phe**244 (28 cases from 30), **Cys**473 (28 cases from 30), **Trp**354 (27 cases from 30), **Ser**358 (23 cases from 30), **Gln**241 (22 cases from 30), **Asp**357 (22 cases from 30), **Cys**36 (22 cases from 30), **Phe**32 (18 cases from 30), **Trp**217 (17 cases from 30), with other AMKs under 15 cases from 30. The inward-facing state binding site contains **Ser**213 (24 cases from 30), **Ser**163 (22 cases from 30), **Trp**217 (22 cases from 30), **Phe**159 (20 cases from 30), **Phe**244 (22 cases from 30), **Leu**160 (17 cases from 30) with other AMKs under 15 cases from 30. **Tyr**279 and **Tyr**278 are noticed in 9 and 10 cases from 30. Only Tyr279 or Tyr278 from the AMKs were present in some ligands. Other AMKs interacting with those ligands were specific for them. Ligands with this specificity are Noroxycodone, Sumatriptan, N-Methyl-Quinidine, Prostaglandin F2alpha, Corticosterone and Clonidine.

Diversity in the representation of amino acid residues can be observed in the case of the "inward" conformation, in contrast to the outward conformation, in which two of the same amino acid residues are represented for each ligand, as described above. The OCT1 protein model in the inward conformation predicted by the SWISS-MODEL was based on PepTso from the POT family, whereas the outward conformation was based on OCT3 from the SLC22 family and is more closely related to the original OCT1, a fact that may weigh in the final evaluation of the binding pocket. Different binding sites for certain groups of ligands can also be considered; however, once

divided into groups, the distribution of amino acid residues that bind the ligands is completely random and shows statistically insignificant and only small matches. An example of the binding site for the OCT1 on Figure 13.

5.2.2. Binding pocket of OCT2

The binding pocket of the OCT2 protein model predicted by SWISS-MODEL shows different AMKs representation, besides the remarkable similarity in the sequence homology with OCT1, which is around 80 %, and with identity around 60 % (Gründemann et al. 1997). OCT2 had 22 natural ligands found by ChEMBL (Attachment 5 and 6). The binding pocket analysis shows AMKs **Tyr**447 and **Thr**444 for nearly all the binding pockets of the ligands in outward conformation. There are also some exceptions. In the dataset for OCT2 ligands in outward conformation can be found two anomalies, a couple of Histamine and Choline and a couple of Amantadine and Memantine have very similar AMKs in their binding pockets and cannot be found in other ligands binding pockets. The Histamine and Choline share in their binding pockets amino acid residues represented by **Phe**45, **Gly**44, **Leu**43, **Phe**42, **Ala**256, **Tyr**377, **Pro**260, **Trp**262, **Ser**133, and **Ser**134, which are specific for these two ligands. The similarity in amino acid residues in binding sites also had Amatandine and Memantine. **Val**255, **Leu**252, **Val**251, **Pro**35, **Thr**32, **Ala**31, **Pro**272, **Val**269, and **Thr**268 represent the binding sites of those ligands. These AMKs can be found just in cases of Memantine and Amatandine.

Amantadine and memantine are similar in geometry and with the resence of specific functional groups, both of them are antagonists at the NMDA (N-methyl-D-aspartate) receptors in the brain, where OCT2 is expressed in small amounts (Engelhart et al. 2020; Kornhuber et al. 1994). On the other hand, choline and histamine share no similarities. Besides the fact that both act as neurotransmitters, they differ in chemical nature, synthesis and function (Engelhart et al. 2020; Maintz & Novak 2007; Zeisel & Caudill 2010). The presence of specifical AMKs in binding sites of those ligands can also point to new possibilities of specific inhibition.

Continuing with the description of the binding site in inward conformation for the protein model of OCT2, the mostly representing amino acid residues are **Tyr**245 and **Leu**180 (11 cases from 22), and other AMKs are under 8 cases from 22. The analysis of the binding site in inward conformation shows a more random presence of amino acid residues in OCT2 than in OCT1. This can also be a problem of

the protein model itself, when the PepTso from Pot family was not the best choice to be a model for the OCT group but was the most similar among the other choices in SWISS-MODEL. The representation of the binding pocket in inward and outward conformation can be found in Figure 14.

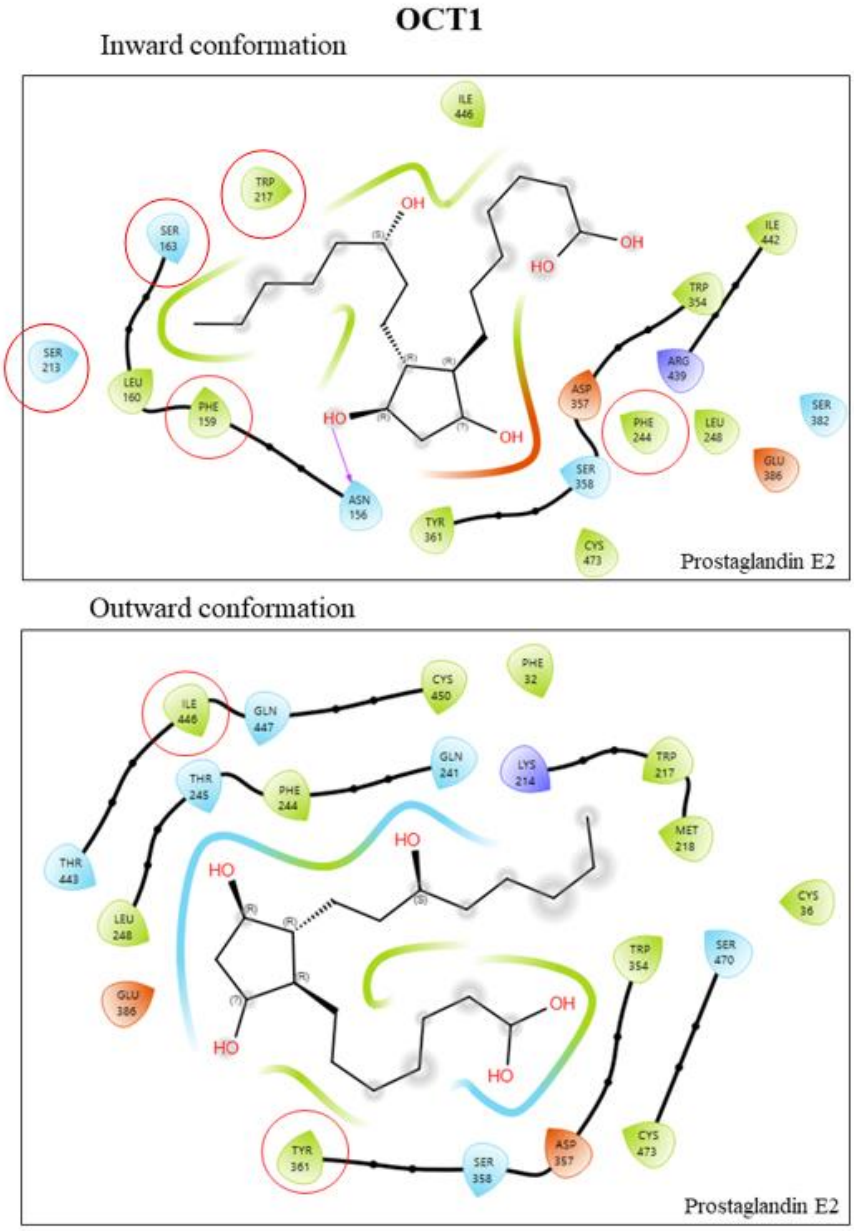


Figure 13: Example of a binding site for the OCT1 protein created using Schrodinger Maestro 13.6. Prostaglandin E2 is shown as the model ligand in the binding site of the model protein in the inward-facing conformation model based on the 4ASP template and in the outward-facing conformation model based on the 7ZH0.

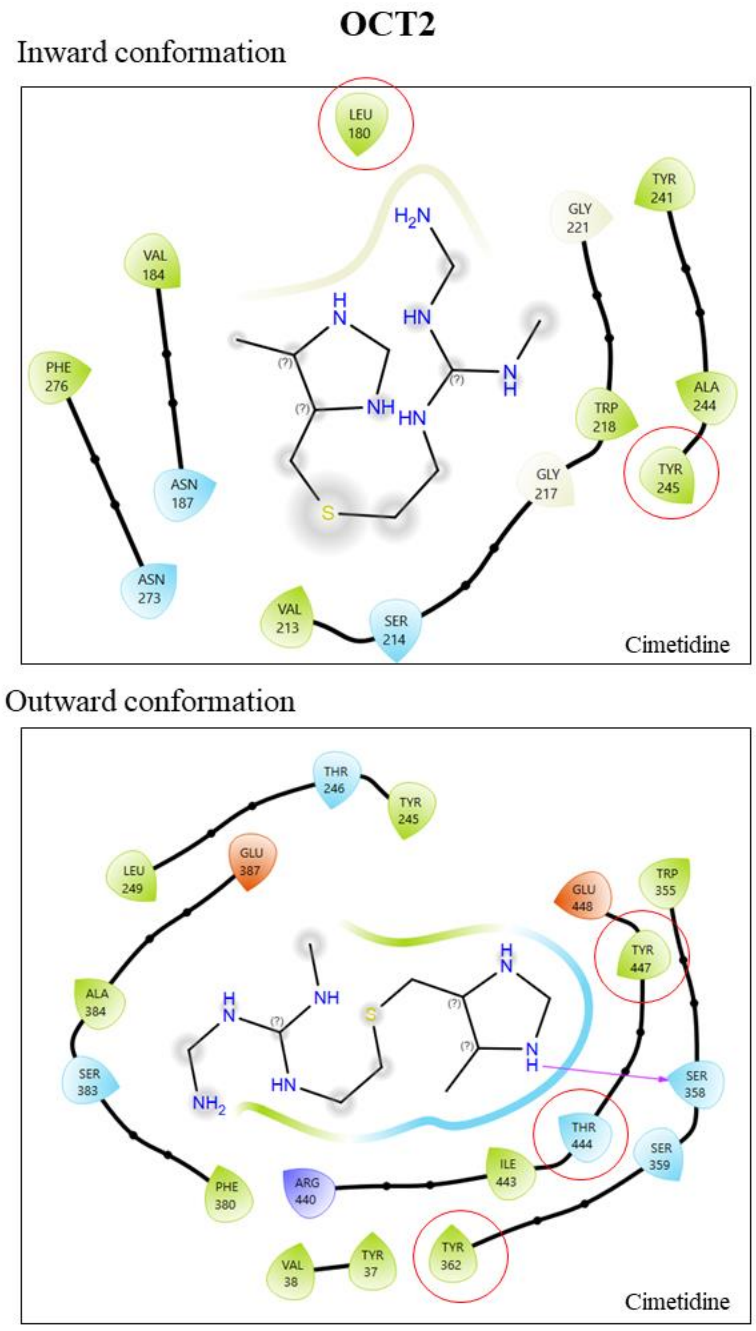


Figure 14: Example of a binding site for the OCT2 protein created using Schrodinger Maestro 13.6. Cimetidine is shown as the model ligand in the binding site of the model protein in the inward-facing conformation model based on the 4ASP template and in the outward-facing conformation, model based on the 7ZH0.

5.2.2.1. Comparison with recent findings on OCT1 and OCT2 structures

As for the protein models, they can be used to evaluate the binding sites for different ligands and to align these results with the OCT1 and OCT2 structures from cryo-EM, which has been done thanks to the two articles from 2022 and 2023 (Khanppnavar et al. 2022; Suo et al. 2023).

The first article (Suo et al. 2023) shows the proteins cryo-EM structure of the organic cation transporter 1 (OCT1) in the apo state with resolution 3,57 Å, and PDB ID is 8ET6, then the cryo-EM structure of the OCT1 in complex with verapamil with resolution 3,45 Å and PDB ID is 8ET8 and cryo-EM structure of the OCT1 in complex with diphenhydramine with resolution 3,77 Å and PDB ID is 8ET7. Another is the Cryo-EM structure of the organic cation transporter 2 (OCT2) in complex with 1-methyl-4-phenylpyridinium (MPP) with resolution 3,61 Å and PDB ID is 8ET9.

As was discussed before, because OCT1 and OCT2 have similarities in homology, around 80 % (Gründemann et al. 1997) showed different amino acid residues in their binding pockets. For OCT1, in outward-facing conformation was found **Ile**446 (in all 30 available cases) and **Tyr**361 (in 29 cases from 30 available cases), while in inward-facing conformation it was **Ser**213 (24 cases from 30 available). For OCT2 in outward conformation, it was **Tyr**447, **Thr**444 for all cases, with anomalies for couples of histamine with choline and amantadine with memantine, and inward conformation, **Tyr**245 and **Leu**180 (11 cases from 22).

Firstly, OCT1 bind with verapamil when the verapamil is a well-established OCT1 inhibitor. Based on the article, verapamil was bound to the central cavity with the help of **Trp**217, **Trp**354, **Phe**244 and **Phe**446. The verapamil cationic tertiary amine group formed a salt bridge with **Glu**386 (Suo et al., 2023). Comparing these findings, we can conclude that verapamil also binds in the OCT1 model generated by SWISS-MODEL with the help of **Trp**217 and **Phe**244 in an inward-facing conformation, and these amino acid residues in this type of conformation were also observed in other ligands. The same is also the case for **Trp**354, which is observed for most ligands in the outward-facing conformation, on top of which **Phe**446 is also mentioned here as well, while **Ile**446 was observed in the results. The position of the amino acid residue is the same, but the type does not match, and the reason is the difference between the wild type (on what is the model from SWISS-MODEL based) and the mutated type (from the article). Another replacement was done in **Cys**36 from the wild type, which

Tyr36 replaced. It is important to mention that Ile446 is also found in other ligands. Glu386 was also observed in the outward-facing conformation of the protein in the case of the results of this work for verapamil but was not found in the other ligands and was specific for verapamil. Authors of the article also suggested that the acidic residue of Glu386 will have a general role in stabilizing the charge and the remaining aromatic/aliphatic residues, and this binding has been termed the orthosteric site, with verapamil still having a 3,4-dimethoxyphenyl group pointing towards the extracellular side of OCT1. The ligand binds via 3-methoxy to Ser382, and the phenyl group interacts with Tyr361, and they called this site opportunistic, whereas it would likely only be occupied by larger substrates and inhibitors (Suo et al., 2023). In the case of the results of this work, only Tyr361 was observed in the case of verapamil binding.

Moving to the next ligand detected for OCT1, diphenhydramine, binds to the Trp217, Trp354, Phe244 and Phe446. The authors described these residues as four residues forming the opposite "walls" of the binding pocket, which the thesis results can prove. The authors determined other sides of the binding pocket where Glu386 and Asp474 form another two sides of the binding pocket (Suo et al. 2023). Both can be found in the results from this work, so the OCT1 protein model from SWISS-MODEL shows very good precision in prediction. Also, the binding pocket of the OCT1 protein model predicted by AlphaFold was compared. The results match the description of the binding pocket exactly to the article, but the inconvenience is that the original AlphaFold prediction cannot select in which state the protein should be modelled. It can be concluded that predicted models from AlphaFold and SWISS-MODEL are suitable for identifying binding sites in conjunction with molecular docking. The model used in this work showed the same (or similar) amino acid residues as the protein structures captured by cryo-EM.

MPP is a well-established substrate for all types of OCTs. The protein-ligand structure of OCT2 with MPP was more compact and semi-closed, according to the article, when it transitioned from the outward-facing conformation to the occluded one. The 4-phenyl group of MPP showed interaction with residues Trp218, Phe245, Trp355, and Phe447, and the 1-methylpyridinium group interacted with the Glu387s charged nitrogen from this acidic residue. It is noteworthy that OCT2, in contrast to OCT1, has a lining of its central cavity made of acidic residues, together with Glu448. Also mentioned in the paper was an interesting observation where the binding position of MPP was consistent with what is predicted by in silico docking with OCT1.

In the case of MPP, the amino acid residues **Tyr**37 and **Tyr**362 are also observed to form a sort of bottleneck (thin gate) when transitioning from the outward-facing conformation to the occluded one. This has not been observed with larger ligands (inhibitors) that bind or block these residues (Suo et al., 2023).

According to the results of this work, **Trp**218 can be found in the binding pocket of OCT2-MPP in inward-facing conformation, while **Trp**355 is not present. Wild-type **Phe**447 is replaced in the experimental structure by **Tyr**447 in a ligand-binding pocket. Both types of protein models present Tyr37 with Tyr362 near the potential "gate". We can thus discuss that our models (especially SWISS-MODEL one) were validated by the recent experimental structure quite well.

5.2.2.2. Possible mechanism of inhibition/activation

Another point that can be taken from the article (Suo et al. 2023) is that this description of the binding pocket can be used for further pharmacological studies when the ligands can be specifically designed for binding to the amino acid residues mentioned before. However, the two distinctive groups of OCT2 ligands remain - histidine and choline or memantine and amantadine. Both groups can be considered as OCT2 inhibitors/activators binding to the specific place within the protein cavity to change the conformation of the protein to either block its transport function or induce it.

As mentioned, amantadine and memantine are drugs used to treat neurological diseases (Kornhuber et al. 1994), whereas the mechanism of action is unknown to this date for amantadine (Okigbo et al. 2019). On the other hand, memantine is known as the inhibitor of calcium influx in brain tissue, which is caused by chronic NMDA receptors activation by glutamate (Kornhuber et al. 1994), so with no connection to the SLC.

In the case of choline and histamine work as neurotransmitters and as substrates in liver and stomach metabolism, where OCT2 is not expressed (Maintz & Novak 2007; Zeisel & Caudill 2010). Hence, we may hypothesize that their metabolic activity may influence the expression pattern and function of OCT2. However, we have no further data for its testing.

5.2.2.3. Comparison with OCT3 experimental structure

Lastly will be discussed similarities with the described protein structure of OCT3 with previous OCT1 and OCT2. This structure in the outward-facing state was also used as a model for SWISS-MODEL predictions of OCT1 and OCT2 in outward-facing states. The structure of OCT3 studied by cryo-EM had two ligands bounded, decynium-22 and corticosterone. The full description can be found in the " Methods " section because this experimental structure was used as a template for OCT1 and OCT2 in SWISS-MODEL. These two ligands are inhibitors of OCT3 when the inhibition is done by the occupation of substrate translocation pathway when corticosterone's (CORT) and decennium-22's (D22) binding sites are slightly overlapping from the protein cavity, but their positions differ. D22 binds mostly to the outer vestibule of OCT3, but in contrast, CORT interacts more in the binding pocket. Although CORT and D22 share similar binding residues, how they interact with them differs. Also, the poses of those ligands' different mechanisms of inhibition when the CORT occupied the binding pocket as a prevention of outward-to-inward rearrangements (like a "clog"), and the D22 position is perpendicular to the membrane, where the binding site is. The result of this position is blockage of the outward-to-inward rearrangements (Khanppnavar et al. 2022). This observation was compared to the result from this work with the CORT-OCT1, CORT-OCT2, D22-OCT2 and D22-OCT1 ligand-proteins. The corticosterone interacts with four main AMKs, **Ile**446, **Phe**244 (which creates two sides of the binding pocket), **Glu**386 and **Asp**474 (located on the poles of the binding pocket), with the presence of **Ser**386 in OCT1. **Ser**386 was also described in verapamil, which inhibits OCT1 activity by blocking the binding pocket. In OCT2, CORT were present in the AMKs of the "gate" **Tyr**37 and **Tyr**362, but none of the AMKs interacted with MPP or the other ligands. The **Tyr**377 is present, which can also be found in histamine and choline cases, and interacts directly with Ile376 through CORT hydroxy- group, which was not observed in other ligands cases. The D22 binds in OCT1 to the described AMKs forming the binding pocket except for **Trp**217, showing interaction directly in the binding pocket. In OCT2, the binding site of D22 cares **Tyr**447 and the "gate" AMKs **Tyr**37 and **Tyr**362 positioned D22 more in front of the binding cavity, corresponding to the D22-OCT3 example.

In conclusion, corticosterone's "clogging" mechanism is very much the same as the verapamil in OCT1 interacts with **Ser**386. OCT2 shows slight similarities with the histamine and choline-binding sites; also, the "gate" AMKs can be observed, but

the clogging mechanism cannot be proven. On the other hand, D22 shows interaction inside the binding cavity of OCT1, but OCT2 interacts with AMKs present more in front of the binding cavity, which corresponds to the article's findings.

On a molecular basis, the binding pocket of OCT3 shows largely conserved AMKs among three OCTs. OCT3 structure was compared in the article with the structures of OCT1 and OCT2 from AlphaFold. According to the article, more similarities can be found between OCT2 and OCT3 when both carry negatively charged residues in binding pockets, which are lacking in the binding pocket of OCT1, which results from this work can confirm. The similarities continue with the comparison of tyrosine residues **Tyr**447 and **Tyr**245 of OCT2, corresponding to **Phe**250 and **Phe**450 of OCT3. These residues provide two additional hydrogen bond donors, also increasing the hydrophilicity of the substrate binding pocket (Khanppnavar et al. 2022). This can be a reason why D22 interacts similarly in OCT2 and OCT3. The inhibition of corticosterone can be proven just in the case of OCT1 with comparison to the verapamil, in OCT2 CORT interact differently and is positioned from the very top of the cavity to the very bottom, where it binds directly to the **Ile**376 through CORT hydroxy- group. In the results, the **Tyr**245 is not detected on the SWISS-MODEL protein structures as one of the amino acid residues in a binding pocket but can be seen in AlphaFold protein models. The binding of corticosterone and decynium-22 can be seen in Figures 19 and 20.

Another comparison with the other OATs and OCTs concludes that the electrostatic properties of OCTs and OATs are consistent with their function as anion or cation transporters, with differences in the charge distribution in the binding pockets for the substrate of these transporters being consistent with their substrate preference (Khanppnavar et al. 2022).

6. Discussion

This thesis is focused on structure analysis and characterization of Organic cation transporters (OCTs) in the Solute carrier family 22 (SLC22). Facilitated transport systems across the plasma membrane are essential for various endogenous and xenobiotic organic ions. The focus was on representatives OCT1 and OCT2.

In this work, the OCT1 and OCT2 models from AlphaFold and SWISS-MODEL generate models based on their cognate single protein in a specific conformation (Waterhouse et al. 2018). Thus, for this work, 3 structures in inward-, occluded and outward-facing states were created using SWISS-MODEL.

From the result, the first docking with the usage of MolMeDB founded ligands showed good dependence between K_m/K_i values and binding energy of ligands when was applied difference (Diff) between binding energy (affinity) values in a cavity and out of the cavity with R^2 around 0,5 for both proteins, which leads to the hypothesis that the frontier between ligand which are substrates and those which are not substrates has been established on value ≥ -1,5 kcal/mol in case of binding energy (affinity) and on the $pK_m \geq 5$ in case of pK_m/pK_i with correlation to the equation (1). This hypothesis needed to be proven, so a bigger dataset was chosen (Koepsell 2013) with around 100 ligands with *Km* and *IC50* described. The hypothesis was not proven due to the weak fit in both cases of OCT1 and OCT2 docked by AutoDock Vina, with the same result also with the AutoDock 1.2.0., when the R^2 was under 0,3. The same results were also observed in the case of the protein models of OCT1 and OCT2 from SWISS-MODEL. This type of predicted model showed low dependences between K_m/K_i values and binding energy of ligands when the difference (Diff) was applied between binding energy (affinity) values in a cavity and out of the cavity with R2 under 0,3 for both proteins and prediction methods. In the case of this data divergence, several factors can be said to be involved, among them different ligand sizes, geometries, and binding properties. After sorting the ligands into classes according to their function and similar size, few ligands and groups themselves remained for further analysis. An increase in the coefficient of determination is shown, but due to the limited size of the ligand groups, no definitive conclusion can be drawn from this observation.

The problem is also mixed values of K_m and *IC50* into one column in the review itself. The online article from 2023 (RDKit Blog 2023) shows the problem with mixing

different assays (IC50, Ki and others), showing us that mixed assays always have low R2 unless the same assays are comparable under the same circumstances.

The dataset taken from the article also lacks information about assays if they have been done under the same circumstances, which could lead to similar studied values for ligands. To prove this, K_m and K_i values from ChEMBL were taken. The database MolMeDB also retrieved those constants from this database. Surprisingly, the R^2 values showed no improvement 0 (K_m) or around 0,2 (K_i) for OCT1 and around 0,3 (K_m) and 0,5 (K_i) for OCT2. A possible explanation is that ChEMBL K_m and K_i values are also mixed up.

The analysis of the binding sites of OCT1 and OCT2 identified the key amino acid residues in their binding cavities. For the outward-facing conformation, **Ile**446 was present in all ligands binding pockets, and **Tyr**361 was present in 29 out of 30 cases within 5 Å of the ligands for OCT1. For OCT2, **Tyr**447 and **Thr**444 were the most common residues for almost all the ligands binding pockets. For the inward-facing conformation, **Ser**213, **Ser**163, **Trp**217, **Phe**159 and **Phe**244 were the most frequent residues for OCT1, occurring in 20 to 24 out of 30 cases. For OCT2, **Tyr**245 and **Leu**180 were the most frequent residues in 11 out of 22 cases. It can be noticed that the AMKs present in the binding sites are mostly aromatic or polar, so that can be discussed that OCT1 and OCT2 would form π-π stacking interactions and hydrogen bonds with ligands (Lanzarotti et al. 2020).

In comparison with the result from the newest article from 2023 discussing the OCT1 and OCT2 in binding with verapamil (OCT1) and MPP (OCT2) (Suo et al. 2023), the binding amino acids residues in predicted protein structures are the same. The positioning of the ligands and the inhibition activity were also confirmed. The main six amino acid residues orienting the ligands in OCT1 are **Trp**217, **Trp**354, **Phe**244, **Phe**446, **Glu**386 and **Asp**474. Also, this work can confirm the role of **Ser**386 interacting with verapamil during its inhibition activity when **Ser**386 is specifical just for verapamil and was not noticed in other ligands cases. Acidic residue **Glu**386 in the binding site have a general role in stabilizing the charge and shape of the cavity. This binding has been termed the orthosteric site (Suo et al., 2023), and it was noticed in cases of larger ligands with a “clogging” effect during inhibition, like corticosterone’s binding to the OCT3 (Khanppnavar et al., 2022).

Decynium-22 is a planar molecule that binds on the upper part of the binding cavity of OCTs in remembrance of a “lid” when **Tyr**37 with **Tyr**362 create the “gate” on

the upper part of the binding cavity. This binding motif was also noticed after comparison with the binding in OCT3 (Khanppnavar et al. 2022).

The general binding sites of OCT2 were also concluded according to the article from authors Suo et al. (2023). It is consisting of **Trp**218, **Phe**245, **Trp**355, **Phe**447, **Glu**387 and **Glu**448, and also with the gating residues **Tyr**37 and **Tyr**362. These amino acid residues were noticed in 18 ligands from 22, and the anomalies have been shown in couples of choline with histamine and memantine with amantadine. According to the article from last year (Khanppnavar et al. 2022), more similarities can be found between OCT2 and OCT3 when both carry negatively charged residues in binding pockets, which are lacking in the binding pocket of OCT1, which results from this work can also confirm.

The final point is that the description of binding pockets can be used for further pharmacological studies when the ligands can be specifically designed to bind to the amino acid residues mentioned above. As was mentioned in “Introduction”, many SLCs are connected to drugs and xenobiotics uptake and play an important role in the organism's metabolism. The human OCT1 and OCT2 play a key role in pharmacokinetics, pharmacodynamics and drug-drug interactions (DDIs), but despite their importance, the mechanism of polyspecific recognition of cationic drugs and the mechanism of alternating OCT access remains a mystery (Khanppnavar et al. 2022; Suo et al. 2023). Further research on the binding sites of those transporters can bring light to this matter and shows new types of inhibition, which can be used in the next drug development.

7. Conclusion

This work, entitled "Structure analysis and characterization of Organic cation transporters (OCTs) in the Solute carrier family 22 (SLC22)", discusses the issue of SLC transporters, more specifically the OCT group falling under the SLC22A subfamily. The OCT group has three main representatives OCT1 (SLC22A1), OCT2 (SLC22A2) and OCT3 (SLC22A3) (Bai et al., 2017; Engelhart et al., 2020). Protein models predicted using AlphaFold software, and SWISS-MODEL were used to study the first two representatives. These prediction models were used due to the lack of OCT structures at the beginning of the work. The OCT group ligands were studied using molecular docking with two versions of AutoDock Vina.

Starting dataset was small, with ligands found in the MolMeDB database, where the K_m and K_i were known for each molecule. These ligands were docked by molecular docking into the protein models showing R^2 around 0,5 in case of difference between docking ligands in the cavity and outside the cavity near the potential extracellular domain of protein in models done by AlphaFold. This was taken as a good result when the hypothesis was that the frontier between ligands which are substrates and those which are not substrates is value ≥ -1,5 kcal/mol in case of binding energy (affinity) and $pK_m \geq 5$ in case of pK_m/pK_i with correlation to the equation (1). In SWISS-MODEL predicted protein structures, the R^2 was under 0,3, which is interesting mainly because SWISS-MODEL predicts protein structures based on their most closely related known protein structures from the PDB, whereas AlphaFold is used to predict the similarity of amino acid sequences across a range of much more related proteins. Therefore, it was expected that the results for the SWISS-MODEL would show better R^2, but the results were the opposite.

The method was also tested on the bigger dataset of substrates taken from (Koepsell 2013), but the spread of ligand sizes and types was more random, and the result from the smaller dataset was not replicated. The R^2 values were under 0,3 or around zero in AlphaFold and SWISS-MODEL models. Since individual subsets have shown much bigger correlations, we can conclude that the molecular docking with Autodock Vina cannot cope with mixing *IC50, Km* and *Ki* values. Nevertheless, from the comparison with recent structures from 2023, we can conclude that molecular docking can predict the binding site composition.

The analysis of the binding sites confirmed the new findings from the latest studies (Khanppnavar et al. 2022; Suo et al. 2023). For the inward-facing conformation, **Ser**213, **Ser**163, **Trp**217, **Phe**159 and **Phe**244 were the most frequent residues for OCT1, occurring in 20 to 24 out of 30 cases, and for OCT2, **Tyr**245 and **Leu**180 were the most frequent residues. For the outward-facing conformation, **Ile**446 was present in all ligands binding pockets, and **Tyr**361 was present in 29 out of 30 cases for OCT1 and OCT2; **Tyr**447 and **Thr**444 were the most common residues for almost all the ligands binding pockets pointing on the π-π stacking interactions and hydrogen bonds as the main interactions of OCT1 and OCT2 with ligands. The main six amino acid residues positioning ligands in OCT1 are **Trp**217, **Trp**354, **Phe**244, **Phe**446, **Glu**386 and **Asp**474, which "builds" the binding pocket. The analogies for the OCT2 are **Trp**218, **Phe**245, **Trp**355, **Phe**447, **Glu**387 and **Glu**448, when **Tyr**37 with **Tyr**362 create the "gate" on the upper part of the binding cavity.

This work also confirmed on theoretical bases of different inhibition mechanisms of OCT. While a larger compound shows a "clogging" effect during inhibition, effectively blocking the whole cavity, the planar binding of decynium-22 on the upper part of the binding cavity of OCTs makes another possible type of inhibition, which can be called "lid" as it blocks the movement of the OCT's helices from outward-facing state to inward-facing state. Finally, we pointed out the role of anomaly-binding ligands from two groups choline and histamine or memantine and amantadine.

While more similarities can be found between OCT2 and OCT3 as they both carry more negatively charged residues in binding pockets, it is good to mention that both transporters, OCT1 and OCT2, have in the binding pockets aromatic AMK. This finding can be important in further substrates studies of OCTs.

Ultimately, this work brings a new approach to studying membrane transporters, specifically the OCTs. The description of their binding sites can be used in new drug development and targeting. This work also confirms the latest findings and quite well described the binding sites using the SWISS-MODEL and AlphaFold when the predicted models showed accurate features compared with experimental ones, and we can recommend them for further research.

8. List of references

Ahn S.-Y., Nigam S. K. (2009) Toward a Systems Level Understanding of Organic Anion and Other Multispecific Drug Transporters: A Remote Sensing and Signaling Hypothesis. *Molecular Pharmacology* **76**, 481–490.

Bahn A., Hagos Y., Reuter S., Balen D., Brzica H., Krick W., Burckhardt B. C., Sabolić I., Burckhardt G. (2008) Identification of a New Urate and High Affinity Nicotinate Transporter, hOAT10 (SLC22A13) *. *Journal of Biological Chemistry* **283**, 16332–16341.

Bai X., Moraes T. F., Reithmeier R. A. F. (2017) Structural biology of solute carrier (SLC) membrane transport proteins. *Molecular Membrane Biology* **34,** 1–32.

Bennett D. I. G., Amarnath K., Park S., Steen C. J., Morris J. M., Fleming G. R. (2021) Models and mechanisms of the rapidly reversible regulation of photosynthetic light harvesting. *Open Biology* **9**, Article 4.

Bien S. A., Su Y. R., Conti D. V., Harrison T. A., Qu C., Guo X., Lu Y., Albanes D., Auer P. L., Banbury B. L., Berndt S. I., Bézieau S., Brenner H., Buchanan D. D., Caan B. J., Campbell P. T., Carlson C. S., Chan A. T., Chang-Claude J., Chen S., Connolly C. M., Easton D. F., Feskens E. J. M., Gallinger S, Giles G. G., Gunter M. J., Hampe J., Huyghe J. R., Hoffmeister M., Hudson T. J., Jacobs E. J., Jenkins M. A., Kampman E., Kang H. M., Kühn T., Küry S., Lejbkowicz F., Le Marchand L., Milne R. L., Li L., Li C. I., Lindblom A., Lindor N. M., Martín V., McNeil C. E., Melas M., Moreno V., Newcomb P. A., Offit K., Pharaoh P. D. P., Potter J. D., Qu C., Riboli E., Rennert G., Sala N., Schafmayer C., Scacheri P. C., Schmit S. L., Severi G., Slattery M. L., Smith J. D., Trichopoulou A., Tumino R., Ulrich C. M., van Duijnhoven F. J. B., Van Guelpen B., Weinstein S. J., White E., Wolk A., Woods M. O., Wu A. H., Abecasis G. R., Casey G., Nickerson D. A., Gruber S. B., Hsu L., Zheng W., Peters U. (2019) Genetic variant predictors of gene expression provide new insight into risk of colorectal cancer. Human Genetics 138, 307–326. (2019) Genetic variant predictors of gene expression provide new insight into risk of colorectal cancer. *Human Genetics* **138**, 307–326.

Drew D., North R. A., Nagarathinam K., Tanabe M. (2021) Structures and General Transport Mechanisms by the Major Facilitator Superfamily (MFS). *Chemical Reviews* **121**, 5289–5335.

Du X., Li Y., Xia Y.-L., Ai S.-M., Liang J., Sang P., Ji X.-L., Liu S.-Q. (2016) Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *International Journal of Molecular Sciences* **17**, 144.

Eberhardt J., Santos-Martins D., Tillack A. F., Forli S. (2021) AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *Journal of Chemical Information and Modeling* **61**, 3891–3898.

Engelhart D. C., Granados J. C., Shi D., Saier M. H., Baker M. E., Abagyan R., Nigam S. K. (2020) Systems Biology Analysis Reveals Eight SLC22 Transporter Subgroups, Including OATs, OCTs, and OCTNs. *International Journal of Molecular Sciences* **21**, 1791.

Eraly S. A., Bush K. T., Sampogna R. V., Bhatnagar V., Nigam S. K. (2004) The molecular pharmacology of organic anion transporters: From DNA to FDA? *Molecular Pharmacology* **65**, 479–487.

Fan J., Xiao Y., Quick M., Yang Y., Sun Z., Javitch J. A., Zhou X. (2021) Crystal structures of LeuT reveal conformational dynamics in the outward-facing states. *Journal of Biological Chemistry* **296,** 100609.

Fernandez-Fernandez B., Montoya-Ferrer A., Sanz A. B., Sanchez-Niño M. D., Izquierdo M. C., Poveda J., Sainz-Prestel V., Ortiz-Martin N., Parra-Rodriguez A., Selgas R., Ruiz-Ortega M., Egido J., Ortiz A. (2011) Tenofovir Nephrotoxicity: 2011 Update. *AIDS Research and Treatment* **2011**, e354908.

Gaulton A., Bellis L. J., Bento A. P., Chambers J., Davies M., Hersey A., Light Y., McGlinchey S., Michalovich D., Al-Lazikani B., Overington J. P. (2012) ChEMBL:

A large-scale bioactivity database for drug discovery. *Nucleic Acids Research* **40**, 1100– 1107.

Gründemann D., Babin-Ebell J., Martel F., Örding N., Schmidt A., Schömig E. (1997) Primary Structure and Functional Expression of the Apical Organic Cation Transporter from Kidney Epithelial LLC-PK1Cells *. *Journal of Biological Chemistry* **272**, 10408–10413.

Hendrickx R., Johansson J. G., Lohmann C., Jenvert R.-M., Blomgren A., Börjesson L., Gustavsson L. (2013) Identification of Novel Substrates and Structure–Activity Relationship of Cellular Uptake Mediated by Human Organic Cation Transporters 1 and 2. *Journal of Medicinal Chemistry* **56**, 7232–7242.

Chicco D., Warrens M. J., Jurman, G. (2021) The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science* **7**, e623.

Jamkhande P. G., Ghante M. H., Ajgunde B. R. (2017) Software based approaches for drug designing and development: A systematic review on commonly used software and its applications. *Bulletin of Faculty of Pharmacy, Cairo University* **55**, 203–210.

Jardetzky O. (1966) Simple Allosteric Model for Membrane Pumps. *Nature* **211**, 5052.

Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O., Tunyasuvunakool K., Bates R., Žídek A., Potapenko A., Bridgland A., Meyer C., Kohl S. A. A., Ballard A. J., Cowie A., Romera-Paredes B., Nikolov S., Jain R., Adler J., Back T., Petersen S., Reiman D., Clancy E., Zielinski M., Steinegger M., Pacholska M., Berghammer T., Bodenstein S., Silver D., Vinyals O., Senior A. W., Kavukcuoglu K., Kohli P., Hassabis D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 7873.

Juračka J., Šrejber M., Melíková M., Bazgier V., Berka K. (2019) MolMeDB: Molecules on Membranes Database. *Database: The Journal of Biological Databases and Curation* **2019**, 078.

Khanppnavar B., Maier J., Herborg F., Gradisch R., Lazzarin E., Luethi D., Yang J.-W., Qi C., Holy M., Jäntsch K., Kudlacek O., Schicker K., Werge T., Gether U., Stockner T., Korkhov V. M., Sitte H. H. (2022) Structural basis of organic cation transporter-3 inhibition. *Nature Communications* **13**, 1.

Kirkpatrick P. (2004) Gliding to success. *Nature Reviews Drug Discovery* **3**, Article 4.

Koepsell H. (2013) The SLC22 family with transporters of organic cations, anions and zwitterions. *Molecular Aspects of Medicine* **34**, 413–435.

Koepsell H., Lips K., Volk C. (2007a) Polyspecific organic cation transporters: Structure, function, physiological roles, and biopharmaceutical implications. *Pharmaceutical Research* **24**, 1227–1251.

Koepsell H., Lips K., Volk C. (2007b) Polyspecific Organic Cation Transporters: Structure, Function, Physiological Roles, and Biopharmaceutical Implications. *Pharmaceutical Research* **24**, 1227–1251.

Kornhuber J., Weller M., Schoppmeyer K., Riederer P. J. (1994) Amantadine and memantine are NMDA receptor antagonists. *Journal of neural transmission. Supplementum* **43**, 91–104.

Lalezari J. P., Stagg R. J., Kuppermann B. D., Holland G. N., Kramer F., Ives D. V., Youle M., Robinson M. R., Drew W. L., Jaffe H. S. (1997) Intravenous Cidofovir for Peripheral Cytomegalovirus Retinitis in Patients with AIDS. *Annals of Internal Medicine* **126**, 257–263.

Lamhonwah A. M., Hawkins C. E., Tam C., Wong J., Mai L., Tein I. (2008) Expression patterns of the organic cation/carnitine transporter family in adult murine brain. *Brain and Development* **30**, 31–42.

Lanzarotti E., Defelipe L. A., Marti M. A., Turjanski A. G. (2020) Aromatic clusters in protein–protein and protein–drug complexes. *Journal of Cheminformatics* **12**, 30.

Lips K. S., Volk C., Schmitt B. M., Pfeil U., Arndt P., Miska D., Ermert L., Kummer W., Koepsell H. (2005) Polyspecific cation transporters mediate luminal release of

acetylcholine from bronchial epithelium. *American Journal of Respiratory Cell and Molecular Biology* **33**, 79–88.

Liu H. C., Goldenberg A., Chen Y., Lun C., Wu W., Bush K. T., Balac N., Rodriguez P., Abagyan R., Nigam S. K. (2016) Molecular Properties of Drugs Interacting with SLC22 Transporters OAT1, OAT3, OCT1, and OCT2: A Machine-Learning Approach. *The Journal of Pharmacology and Experimental Therapeutics* **359**, 215–229.

Liu X. (2019. SLC Family Transporters. *Advances in Experimental Medicine and Biology* **1141**, 101–202.

Long T., Hicks M., Yu H.-C., Biggs W. H., Kirkness E. F., Menni C., Zierer J., Small K. S., Mangino M., Messier H., Brewerton S., Turpaz Y., Perkins B. A., Evans A. M., Miller L. A. D., Guo L., Caskey C. T., Schork N. J., Garner C., Spector T. D., Venter J. C., Telenti A. (2017) Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nature Genetics* **49**, 4.

Maintz L., Novak N. (2007) Histamine and histamine intolerance. *The American Journal of Clinical Nutrition* **85**, 1185–1196.

Marconnet A., Michon B., Le Bon C., Giusti F., Tribet C., Zoonens M. (2020) Solubilization and Stabilization of Membrane Proteins by Cycloalkane-Modified Amphiphilic Polymers. *Biomacromolecules* **21**, 3459–3467.

Maruyama S., Ito M., Ikami Y., Okitsu Y., Ito C., Toshimori K., Fujii W., Yogo K. (2016) A critical role of solute carrier 22a14 in sperm motility and male fertility in mice. *Scientific Reports* **6**, 1.

Matsson P., Karlgren M. (2021) *O*ATs and OCTs: The SLC22 Family of Organic Anion and Cation Transporters. In: *The Medicinal Chemist's Guide to Solving ADMET Challenges.* (P. Schnider eds.), The Royal Society of Chemistry, Croydon, U.K., 128-142.

Mendez D., Gaulton A., Bento A. P., Chambers J., De Veij M., Félix, E., Magariños M. P., Mosquera J. F., Mutowo P., Nowotka M., Gordillo-Marañón M., Hunter F., Junco L., Mugumbate G., Rodriguez-Lopez M., Atkinson F., Bosc N., Radoux C. J., Segura-Cabrera A., Hersey A., Leach A. R. (2019) ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Research*, **47**, D930–D940.

Minuesa G., Purcet S., Erkizia I., Molina-Arcas M., Bofill M., Izquierdo-Useros N., Casado F. J., Clotet B., Pastor-Anglada M., Martinez-Picado J. (2008) Expression and Functionality of Anti-Human Immunodeficiency Virus and Anticancer Drug Uptake Transporters in Immune Cells. *Journal of Pharmacology and Experimental Therapeutics* **324**, 558–567.

Monte J. C., Nagle M. A., Eraly S. A., Nigam S. K. (2004) Identification of a novel murine organic anion transporter family member, OAT6, expressed in olfactory mucosa. *Biochemical and Biophysical Research Communications* **323**, 429–436.

Nies A. T., Koepsell H., Damme K., Schwab M. (2011) Organic cation transporters (OCTs, MATEs), in vitro and in vivo evidence for the importance in drug therapy. *Handbook of Experimental Pharmacology* **201**, 105–167.

Nigam S. K. (2018) The SLC22 Transporter Family: A Paradigm for the Impact of Drug Transporters on Metabolic Pathways, Signaling, and Disease. *Annual review of pharmacology and toxicology* **58**, 663–687.

Okigbo A. A., Helkowski M. S., Royes B. J., Bleimeister I. H., Lam T. R., Bao G. C., Cheng J. P., Bondi C. O., Kline A. E. (2019) Dose-dependent neurorestorative effects of amantadine after cortical impact injury. *Neuroscience Letters* **694**, 69–73.

Owoloye A. J., Ligali F. C., Enejoh O. A., Musa A. Z., Aina O., Idowu E. T., Oyebola K. M. (2022) Molecular docking, simulation and binding free energy analysis of small molecules as PfHT1 inhibitors. *PLoS ONE* **17**, e0268269.

Perez C., Koshy C., Yildiz Ö., Ziegler C. (2012) Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP. *Nature* **490**, 7418.

Quistgaard E. M., Molledo M. M., Löw C. (2017) Structure determination of a major facilitator peptide transporter: Inward facing PepTSt from Streptococcus thermophilus crystallized in space group P3121. *PLOS ONE* **12**, e0173126.

RDKit blog—The hazards of combining data from IC50 assays: https://greglandrum.github.io/rdkit-blog/posts/2023-06-12-overlapping-ic50-assays1.html (12. 6. 2023).

Rose Y., Duarte J. M., Lowe R., Segura J., Bi C., Bhikadiya C., Chen L., Rose A. S., Bittrich S., Burley S. K., Westbrook J. D. (2021) RCSB Protein Data Bank: Architectural Advances Towards Integrated Searching and Efficient Access to Macromolecular Structure Data from the PDB Archive. *Journal of Molecular Biology* **433**, 166704.

Roy K., Kar S., Das R. N. (2015) Other Related Techniques. In: *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*. (V K. Roy, S. Kar, R. N. Das eds.), Academic Press, U.S.A., 357–425.

Sager G., Smaglyukova N., Fuskevaag O.-M. (2018) The role of OAT2 (SLC22A7) in the cyclic nucleotide biokinetics of human erythrocytes. *Journal of Cellular Physiology* **233**, 5972–5980.

Shin S.-Y., Fauman E. B., Petersen A.-K., Krumsiek J., Santos R., Huang J., Arnold M., Erte I., Forgetta V., Yang T.-P., Walter K., Menni C., Chen L., Vasquez L., Valdes A. M., Hyde C. L., Wang V., Ziemek D., Roberts P., Xi L., Grundberg E. Waldenberger M., Richards J. B., Mohney R. P., Milburn M. V., John S. L., Trimmer J., Theis F. J., Overington J. P., Suhre K., Brosnan M. J., Gieger C., Kastenmüller G., Spector T. D., Soranzo N. (2014) An atlas of genetic influences on human blood metabolites. *Nature Genetics* **46**, 6.

Shiraya K., Hirata T., Hatano R., Nagamori S., Wiriyasermkul P., Jutabha P., Matsubara M., Muto S., Tanaka H., Asano S., Anzai N., Endou H., Yamada A., Sakurai H., Kanai Y. (2010) A Novel Transporter of SLC22 Family Specifically Transports Prostaglandins and Co-localizes with 15-Hydroxyprostaglandin Dehydrogenase in Renal Proximal Tubules *. *Journal of Biological Chemistry* **285**, 22141–22151.

Schmidpeter P. A. M., Sukomon N., Nimigean C. M. (2020) Reconstitution of Membrane Proteins into Platforms Suitable for Biophysical and Structural Analyses. In: *Expression, Purification, and Structural Biology of Membrane Proteins* (V C. Perez, T. Maier eds.), Springer US, New York, U.S.A., 191–205.

Schumann T., König J., Henke C., Willmes D. M., Bornstein S. R., Jordan J., Fromm M. F., Birkenfeld A. L. (2020) Solute Carrier Transporters as Potential Targets for the Treatment of Metabolic Disease. *Pharmacological Reviews* **72**, 343–379.

Solcan N., Kwok J., Fowler P. W., Cameron A. D., Drew D., Iwata S., Newstead S. (2012) Alternating access mechanism in the POT family of oligopeptide transporters. *The EMBO Journal* **31**, 3411–3421.

Srikant S., Gaudet R. (2019) Mechanics and pharmacology of substrate selection and transport by eukaryotic ABC exporters. *Nature Structural & Molecular Biology* **26**, 9.

Srinivasan B. (2022) A guide to the Michaelis–Menten equation: Steady state and beyond. *The FEBS Journal* **289**, 6086–6098.

Sun S., Arosio P., Levi S., Chasteen N. D. (1993) Ferroxidase kinetics of human liver apoferritin, recombinant H-chain apoferritin, and site-directed mutants. *Biochemistry* **32**, 9362–9369.

Suo Y., Wright N. J., Guterres H., Fedor J. G., Butay K. J., Borgnia M. J., Im W., Lee S.-Y. (2023) Molecular basis of polyspecific drug binding and transport by OCT1 and OCT2. *bioRxiv*, s. 2023.03.15.532610.

Swinney D. C. (2011) Molecular Mechanism of Action (MMoA) in Drug Discovery. In: *Annual Reports in Medicinal Chemistry*. (V J. E. Macor eds.), Academic Press, Cambridge, U.S.A., 301–317.

Trott O., Olson A. J. (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, **31**, 455–461.

Uhl G. R., Drgon T., Li C.-Y., Johnson C., Liu Q.-R. (2009) Smoking and smoking cessation in disadvantaged women: Assessing genetic contributions. *Drug and Alcohol Dependence* **104**, S58–S63.

Uhl G. R., Liu Q.-R., Drgon T., Johnson C., Walther D., Rose J. E., David S. P., Niaura R., Lerman C. (2008) Molecular Genetics of Successful Smoking Cessation: Convergent Genome-Wide Association Study Results. *Archives of General Psychiatry* **65**, 683–693.

Uhlén M., Fagerberg L., Hallström B. M., Lindskog C., Oksvold P., Mardinoglu A., Sivertsson Å., Kampf C., Sjöstedt E., Asplund A., Olsson I., Edlund K., Lundberg E., Navani S., Szigyarto C. A., Odeberg J., Djureinovic D., Takanen J. O., Hober S., Alm T., Edqvist P. H., Berling H., Tegel H., Mulder J., Rockberg J., Nilsson P., Schwenk J. M., Hamsten M., von Feilitzen K., Forsberg M., Persson L., Johansson F., Zwahlen M., von Heijne G., Nielsen J., Pontén F. (2015) Proteomics. Tissue-based map of the human proteome. *Science (New York, N.Y.)* **347**, 1260419.

UniProt Consortium (2021) UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research* **49**, D480–D489.

Vialou V., Balasse L., Callebert J., Launay J.-M., Giros B., Gautron S. (2008) Altered aminergic neurotransmission in the brain of organic cation transporter 3-deficient mice. *Journal of Neurochemistry* **106**, 1471–1482.

Volk C. (2014) OCTs, OATs, and OCTNs: Structure and function of the polyspecific organic ion transporters of the SLC22 family. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling* ***3***, 1–13.

Wan-Jie C., Wei-Yi O., Gavin D. (2016) Lipocalin 2 and lipocalin 2 receptor in neuron-glia interactions following injury. *Frontiers in Cellular Neuroscience* **10**.

Waterhouse A., Bertoni M., Bienert S., Studer G., Tauriello G., Gumienny R., Heer F. T., de Beer T. A. P., Rempfer C., Bordoli L., Lepore R., Schwede T. (2018) SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research* **46**, W296–W303.

Wessler I., Kirkpatrick C. J. (2008) Acetylcholine beyond neurons: The non-neuronal cholinergic system in humans. *British Journal of Pharmacology* **154**, 1558–1571.

Wittwer M. B., Zur A. A., Khuri N., Kido Y., Kosaka A., Zhang X., Morrissey K. M., Sali A., Huang Y., Giacomini K. M. (2013) Discovery of Potent, Selective Multidrug and Toxin Extrusion Transporter 1 (MATE1, SLC47A1) Inhibitors Through Prescription Drug Profiling and Computational Modeling. *Journal of Medicinal Chemistry* **56**, 781–795.

Wu W., Bush K. T., Liu H. C., Zhu C., Abagyan R., Nigam S. K. (2015) Shared Ligands Between Organic Anion Transporters (OAT1 and OAT6) and Odorant Receptors. *Drug Metabolism and Disposition* **43**, 1855–1863.

Yee S. W., Stecula A., Chien H.-C., Zou L., Feofanova E. V., Borselen M. van, Cheung K. W. K., Yousri N. A., Suhre K., Kinchen J. M., Boerwinkle E., Irannejad R., Yu B., Giacomini K. M. (2019) Unraveling the functional role of the orphan solute carrier, SLC22A24 in the transport of steroid conjugates through metabolomic and genome-wide association studies. *PLOS Genetics* **15**, e1008208.

Yung-Chi C., Prusoff W. H. (1973) Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochemical Pharmacology* **22**, 3099–3108.

Zeisel S. H., Caudill M. A. (2010) Choline. *Advances in Nutrition* **1**, 46–48.

Zhu C., Nigam K. B., Date R. C., Bush K. T., Springer S. A., Saier M. H., Wu W., Nigam S. K. (2015) Evolutionary Analysis and Classification of OATs, OCTs, OCTNs, and Other SLC22 Transporters: Structure-Function Implications and Analysis of Sequence Motifs. *PLoS ONE* **10**, e0140569.

Attachments

Attachment 1: Table of compounds sorted into two categories, "Receptor antagonist" and "Receptor agonist", with Km and IC50 for OCT1 and OCT2 taken from various articles compiled into a review (taken from Koepsell 2013).

Class	**Compound**	***Km* or (*IC50*) [μM]*a, b***	
		hOCT1	**hOCT2**
Receptor antagonists			
Acetylcholine receptor (muscarinic)	Atropine	1,2	-29
Acetylcholine receptor (muscarinic)	Mepiperphenidol		**4,8**
α-Adrenoceptor	Phenoxybenzamine	-2,7	-4,9
α-Adrenoceptor	Prazosin	-1,8	(>100)
β-Adrenoceptor	Acebutolol	-96	
Histamine H1 receptor	Diphenhydramine	-3,4	-15
Histamine H1 receptor	Pyrilamine		
Histamine H2 receptor	Cimetidine	**-166**	**8.6, 73** (**70**)
Histamine H2 receptor	Famotidine		**204** (**111**)
Histamine H2 receptor	Ranitidine	-28	**265** (**40**)
NMDA receptor	Amantadine	-236	**20, 27** (**28**)
NMDA receptor	Ketamine	-115	-23
NMDA receptor	Memantine	-3,7	**34** (**7.3**)
NMDA receptor	Phencyclidine	-4,4	-25
Receptor agonists			
Acetylcholine receptor (nicotinic)	Nicotine	-186	-42
α-Adrenoceptor	Clonidine	(0.6 – 6.5)	-23
α-Adrenoceptor	Etilefrine	-447	(4,009)
β-Adrenoceptor	O-Methylisoprenaline	(>100)	-570
NMDA receptor	Dizocilpine	-81	-22

Attachment 2: Table of compounds sorted into five categories. First are uncategorized compounds, then "Metabolites", "Neurotransmitters", "Hormones", and "Miscellaneous" with Km and IC50 for OCT1 and OCT2 taken from various articles compiled into a review (taken from Koepsell 2013).

Compounds	***Km* or (*IC50*) [µM]*a, b***	
	hOCT1	**hOCT2**
MPP	**15, 32**	**19, 78** (**24**)
TEA	**229 (158,173)**	**76** (**48–270**)
N1-methylnicotinamide	(7,700)	**340** (**270**)
Tetramethylammonium	(12,400)	(180,525)
Tetrapropylammonium	-102	(20, 128)
Tetrabutylammonium	-30	(20, 120)
Tetrapentylammonium	(1.5, 7.5)	(1.5, 11)
Decynium 22	(2.7–4.7)	(0.10–1.1)
Disprocynium 24		
Metabolites		
Acetyl-L-carnitine		
Betaine		
Choline	(16,700)c	210 (381)
Creatinine	(>20,000)	(<2,000)c
D-Carnitine		
L-Carnitine	(12,400)c	(13,000)c
Guanidine	(5,030)c	(2,300)c
Thiamine		
Neurotransmitters		
Acetylcholine	-580	117 (149)
Dopamine	(>20,000)c	390–1,400
Epinephrine	(>30,000)c	400
Histamine	(>20,000)c	940, 13,00
Norepinephrine	(7,100)c	1,500, 1,900
Serotonin	(>20,000)c	80, 290 (310)
Hormones		
Aldosterone		
Corticosterone	(7, 22)	-34
Progesterone	-3,1	-27
Prostaglandin E2	0,66	0,03
Prostaglandin F2αF2α	0,48	0,33
Testosterone	(10)c	(3)c
Miscellaneous		
Agmatine	(24,000)	1,400 (3,251)

Attachment 3: Table of compounds sorted into one category, "Drugs and xenobiotics", with Km and IC50 for OCT1 and OCT2 taken from various articles compiled into a review (taken from Koepsell 2013).

Class	**Compound**	***Km* or (*IC50*) [µM]*a, b***	
		hOCT1	**hOCT2**
Drugs and Xenobiotics			
Anaesthetic	Midazolam	-3,7	
Antiasthmatic	Beclometasone		-4,4
Antiasthmatic	Budesonide		-7,3
Antibiotic	Cefepine		
Antibiotic	Cefoselis		
Antibiotic	Cefsulodine		
Antibiotic	Cephaloridine		
Antibiotic	Levofloxacine		
Antibiotic	Ofloxacine		
Antidepressant	Desipramine	-5,4	-16
Antidepressant	Imipramine		
Antidiabetic	Metformin	**1,470** (**2,010**)	(**339, 1,700**)**990**
Antidiabetic	Phenformin	(10, 153)	-65
Antiemetic	Emetine		
Antiemetic	Granisetron	(<100)*c*	(<100)*c*
Antiemetic	Odansetron	(<100)*c*	(<100)*c*
Antiemetic	Tropisetron	(<100)*c*	(<100)*c*
Antihypertensive	Debrisoquine		(**7.3**)*c*
Antimalarial	Quinine	(**13, 23**)	(**23, 34**)
Antioxidant	Ergothioneine		
Antiviral	Aciclovir	**151**	
Antiviral	Ganciclovir	**516**	
Antiviral (HIV)	Aquinavir	-8,3	
Antiviral (HIV)	Indinavir	-62	
Antiviral (HIV)	Nelfinavir	-22	
Antiviral (HIV)	Ritonavir	-5,2	
Cytostatic	Actinomycin D		
Cytostatic	Cisplatin	(>100)	-1,5
Cytostatic	Doxorubicin		
Cytostatic	Daunorubicin		
Cytostatic	Mitoxantrone	(16)*c*	(800)*c*
Muscle relaxant	Vecuronium	(127, 232)	
Psychostimulant	3,4-Methylenedioxy-methamphetamine	-24	-1,6
Psychostimulant	D-Amphetamine	-202	-11
Serine protease inhibitor	Nafamostatmesilate		**-20**

Attachment 4: Table of compounds sorted into one category, "Ion channel and transporter blockers," with Km and IC50 for OCT1 and OCT2 taken from various articles compiled into a review (taken from Koepsell 2013).

Class	**Compound**	***Km* or (*IC50*) [µM]*a, b***	
		hOCT1	**hOCT2**
Ion channel and transporter blockers			
Ca^{2+}channel	Verapamil	-2,9	(206)*c*
Na^{+}channel	r (−) Disopyramide	-15	
Na^{+}channel	s (+) Disopyramide	-30	
Na^{+}channel	Procainamide	-74	(50, 58)
Na^{+}channel	Quinidine	**-18**	
Noradrenaline transporter	Cocaine	-85	(113, 277)*c*
Serotonin transporter	Citalopram	(2.8)*c*	(21)*c*

Attachment 5: Table of findings by ChEMBL database for OCT1 and OCT2 with Km taken from various articles (ChEMBL 2023; https://www.ebi.ac.uk/chembl).

OCT1		**OCT2**	
	Km [nM]		*Km* [nM]
Prostaglandin E2	657	Norepinephrine	1900000
Tetraethylammonium	229000	Prostaglandin E2	28,9
Azidoprocainamide methoiodide	101000	Memantine	34000
TEA	2320	Agmatine	1400000
MPP	14600	Choline	102000
Acyclovir	151200	Choline	210000
Tributylmethylammonium	53000	MPP	22200
Noroxycodone	20100	MPP	16000
Ipratropium	9000	Cimetidine	60000
Sumatriptan	55000	Serotonin	80000
N-Methyl-Quinidine	11500	Quinine	2600
N-Methyl-Quinidine	20000	Tetraethylammonium	63000
Hydromorphone	56100	Tetraethylammonium	31000
Prostaglandin F2alpha	477	Amantadine	27000
N-Methyl-Quinidine	19500	Dopamine	520000
Norfentanyl	7700	MPP	19000
Ganciclovir	516200	Dopamine	330000
Methylnaltrexone	20000	Dopamine	390000
N-methyl-quinidine	12000	Tetraethylammonium	431000
		Prostaglandin F2alpha	334
		Choline	190000
		N1-methylnicotinamide	300000
		Tetraethylammonium	76000
		MPP	1200
		Histamine	1300000

Attachment 6: Table of findings by ChEMBL database for OCT1 and OCT2 with Ki taken from various articles (ChEMBL 2023; https://www.ebi.ac.uk/chembl/).

OCT1		**OCT2**	
	Ki [nM]		*Ki* [nM]
Midazolam	3700	MPP	2400
Corticosterone	7020	Procainamide	50000
Verapamil	2900	Mepiperphenidol	4800
Clonidine	550	o-Methylisoprenaline	570000
Vecuronium	120000	Cimetidine	57000
Tetramethylammonium	12400000	Quinine	3400
Decynium-22	2730	Cyanine-863	210
Acebutolol	95800	Decynium-22	100
Vecuronium	232000	Famotidine	1800000
Tetraethylammonium	163000	Desipramine	16000
Agmatine	24000000	Tetramethylammonium	180000
N1-methylnicotinamide	7700000	Tetrapentylammonium	1500
Eltrombopag	103000	N1-methylnicotinamide	266000
Cimetidine	166000	Decynium-22	100
Decynium-22	4400	Famotidine	1800000
Tetraethylammonium	161000	Desipramine	16000
Procainamide	73900	Tetramethylammonium	180000
Tetrapropylammonium	102000	Tetrapentylammonium	1500
MPP	12300	N1-methylnicotinamide	266000
Tetrabuthylammonium	29600		
Tetrapentylammonium	7460		
Quinine	22900		
Desipramine	5360		
Tetraethylammonium	158000		
Quinidine	17500		

Attachment 7: Ligands found by MolMeDB with results of docking done in a possible binding cavity and out of it, near possible intra or extracellular domain, with the difference (Diff) of these two values for protein OCT1 from AlphaFold.

OCT1				
	pKm	Vina Scoring cavity	Vina Scoring near the domain	**Diff**
Tetrylammonium	3,64	-4,2	-3,6	**-0,6**
Azidoprocainamide methoiodide	4	-7,5	-6,4	**-1,1**
Ganciclovir	3,29	-6,3	-6	**-0,3**
Acyclovir	3,82	-5,8	-5,8	**0**
Ipratropium	5,05	-7,9	-7	**-0,9**
Dinoprostone	6,18	-7,4	-6,3	**-1,1**
Dinoprost tromethamine	6,32	-7,5	-6	**-1,5**
Tetrylammonium	5,63	-4,2	-3,6	**-0,6**

Attachment 8: Ligands found by MolMeDB with results of docking done in a possible binding cavity and out of it, near possible intra or extracellular domain, with the difference (Diff) of these two values for protein OCT2 from AlphaFold.

OCT2				
	pKm	Vina Scoring cavity	Vina Scoring near the domain	**Diff**
Nicotinyl methylamide	5,45	-5,6	-5,6	**0**
Memantine	5,34	-7,2	-6	**-1,2**
Tetrylammonium	5,36	-4,5	-3,5	**-1**
Tetrylammonium	5,42	-4,5	-3,5	**-1**
Cimetidine	5,37	-6,1	-5,9	**-0,2**
Choline	5,43	-3,5	-3,1	**-0,4**
Choline	5,40	-3,5	-3,1	**-0,4**
Norepinephrine	5,56	-6	-5,3	**-0,7**
Agmatine	5,54	-4,8	-4,7	**-0,1**
Serotonin	5,38	-6,9	-5,6	**-1,3**
Dopamine	5,47	-6	-5,3	**-0,7**
Quinine	5,25	-8,3	-6,9	**-1,4**
Dinoprostone	5,12	-7,6	-5,8	**-1,8**
Dinoprost tromethamine	5,18	-7,8	-5,4	**-2,4**

Attachment 9: Ligands found by MolMeDB with results of docking done by AutoDock Vina in binding cavity of inward-facing, occluded and outward-facing state in protein OCT1 predicted by SWISS-MODEL.

OCT1				
	pKm/IC50	Inward conf.	Occluded conf.	Outward conf.
Tetrylammonium	3,64	-3,4	-3,4	-3,1
Azidoprocainamide methoiodide	4	-6,2	-6,5	-6
Ganciclovir	3,29	-5,3	-5,6	-5,4
Acyclovir	3,82	-5,1	-5,1	-5,1
Ipratropium	5,05	-7,1	-7,5	-7,4
Dinoprostone	6,18	-6,3	-6,5	-6,8
Dinoprost tromethamine	6,32	-6,4	-6,5	-7,6

Attachment 10: Ligands found by MolMeDB with results of docking done by AutoDock Vina in binding cavity of inward-facing, occluded and outward-facing state in protein OCT2 predicted by SWISS-MODEL.

OCT2				
	pKm/IC50	Inward conf.	Occluded conf.	Outward conf.
Nicotinyl methylamide	5,45	-4,9	-5,2	-5,7
Memantine	5,34	-6,4	-6,3	-5,9
Tetrylammonium	5,36	-3,3	-3,8	-3,5
Tetrylammonium	5,42	-3,3	-3,8	-3,5
Cimetidine	5,37	-4,9	-5,7	-6,2
Choline	5,43	-3,1	-3,3	-3,6
Choline	5,40	-3,1	-3,3	-3,6
Norepinephrine	5,56	-5,3	-5,5	-6,3
Agmatine	5,54	-4,1	-4,2	-4,9
Serotonin	5,38	-5,5	-6,3	-6,9
Dopamine	5,47	-5	-5,4	-6,2
Quinine	5,25	-7,6	-7,7	-7,7
Dinoprostone	5,12	-6,1	-6,5	-7,2
Dinoprost tromethamine	5,18	-6,2	-6,7	-8,1