University of South Bohemia, České Budějovice
Faculty of Science

# Domain conformations of the motor subunit of EcoR124I involved in ATPase activity and dsDNA translocation

Ph.D. Thesis

**Vitali Bialevich, MSc.**

Supervisor: Prof. RNDr. Rüdiger Ettrich Ph.D.
University of South Bohemia, České Budějovice
and
Center for Nanobiology and Structural Biology
Institute of Microbiology
Academy of Sciences of the Czech Republic

**České Budějovice 2016**

This thesis should be cited as:
**Bialevich V.,** 2016: Domain conformations of the motor subunit of EcoR124I involved in ATPase activity and dsDNA translocation. Ph.D. Thesis Series, № 19. University of South Bohemia, Faculty of Science, České Budějovice, Czech Republic, 111 pp.

## Annotation

Bacterial type I restriction-modification systems are composed of three different subunits: one HsdS subunit is required for identification of target sequence and anchoring the enzyme complex on DNA; two HsdM subunits in the methyl-transferase complex serve for host genome modification accomplishing a protective function against self-degradation; two HsdR (or motor) subunits house ATP-dependent translocation and consequent cleavage of double stranded DNA activities.

The crystal structure of the 120 kDa HsdR subunit of the Type I restriction-modification system EcoR124I in complex with ATP was recently reported. HsdR is organized into four approximately globular structural domains in a nearly square-planar arrangement: the N-terminal endonuclease domain, the RecA-like helicase domains 1 and 2 and the C-terminal helical domain. The near-planar arrangement of globular domains creates prominent grooves between each domain pair. The two helicase-like domains form a canonical helicase cleft in which double-stranded B-form DNA can be accommodated without steric clash. The helical domain, probably involved in complex assembly, exhibits only a few specific interactions with helicase 2 domain.

Molecular mechanism of dsDNA translocation, cleavage and ATP hydrolysis has not been yet structurally investigated. Here we propose a translocation cycle of the restriction-modification system EcoR124I based on analysis of available crystal structures of superfamily 2 helicases, strutural modeling and complementary biochemical characterization of mutations introduced in sites potentially inportant for translocation in the HsdR motor subunit. Also a role of the extended region of the helicase motif III in ATPase activity of EcoR124I was probed.

**Declaration [in Czech]**

Prohlašuji, že svoji disertační práci jsem vypracoval samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své disertační práce, a to v úpravě vzniklé vypuštěním vyznačených částí archivovaných Přírodovědeckou fakultou elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

České Budějovice, 09.12.2016 .....

Vitali Bialevich

This thesis originated from a partnership of the Faculty of Science, University of South Bohemia, and Center for Nanobiology and Structural Biology Institute of Microbiology Academy of Sciences of the Czech Republic, supporting doctoral studies in the Biophysics study programme.

**Financial support**

This work was funded by the Czech Science Foundation (P207/12/2323 to RE), http://www.gacr.cz and the grand agency of the University of South Bohemia via the doctoral team grands №170/2010/P and №141/2013/P.

**List of papers and author's contribution**

The thesis is based on the following publications (listed chronologically):

1. D. Sinha*, K. Shamayeva*, V. Ramasubramani, D. Řeha, **V. Bialevich**, M. Khabiri, N. Milbar, A. Guzanova, M. Weiserova, E. Csefalvay, J. Carey, R. Ettrich (2014) Interdomain communication in the endonuclease/motor subunit of Type I restriction-modification enzyme EcoR124I. *Journal of Molecular Modeling* 20 (7):2334. (* - contributed equally) (IF=1.867)
   *VB tested purified enzymes in vitro, analyzed experimental data.*

2. E. Csefalvay, M. Lapkouski, A. Guzanova, L. Csefalvay, T. Baikova, I. Shevelev, **V. Bialevich**, K. Shamayeva, P. Janscak, I. Kuta Smatanova, S. Panjikar, J. Carey, M. Weiserova, R. Ettrich (2015) Functional coupling of duplex translocation to DNA cleavage in a Type I restriction enzyme. *PLoS ONE* 10 (6):e0128700. (IF=3.534)
   *VB expressed and purified methyltransferase, WT and mutant HsdR subunits, performed in vitro restriction activity assays and analized data.*

3. **V. Bialevich***, D. Sinha*, K. Shamayeva, A. Guzanova, D. Řeha, E. Csefalvay, J. Carey, M. Weiserova, R. Ettrich (2016) The helical domain of the motor subunit of EcoR124I participates in ATPase activity and dsDNA translocation. *PeerJ*, accepted 8.12.2016. (* - contributed equally) (IF=2.183)
   *VB performed site-directed mutagenesis, expressed and purified WT and mutant proteins, tested purified enzymes in vitro and in vivo, analyzed experimental data, participated in writing of the manuscript.*

4. Sinha D***, Bialevich V***, Shamayeva K, Řeha D, Guzanova A, Sisakova A, Csefalvay E, Krejci L, Carey J, Weiserova M and Ettrich R The role of motif III and its extended region in positioning the two helicase domains in the motor subunit of the restriction–modification system EcoR124I. *Manuscript* (* - contributed equally)
   *VB performed site-directed mutagenesis, expressed and purified WT and mutant proteins, tested purified enzymes in vitro and in vivo, analyzed experimental data, participated in writing of the manuscript.*

## ACKNOWLEGMENTS

I would like to thank **Prof. Rüdiger Ettrich** who was my supervisor during all these years, for his support and help not only in solving scientific problems but also in difficult life circumstances.

Special thanks go to **Prof. Jannette Carey** for sharing her experience in laboratory work and for her patience during revision of manuscripts.

This work would not be possible without help of my colleagues from Prague **Dr. Marie Weiserova**, **Alena Guzanova** and **Anicka Stahova**.

Also I thank **Dr. Jost Ludwig** for experience I obtained in his laboratory in Bonn during two month stay there as a visiting student.

Many tnanks to my colleagues from computational group, especially to **Dr. David Reha** who contributed substantially in data analysis and preparation of manuscripts.

I would like to thank **Dr. Eva Csefalvay** for sharing her experience in the field of molecular biology and for good practical advices during my work in the laboratory.

Also I thank our lab technician **Kamila Blafkova** for keeping order in our laboratory and especially for sharing practical information regarding live in the Czech Republic.

I would like to acknowledge everyone who was not named here and has contributed to my research work in **Center for Nanobiology and Structural Biology**.

I'm very grateful to my schoolteacher **Grigorcevich Inna Stanislavovna** who introduced me in the field of biology and helped to determine my life choice.

And, of course, I would like to express my feeling of gratitude to my mother **Yania** and my brother **Andrej** for their love and support during my study here in Nove Hrady.

Finally, I would like to tell words of gratitude to the most important person in my life, to my beloved wife. Thank you, **Katsiaryna**, for your love and patience during all these years.

*I dedicate this work to my family*

## TABLE OF CONTENTS

## LIST OF ABREVIATIONS

| | |
|---|---|
| **Å** | **Angstrom** |
| **Amp** | **Ampicillin** |
| **ATP** | **Adenosine triphosphate** |
| **bp** | **base pair** |
| **cccDNA** | **Covalently closed circular DNA** |
| **DEAE** | **Diethylaminoethyl** |
| **DNA** | **Deoxyribonucleic acid** |
| **dsDNA** | **Double-stranded deoxyribonucleic acid** |
| **DTT** | **Dithiothreitol** |
| **E. coli** | **Escherichia coli** |
| **e.o.p.** | **Efficienty of plating** |
| **EDTA** | **Ethylenediaminetetraacetic acid** |
| **FRET** | **Fluorescence resonance energy transfer** |
| **GTP** | **Guanosine triphosphate** |
| **HJ** | **Holliday junction** |
| **IPTG** | **Isopropyl β-D-1-thiogalactopyranoside** |
| **Kd** | **dissociation constant** |
| **kDa** | **kilodalton** |
| **l** | **Litr** |
| **LB** | **Luria-Bertani** |
| **M** | **Molar** |
| **MD** | **Molecular dynamics** |
| **OD** | **Optical density** |
| **PCR** | **Polymerase chain reaction** |
| **PCNA** | **Proliferating cell nuclear antigen** |
| **pdb** | **Protein data bank** |
| **PFU** | **Plaque forming units** |
| **Pi** | **Inorganic phosphate** |
| **R-M** | **Restriction-modification** |
| **SAH** | **S-adenosil homocysteine** |
| **SAM** | **S-adenosyl methionine** |
| **SF2** | **Superfamily 2** |

| | |
|---|---|
| **SDS** | **Sodium dodecyl sulfate** |
| **SDS-PAGE** | **Sodium dodecyl sulfate polyacrylamide gel electrophoresis** |
| **STE** | **Sodium chloride-Tris-EDTA** |
| **TAE** | **Tris-acetate-EDTA** |
| **TBS** | **Triplex binding site** |
| **Tc** | **Tetraciclin** |
| **TFO** | **Triplex forming oligonucleotide** |
| **TRD** | **Target recognition domain** |
| **UV** | **Ultraviolet** |
| **V** | **Volt** |
| **WT** | **Wild type** |
| **X-ray** | **X-radiation** |

# PROLOGUE AND AIMS OF THE WORK

The phenomenon of restriction and modification in bacteria was first noticed in the 1950s. While the growth of viruses in some bacteria strains was successful, plating of propagated viruses in other strains lead to inhibition (restriction) of their growth. In 1953 this effect was linked with sequence-specific restriction enzymes (Bertani & Weige, 1953). The restriction-modification systems in bacteria were described to protect the bacterial cell from foreign DNA, such as invading DNA coming from bacteriophages (Redaschi & Bickle, 1996). Based on differences in cofactor requirements, subunit composition, sequence recognition specificity and cleavage position all restriction-modification systems are divided into four Types (I-IV) (Loenen et al. 2014a).

The restriction-modification enzyme EcoR124I from *E. coli* belongs to the so-called Type I restriction-modification systems, which differ from the general definiftion by providing also the ability of translocating DNA prior to restriction. EcoR124I exhibits various consequently coordinated enzyme functions determined by its multi-subunit organization. If EcoR124I recognizes its own (hemi-methylated) DNA then it switches to modification mode and methylates the DNA recognition sequence to protect its own DNA from restriction (Makovets et al. 2004). If EcoR124I recognizes an unmethylated specific sequence of (foreign) DNA then it switches to restriction mode. The enzyme being tighly anchored on target sequence starts to translocate DNA towards itself in an ATP-dependent manner extruding long DNA loops (van Noort et al., 2004) and, when translocation is stalled, cleaves the DNA duplex up to thousand base pairs away from the recognition sequence. This results in large pieces of DNA (Janscak et al., 1996), and a role of these enzymes in horizontal gene transfer is predicted (Waldron et al., 2006). Translocation and endonuclease activities are housed on the motor subunit HsdR. The HsdR subunit of EcoR124I was structurally characterized (Lapkouski et al., 2009). The HsdR subunit is a symbiotic fusion of a RecA-like helicase fold characteristic for SF2 helicases and an endonuclease unit with an auxiliary helical domain at the C-terminal of the single polypeptide chain. ATPase and translocation activities of the motor subunit are determined by two

helicase/translocase domains clustering helicase/translocase characteristic motifs, which are conserved throughout Type I and the SF2 superfamily of DEAD-box helicases (Gorbalenya & Koonin, 1993). The endonuclease domain shares a common core with Type II restriction endonuclease of the PD-(E/D)xK superfamily and its catalytic residues are involved in DNA cleavage (Niv et al., 2007).

In SF2 helicases no additional domains or complex formation with other proteins is needed for translocation activity. The same is true for endonuclease activity by Type II restriction endonucleases. They act as isolated proteins and a single enzyme bearing one activity only carries out translocation or restriction function. For example, isolated Rad54 translocase from SWI2/SNF2 family (member of the SF2 helicase superfamily) displays branch migration, translocation on dsDNA and displacement of Rad51 activities in an ATP-dependent manner (Burgess et al., 2013; Zhang et al., 2013). Contrary, in Type I restriction-modification systems these two activities are located on the array of functionally integrated domains. These exhibit mutual specific contacts in form of hydrogen bonds at the helical-helicase 2 domain interface in the available crystal structure. Likely, EcoR124I translocates DNA similarly as Rad54 does, where DNA advancement occurs because of a rotational movement of the helicase 2 domain. This would predict the possibility of free rotation for the helicase 2 domain. However, at least in the conformation presented in the crystal structure, the helicase 2 domain exhibits specific interdomain contacts that thus might be functionally relevant and stabilizing a specific conformation in the translocation cycle.

In this work the potential contribution of the residues participating in helical-helicase 2 domain interactions either to the motor activity of the helicase domains, or the stabilization of a particular translocation stage was probed. These aims required:

1. Site-directed mutagenesis of key residues at the helical-helicase 2 domain interface;
2. Testing of expression conditions for mutant HsdR subunits;
3. Expression and purification of WT and mutant HsdR subunits;
4. Expression and purification of the methyltransferase;

5. Determination of restriction phenotype of host cells expressing mutant HsdRs (*in vivo* restricton assay);
6. Examining of ATPase and translocation activity of mutant enzymes *in vitro*;
7. Characterization of the endonuclease activity of mutant enzymes *in vitro*.
8. Computational modeling to interpret the gained experimental data, including bioinformatics approaches (alignments, data mining), homology modeling (including a Frankenstein monster approach) and molecular dynamics.

# 1. INTRODUCTION

## 1.1. Phenomenon of restriction and modification

Usually DNA of bacteria is labeled specifically and when transfer of genomic material between bacterial cells from different strains within one species occurs, incoming DNA without identity information of the host cell is determined as suspicious and dangerous for cell safety and to be destroyed (Murray, 2000). For the first time, this phenomenon was demonstrated by Bertani and Weige in 1953. The temperate phages P2 and λ were propagated in *E. coli* strains B and C and then used for infection of *E. coli* K-12 strain. Thus, unlabeled DNA was transferred into the cell and it triggered activation of the defense mechanism within the cell (Bertani & Weigle, 1953). By chance, defined bases of phage DNA can be modified at specific sites by host modification enzymes. After that methylated phage DNA is recognized as host own DNA and protected against degradation by endonuclease. Such properties give a phage an opportunity to survive and propagate further (Arber & Dussoix, 1962; Smith et al., 1972). Later it was shown *in vitro* that this phenomenon has enzymatic origin. EcoKI was the first Type I restriction-modification (R-M) system isolated from *E. coli* K-12 and its ability to degrade DNA was demonstrated *in vitro* (Meselson & Yuan, 1968; Linn & Arber, 1968).

## 1.2. Classification of R-M systems

Many useful features of restriction enzymes were characterized and this group of proteins gained popularity among molecular biologists all over the world. Restriction enzymes were found with specificity for certain viruses and only within prokaryotes. Analysis of bacterial genomes suggests that all free-living bacteria and the Archaea apparently have genes to encode for restriction enzymes (Murray, 2000; Roberts et al., 2003). Due to the property to cleave close or within the recognition site endonucleases from Type II restriction enzymes became popular with molecular biologists, and this branch of science was developing rapidly and by 1982 about 350 restriction enzymes and their 90 target sequences were characterized (Linn & Roberts, 1982). Nowadays REBASE numbers 3945 characterized restriction enzymes which include 3834 Type II restriction enzymes with 299 distinct known specificities.

Biotechnological companies offer 641 restriction enzymes with 235 distinct specificities for molecular biology needs (Roberts et al., 2010).

**Table 1.** Types of R-M enzymes (adapted from Loenen et al., 2014a).

| *Feature* | *Type I* | *Type II* | *Type III* | *Type IV* |
|---|---|---|---|---|
| **Structural:** Subunit | Tree different | Two identical | Two different | Two different |
| Enzyme activity | Endonuclease, Methyltransferase ATPase | Endonuclease or Methyltransferase | Endonuclease, Methyltransferase ATPase | Endonuclease, GTPase |
| **Biochemical:** Cofactor for DNA cleavage | ATP, SAM, $Mg^{2+}$ | $Mg^{2+}$ | ATP, $Mg^{2+}$ (SAM) | $Mg^{2+}$ , GTP |
| Methylation | SAM, $Mg^{2+}$ | SAM | SAM, $Mg^{2+}$ | — |
| Recognition sequence | Asymmetric, bipartite | Usually symmetric | Asymmetric | Bipartite, methylated |
| Cleavage site | Random, at least 1000 bp from recognition site | At or near recognition site | 25-27 bp from recognition site | Between methylated bases at multiple positions |
| **DNA translocation** | Yes | No | Yes | Yes |

Differences in cofactor requirements, subunit composition, sequence recognition specificity and cleavage position of R-M enzymes were taken into account and nowadays R-M systems are divided into four Types (I-IV) (Table 1). R-M enzymes, which first were found in *E. Coli* K-12 and *E. coli* B, are included in Type I and those restriction enzymes which are used as molecular tools in modern biology belong to Type II (Loenen et al., 2014a).

### 1.2.1. Type I

Type I R-M systems are composed of three different subunits. One S (Specifity) subunit is required for identification of the target sequence on dsDNA and tightly anchors the enzyme complex on dsDNA substrate. Two M (Methylation) subunits in cooperation with the S subunit form the methyltransferase (MTase) complex and are responsible for host genome modification, after which labeled host DNA is not degraded by host R-M enzymes anymore. Two R (Restriction) subunits assemble with the MTase and house DNA translocation and consequent cleavage of dsDNA substrate activities (Murray, 2000; Dryden et al., 2001). The methylation status of DNA determines whether Type I R-M enzyme carries out restriction or modification function. Unmethylated DNA is considered as foreign and the restriction mode is activated, which results in DNA translocation and consequent cleavage; hemi-methylated DNA is recognized as host own DNA and is then fully methylated by the complex, and fully methylated DNA doessn't lead to any enzymatic activity of Type I R-M complex (Vovis et al., 1974). S-adenosyl methionine (SAM) and $Mg^{2+}$ are required for MTase activity to methylate the $N_6$ position of adenine within specific bipartite asymmetric target sequence. Restriction event requires the presence of SAM, ATP and $Mg^{2+}$. The type I R-M complex is tightly bound to recognition site on DNA and cleaves DNA substrate distantly from recognition site after dsDNA is translocated by R-subunits in ATP-dependent manner. Cleavage of dsDNA occurs when further translocation is impossible, for example, collision of two translocating enzymes can provoke it (Janscak et al., 1999).

### 1.2.2. Type II

Type II restriction enzymes recognize short target sequences (usually 4-8 bp) and introduce a double strand cut within the recognition site or in close proximity to it. Type II restrictions enzymes are homodimeric, and $Mg^{2+}$ is required for their activity. Based on their structural and functional specificities they are divided into subtypes. Type II systems consist of two separate enzymes R (restriction function) and M (methylation function), which act independently in target recognition and catalysis. These enzymes have attracted huge attention

in protein-DNA interaction studies and are widely used as a powerful tool in recombinant DNA technology due to their particular manner of DNA cleavage (Roberts et al., 2003; Pingoud et al., 2005).

### 1.2.3. Type III

Type III R-M systems are encoded by *res* and *mod* genes and share some degree of similarity with Type I R-M enzymes but are less complex. The mod-subunit is required for specific binding and methylates $N_6$ of adenine in recognition sequences using SAM as donor of methyl groups. Res-subunit consumes ATP in the presence of $Mg^{2+}$ and cleaves DNA. Both activities are possible only when the Res-subunit and Mod-subunit are assembled in a complex with stoichiometry $Res_2Mod_2$ (Janscak et al., 2001). Type III R-M enzymes recognize two inverted copies of specific sequence 5-6 bp in length. During ATP-dependent translocation the $Res_2Mod_2$ complex stays bound to the recognition sequence. A stalling event or collision of two $Res_2Mod_2$ complexes provokes DNA cleavage 25-27 bp away from recognition site, each complex cuts one DNA strand (Tock & Dryden, 2005).

### 1.2.4. Type IV

Methylated, hydroxymethylated and glucosyl-hydroxymethylated DNA species are substrates for Type IV R-M systems, which numbers 227 putative enzymes. These enzymes require $Mg^{2+}$ and GTP for translocation activity. Type IV R-M enzymes recognize two copies of dinucleotide sequence, which includes a purine followed by a cytosine methylated at either the $N_4$ or the $C_5$ position, separated by 40 to 3000 nucleotides. The enzyme remains attached to recognition sequence during translocation. A stalling event results in DNA cleavage, which takes place preferentially 30 bp away from either of two sites (Tock & Dryden, 2005).

## 1.3. Type I R-M systems. Classification

Based on cross-hybridization of genes and antibody cross-reactivity Type I R-M systems are divided into four families (Table 2): Type IA (e.g.

EcoKI), Type IB (e.g. EcoAI), Type IC (e.g. EcoR124I) and Type ID (e.g. StySBLI). Also EcoR124II, EcoDXXI and EcoprrI belong to the Type IC family of R-M enzymes (Piekarowicz et al., 1985; Price et al., 1987; Skrzypek & Piekarowicz, 1989; Tyndall et al., 1994; Barcus et al., 1995, Loenen et al., 2014b).

**Table 2.** Type I restriction-modification systems (adapted from Loenen et al., 2014b).

| Family | Emzyme | Recognition sequence* | Organism | Salient features |
|---|---|---|---|---|
| **A** | EcoKI<br>EcoAI<br>StySBI<br>EcoID | AAC(N)$_6$GTGC<br>TGA(N)$_8$TGCT<br>GAG(N)$_6$RTAYG<br>TTA(N)$_7$GTCY | *E. coli*, *Klebsiella* and *Salmonella* species | Chromosomally enconded, allelic |
| **B** | EcoAI<br>CfrAI<br>StySKI | GAG(N)$_7$GTCA<br>GCA(N)$_6$GTGG<br>CGAT(N)$_7$GTTA | *E. coli*, *C. freundii* | Chromosomally enconded, allelic |
| **C** | EcoR124I<br>EcoR124/3<br>EcoDXXI<br>KpnBI<br>NgoAV | GAA(N)$_6$RTCG<br>GAA(N)$_7$RTCG<br>TCA(N)$_7$RTTC<br>CAAA(N)$_6$RTCA<br>GCA(N)$_8$TGC | *E. coli*, *Klebsiella* and *Neisseria* species | *E. coli* conjugative plasmids, nonallelic |
| **D** | StySBLI<br>KpnAI<br>EcoR9I<br>LIdI | CGA(N)$_6$TACC<br>CAA(N)$_6$TGCC | *E. coli*, *Klebsiella* and *Salmonella* species<br>*Lactococcus lactis* | Novel specificity |

*N, any nucleotide; R, either purine; Y, either pirimidine

Three subunits HsdR, HsdM and HsdS are encoded by three genes ***hsdR***, ***hsdM*** and ***hsdS***, respectively, which are expressed from two promoters: initiation of ***hsdR*** gene transcription occurs from $P_{RES}$ promoter, ***hsdM*** and ***hsdS*** genes have common $P_{MOD}$ promoter. For the first time restriction activity of R-M system encoded by plasmid R124 was noticed by Bannister, she demonstrated negative influence of R-M system expressed from this plasmid on growth of bacteriophages T1, *Φ*80, P1 and λ. Later isolated enzyme expressed

from R124 plasmid was named EcoR124, but originally *Salmonella enterica* serovar Typhimurium was the source of R124 plasmid and the enzyme should have been named StyR124 (Youell & Firman, 2008).

Correct functionality of Type I R-M enzymes is determined by complex assembly and their localization in the bacterial cell (Firman et al., 2000). Restriction activity of EcoR124I can be turned off by dissociation of $R_2M_2S_1$ complex with release of one R subunit into solution. It was proposed that it may be special way to control modification and restriction functions of EcoR124I and also it could be important for transmissibility of the plasmid R124 (Firman et al., 2000; Janscak & Bickle, 2000). The restriction alleviation phenomenon was considered as another mechanism of control of restriction and modification activities through proteolysis of the HsdR subunit. Being family dependent restriction alleviation was shown for EcoKI R-M system but not detected for EcoR124I (Doronina & Murray, 2001; Blackely & Murray, 2006). The level of restriction alleviation for each family of Type I R-M enzymes might depend on enzyme localization in the host cell (Holubova et al., 2000, 2004). Thus, part of EcoR124I complexes is localized on the periplasmic side of the cytoplasmic membrane, part of EcoKI is present in periplasm and EcoAI was not detected there. The presence of R-M enzymes at the cell periphery may separate R-M functions and the periplasm can be the first place where invading DNA is recognized and degraded by Type I R-M enzymes. It was observed that EcoKI binds to the first 850 bp of bacteriophage T7 DNA and acts as molecular motor translocating DNA into the cell (Garcia & Molineux, 1999).

### 1.3.1. S-subunit of EcoR124I

EcoR124/3 R-M enzyme encoded by plasmid R124/3 belongs to the same family as EcoR124I and shows alternative specificity. S-subunits of both enzymes recognize identical target sequences, but non-specific spacer of EcoR124/3 target sequence includes one additional non-defined nucleotide (Price et al., 1987a).

EcoR124I: 5' – GAANNNNNNRTCG – 3' (or $GAAN_6RTCG$),

EcoR124/3: 5' – GAANNNNNNNRTCG – 3' (or $GAAN_7RTCG$),

where R is either purine.

DNA sequences of genes encoded for S-subunits of EcoR124I and EcoR124/3 are almost identical with the only differences being in the central region, where a 12 bp sequence is repeated two times in the *hsdS* gene of EcoR124I and the *hsdS* gen of EcoR124/3 has three 12 bp repeats (Fig. 1) (Price et al., 1987b).

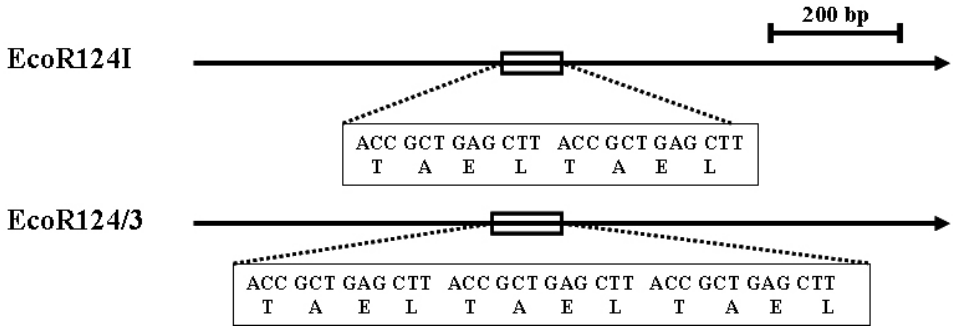



**Figure 1.** The central region of *hsdS* genes of EcoR124I and EcoR124/3. There are two 12 bp repeats in *hsdS* gen of EcoR124I and one 12 bp extra repeat is in *hsdS* gen of EcoR124/3, those repeats are reflected in T-A-E-L sequence in primary protein structure of S-subunit.

The 46 kDa HsdS subunit of EcoR124I is insoluble by itself and its normal functionality is possible only when bound to two HsdM subunits within the MTase complex what causes difficulties for structural studies of the S-subunit alone (Patel et al., 1992, Taylor et al., 1992). Sequence analysis of HsdS subunits of Type I R-M enzymes has shown that there are two variable regions (Fig. 2) which are organized into two target recognition domains (TRD1 and TRD2) separated by a central conserved region, each domain contacts one-half of specific sequence (Taylor et al., 2010). Trinucleotide at the 5'-end of bipartite sequence is recognized by N-terminal TRD and C-terminal TRD interacts with tetranucleotide at the 3'-end. Independent manner of functionality of each TRD allows them to interchange, and a novel hybrid specificity can be generated (Fuller-Pace et al., 1984; Nagaraja et al., 1985; Cowan et al., 1989). Indeed, modified S-subunit, containing only one TRD, is unable to recognize target sequence. Dimerization of such truncated HsdS subunits allows creating a complex with functional specificity to a symmetric bipartite target sequence (O'Neill et al., 1998). In the central region and at N-/C-termini of S-subunit are regions conserved within a given family, their sizes

vary from family to family. Since the pentameric Type I R-M complex can interchange with its S-subunits within one family, it was hypothesized that these conserved regions are necessary for protein-protein interactions of S-subunit with the other subunits (Janscak et al., 1998, Abadjieva et al., 2003). The sequence of the HsdS subunit is "capped" at the N- and C-terminal by conserved regions and that link the HsdS subunit to one M-subunit (Fig. 3). The central conserved region would then interact with the second M-subunit and a structure with a 2-fold rotational symmetry is formed (Kneale, 1994). This idea is supported experimentally by closing of the sequence in a circle at the N- and C-terminal conserved regions of the HsdS subunit of EcoAI (Janscak & Bickle, 1998) and MG438 (Calisto et al., 2005).

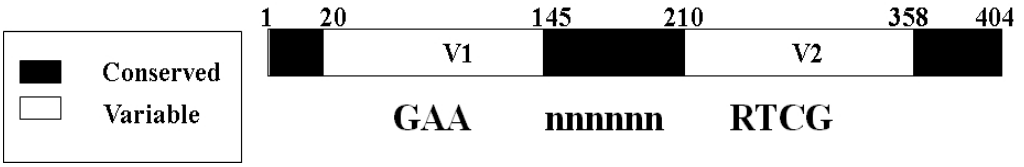



**Figure 2.** The arrangement of conserved and variable domains in HsdS subunit of EcoR124I. Tree-nucleotide and tetra-nucleotide of target sequence of EcoR124I are recognized by variable region 1 and 2, respectively. R is either purine.

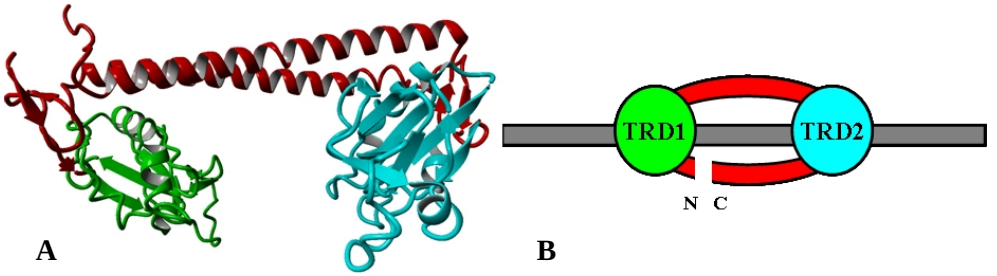



**Figure 3.** HsdS subunit of EcoR124I. A) Predicted 3D structure of HsdS subunit with mapped conserved (red) regions. Variable domain 1 and 2 form target recognition domains 1 (TRD1, green) and 2 (TRD2, cyan), respectively (M.S.EcoR124I.DNA.pdb, ftp://genesilico.pl/iamb/models/MTases/M.EcoR124I, Obarska et al., 2006. Figure prepared in YASARA). B) Schematic representation of HsdS subunit reflects its approximately 2-fold symmetry: grey bar represents DNA with bound TRD1 and TRD2, central and distal conserved regions are closed in "circle" through N- and C-terminal.

It has been shown that central conserved region of S-subunit of EcoR124I is involved in interaction with second M-subunit. Change of amino acid residues in this region gives unpredictable alteration of recognition properties suggesting that this central domain is not just α-helical linker between TRDs but plays a more important role (Gubler & Bickle, 1991). The non-classical point mutation $Trp^{212}$ to Arg at the distal border of the central conserved region near TRD2 of HsdS subunit of EcoR124I significantly alters the ability of the S-subunit in the MTase complex for binding to target sequence. Interestingly, HsdR subunits are able not only to restore binding function of the S-subunit with such mutation but also the $Mod^+$ phenotype (Weiserova & Firman, 1998; Weiserova et al., 2000). Mutagenesis of the *hsdS* gene of EcoKI in the critical interval between TRD2 and the central conserved region revealed temperature-sensitive mutations, which displayed a similar phenotype, and demonstrated the importance of this region of the S-subunit in subunit assembly (Zinkevich et al., 1992; Janscak et al., 2000).

### 1.3.2. M-subunit of EcoR124I

One copy of the HsdS subunit (46 kDa) and two copies of the HsdM subunit (58 kDa) assemble in solution into a trimeric enzyme, the methyltransferase or M.EcoR124I or MTase complex (162 kDa) with subunit stochiometry $M_2S_1$. HsdM subunits of Type I R-M systems contain a motif that is common for all methyltransferases, and mutational analysis shows that this motif is essential for modification activity (Willcock et al., 1994; Patel et al., 1992). M.EcoR124I binds to target DNA sequence $GAAN_6RTCG$ with affinity of $10^8$ $M^{-1}$ (Taylor et al., 1992) what is four orders higher than affinity to non-specific DNA sequence. Methylation of adenine at either side of bipartite sequence reduces the binding affinity to cognate recognition site by approximately 30-fold and negligible influence on methylation activity is detected when the recognition sequence is methylated in other positions. Hemi-methylated DNA substrate increases the rate of the methylation reaction of the enzyme despite the decrease in binding affinity (Taylor et al., 1993).

X-ray solution scattering experiments showed that M.EcoR124I shrinks in all dimensions when the enzyme binds to DNA with a target sequence, and SAM as a cofactor has no influence on DNA-driven compactization. It was

proposed that the structural change in conformation might be caused by a large rotation of HsdM subunit during binding to DNA target sequence. This would allow to grab the DNA at the target sequence and then to wrap the protein around via non-specific interactions of the MTase outside the target sequence. Together with changes in the enzyme arrangement, the increased circular dichromism signal also suggests an alteration of DNA conformation (Taylor et al., 1994).

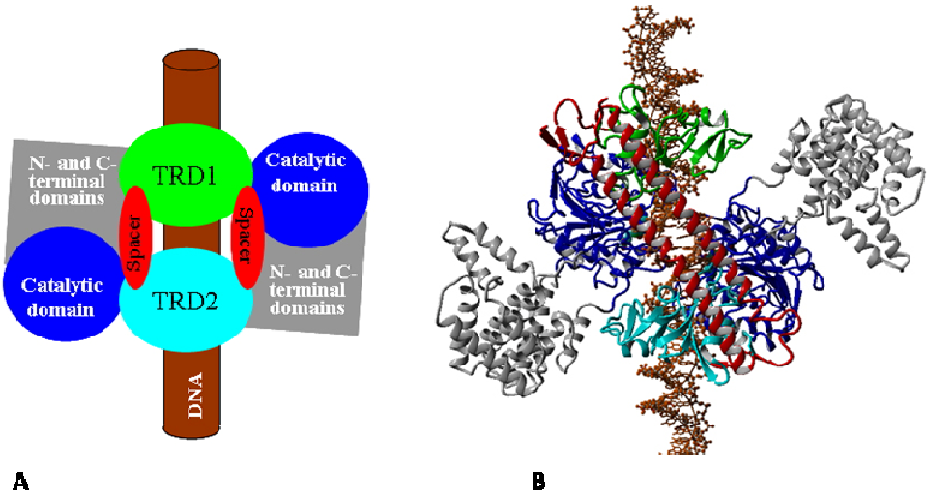



**Figure 4.** Trimetric complex of methyltransferase. A) Cartoon model of Type I MTase bound to DNA. S-subunit is comprised of two target recognition domains (TRDs) joined by two spacers formed by distal and central conserved regions. M-subunits are placed on either side of the DNA helical axis (brown tube) and face S-subunit. Catalytic domains of each M-subunit are associated with TRDs. B) Predicted 3D model of methyltransferase of EcoR124I from Type IC family (colors as on panel A)
(M.S.EcoR124I.DNA.pdb, ftp://genesilico.pl/iamb/models/MTases/M.EcoR124I, Obarska et al., 2006. Figure prepared in YASARA).

It was shown using proteolysis analysis that M.EcoKI from Type IA family has a domain structure, and HsdS and HsdM subunits are equally sensitive to proteases (Cooper & Dryden, 1994). In contrast to M.EcoKI, susceptibility of the S-subunit of M.EcoR124I to limited proteolysis is much higher than that for the M subunits. Proteolysis is not affected by the presence of the cofactor SAM, however susceptibility to proteolysis is reduced when

MTase is bound to DNA. Trypsinated MTase contains S-subunit with proteolised domains and two intact M-subunits preserving the trimeric complex assembly, but binding properties of such complex become worse due to a decrease in DNA sequence specificity. DNA bound to the MTase complex gives significant protection from digestion by trypsin and chymotrypsin. DNA-protein complex of M.EcoR124I is more compact in comparison to free MTse (Webb et al., 1995). Binding of M.EcoKI to DNA does not lower significantly susceptibility to proteases. Limited proteolysis suggests that there are no conformational changes associated with cofactor binding in M.EcoR124I (Cooper & Dryden, 1994). This was confirmed by the small angle X-ray scattering data, SAM and its analogues have no influence on structure arrangement of M.EcoR124I (Taylor et al., 1994). In comparison to M.EcoR124I, binding of SAM provokes considerable changes in M.EcoKI structure. M.EcoKI with bound SAM has an increased affinity to its cognate target sequence (Powell et al., 1993), but such effect was not detected for M.EcoR124I (Webb et al., 1995).

Symmetry of the MTase complex is influenced by interactions between two M-subunits (Fig. 4) (Abadjieva et al., 1994), free M-subunits do not interact, and dimerization of M-subunits was not detected without S-subunit (Taylor et al., 1992). Thus, the required symmetry is most likely provided by the S-subunit and taking into account the presence of single S- and two M-subunits (Kneale, 1994). Crystallographic studies of S-subunit of hypothetical Type I R-M systems MgeORF438P (ORF MG3435) from *Mycoplasma genitalium* (Calisto et al., 2005) and MjaXIP (ORF MJ0130m) from *Methanococcus jannaschii* (Kim et al., 2005) also support this hypothesis.

### 1.3.3. R-subunit of EcoR124I

Expressed HsdR subunit of EcoR124I (~120 kDa) is located in soluble fraction. Gel filtration shows that the HsdR subunit is present in solution as a monomeric protein and no oligomerization is detected. It was shown that the isolated HsdR subunit is able to hydrolyze ATP, however the rate of ATP hydrolysis is negligible and requires the presence of DNA and $Mg^{2+}$. Also purified HsdR subunit was found to possess a weak nuclease activity, it cleaves fully modified pDRM-1R (circular plasmid DNA with a single EcoR124I

recognition site) and non-specific cccDNA (Zinkevich et al., 1997). Despite the fact that the R-subunit alone shows weak, residual nuclease activity and is able to hydrolyze ATP, site-specific binding and DNA translocation with consequent DNA cleavage are possible only when two R-subunits together and MTase form EcoR124I R-M complex with stoichiometry of $R_2M_2S_1$ (Janscak et al., 1996; Seidel et al., 2005). The HsdR subunit in the EcoR124I R-M complex displays a by far higher rate of ATP hydrolysis than the isolated subunit and ATP hydrolysis is coupled to DNA translocation, ~1 ATP molecule is hydrolyzed for 1 bp to be advanced (Seidel et al., 2008). EcoR124I forms long DNA loops as the HsdR motors can translocate dsDNA along the 3' – 5' strand up to several thousand base pairs prior to introducing a double strand break (Firman & Szczelkun, 2000; Seidel et al., 2004; van Noort et al., 2004; Stanley et al., 2006). Usually endonuclease activity is triggered by collision with another translocating complex (Janscak et al., 1996; Szczelkun et al., 1996, 1997). Thus, full functionality of HsdR subunit requires assembly of a complete pentameric protein complex. That might be a way of nuclease activity control *in vivo* (Seidel et al., 2005).

Prior to X-ray crystal structure determination of the HsdR subunit of EcoR124I four distinct domains were identified based on sequence analysis (Fig. 5): a nuclease domain at the N-terminal or “Region X”; a helicase domain that contains two RecA-like folds and motifs typical for Superfamily 2 (SF2) helicases; and a C-terminal α-helical domain. Hereby, “region X” of HsdR of EcoR124I according to biochemical and bioinformatics analysis was considered partially to be a RecB-like nuclease from the PD-(E/D)xK superfamily. Motifs similar to motifs I, II, III and QxxxY, characteristic of RecB-like nucleases, were found in repair and recombination nucleases. HsdR of EcoR124I and many others Type I R-M enzymes have similar amino acid arrangement (Sisáková et al., 2008a).

Importance of conserved residues in motifs II and III was demonstrated earlier for Type I R-M enzymes EcoAI and EcoKI. Substitution of conserved residues D161 (motif II) and E76 and K78 (motif III) in EcoAI to an alanine knocks out endonuclease activity of the R-M complex and the presence of such HsdR mutants in the cell leads to an $r^-$ phenotype. A K78A mutation results in nicking activity *in vitro*, while ATPase and translocation activities are preserved

(Janscak et al., 1999; McClelland et al., 2005). The HsdR subunit of EcoKI with mutation D298E (motif II), E312D, E312H or K314A (motif III) was found to be restriction deficient *in vivo* and *in vitro*. Nevertheless, ATP hydrolysis rate and translocation activity were similar to WT, except for an E312H mutant enzyme, which exhibited a large decrease in ATPase rate and it was not possible to measure translocation activity (Davies et al., 1999; McClelland et al., 2005).

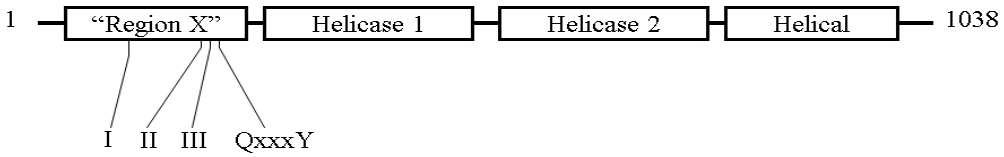



**Figure 5.** Domain organization of HsdR subunit of EcoR124I. PD-(E/D)xK superfamily fold of endonuclease domain or "Region X" clusters motifs I, II, III and QxxxY. Two RecA-like domains (helicase domain 1 and 2) are involved in ATPase/translocation activity. The α-helix rich C-terminal domain is necessary for interactions with methyltransferase of EcoR124I. Based on properties of HsdR subunit EcoR124I belongs to SF2 helicases and defined as SF2Aβ helicase: "A" indicates polarity of movement on DNA in 3' - 5' direction and "β" denotes double stranded character of substrtate for translocation.

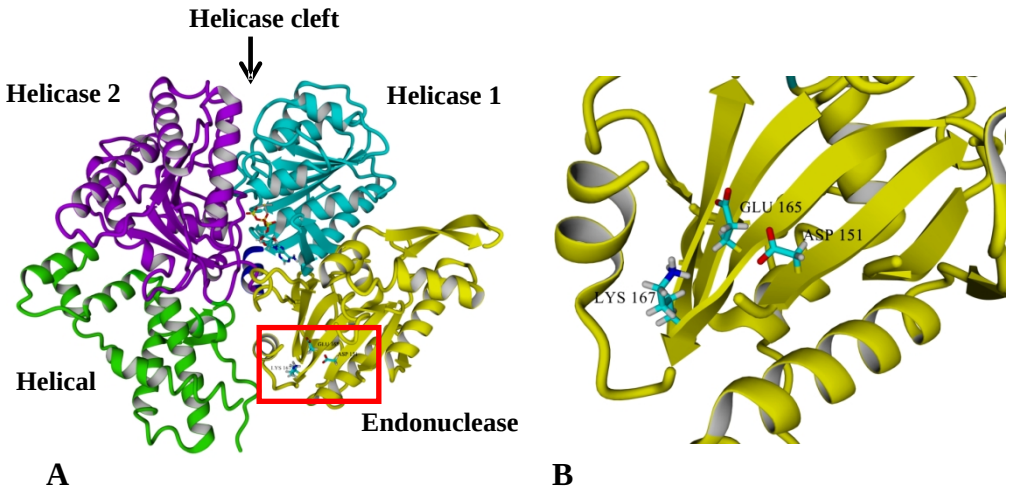



**Figure 6.** The crystal structure of HsdR subunit of EcoR124I R-M enzyme (pdb id: 2W00, Lapkouski et al., 2009). A) N-terminal endonuclease domain is in yellow; two RecA-like helicase domain 1 and 2 are in cyan and magenta, respectively; C-terminal helical domain

is in green (Front view). 886-1038 residues are unrelolved. B) Zoom in on panel A (red rectangle): The αβα core, typical for Type II restriction enzymes, forms the endonuclease domain. Side chains of residues D151, E165 and K167 forming endonuclease active site are shown (nitrogen is in blue, oxygen is in red and carbon is in cyan). Figure prepared in YASARA.

Conserved amino acid residues D151, E165 and K167 from II and III motifs of HsdR of EcoR124I were proposed to play a key role in dsDNA cleavage. Mutational analysis of amino acid residues from the endonuclease domain displays significant alteration of dsDNA translocation despite the distant location from helicase domains (Sisáková et al., 2008b).

Recently the HsdR subunit of EcoR124I was crystallized, the crystal structure was solved and described (Lapkouski et al., 2007, 2009). The HsdR subunit is comprised of four distinct globular domains in a square-planar arrangement (Fig. 6): the endonuclease domain (residues 13-260), the helicase domain 1 and 2 (261-461 and 470-731), and the helical domain (732-892), which are separated by prominent grooves. The C-terminal part from 893 to 1038 remains unresolved, however was confirmed to be present in the crystal. Two RecA-like helicase domains of HsdR form the helicase cleft, similarly as in related SF2 helicases co-crystalized with DNA (Dür et al., 2005). It was proposed that the endonuclease and the helicase activities are potentially coupled though K220 residue from α8-helix in the endonuclease domain, which has a novel contact to ATP within hydrogen-bonding distance (3.1 Å) what is not peculiar for other helicase-like proteins with known structure (Lapkouski et al., 2009).

## 1.4. EcoR124I R-M complex

### 1.4.1. Subunit assembly and activity control

The EcoR124I R-M complex consists of three different subunits: HsdS, HsdM and HsdR encoded by ***hsdS***, ***hsdM*** and ***hsdR*** genes, respectively. MTase of EcoR124I was expressed and purified (Patel et al., 1992; Taylor et al., 1992) and together with purified HsdR subunit can be used for ***in vitro*** reconstitution of EcoR124I R-M complex ***in vitro***. Subunit assembly analysis detected just the $M_2S_1$ form of MTase, no complex with a stoichiometry of $M_1S_1$ has been shown

to exist (Taylor et al., 1992). In contrast to MTase of EcoR124I, MTase of EcoKI with stoichiometry $M_2S_1$ is weak and dissociates into $M_1S_1$ form with release of free HsdM subunit losing ability for methylation function (Meselson et al., 1972; Dryden et al., 1993, 1997).

ATPase and restriction activities require the fully assembled pentameric complex of EcoR124I with a stoichiometry of $R_2M_2S_1$ (Fig. 7). The intermediate complex with a stoichiometry of $R_1M_2S_1$ is formed first in the cell. Affinity of the first HsdR subunit to the MTase is much stronger than the affinity of the second HsdR subunit. The $R_1$ complex is not able to cleave dsDNA. The fully assembled complex with stoichiometry of $R_2M_2S_1$ dissociates into $R_1M_2S_1$ and $R_1$ with Kd of ~$2.4 \times 10^{-7}$ M (Janscak et al., 1998). Surface plasmom resonance analysis suggests that MTase of EcoR124I has two non-equivalent binding sites for HsdR subunit (Mernagh et al., 1998). In comparison to EcoR124I, EcoKI endonuclease is presented by stable $R_2M_2S_1$ complex, each HsdR subunit has similar binding affinity to MTase (Dryden et al., 1997). The EcoBI endonuclease has a variety of stoichiometric forms including $R_2M_2S_1$, $R_1M_2S_1$ and $R_1M_1S_1$ (Eskin & Linn, 1972). Purified EcoAI endonuclease is weakly assembled and dissociates into MTase and HsdR subunits, only reconstitution of whole enzyme complex *in vitro* gives active endonuclease suitable for biochemical characterization (Suri et al., 1984).

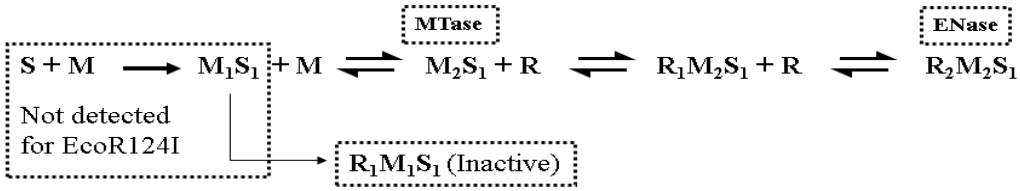



**Figure 7.** Subunit assembly for Type I R-M enzymes. First HsdR subunit of EcoR124I binds to MTase with higher affinity and forms pretty stable intermediate $R_1$ complex which is not able to cleave DNA substrate. Second motor subunit interacts with $R_1$ complex weaker and dissociates from $R_2$ complex with Kd of ~$2.4 \times 10^{-7}$ M.

Since binding affinity of each HsdR subunit of EcoR124I is different, the ratio of MTase and the HsdR subunit in the cell defines whether $R_2M_2S_1$ or $R_1M_2S_1$ complex is formed. The $R_2M_2S_1$ is formed only if HsdR:MTase ratio is >1. It was confirmed that production of endonuclease from *hsd* genes,

controlled by their natural promoters, results in the presence of two forms $R_1M_2S_1$ and $R_2M_2S_1$. Since the $R_1M_2S_1$ complex is not able to cleave DNA, this deficiency might be considered as a way of control of endonuclease activity of EcoR124I R-M complex (Janscak et al., 1998). The conjugative R124 plasmid is the original source of the EcoR124I endonuclease (Hedges & Datta, 1972). Its transfer into a new cell with unmodified DNA would lead to expression of the active endonuclease with consequent degradation of host DNA. However, instead to be hydrolyzed, host DNA undergoes modification and no mechanism of control for this phenomenon for the EcoR124I and other type I R-M systems was observed at the gene level (Kulik & Bickle, 1996). Endonuclease activity control based on subunit assembly was proposed by Janscak (1998). In the new host first a stable complex of MTase is formed. High binding affinity of the first HsdR subunit to MTase prevents formation of weak restriction proficient $R_2M_2S_1$ complex for the first time after *hsd* genes were introduced into a new host. The $R_2M_2S_1$ complex will be formed when no free MTase is left for assembly of $R_1M_2S_1$ complex. Thus, there is time for MTase and/or $R_1M_2S_1$ to modify host DNA, and, as result, cell DNA is protected against degradation by EcoR124I. A lethal effect of excess of HsdR over MTase was displayed by using different promoters and vectors (Janscak et al., 1996). EcoKI was proposed to have a mechanism of activity control via subunit assembly similar to EcoR124I (Dryden et al., 1997). Considering the fact that modification activity of EcoR124I with unmethylated DNA is relatively high, unmodified host DNA with higher probability would be protected by modification activity than hydrolyzed after conjugal transfer of the *hsd* genes of EcoR124I into the cell (Janscak et al., 1996).

### 1.4.2. Cofactor requirements

Previously ATP, SAM and $Mg^{2+}$ were described as cofactors required for DNA cleavage by Type I R-M systems, however it was not fully confirmed for EcoR124I (Price et al., 1987a). Later, the dependence of EcoR124I activities on ATP, SAM and $Mg^{2+}$ was studied by Janscak (1996). It was shown that ATP and $Mg^{2+}$ are absolutely necessary for ATPase and restriction activities. No effect on DNA cleavage was detected in the presence of *S*-adenosil homocysteine (SAH), competitive inhibitor of EcoKI (Burckhardt et al., 1981),

both activities were present in the absence of SAM. In contrast, EcoKI was found to require SAM as an allosteric effector for enzyme activation prior to DNA binding (Hadi et al., 1975). While SAM is not required for EcoR124I as a cofactor, it still stimulates ATP hydrolysis and restriction activity and is considered as a positive effector. Surface plasmom resonance analysis of DNA binding by EcoR124I in the presence and the absence of SAM suggests that SAM stimulates specific DNA binding and leads to slower dissociation of bound enzyme from DNA what consequently has a positive effect on ATPase and restriction activities (Janscak et al., 1996).

HsdM subunit of MTase uses SAM as donor of methyl groups to modify adenine bases within the recognition sequence (Bianco and Hurley, 2005). Hemi-methylated DNA is the most preferred substrate for methylation for EcoKI and EcoR124II endonucleases (Price et al., 1987b), and for methyltransferase of EcoR124I (Taylor et al., 1993). Janscak (1996) has shown that non-methylated DNA is the preferable substrate for EcoR124I binding, however the enzyme methylates hemi-methylated DNA with a five fold higher rate than unmodified DNA. While β,γ-imido ATP, a non-hydrolysable ATP analog, was shown to have no effect on the methylation rate of unmethylated DNA, the presence of ATP reduces the methylation activity and switches EcoR124I to DNA cleavage (Janscak et al., 1996). The same inhibitory effect was observed for EcoKI methylation of unmodified DNA (Suri & Bickle, 1985).

### 1.4.3. Translocation activity of EcoR124I

Despite the fact that the HsdR subunit of EcoR124I has two RecA-like domains which are characteristic for SF2 helicases, and the endonuclease active site (Davies et al., 1999; Janscak et al., 1999), full set of motor subunit activities such as ATPase (Janscak et al., 1998), DNA translocation (Firman & Szczelkun, 2000; Seidel et al., 2004) and the endonuclease (Janscak et al., 1998) is possible only within fully assembled $R_2M_2S_1$ complex (Dryden et al., 1997). Two HsdR subunits associated with MTase translocate dsDNA bidirectionally toward enzyme complex anchored on target sequence and with a rate of several hundred base pairs per second (McClelland et al., 2005) extruding two DNA loops (Yuan et al., 1980). Translocation rates for HsdR in

restriction deficient $R_1M_2S_1$ (Firman & Szczelkun, 2000) and restriction proficient $R_2M_2S_1$ complexes are the same, however, the first complex is three times less efficient (Seidel et al., 2004). Initiation of translocation begins with a short ~8 nm initial DNA loop formation in the presence of ATP (van Noort et al., 2004). Detailed characterization of initiation, termination and re-initiation of translocation was done using bulk stopped-flow experiments and single-molecule experiment using magnetic tweezers (Seidel et al., 2005).

Stopped-flow experiments (Fig. 8 A) imply the use of a fluorescently labeled triplex forming oligonucleotide (TFO) which is bound to DNA at specific triplex binding site (TBS) on the double stranded DNA. Translocation activity of Type I enzyme leads to release of TFO from DNA. A variety of DNA substrates with determined distances between TBS and EcoR124I target sequence together with observed delays of an increase in fluorescence signal of released TFO allow to calculate the translocation rate (Firman & Szczelkun, 2000; McClelland et al., 2005). Speed of translocation of EcoR124I measured by this method is 400±32 bp/sec at 20 °C for each HsdR subunit (Firman & Szczelkun, 2000).

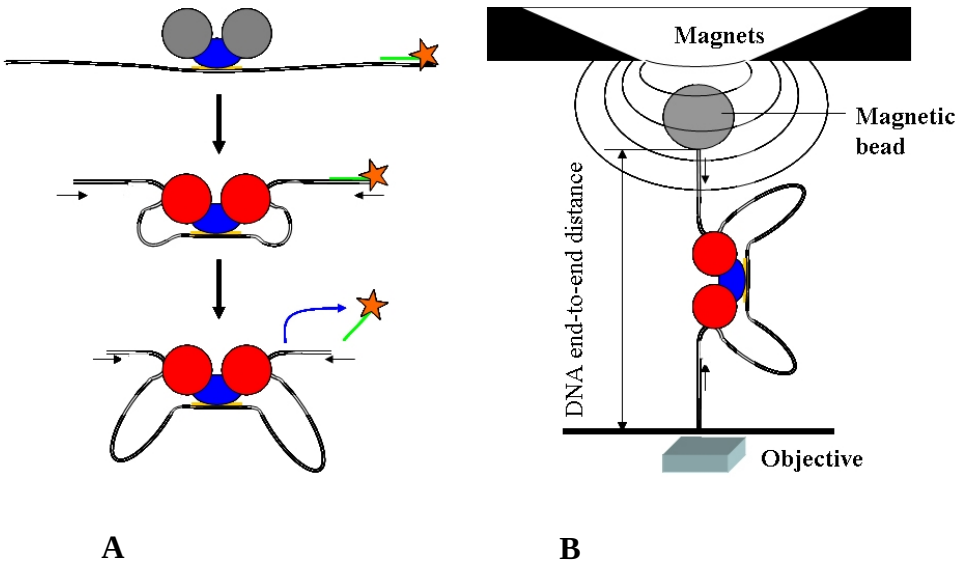



**Figure 8.** Schematic representation of two methods to measure translocation rates of Type I R-M enzymes. A) Triplex displacement assay: a triplex labeled with tetramethylrhodamine is bound to DNA at known distance from EcoR124I target site. Translocation activity makes possible enzyme to displace TFO from DNA and thus an increase of fluorescence of released

TFO is detected by stopped-flow device. B) Magnetic tweezers: one end of dsDNA with single EcoR124I target site is anchored to a glass slide and magnetic bead is attached to opposite DNA end. Magnets stretch DNA molecule. When Type I enzyme translocates dsDNA, decrease in end-to-end DNA molecule distance is observed in real time using video microscopy.

Magnetic tweezers (Fig. 8 B) (Strick et al., 1998) permit to register end-to-end shortening of DNA during translocation of DNA by EcoR124I and distinguish between $R_1$ and $R_2$ translocating complexes. The time between two translocation events provides information about time required for re-initiation. EcoR124I complex with $R_1M_2S_1$ stoichiometry moves dsDNA with speed of 550±30 bp/sec at 4 mM ATP concentration without an influence of force up to 4 pN. Together two motor subunits pull DNA independently with doubled rate what is compatible with bulk measurement. The translocation rate varies with ATP concentration accordingly to Michaelis-Menten kinetics, the $R_1M_2S_1$ complex has $V_{max}$ = 560±20 bp/sec and $K_m$ = 88±7 μM (Seidel et al., 2004). Re-initiation of DNA translocation by EcoR124I is in direct dependence on HsdR subunit concentration in solution, while MTase concentration has no influence on it. HsdR dissociation is considered as final step in termination of translocation, MTase remains attached to target site and free motor subunits from solution are recruited by MTase to re-initiate translocation (Seidel et al., 2005). The presence of ATP is required for initiation of DNA translocation by EcoR124I and it occurs with a $K_m$ of 200-400 μM (van Noort et al., 2004; Seidel et al., 2005).

High DNA translocation processibility of EcoR124I can explain how endonuclease introduce dsDNA break >1000 bp away from recognition site (Janscak et al., 1996; Szczelkun et al., 1997). For example, *in vivo* translocation for EcoKI occurs at the rate of ~100 bp/sec at 30 °C (García & Molineux, 1999), *in vitro* EcoAI translocation speed is ~275 bp/sec (McClelland et al., 2005), DNA unwinding rate of RecBCD helicase *in vitro* is 470±80 bp/sec at 25 °C and ~900 bp/sec at 37 °C (Roman & Kowalczykowski, 1989). In comparison, velocities of single stranded translocation by PriA (Lee & Marians, 1990) and PcrA (Dillingham et al., 2000) are 90 and 50 nt/sec, respectively.

### 1.4.4. Mechanism of DNA cleavage

High excess in molar concentration of Type I endonuclease over linear DNA with a single target site brings to hydrolysis only a small portion of such DNA substrate (Murray et al., 1973; Szczelkun et al., 1996), and there is no cleavage at standard conditions (Rosamond et al., 1979; Dreier et al., 1996). Type I R-M enzymes cleave efficiently single site circular DNA substrate and linear DNA with more than one target site (Rasamond et al., 1979; Studier & Bandyopadhyay, 1988; Drier et al., 1996; Janscak et al., 1996; Szczelkun et al., 1997). Firstly, it was proposed that the endonuclease introduces a double strand break at the meeting point of two translocating enzymes bound specifically to the substrate with two target sites or that a non-specifically bound complex coming from solution is required for a single-site DNA substrate, and the cleavage event is provoked by specific protein-protein interactions of these two complexes (Studier & Bandyopadhyay, 1988). EcoR124II and EcoDXXI R-M enzymes demonstrate such cooperation in the experiments with two-site linear DNA substrate carrying one site for each enzyme (Dreier et al., 1996). Both enzymes are from the same type IC family, show a high degree of similarity and are able to interchange with their HsdR subunits (Janscak et al., 1999). Similar experiments were done with endonucleases from different families. EcoAI(IB)-EcoKI(IA) and EcoKI(IA)-EcoR124I(IC) mixtures suggest that collision of two translocating enzymes triggers DNA cleavage and specific protein-protein interactions are not required.

More detailed investigation of the mechanism of DNA cleavage in connection with translocation was done using Holliday junctions and DNA substrates with a higher degree of supercoiling which potentially present difficulties for DNA translocation. Higher degrees of positive and negative supercoiling of substrates were found to reduce rate of DNA cleavage by EcoAI and EcoR124I, thus the potential influence of supercoiling on DNA translocation rate is not excluded (Janscak et al., 1996, 1999). However, a Holliday junction creating a physical barrier for DNA translocation by Type I enzymes provokes the endonuclease to introduce a double strand cut either on circular (α-structure) or linear (χ-structure) substrates suggesting that only a

stalling event is necessary for triggering of DNA cleavage (Janscak et al., 1999).

Based on the above mentioned experiments a model for DNA cleavage by Type I R-M enzymes was proposed (Fig. 10) (Janscak et al., 1999). Any physical barrier that stalls or impedes DNA translocation promotes cleavage of DNA substrate by Type I R-M enzymes. Protein-protein interactions are not required as it was proposed in the collision model for DNA cleavage by Studier and Badyopadhyay in 1988. Other proteins bound non-specifically to DNA do not lead to DNA cleavage and easily can be displaced by Type I translocating enzymes. Linear DNA with placed nick between the two recognition sites was cleaved by EcoR124II preferentially at or in close proximity of the nick. However, single-site DNA with site-specific nick is not cleaved by EcoR124II suggesting that short interruption of translocation by the nick does not trigger DNA cleavage (Dreier et al., 1996).

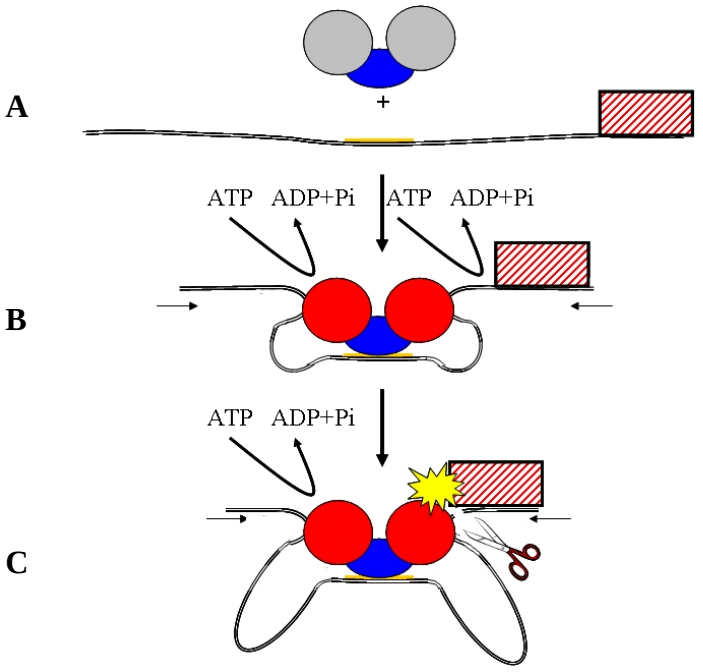



**Figure 10.** A model for DNA cleavage by Type I R-M enzymes. A) MTase with stoichiometry of $M_2S_1$ (blue oval) with two R-subunits (grey circles or red when translocation is initiated as on panel B and C) binds to an unmethylated target site of Type I R-M enzyme (yellow box) on DNA (black line) and hatched box is a barrier that can stall translocation. B) Pentameric enzyme tightly bound on substrate moves dsDNA toward itself, R-subunits hydrolyze ATP and produce long DNA loops. C) Collision of Type I complex with a barrier which does not allow further translocation, it could be another translocating Type I enzyme or when circular substrate

is fully translocated, provokes R-subunit to introduce double-strand break in close proximity to the collision site (indicated by scissors) while second motor subunit continues to translocate.

### 1.4.5. DNA substrates for EcoR124I

Type I R-M enzymes display a complicated mechanism of DNA hydrolysis, they introduce a double strand cut thousands base pairs away from the recognition sequence (Firman & Szczelkun, 2000). Janscak (1996) used circular DNA with one and two specific sites for EcoR124I to study DNA cleavage. It was shown that nicked DNA is accumulated during the first minutes of the clevage reaction with both substrates and is then further degraded into linear DNA. The linearized two-site DNA substrate undergoes secondary cleavage, and thus DNA fragments of random size appear on an agarose gel as a smear (Janscak et al., 1996). In comparison, linear DNA requires two specific sites for cleavage; endonuclease introduces double-strand cut randomly between recognition sites producing DNA fragments of different size. One-site linear substrate remains uncleaved (Eskin & Linn, 1972; Studier & Bandyopadhyay, 1988; Dreier et al., 1996).

Jindrova (2005) studied the character of DNA ends produced by restriction activity of endonucleases from different Type I families. Both 3'- and 5'- single stranded overhangs of different length without any dependence on specific base composition around the cleavage site are produced by enzymes, while generation of blunt ends was not observed. Cleavage of dsDNA by EcoKI (Type IA) results preferentially in 5'-ovehangs of 6-7 nucleotides in length. 3'-overhangs of 2-3 nucleotides and 5'-ovehangs of 3-5 are observed in most cases for EcoAI (Type IB) and EcoR124I (Type IC), respectively (Jindrova et al., 2005). Mechanism of dsDNA cleavage by Type I with short overhangs formation appears to be similar to mechanism of Type II enzymes where two nicks are introduced in each strand adjacently, however long overhangs detected for all three Type I endonucleases do not support this mechanism (Pingoud & Jeltsch, 1997).

### 1.4.6. Recycling of enzyme subunits

Type I R-M enzymes cleave dsDNA randomly and far from the recognition site, enzyme collision and stalling of translocation initiates the cleavage event (Studier & Bandyopadhyay, 1988; Ssczelkun et al., 1997). The second HsdR subunit of EcoR124I has a lower binding affinity to MTase and can dissociate from $R_2M_2S_1$ complex ($K_D$ = 240 nM), $R_1M_2S_1$ is a quite stable complex ($K_D$ < 1 nM) (Janscak et al., 1996, 1998). Releasing of the HsdR subunit from the translocating pentameric EcoR124I R-M complex provokes termination of translocation by this HsdR subunit; the dissociated subunit could interact with another MTase bound specifically to DNA and, thus, the HsdR subunit is used again (Seidel et al., 2005). Weak interactions of the HsdR subunit with the MTase core of EcoR124I play an important role in regulation of restriction activity *in vivo* (Dryden et al., 1997a; Firman et al., 2000). EcoR124I and EcoKI show the same behaviour on linear DNA, HsdR subunits dissociated via free DNA ends (Fig. 11 B) can bind to another MTase attached specifically to DNA (Simons & Szczelkun, 2011). Ability of the motor subunits to assemble/disassemble is a characteristic property of Type I R-M enzymes, however the $R_1$ complex of EcoR124I is quite stable when ATP is absent, it quickly dissociates in the presence of ATP (Janscak et al., 1998; Seidel et al., 2005). In comparison to EcoR124I, the relatively stable $R_2M_2S_1$complex of EcoKI (Dryden et al., 1997b) dissociates in the presence of ATP (Simons & Szczelkun, 2011).

In contrast, it was not possible to measure the level of HsdR turnover following termination of translocation on circular DNA substrate for EcoR124I and EcoKI. This difference, if compared with linear DNA substrate, was caused, probably, by a DNA topology problem or absence of free DNA ends which are required for the HsdR subunit to escape (Fig. 11 C). Despite the presence of free ends on circular substrate after its cleavage and the turnover of the MTase, the HsdR subunit does not turn over under these conditions (Simons & Szczelkun, 2011). Normally all recognition sites *in vivo* are modified and sometimes the DNA reparation processes could lead to appearance of unmethylated sites causing initiation of DNA translocation. As the HsdR subunit has no turnover following termination of translocation, this subunit becomes trapped on host circular DNA and excluded from pool of free HsdRs

without possibility to initiate new translocation. Thus, ability for EcoR124I to cleave DNA is reduced despite the fact that the HsdR pool is constant (Makovets et al., 2004). Such maintenance of endonucleases in disassembled form *in vivo* was proposed to regulate the level of restriction activity (Firman et al., 2000; Seidel et al., 2005).

In the magnetic tweezers apparatus, DNA has no free ends (Fig. 11 A), one end is attached to cover slip and second to magnetic bead. HsdR, released from MTase after termination of translocation, may further travel along dsDNA until reaches cover slip or magnetic bead and no turnover for HsdR subunit occurs in this case. At the same time, free HsdR from solution can bind to MTase and again re-initiate DNA translocation (Simons & Szczelkun, 2011).

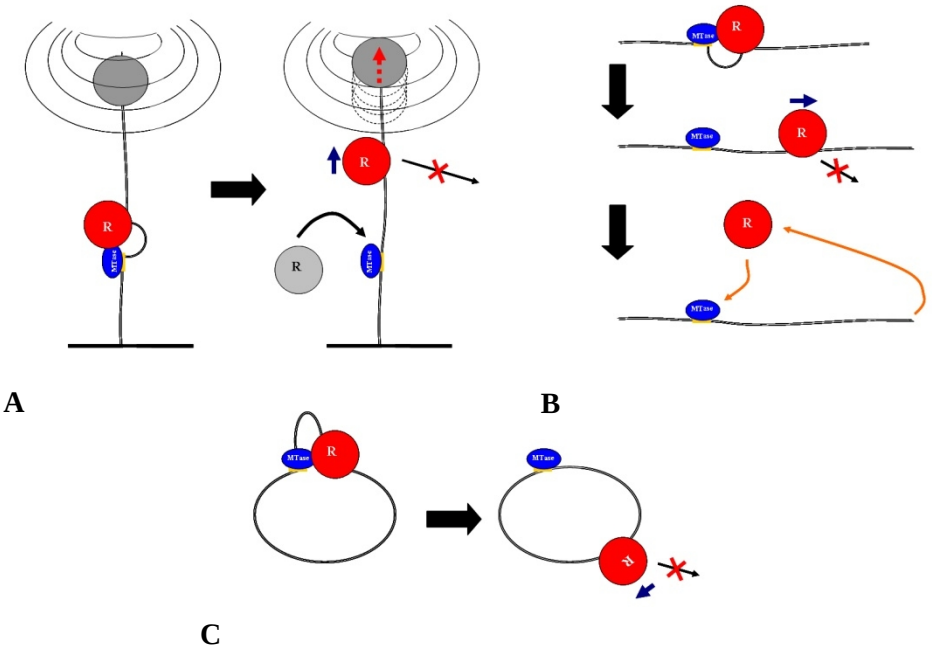



**Figure 11.** Different fates of HsdR subunits on different DNA substrates *in vitro* after their dissociation from MTase core. DNA is perpesented as a black line, MTase is a blue oval, HsdR subunit is a red circle. A) DNA substrate attached to the coverslip surface (black horizontal line) with magnetic bead (gray circle) at the opposite end in a magnetic tweezers apparatus. B) Linear DNA substrate in solution. C) Circular DNA substrate in solution.

### 1.4.7. EcoR124I as SF2 helicase

Helicases are enzymes capable to translocate and unwind DNA or RNA duplexes in an ATP-dependent manner. Unwinding activity is essential in every cellular process such as DNA replication, transcription, translation, repair, synthesis of ribosome, RNA splicing, etc. (Singleton et al., 2007). Based on the primary sequence of seven helicase characteristic motifs, described by Gorbalenya and Koonin, all helicases can be divided into six superfamilies – SF1, SF2, SF3, SF4, SF5, SF6; where SF2 is the largest (Gorbalenya & Koonin, 1993; Hall & Matson, 1999). Structural and biochemical characterization of many enzymes revealed that many NTP-dependent nucleic acid motor enzymes with a collection of helicase characteristic motifs clustered in RecA-like folds do not have unwinding activity and some of them even have no translocation activity (Lohman et al., 2008). SF2 members contain the motif DEAD or its modification and are often mentioned as DEAD-box helicases (Murray, 2000; McClelland & Szczelkun, 2004), which include nine conserved motifs Q, I, Ia, Ib, II-VI, a conserved GC between Ia and Ib and QxxR between IV and V. The Q motif is required for recognition of adenine of ATP. Motifs I and II, also known as Walker A and B, interact with α and β phosphates of ATP molecule. Motifs Ia, Ib, IV and V are involved in nucleic acid binding. Motifs III, V and VI (Arginine fingers) are required for coordination of α, β, and γ phosphates of ATP (Tanner & Linder, 2001; Caruthers & McKay, 2002; Cordin et al., 2006; Bleichert & Baserga, 2007).

Based on the primary structure of the HsdR subunit (motifs, characteristic for SF2 superfamily helicases (Fig. 12), are clustered in the RecA-like helicase domains of the motor subunit) the EcoR124I R-M enzyme from *E. coli* is classified as a SF2 helicase which is able to translocate dsDNA however without duplex unwinding activity (McClelland & Szczelkun, 2004; Firman & Szczelkun, 2000; Seidel et al., 2004; Stanley et al., 2006). Duplex DNA was proposed to be accommodated in the helicase cleft between two RecA-like helicase domains of HsdR subunit without any steric clash. Helix 29 in the helicase domain 2 would contact the 3' – 5' strand via motif IVa, the helicase domain 1 would contact 5' – 3' strand via motif IIa. The DNA duplex was proposed to bent about 80° leaving the helicase cleft and to passes across

motor subunit along the groove formed between the helical and the endonuclease domains. The major groove of the 5' – 3' strand would then face to the endonuclease active site, as was suggested by structural analysis of the endonuclease core domain and Type II restriction enzymes (Lapkouski et al., 2009). Interactions of motor subunit with 3' – 5' DNA strand were suggested to be more important for DNA translocation and contacts with 5' – 3' imply to stabilize the enzyme on the substrate (Stanley et al., 2006). Rad54 translocase from *Sulfolobus solfaticus* related to DExx helicases from SWI2/SNF2 family (SF2 member) (Richmond & Peterson, 1996; Caruthers & McKay, 2002) and its helicase domain 1 purified separately were shown to have high binding affinity to dsDNA. However, an isolated helicase 2 domain of Rad54 translocase displays no affinity to DNA. Rad54 was suggested to interact with DNA all the time during translocation via Ia motif in the helicase 1 domain walking on DNA in 3' – 5' direction (Dürr et al., 2005).

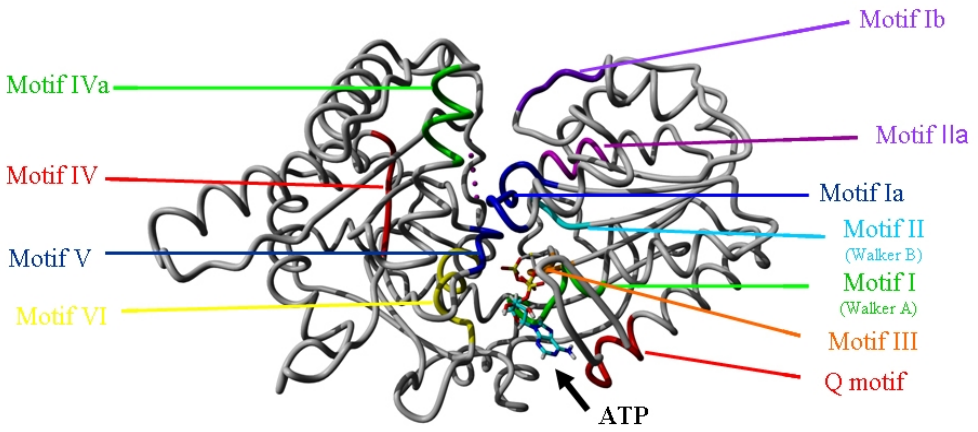



**Figure 12.** Excised helicase 1 (right) and 2 (left) domains of HsdR subunit of EcoR124I with ATP molecule in the ATP binding pocket (tube representation). All helicase characteristic motifs in two RecA-like helicase domains are color labeled. (pdb id: 2W00, Lapkouski et al., 2009). Figure prepared in YASARA.

According to FRET experiments, the Rad54 translocase from *Sulfolobus solfaticus* is present in solution in an open conformation. The helicase 2 domain undergoes rotation about 180° relative to the helicase 1 domain upon DNA binding resulting in a closed conformation of the enzyme. FRET experiments

confirmed the relative rotation of the helicase 2 domain of Rad54 which was proposed earlier based on its crystal structure and structures of other related helicase (Dürr et al., 2005; Lewis et al., 2008). Once the helicase cleft of Rad54 is closed after DNA binding, further rotational degrees of the helicase 2 domain depend on nucleotide binding, hydrolysis and release. A rotation bringing the enzyme from the closed conformation in the ATP bound state to a semi-open conformation is provoked by rapid hydrolysis of ATP or by release of inorganic phosphate after ATP hydrolysis. The helicase domain 2 via motif IV rebinds downstream relative to the binding position of the helicase domain 1 (Lewis et al., 2008).

The crystal structure of HsdR of EcoR124I (2W00) where the helicase cleft is closed is suggested to be in ATP bound competent state (HsdR was co-crystallized with ATP) (Lapkouski et al., 2009). Recent studies of EcoR124I R-M complex using electron microscopy in combination with small-angle X-ray scattering and detailed molecular modeling revealed that an relaxed conformation in the absence of DNA appears to be preferable for EcoR124I R-M complex and DNA binding triggers a large contraction of the enzyme resulting in a compact form (Kennaway et al., 2012).

The high degree of similarity between the translocase domains in the HsdR subunit of EcoR124I and Rad54 translocase would imply an analogous mechano-chemical cycle for HsdR upon DNA binding where nucleotide binding/hydrolysis competent states would determine rotational positions of the helicase 2 domain during translocation cycle.

# 2. MATERIALS AND METHODS

## 2.1. Preparation of competent cells

5 ml of fresh Luria-Bertani (LB) media was inoculated with stock of competent cells and grown overnight at 37 °C with shaking. Next morning 100 ml of fresh LB media in 250 ml flask was inoculated with 1 ml of overnight culture and incubated at 37 °C at 180 RPM until $OD_{600}$ (optical density at 600 nm) was ~0.4. Cells were cooled on ice, harvested in pre-cooled centrifuge at 4000 RPM for 15 min, resuspended in 50 ml of 100 mM $CaCl_2$ (pH 8.0) and hold on ice for 20 min. Then cells were harvested by centrifugation at 4000 RPM for 10 min (4 °C) and resuspended in 1.75 ml of 100 mM $CaCl_2$ + 30 % glycerol solution (pH.8.0), kept on ice for 30 min, then 50 μl aliquots were pipetted in 1.5 ml Eppendorf tubes and stored at -70 °C.

*E. coli* stains DH5α (Novagen) and NEB5α (New England Biolabs) were used for multiplication of plasmid DNA; BL21(DE3)Gold (Stratagene) and JM109(DE3) (Promega) were used for protein expression. In case of BL21(DE3)Gold cells LB media was supplemented with 12.5 μg $ml^{-1}$ of tetracycline (Tc).

## 2.2. Site-directed mutagenesis

Plasmid pTrcR124 (Janscak et al., 1996) carrying ***hsdR*** gene was used for site-directed mutagenesis. Oligonucleotide primers were used in polymerase chain reaction (forward primer sequence is given, reverse primer is its complement; mutated codon is underlined, changed nucleotides are in bold):

5‘–GGTTTTACC**GCC**ACCCCTATTTTCC–3’ for G436A,
5‘–CTATTTTCCCG**GCA**AACGCCTTAG–3’ for E442A,
5‘–GAAAACGCC**GCA**GGCTCAGAAAC–3’ for L445A,
5‘–GAAAACGCC**GAA**GGCTCAGAAAC–3’ for L445E,
5’-TTTCCGCCAG**GCA**ACCCACCGTACC-3’ for K527A
5‘-GAGTTATACAGAG**GCT**ATGGAAGGC-3‘ for Y736A
5‘-GCAAAACTAT**GCT**GAATTTGCCACG-3‘ for D796A.

Double mutant K527A_Y736A was obtained from plasmid DNA pTrcR124 already containing Y736A mutation.

Polymerase chain reaction was performed in Eppendorf (Hamburg, Germany) mastercycler. 200 µl PCR tube contained 20 µl of reaction mixture: 1.5 U of Expand Long Range polymerase (Roche, Basel, Switzerland); 5× Expand Long Range Buffer with 12.5 mM $MgCl_2$; 100 ng of plasmid DNA pTrcR124; 0.3 µM of Forward and Reverse primers; 500 µM PCR Nucleotide Mix (dATP, dCTP, dGTP, dTTP each at 10 mM concentration); DMSO 3-12 % and $H_2O$. PCR amplification conditions were following: 2 min 96 °C, 30 x (0.5 min 96 °C, 1.5 min 55 °C, 8 min 68 °C), 7 min 68 °C. DpnI (20 U) was used to digest (1h at 37 °C) the parental supercoiled DNA. Reaction mixture was transformed into *E. coli* competent cells DH5α (Novagen) or NEB5α (New England Biolabs) and cells were grown on LB/agar plates with ampicillin (Amp). Single colonies were used to inoculate 3 ml of LB media with Amp to a final concentration 100 µg $ml^{-1}$, cells were grown overnight at 37 °C with shaking. Zyppy Plasmid Miniprep KIT (Zymo Research, Irvine, CA, USA) was used to isolate plasmid DNA from cell pellet. The presence of desired mutation was confirmed by sequencing. All mutants with introduced mutations were fully sequenced.

## 2.3. Protein induction test

Protein induction test served to find appropriate expression conditions for production of mutant HsdR subunits in large scale. 3 ml of LB medium supplemented with Amp to a final concentration 100 mg $ml^{-1}$ was inoculated with single colony and grown overnight at 37 °C with shaking. Next morning the overnight culture was diluted 1:100 in 20 ml of fresh LB media with Amp in 100 ml flasks and cells were grown at 37 °C with shaking. Where it was necessary the culture was inducted with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) after $OD_{600}$ reached ~0.4 or ~0.6. Induced and non-induced cultures were grown additional 4 h at 37 °C or 30 °C and then harvested by centrifugation for 1 min at 14000 RPM. Cell pellet from 2 ml of the culture was resuspended in 200 µl of STE buffer (0.1 M NaCl, 10 mM Teis-HCl (pH 8.0), 1 mM ethylenediaminetetraacetic acid (EDTA) (pH 8.0)) with 1

μl of 100 mg $ml^{-1}$ lysozyme and kept on ice for 30 min. Then cells were disrupted by sonication on ice 3 × 1 min bursts with 1 min cooling in between and cell lysate was clarified by centrifugation for 1 min at 14000 RPM. The supernatant was collected for analysis and the pellet was resuspended in lysis buffer (100 mM $NaH_2PO_4$, 10 mM Tris-HCl, 8 M Urea, pH 8.0), kept for 20 min at room temperature and used for analysis. The presence of expressed protein in soluble fraction and pellet was visualized by 12 % sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (SDS-PAGE).

## 2.4. Expression and purification of WT and mutant HsdRs

WT HsdR of EcoR124I was expressed from the plasmid pTrcR124 (Janscak et al., 1996) in BL21(DE3)Gold *E. coli* (Lapkouski et al., 2007). Mutated HsdRs were produced from plasmid pTrcR124 containing mentioned above mutations. The overnight culture was diluted 1:100 in 0.5 l of fresh LB medium (in 3 l flask) supplemented with Amp to a final concentration of 100 μg $ml^{-1}$ and grown with shaking (180 RPM) at 37 °C until the $OD_{600}$ was ~0.5–0.6. Expression WT HsdR was induced by addition of IPTG to a final concentration of 1 mM and the culture was incubated additional 4 h at 37°C with shaking (180 RPM). G436A, E442A, K445A and L445E mutatant HsdRs were expressed as WT HsdR. K527A, Y736A, D796A and K527A_Y736A mutatant HsdRs were expressed at 30 °C after induction with IPTG. Cells were harvested by centrifugation at 4000 RPM for 20 min, washed in STE buffer (pH 8.0), stored at -20°C and used for purification.

Cell pellet (obtained from 0.5 of the culture) was defrosted on ice and resuspended in 5 ml of buffer A (20 mM Tris; 50mM NaCl; 0.1 mM EDTA;1 mM DTT; pH 8.0) supplemented with 100 μg $ml^{-1}$ of lysozyme. After 30 min of incubation on ice cells were disrupted using French press (THERMO electron corporation, 40K cell, USA) at 100 MPa. The cell lysate was clarified by ultracentrifugation at 45 000 g for 1.5 h at 4 °C.

The supernatant was applied onto a 25 ml DEAE Sepharose Fast Flow column (Sigma-Aldrich) pre-equilibrated in buffer A. Bound proteins were eluted with buffer B (20 mM Tris; 1M NaCl; 0.1 mM EDTA; 1 mM DTT; pH 8.0) applying a linear gradient of NaCl (0.05 – 1 M; 250 ml). ÄKTA Purifier

(GE Healthcare) was used for purification. Elution of proteins was recorded at wavelength 280 nm. WT and each mutatant HsdR were eluted from DEAE Sepharose Fast Flow column as a single peak. Collected fractions were analyzed by 12 % SDS–PAGE. Fractions contained protein of interest were pooled together and concentrated using 50 kDa centrifugal filter unit (Millipore). Concentrated protein as 50 % glycerol solution was stored at -20 °C. Protein concentration was determined by UV absorption spectroscopy (Eppendorf BioPhotometer Plus) at 280 nm and calculated using a molar extinction coefficient (98 225 $M^{-1}$ $cm^{-1}$) derived from the amino-acid sequence (Gill & von Hippel, 1989; Obarska-Kosinska et al., 2008).

## 2.5. Expression and purification of MTase

Plasmids DNA pJS491 (Amp resistance) (Patel et al., 1992) encoding HsdS subunit of EcoR124I and pAC15M (Tc resistance) (Holubova, 2004) encoding HsdM subunit of Ecor124I were used to express methyltransferase of EcoR124I in *E. coli* JM109(DE3). 20 ml of LB-media supplemented with Amp 100 μg $ml^{-1}$ and Tc 12.5 μg $ml^{-1}$ was inoculated with 200 μl of cell stock (*E. Coli* JM109(DE3)[pAC15M, pJS491] stored in glycerol (50 %) at -80 °C) and grown overnight. Next morning 0.5 l of 2xYT media supplemented with Amp 100 μg $ml^{-1}$ and Tc 12.5 μg $ml^{-1}$ in 2 l flask was inoculated with 5ml (1:100) of overnight culture and shacked at 200 RPM at 37 °C until OD reached ~0.4. Then the culture was induced with IPTG to a final concentration 0.5 mM and incubated overnight at 37 °C at 200 RPM. Cells were cooled on ice for 15-30 minutes, harvested by centrifugation at 4000 RPM for 30 min at 4 °C, washed 2 times with ice-cold STE. Efficiency of induction was analyzed by 12 % SDS-PAGE.

1.13 g of cell paste was resuspended in Lysis buffer (see below) (10 ml of buffer per 1 g of cell paste). Cells were broken by sonication on ice for 5 × 2 min bursts with 1 min cooling in between. Cell lysate was clarified by centrifugation at 30000 g for 20 min at 4 °C. Ammonium sulphate precipitation was performed at 4 °C for ~4 h and precipitated proteins were harvested by centrifugation at 30000 g for 30 min at 4 °C. Pellet was resuspended in 15 ml of

buffer L (see below) and dialyzed against 4 l of buffer L overnight. Dialyzed sample was clarified by ultracentrifugation at 30000 g for 1 h at 4 °C.

The supernatant was applied on DEAE Sepharose Fast Flow column (20ml, pre-packed, GE Healthcare) pre-equilibrated in buffer L. Bound proteins were eluted with a linear gradient of NaCl (0.05 – 0.5 M; 200 ml) using Bio-Rad Econo System. Eluted fractions (5 ml each) were visualized by 12 % SDS-PAGE. Fractions contained methyltransferase were pooled together and diluted with buffer O to get 50 mM final NaCl concentration. Then sample was applied directly on 5 ml Heparin column (GE Healthcare) pre-equilibrated with buffer L. Bound proteins were eluted with a linear gradient (0.05 – 1 M NaCl, 50 ml). Collected fractions were analyzed by 12 % SDS–PAGE. Fractions contained protein of interest were pooled together and concentrated using 50 kDa centrifugal filter unit (Millipore). Concentrated protein as 50 % mixture with glycerol was stored at -20 °C. Protein concentration was calculated from the absorbance at 280 nm and 320 nm using a molar extinction coefficient 160 400 $M^{-1}$ $cm^{-1}$ (Taylor et al., 1992; Webb et al., 1995):

$$C = (A_{280} - A_{320})/ \varepsilon_{MTase}$$

Lysis buffer: 50mM Tris-HCl, pH 8.0; 25 % w/v sucrose; 5mM EDTA, pH 8.0; 3 mM DTT (added before use), complete protease inhibitor cocktail tablet from Roche;
Buffer L: 10 mM Tris-HCl, pH 8.0; 50 mM NaCl; 0.1 mM EDTA, pH 8.0; 1 mM DTT;
Buffer H: 10 mM Tris-HCl, pH 8.0; 2 M NaCl; 0.1 mM EDTA, pH 8.0; 1 mM DTT;
Buffer O: Tris-HCl, pH 8.0; 0.1 mM EDTA, pH 8.0; 1 mM DTT.

## 2.6. *In vitro* endonuclease activity assay

Endonuclease activity was assayed *in vitro* using two types of DNA substrates: linear DNA pDRM-2R with two specific sites for EcoR124I and circular pRK with single specific site for EcoR124I (Taylor et al., 1992; Janscak et al., 1996). Plasmid DNA pDRM-2R was linearized with ScaI

restriction enzyme (Fermentas), purified from 1 % agarose/TAE gel using DNA Clean and Concentrator KIT (Zymo Research, Irvine, CA, USA) and used for analysis. Both pRK and pDRM-2R were isolated from *E. coli* DH5α or NEB5α harboring the plasmid using Zyppy Plasmid Maxiprep Kit (Zymo Research, Irvine, CA, USA). Endonuclease activity was assayed at 37 °C. EcoR124I R-M complex for *in vitro* restriction studies was reconstituted by simple mixing of separately purified MTase and HsdR subunits in ration 1:6. Such a way of assembly does not impair enzymatic activity of the complex which, as was shown for Type I R-M systems (Dryden et al., 1997; Janscak et al., 1998; Seidel et al., 2005).

40 μl reaction mixture contained buffer R (50 mM Tris-HCl, pH 8.0; 1 mM DTT; 10 mM $MgCl_2$; 50 mM NaCl), 5 nM of linear DNA pDRM-2R, 50 nM MTase and 300 nM HsdR or 15 nM of circular pRK DNA, 15 nM MTase and 90 nM HsdR. After 1 min of pre-incubation at 37 °C, reaction was initiated by addition of ATP and SAM to a final concentration 4 mM and 0.2 mM, respectively. Reactions were carried out as a time course and aliquots of 40 μl were withdrawn at the time points 0, 30, 60, 90, 120, 300, 600, 1800 sec and immediately quenched in 13 μl of stop solution (3 % SDS; 0.15 M EDTA; 10 % glycerol; 0.1 % bromophenol blue) and heated at 65 °C for 5 minutes. Samples were loaded into 1 % agarose/TAE gel and electrophoresed at 1.25 V/cm for 5 h and 5 V/cm for 130 min for linear pDRM-2R DNA and circular pRK DNA, respectively. Then gels were stained in 2 μg $ml^{-1}$ ethidium bromide solution for 20 min, washed in distillate water during 1 h and photographed under UV illumination. The percentage of DNA in each band was evaluated by densitometry of the ethidium bromide fluorescence using ImageJ software (Abramoff et al., 2004).

## 2.7. *In vitro* ATPase activity assay

### 2.7.1. Malachite green method

Malachite green method was used to test ATPase activity of mutant enzymes of EcoR124I *in vitro*. 20 μl of reaction mixture contained buffer R (50 mM Tris-HCl, pH 8.0; 10 mM $MgCl_2$; 50 mM NaCl; 1 mM DDT), 15 nM MTase, 90 nM HsdR subunit and 90 nM circular pRK DNA. *In vitro* ATPase

activity assay was carried out as a time course. Reactions were initiated after 1 min of pre-incubation at 37 °C by addition of ATP to a final concentration 4 mM. Aliquots (20 µl) were withdrawn at time points 0, 15, 30, 45, 60, 120, 300, 600 and 900 sec and immediately quenched in of 20 µl 0.1 M EDTA (pH 8.0). The concentration of inorganic phosphate (Pi) released during ATP hydrolysis was measured by a colorimetric assay based on a spectrophotometric quantification of a phosphomolybdate-malachite green complex (Chan et al., 1986). 40 µl of sample was mixed in ELISA microplates with 150 µl of malachite green reagent (5.72 % ammonium molybdate in 6 M HCl/ 0.0812 % malachite green/ $H_2O$ mixed in ratio of 1:2:3 immediately before use) (Janscak et al., 1996). After 5 min of incubation at room temperature ELISA microplates were scanned at 620 nm in the microplate ELISA reader. The concentration of Pi was determined by using a calibration curve derived from solutions of known Pi concentration ($KH_2PO_4$).

### 2.7.2. [λ-$P^{32}$]ATP method

Radioactive *in vitro* ATPase activity assay was performed by members of our group. Preparation of the cellulose plates. Positions for sample drops on cellulose plates were marked with a pencil: 1.5 cm from the baton of the plate and 1 cm in between dots. 50 ml of running buffer (0.4 M LiCl; 1 M Formic acid) or water was carefully poured in the glass chamber, avoiding drops on the walls. At first cellulose plate was pre-run in running buffer up to the top of the plate and dried. Then the same steps were performed in distillate water. Prepared plates can be stored at room temperature up to one week.

ATPase activity was assayed in 40 µl of reaction mixture containing 15 nM of EcoR124I, 90 nM of circular DNA with one recognition site for EcoR124I (pRK) and NEB2 reaction buffer (10 mM Tris-HCl, pH 7.9; 10 mM $MgCl_2$; 50 mM NaCl; 1 mM DTT). Reaction was initiated by addition of ATP mixture (calculated from 5 µl of 100 mM ATP + 0.2 µl of [γ-$P^{32}$]ATP) containing 0.16 µCi (0.0013 mM) [γ-$P^{32}$]ATP to final concentrations of 2 mM and incubated at 37 °C. Time course was carried out at 37 °C and 4 µl aliquots were withdrawn at 15, 30, 60, 90, 120, 180, 300, 900 sec. To stop reaction, taken aliquots were added immediately into the tubes containing 10 µl of 1 %

SDS. Samples were loaded on the cellulose plate at the marked dots and dried at room temperature. Then the plate was run in running buffer up to the top and dried. Dried plate was wrapped in transparent film and exposed for 10 min. Then screen was scanned and data was processed. The hydrolyzed [$\gamma$-$P^{32}$] was separated from [$\gamma$-$P^{32}$]ATP by TLC (stationary phase - celulose, mobile phase – 0.4 M $LiCl_2$ and 1M formic acid). Portion of [$\gamma$-$P^{32}$]/[$\gamma$-P32]ATP was visualized by Fujitsu 9000 and evaluated by attached software (Randerath, 1964; Marini & Krejci, 2012).

## 2.8. *In vivo* restriction activity assay

*E. coli* strain JM109(DE3) (Yanisch-Perron et al., 1985) served for complementation analysis of restriction function *in vivo*. For positive and negative complementation assays the plasmid DNA pKF650 with all three EcoR124II *hsd* genes and plasmid DNA pACMS with *hsdS* and *hsdM* genes for EcoR124II methyltransferase (Patel et al., 1992) were introduced into JM109(DE3) together with compatible plasmids carrying WT *hsdR* or each individual mutated *hsdR* gene. Fresh overnight culture was inoculated into LB media (1:10) with appropriate antibiotics and incubated at 37 °C with shaking until $OD_{600}$ ~0.5-0.6. Soft agar at 45 °C was mixed with 0.5 ml of grown cell culture, gently mixed, and immediately poured onto agar plates with appropriate antibiotics. The virulent mutant of phage λ was used for testing of restriction phenotype (Jacob & Wollman, 1954). Solidified soft agar was spotted with 30 μl each of tenfold serial dilutions of λvir.0 phage lysate at $10^2$ to $10^6$ plaque-forming units/ml (PFU/ml). The spots were dried at room temperature and the plates incubated overnight at 37 °C (Colson et al., 1965). The solid medium is LB with agar added at 1.5 %. Soft agar overlay is LB with agar added at 0.6 %. Antibiotics concentrations used: Amp at 100 mg $ml^{-1}$, chloramphenicol at 50 mg $ml^{-1}$. The efficiency of plating was determined as the number of plaques on the tested strains compared to the number of plaques on the non-restricting control strain *E. coli* JM109(DE3) (Yanisch-Perron et al., 1985). Values in the range 0.0001–0.01 correspond to the restriction-competent (r+) phenotype, those in the range 0.01–0.1 to the mixed-competence (r±)

phenotype, and those in the range 0.1–1 to the restriction-deficient (r−) phenotype.

## 2.9. Triple-helix displacement assay

Triple-helix displacement assay was performed by our collaborators in Prague. Linear DNA for these assays was generated by ApaI digestion of pLKS5 plasmid DNA (Simons et al., 2014) followed by phenol/chloroform extraction, chloroform extraction and isopropanol precipitation. 5'-end TAMRA labeled oligonucleotide # 125 (5' – TTTCTTCTTCTTTTCTTTTCTT – 3') was supplied and HPLC-purified by Eurofins MWG Operon (Germany). For triplex formation 50 nM linear DNA and 25 nM TAMRA labeled TFO # 125 (Simons et al., 2014) was mixed in buffer MM (10 mM MES pH 5.5, 12.5 mM $MgCl_2$) and incubated at 23 °C overnight in the darkness. Following overnight incubatiom, the triplex was stored on ice, and then diluted 1/10 in reaction buffer before use. 5 nM triplex was pre-incubated with 40 nM of MTase and 120 nM of WT or mutant HsdR at room temperature. Translocation reaction was initiated with 4 mM ATP and after 10 minutes triplex displacement was measured by fluorescence intensity (McClelland et al., 2005). Steady state fluorescence measurements were performed at 23 °C using Tecan Safire 2 Multi-detection Microplate Reader. Excitation and emission wavelengths used were: 544 nm and 576 nm, bandwidths were kept < 10 nm.

## 2.10. Molecular dynamics simulations

Molecular dynamics simulations were performed by members of our group (D. Sinha and others). The WT crystal structure of the motor subunit HsdR from the restriction-modification system EcoR124I (pdb id: 2W00) and a recent crystal structure of mutant HsdR Lys220Ala (pdb id: 4BEC) were used for preparing the structural model for all simulations. Three missing segments from the WT crystal structure (residues 142-147, 585-590, and 859-869) were built using standard loop modeling in YASARA (Krieger et al., 2002; Konagurthu et al., 2006) and added to the WT crystal structure. The missing segment from residue 182 to 189 is resolved in the Lys220Ala mutant crystal

structure and thus was built into the WT structure by adding the coordinates from the mutant to the above modeled structure, followed by steepest-descent energy minimization. All classical MD simulations were performed using GROMACS 4.64 (van Der Spoel et al., 2005; Berendsen et al., 1995; Pronk et al., 2013) with the AMBER99SB forcefield (Hornak et al., 2006). ATP was parameterized by applying the standard RESP procedure using Antechamber (Wang et al., 2004), where charges for free Mg-ATP were derived from HF/6-31G* calculation in Gaussian03 (Frisch et al., 2004). Histidine was assumed to be charged, with the ND and NE atoms protonated; arginine and lysine residues were assumed to be protonated. Mutants were prepared *in silico* by replacing the respective side-chain in YASARA. All systems were solvated in explicit TIP3P water (Jorgensen et al., 1983) in a cubic box with a margin of 10 Å and neutralized by adding sodium counterion. The particle-mesh Ewald method (Darden et al., 1993) was applied to calculate long-range electrostatic interactions with a cutoff distance of 10 Å and a Lennard-Jones 6-12 potential was used to evaluate van der Waals interactions within 10 Å cutoff distance. The LINCS algorithm of fourth order expansion was used to constrain bond lengths (Hess et al., 1997). After solvation and neutralization each system was energy minimized for 10000 steps using steepest descent optimization method to remove poor van der Waals contacts in the initial geometry. After minimization two stages of equilibration were conducted. Initially the system was equilibrated for 1 ns with position restraints of 10000 KJ/mol on all heavy atoms. A constant temperature of 300 K was maintained using the V-rescale algorithm (Bussi et al., 2007) with a coupling time of 0.1 ps and separate baths for the solute and the solvent. The pressure was kept constant at 1 bar using the Parrinello-Rahman pressure coupling scheme (Parrinello & Rahman, 1981) with a time constant of 2 ps. Initial velocities were generated randomly using a Maxwell-Boltzmann distribution corresponding to 300 K. Neighbor lists were updated every 10 fs using a group cut-off scheme. Finally, the production run was performed for 100 ns without restraints at 300 K in the isothermal-isobaric ensemble.

Principal-components analysis (Amadei et al., 1993) was used using g_covar and g_anaeig tools in the GROMACS package to identify the global motions of the system using backbone atom only. The g_rms tool of the

GROMACS package was used to calculate root mean square deviations (RMSD) during the trajectories taking the minimized crystal structure as reference. Root mean square fluctuations (RMSF) of the backbone of each residue were calculated by g_rmsf while atomic distances were measured by g_dist.

## 2.11. Crystallization of G436A mutant HsdR

G436A mutant HsdR subunit was purified as was described previously (Janscak et al., 1996). 50 kDa cutoff spin concentrator (Millipore) was used for protein concentration and buffer exchange (to 20 mM phosphate buffer pH 7.5, 0.1 M KCl). Fresh protein (at concentrations 27 mg $ml^{-1}$, 22 mg $ml^{-1}$, 15 mg $ml^{-1}$, 11 mg $ml^{-1}$) supplemented with ATP to a final concentration 5 mM was used for crystallization experiments. Sitting-drop vapor-diffusion method was used for protein crystallization. Crystals of G436A mutant HsdR were grown at +4 °C within few days from precipitant (500 µl reservoir solution in 24-well Cryschem Plate) containing 8 % PEG 20 000, 8 % PEG 550 MME and 0.2 M $Li_2SO_4$ in 4 µl drops (2 µl of precipitant and 2 µl of G436A mutant HsdR) (Lapkouski et al., 2007). Increased concentration of PEG550 MME in precipitant to 15 % and 20 % and doubled volume of crystallization drop (4 µl of precipitant and 4 µl of protein) allowed obtaining of bigger crystals of G436A mutant HsdR with dimensions of about 0.3 x 0.2 x 0.05 mm. A crystal mounted in a nylon loop (Hampton Research) was transferred for a few seconds to 10 µl of cryosolution (25 % glycerol (in water) or 100 % PEG550 MME) and then was flash-frozen in liquid nitrogen prior measurement at a synchrotron facility.

## 2.12. Proteolysis of EcoR124I

Separately purified methyltransferase and WT HsdR subunit were used for limited proteolysis. EcoR124I was reconstituted in buffer R (50 mM Tris-HCl, pH 8.0; 1 mM DTT; 10 mM $MgCl_2$; 50 mM NaCl) using methyltransferase and HsdR ratio of 1:3 and then was treated with trypsin following Carey (2000). Trypsin powder (Sigma Aldrich) was dissolved in 1

mM HCl to obtain 1 mg/ml stock solution which then was aliquoted and stored at -80 °C. Proteolytic reactions were carried out in 500 µl Eppendorf tubes in R buffer containing ~20 pmoles of methyltransferase and ~60 pmoles of WT HsdR for one reaction (15 µl total volume) on ice. In range-finding experiments 2 µl of $10^1$, $10^2$, $10^3$ and $10^4$ dilutions of trypsin stock solution were used to start reaction and after 30 minutes reactions were stopped by adding of 4x SDS-loading buffer and immediate boiling for 5 min, then whole volume of reaction mixture was loaded in 10 % SDS-PAGE and run at 30 mA for 1 h. PVDF membrane (Bio-Rad) was used for protein blotting. Blotting apparatus was assembled as for Western blotting and buffer 10 mM MES, pH = 6.0, 20% MeOH was use as blotting buffer. PVDF membrane was prewet prior to use with methanol. Transfer was performed in cold box for 1 h at 90 volts. After blotting PVDF membrane was washed in distilled water several times, then stained (in 0.02% Coomassie Brilliant blue in 40% methanol, 5% acetic acid) for 20-30 seconds and distained (in 40% methanol, 5% acetic acid) for up to 1 min, then membrane was rinsed in distilled water for 3-5 min with at least 3 changes and air dried completely. Prepared membrane was placed between two Whatman papers, wrapped in foil and stored in plastic bag with zip at -20°C. Bands of interest were cut out with a clean scalpel or razor blade and sent for Edman degradation protein sequencing.

# 3. RESULTS AND DISCUSSION

## 3.1. Preparation of proteins and substrates

### 3.1.1. Preparation of MTase of EcoR124I

The HsdS subunit of EcoR124I alone is insoluble (Patel et al., 1992; Taylor et al., 1992) and co-expression with the HsdM subunit results in the MTase complex with a stoichiometry $HsdM_2HsdS_1$. MTase was obtained following Taylor et al., 1992. MTase was expressed in *E. coli* JM109(DE3) from pJS491 and pAC15M plasmids encoding for *hsdS* and *hsdM* genes, respectively. Amp and Tc were used to maintain plasmids pJS491 and pAC15M in the host, respectively. 0.5 l of 2xYT media suplemented with appropriate antibiotics was inoculated with fresh overnight culture. After $OD_{600}$ was ~0.4 the culture was induced with 0.5 mM of IPTG. After ~12 h of incubation cells were harvested by centrifugation. Induced culture was compared with culture before induction (Fig. 13).

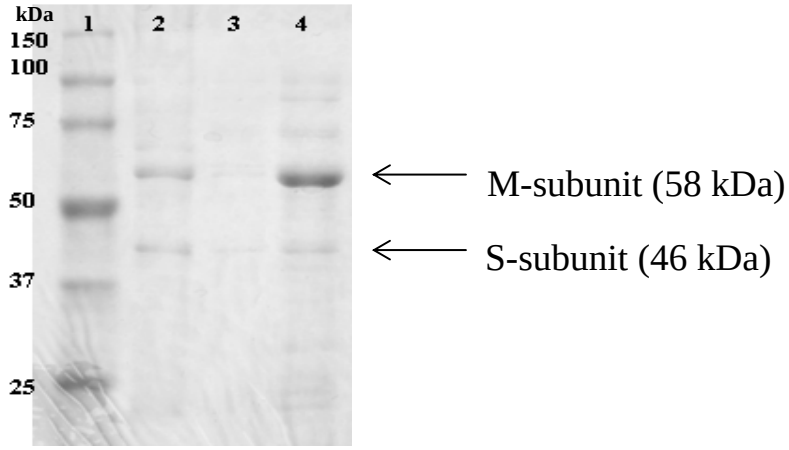
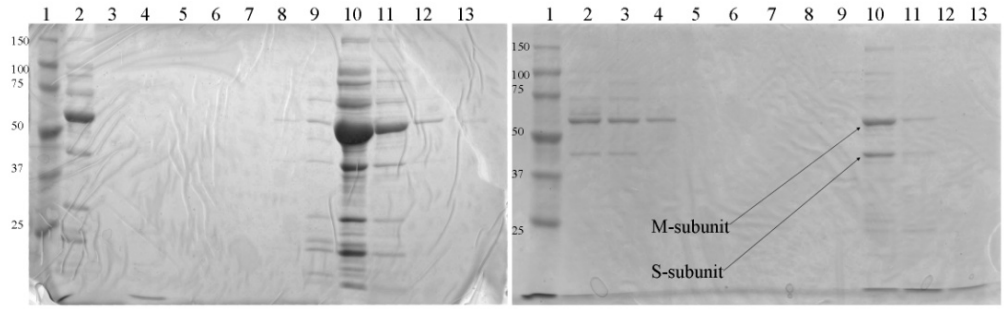



**Figure 13.** 12 % SDS-PAGE showing MTase expression from pAC15M and pJS491 plasmids before and after induction with IPTG: line 1, molecular-weight markers; line 2, purified MTase as control; lines 3, MTase expression before induction with IPTG; line 4, MTase expression after induction with IPTG and overnight incubation. The ratio of HsdM:HsdS on line 4 is greater than 2 indicating that M-subunit is expressed in excess over S-subunit.

Supernatant obtained after cell disruption was used for ammonium sulphate precipitation. Protein pellet was dissolved in buffer L and dialyzed against 4 l of buffer L overnight. Dialyzed sample clarified by centrifugation

was used for purification: 1) anion-exchange chromatography on DEAE Sepharose Fast Flow column; 2) affinity chromatography on Heparin column. In both cases, linear gradient of NaCl was used to elute proteins (Fig. 14). Fractions of interest eluted from DEAE column were pooled together; NaCl concentration in the sample was brought to a NaCl concentration as in buffer L by dilution with buffer O. The second step of purification on the Heparin column lead to approx 90 % MTase purity what is satisfactory for reconstitution of EcoR124I *in vitro*. Purified MTase was stored in 50 % glycerol at -80 °C and used for testing of WT and mutant HsdR.




**A** **B**

**Figure 14.** 12 % SDS-PAGE from purification of MTase of EcoR124I. A) Fractions collected from DEAE column: 1 – molecular-weight markers (kDa); 2 – crude extract; 3 – sample loading; 4 and 5 – equilibrating after loading; 6 to 13 – collected fractions (10 and 11 were used for purification on Heparin column). B) Fractions collected from Heparin column: 1 - molecular-weight markers (kDa); 2 – sample for loading; 3 – loading of sample; 4 – equilibrating after loading; 5 to 13 – collected fractions (10 was concentrated and used for *in vitro* assays). The line 3 indicates some MTase and free M-subunit went through Heparin column, additionally traces of free M-subunit are seen on the line 4 at the beginning of the equilibrating step. MTase interacts with the Heparin column matrix via region on S-subunit, thus the ratio HsdM:HsdS on line 10 is ~2.

### 3.1.2. Preparation of WT and mutant HsdRs

Plasmid pTrcR124 and its derivatives carrying mutations (1$^{st}$ group: K527A, Y736A, D796A, K527A_Y736A; 2$^{nd}$ group: G436A, E442A, L445A and L445E) were used for expression of WT and mutant HsdRs, respectivelly. Before large scale protein production for purification purposes a protein induction test was carried out (20 ml of culture in 100 ml flask) to test

expression conditions. At first WT expression conditions (1 mM IPTG, 37 °C for 4 h at 180 RPM) were applied for all mutant HsdRs. The culture was induced with 1 mM IPTG at $OD_{600}$ ~0.4 or ~0.6, after 4 h of incubation at 37 °C cells were harvested and level of expression was evaluated. All mutant HsdRs from the 1st group at 37 °C incubation temperature were partially soluble and found in insoluble fraction. HsdRs from the 2nd group were well expressed at these expression conditions. Expression conditions determined in protein production test were applied to large scale culture (0.5 l of culture in 3 l flask). Cell paste obtained from 0.5 l of culture was twice washed in STE buffer and stored at -20 °C.

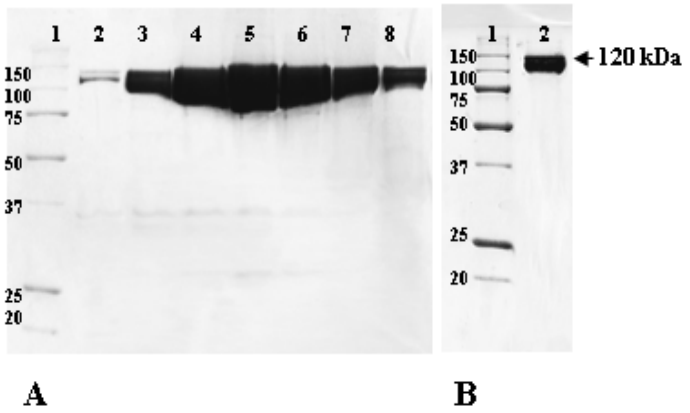



**Figure 15.** Purification of K527A mutant HsdR of EcoR124I. A) 12 % SDS-PAGE showing eluted fractions from DEAE column: line 1, molecular-weight markers (kDa); lanes 2-8, collected fractions. B) 12 % SDS–PAGE showing purified and concentrated K527A in 50 % of glycerol on lane 2, line 1 - molecular-weight markers (kDa).

Anion-exchange chromatography was used for protein purification. The supernatant obtained from cell paste was applied on DEAE Sepharose Fast Flow column (25 ml) pre-equilibrated in with low salt solution (Buffer A). Purification profiles for WT and all mutant HsdRs were similar, linear gradient of NaCl (0.05 – 1 M; 250 ml) was used, all proteins were eluted in range from ~12 to 22 % of high salt solution (Buffer B). Collected fractions were checked by 12 % SDS-PAGE, best fractions were saved as 50 % glycerol solution at -20 °C for further experiments (Fig. 15).

Experimental results for 1st group of mutations (K527A, Y736A, D796A, K527A_Y736A) and 2nd group of mutations (G436A, E442A, L445A and L445E) are reported in **Paper 3** and **Paper 4**, respectively.

### 3.1.3. Preparation of DNA substrates for EcoR124I

Studies of the HsdR subunit with introduced point mutations imply testing of fully assembled EcoR124I R-M complex *in vitro*. Motor subunits fulfill ATPase/translocation and endonuclease activities only if they are part of the fully assembled $R_2M_2S_1$ complex (Janscak et al., 1998; Seidel et al., 2005). Simple mixing of separately purified MTase and HsdR makes it possible to reconstitute EcoR124I *in vitro*. EcoR124I R-M complex assembled in such way has the same activities as EcoR124I purified as holoenzyme, as was shown for Type I R-M systems (Dryden et al., 1997; Seidel et al., 2005).

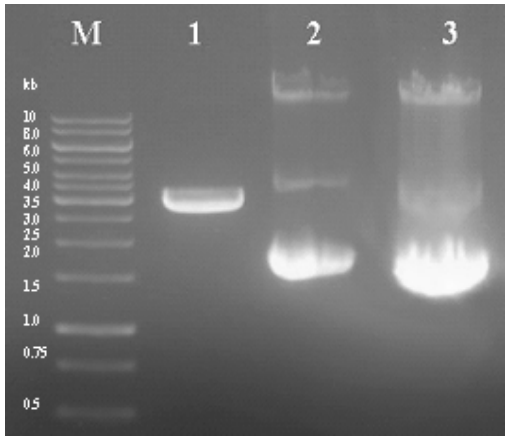



**Figure 16.** DNA substrates of EcoR124I R-M enzyme, 1 % agarose/TAE gel: M – 1 kb DNA Ladder (Fermentas), lines 1 and 2 – linearized by ScaI and circular pDRM-2R, respectively; line 3 – pRK plasmid DNA.

Endonuclease activity of EcoR124I with WT or mutant HsdRs was tested *in vitro* using linear pDRM-2R DNA with two EcoR124I target sites or circular pRK with one specific recognition site. The mechanism of cleavage for both substrates is different, requiring the collision of two translocating enzymes for linear two-site substrates, and steric hindrances on single-site circular DNA due to the fact that on a circle the enzyme finally runs into itself (Janscak et al.,

1999). Also cleavage of each type of substrate gives different specific products. Two-site linear substrate is cleaved in random manner between two target sites resulting in smear on agarose gel. Circular DNA substrate gives two particular cleavage products open circle and linear DNA species (Janscak et al., 1996). Linear pDRM-2R substrate was derived from plasmid DNA pDRM-2R using ScaI restriction enzyme and purified from 1 % agarose/TAE gel (Fig. 16). Both plasmids pDRM-2R and pRK were isolated from *E. coli* cells Dh5α or NEB5α. Concentration of DNA substrates was determined spectrophotometrically.

## 3.2. Residues at the helical-helicase 2 domain interface

Already analysis of protein sequences preceding structural/functional studies allowes grouping of various proteins based on the presence of specific motifs, usually mentioned as helicase characteristic motifs, in SF2 superfamily helicases (Gorbalenya & Koonin, 1993). Studies of enzymes possessing translocation activity such as Rad54 from *Sulfolobus solfaticus* or EcoR124I revealed one common feature, both belong to SF2 helicases and cluster a collection of helicase characteristic motifs in two RecA-like helicase domains (for EcoR124I in HsdR subunit). Also EcoR124I translocates on dsDNA (McClelland & Szczelkun, 2004; Firman & Szczelkun, 2000; Seidel et al, 2004) without strand separation activity (Stanley et al., 2006) so do true DNA translocases from the SF2 helicase family. In the HsdR subunit helicase domain 1 and helicase domain 2 are involved in interactions with dsDNA via motifs Ia, Ib and IIa and IVa, respectively (Lapkouski et al., 2009). Interactions of the motor subunit of EcoR124I with the 3'-5' dsDNA strand during translocation are constant and more effective than with the 5'-3' strand, contacts which stabilize enzyme on DNA (Stanley et al., 2006).

An inchworm mechanism of dsDNA translocation with alternating binding affinity to DNA of the domain 2 was proposed for Rad54 translocase from *Sulfolobus solfataricus* (Dürr et al., 2005). Interestingly, the domain 2 in the crystal structure of Rad54 in complex with DNA duplex is rotated ~180° relative to the domain 1 (Dürr et al., 2005) if compared with the position of equivalent domains in ATP bound DExx box helicases (Caruthers & McKay, 2002). A 180° flip of the domain 2 in Rad54 likely takes place already during

DNA uptake. FRET experiments indicate that Rad54 is in open conformation with corresponding ~180° relative position of the domain 2 as in the crystal structure if DNA is not bound. Binding of Rad54 to DNA triggers re-orientation of the domain 2 with respect to the domain 1 by ~180° resulting in closed form of Rad54 (Dürr et al., 2005; Lewis et al., 2008). Although the binding of the protein to the DNA is strictly speaking not part of the regular mechano-chemical cycle and it remains unclear if this large conformational change is physiologically relevant and occuring *in vivo*, it demonstrates the ability of the domain 2 to undergo large rotational movements (Lewis et al., 2008). FRET intensity upon binding of ATP to the closed form of Rad54 with DNA is not detectable suggesting the absence or minor conformational changes. Increase in FRET intensities was observed after ATP hydrolysis and prior to ADP release. It seems that phosphate release leads to a powerstroke required for re-orientation of the domain 2 in Rad54. Thus, the orientation of the domain 2 in Rad54 during translocation cycle is determined by nucleotide competent states and likely conformational changes of the domain 2 can be interpreted either as a rotation (to a smaller extend than a 180° rotation) or as changes in the interdomain distance or a combination of both (Lewis et al., 2008).

Common, for both Rad54 and EcoR124I, structural features and similar fashion of action on DNA substrate imply that DNA uptake and nucleotide binding/hydrolysis states would determine the orientation of the helicase 2 domain in HsdR of EcoR124I at different translocation stages similarly as in Rad54. Despite the absence of DNA substrate in the helicase cleft of the crystal structure of HsdR (2W00), where a molecule of ATP is bound in the ATP binding pocket, the structure represents one of the nucleotide competent states (Lapkouski et al., 2009). Rad54 translocase from SF2 helicases is the closest relative to HsdR of EcoR124I. Comparison of the crystal structures of related translocases from SF2 helicases reveals that, in comparison with the HsdR subunit with domain organization, enzymes are represented only by their translocation/helicase functional unit which is not covelantly linked to other domains and can act independently. Although HsdR has the translocation/helicase unit represented by the helicase 1 and the helicase 2 domains, it can act as translocase only within EcoR124I R-M complex. Up to now the HsdR subunit in Type I R-M systems was hypotisized to assemble with

methyltransferase via the C-terminal helical domain and experimentally it was confirmed for EcoKI where truncation of 155 residues at the C-terminal prevented interaction with methyltransferase and complex fromation (Davies et al., 1999). Also the crystal structure of the HsdR subunit is peculiar with specific hydrogen bonds at the helical-helicase 2 domain interface. During the translocation cycle large conformational changes occuring in regions with these interactions would change the helical-helicase 2 domain interface drastically. The question to answer is whether these interdomain interactions are important for translocation activity and participate actively in the translocation cycle of the translocase domain, as they could potentially stabilize a particular translocation stage.

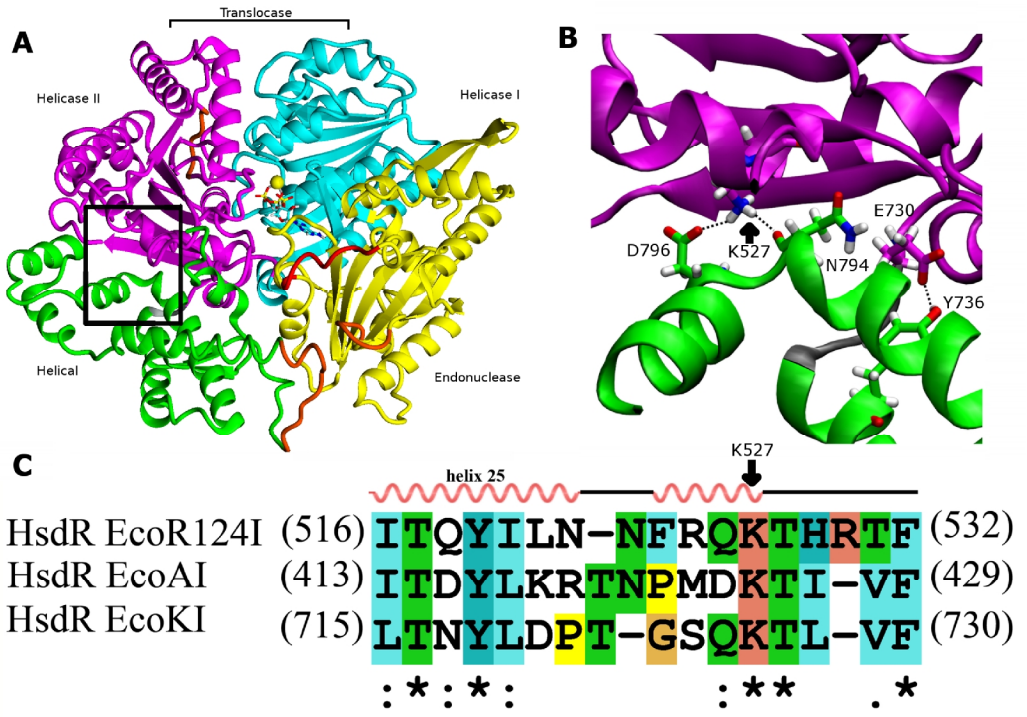



**Figure 17.** Interdomain interactions at the helical-helicase 2 domain interface in the HsdR subunit of EcoR124I restriction-modification complex. **A.** The structure of HsdR consisting of four domains: the endonuclease domain is in yellow, the helicase 1 and the helicase 2 domains are depicted in cyan and magenta, respectively. The C-terminal helical domain is shown in green. ATP is represented in elemental colors as a skeletal model with cyan carbons, and magnesium is indicated by a yellow sphere. The modelled parts of the structure (not resolved in the crystal structure) are in orange and the modeled 180s loop comprising residues 182–189 that is close to the catalytic site on the endonuclease domain is in red just below ATP (coordinates

borrowed from K220A mutant HsdR crystal structure ((PDB id: 4BEC, Csefalvay *et al.*, 2015)). The short linker between helicase 2 domain and the C-terminal helical domain is in grey at the lower right corner of the black rectangle **B.** Enlargement of the region enclosed by the black rectangle in panel A, showing residues involved in interactions at the interface between the helical domain and helicase domain 2. Side chains are shown as skeletal models in atomic colors with carbon atoms colored according to their subunit shown in panel A. Hydrogen bonding interactions discussed in the text are indicated with black dotted lines. **C.** Alignment of the interface region sequence of EcoR124I from family IC of with sequences from HsdR of the archetypical members of the other two type I R-M families, EcoKI from family IA and EcoAI from family IB. Secondary structures of Eco124I HsdR from the crystal strutcure (pdb code: 2w00) are shown on top. Colors and symbols below the sequences indicate residue similarity: '*' fully conserved residues; ':' highly similar residues; and '.' somewhat similar residues.

A closer loot at the helical-helicase 2 domain interface (Fig. 17A) in the HsdR subunit of EcoR124I (Lapkouski et al., 2009) reveals one centrally located interdomain contact from residue K527 at the c-term of α-helix 25 of the helicase 2 domain to residues N794 and D796 in a helix-turn-helix region (end/beginning of α-helices 35/36, respectively): a bifurcated contact of the K527 amino group in which the K527 heavy atom Nζ and the D796 carboxyl Oδ atom are in a 2.9 Å distance, while K527 Nζ is 2.7 Å from the N794 backbone oxygen. The helix-turn-helix region (α-helices 32 and 33) at the boundary of the helicase 2 (α-helix 32) and the helical (α-helix 33) domains are stabilized by a second contact: Y736 (in α-helix 33) contacts E730 (end of α-helix 32) with the phenol O of Y736 and atom Oε2 of E730 being in a 2.6 Å distance (Fig. 17B). Lysine residue K527 is conserved throughout Type I R-M families IA, IB, IC (Fig. 17C) suggesting a similar general function of this region in the helicase 2 domain, and also a similar three-dimensional structure for this helix-turn-helix region in the helical domain is predicted.

Outside the recA-like helicase motifs a reliable sequence alignment of HsdR with other SF2 translocases such as Rad54 is not possible due to the low sequence identity, also the conserved motif Q/D-KT contacting the fourth helical domain in HsdR is not found in Rad54 and other translocases. Despite that, the helical-helicase 2 domain interface might be rearranged during translocation cycle by the rapid conformational changes leading to rotational motion of the helicase domain 2 similarly as in Rad54 translocase (Lewis et al., 2008). Because of domain rotation, these interactions at the helical-helicase cleft might be altered or even lost or interdomain distance would increase in

different translocation stages. This leads to the hypothesis that these interactions might stabilize a particular translocation stage (or at least participate in) and be essential for translocation of DNA. Thus the relevance of the above mentioned interactions for the translocation activity of EcoR124I was studied. Site-directed mutagenesis allowed to construct a R-subunit lacking those interactions, and these mutant enzymes were tested *in vivo* and *in vitro* and also characterized using MD. Mutations introduced at the helical-helicase 2 domain interface and described below provoke a loss of ATPase and translocation activities in EcoR124I. Consequently, the endonuclease activity of mutant enzymes is affected as well, indicating that mutations in domains other than the endonuclease domain might influence DNA cleavage in EcoR124I, however indirectly through altered ATPase/translocation activity. As a result, host cells having such mutant enzymes cannot stop phage infection and do not survive in a phage attack.

Interaction analysis was carried out in the MD simulations (provided by D. Sinha from our group) to determine the persistence of the above mentioned interactions observed in the crystal structure of HsdR. For that purpose, three missing loops (residue 142-147, 585-590 and 858-869) were added by standard loop modeling in YASARA and the coordinates for 180-loop (residue 182-189) were taken from K220A mutant HsdR crystal structure (pdb id: 4BEC) resulting in WT HsdR model used for MD simulations. This model was used for interaction analysis and for generation of alanine point mutations in YASARA at K527, Y736 and D796 positions. MD simulations using WT HsdR model indicate that H-bond between amino group of K527 and backbone oxygen of N794 is lost after equilibration time, while the H-bond between amino group of K527 residue and carboxyl group of D796 became persistent after the initial equilibrating time. MD simulations suggest that the Y736 and E730 interaction is stable throughout the whole simulation. To examine the role of above mentioned interactions for HsdR of EcoR124I in DNA translocation/cleavage activities mutants were designed that would not posses the ability of bonding. The single mutant K527A would not be able in hydrogen bonding to D796 and N794, while D796A would not be able to contact K527 anymore, so half of bifurcated contact is lost. On the other hand, the single mutant Y736A would nullify the contact in the interdomain linker region and

could help to understand the importance of this region, and the double mutant K527A_Y736A would sum up the effect of both these contacts. All mutants were tested *in silico, in vitro* and *in vivo*. All details of computational analysis and also experimental results of section 3.2. are described in **Paper 3**.

Plasmid pTrcR124 (Janscak et al., 1996) carrying ***hsdR*** gene was used for site-directed mutagenesis to introduce K527A, Y736A, D796A and K527A_Y736A mutations. After digestion of methylated DNA template with DpnI PCR mixture was transformed into Dh5α or NEB5α competent cells and plated on LB-agar plate with Amp. From 5 to 10 single colonies from each plate were used to prepare the overnight culture. Then plasmid DNA was isolated and the presence of desired mutation was confirmed by sequencing using specific oligonucleotide primers. Plasmid DNA carrying the introduced mutations were additionally sequenced across ***hsdR*** gene using a collection of specific primers.

As was shown previously HsdR subunit in solution is represented in monomeric form (Zinkevich et al., 1997). An initial effect of introduced mutations was already observed at the protein expression step. Substitution of residues K527, D796, Y736, and both K527 and Y736 to an alanine results in formation of inclusion bodies when mutant HsdRs were expressed at conditions as WT HsdR (OD ~0.6, 1 mM IPTG, 37 °C for 4h). It seems that already at this step it can be predicted that the expression of such mutant HsdRs would lower chances of the cell to survive during phage infection as these mutations might cause folding problems, although uninduced expression of mutant HsdRs at levels close to natural might differ from induced protein expression in the cell. However, if compared with WT induced expression, it already indicates that introduced mutantions in HsdR subunit might lead to partial misfolding and consequent aggregation of mutant HsdRs. Decrease of expression temperature to 30 °C allowed to obtain mutatant HsdRs in the soluble fraction. Usually level of expression for Y736A and D796A mutant HsdRs was similar to that of WT and in case of K527A and double K527A_Y736A mutant HsdRs observed expression level was about 2-3 fold lower.

### 3.2.1. Translocation activity of mutant enzymes

Although the HsdR subunit of EcoR124I R-M enzyme clusters motifs characteristic for SF2 helicases (Gorbalenya & Koonin, 1991) forming the ATPase active site and typical for RecA-like helicase domains (Davies et al., 1999; McClelland & Szczelkun, 2004), EcoR124I does not unwind DNA duplexes but instead hydrolysis of ATP is coupled to translocation of intact duplex DNA through the motor subunit. According to the cleavage model proposed by Janscak (1999) endonuclease activity for Type I enzymes is triggered by blockage of DNA translocation due to steric hindrance on DNA which the enzyme is unable to pass. However, if the steric hindrance is labeled specifically and furthermore can be easily displaced from DNA by the translocating enzyme, then it is possible to monitor whether the enzyme translocates over a specific distance in a given interval of time if the hindrance is placed on DNA at a certain known distance from the recognition sequence. This idea was used for testing of DNA translocation by WT and mutant enzymes. pLKS5 plasmid DNA with a DNA triplex binding site (TBS) for a CT-rich triplex forming oligonucleotide (TFO) located 1517 bp downstream of the recognition sequence for enzyme EcoR124I was used to prepare linear DNA substrate to test DNA translocation (Simons et al., 2014).

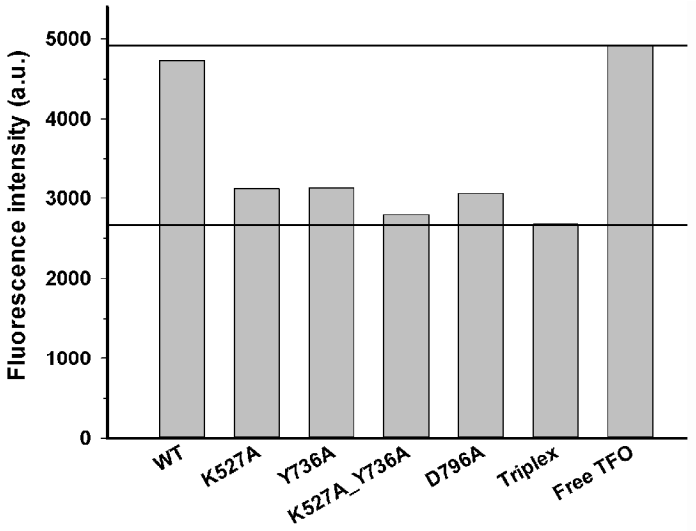



**Figure 18.** DNA translocation analysis by triplex displacement. Triplex-forming oligonucleotide (TFO) fluorescently labeled with TAMRA is hybridized to *linearized* DNA and

translocation is initiated by addition of reconstituted enzyme and ATP as described in Methods. Fluorescence intensities after 10 min reaction time are shown. Bound (triplex) and free (displaced) TFO are indicated by lower and upper horizontal lines, respectively. From left to right, enzyme reconstituted with: wildtype HsdR, mutant K527A HsdR, mutant Y736A HsdR, double mutant K527A_Y736A HsdR, mutant D796A HsdR, unreacted triplex only, and free TFO. No triplex displacement is observed in the absence of ATP or HsdR (not shown). Average values of fluorescence intensities calculated from four or more independent experimental replicates for each individual WT or mutant enzyme species are shown.

A triplex displacement assay was used to distinguish whether introduced mutations alter DNA translocation activity of EcoR124I. Hereby, high fluorescence intensity close to the intensity of free TFO in solution was detected only for WT (Fig. 18). All mutant enzymes K527A, Y736A, D796A and K527A_Y736A showed intensities close to the intensity of the DNA triplex in which the TFO is bound to duplex DNA (Fig. 18), indicating that these enzymes lost their ability to translocate duplex DNA and displace the TFO. In control experiments no triplex displacement was observed in the absence of ATP or HsdR.

### 3.2.2. How mutations at the helical-helicase 2 domain interface affect *in vitro* ATPase activity of EcoR124I

The helicase domains of EcoR124I R-M enzyme require a specific DNA substrate for ATPase activity. ATPase and translocation activities are coupled, ~1 ATP molecule is required for 1 bp advanced by EcoR124I (Seidel et al., 2008), and the DNA-dependent ATPase hydrolysis rate is considered to be related to the rate of translocation (Janscak et al., 1999; Sisáková et al., 2008b). So, it is possible to judge indirectly about DNA translocation of EcoR124I using the rate of ATP hydrolysis and compare whether EcoR124I with mutant HsdRs behaves as WT or not. If introduced mutations in HsdR subunit affect the ATPase activity as well as translocation, then consequent endonuclease activity might be significantly inefficient or lost.

EcoR124I was reconstituted *in vitro* using separately purified methyltransferase and WT or mutant HsdR subunits in mixed in ratio 1:6. The ATPase activity assay was carried out as a time course. 6-fold excess of circular plasmid DNA pRK over enzyme was used to prevent early cleavage of DNA

(Janscak et al., 1996). Control assays containing neither DNA nor methyltransferase indicate that ATP degradation is negligible in the absence of the enzyme. Results obtained using [$\gamma$-$P^{32}$]ATP method are similar to those from the malachite green assay.

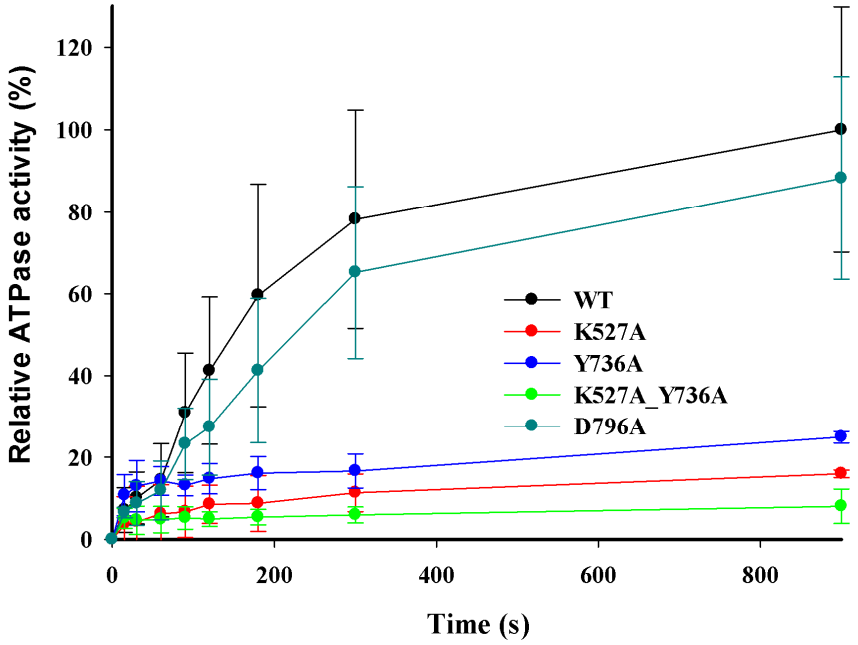



**Figure 19.** DNA-dependent ATPase activity. EcoR124I reconstituted from methyltransferase and HsdRs WT (black), or mutant K527A (red), Y736A (blue), D796A (green) or K527A_Y736A (dark cyan) was incubated at a final concentration of 15 nM with 90 nM circular plasmid DNA containing one recognition site and 2 mM ATP containing 0.16 μCi g-32P-ATP. At the indicated time points ATP and inorganic phosphate were resolved on cellulose TLC, autoradiographed, and scanned to quantify the extent of hydrolysis. The amount of ATP degraded is plotted as a function of time. Average values and standard deviations are calculated from at least three independent replicates for each mutant enzyme and six for WT enzyme. Data for mutants are normalized to the average value for WT at 900 sec taken as 100 %.

A relative ATPase activity of 100% was assigned to the enzyme reconstituted with WT HsdR. The relative ATPase activity of mutants Y736A, K527A and double K527A_Y736A mutant enzymes was ~25%, 16% and 8% (Fih. 19), indicating that these enzymes cannot hydrolyze ATP efficiently, consistent with the fact that these enzymes also lost their ability to translocate duplex DNA. However, mutant enzyme D796A exhibits a relative ATPase

activity of ~88% in this *in vitro* experiment, in apparent contradiction of the fact that this mutant had no translocation activity in the triplex-displacement experiments. Nevertheless, we can exclude that ATPase activity would be uncoupled from translocation in this case, as evaluation of the *in vitro* restriction assays (reported below) demonstrate that the D796A mutant did indeed translocate under the conditions used in the *in vitro* test and that the reason for the apparent contradicting behavior must be in the different conditions used for the assays and the excess of R-subunit in the ATPase and *in vitro* cleavage tests.

### 3.2.3. Influence of mutations at the helical-helicase 2 domain interface on endonuclease activity of EcoR124I *in vitro*

#### Cleavage of two-site linear DNA substrate

Restriction proficient $R_2M_2S_1$ complex reconstituted *in vitro* using 6 molar excess of HsdR subunit over MTase efficiently cleaves single-site and two-site DNA substrates (Janscak et al., 1998, Seidel et al., 2005). DNA to enzyme ratio of 1:10 was used for cleavage reaction with two-site linear DNA substrate, at this ratio each target site is occupied with $R_2M_2S_1$ complex and collision of both translocating enzymes results in effective cleavage (Drier et al., 1996; Janscak et al., 1996; Szczelkun et al., 1997).

Restriction activity assays for WT and mutant HsdRs were carried out as time courses at 37 °C. After a short pre-incubation the time point 0 aliquot was withdrawn and the reaction was initiated by addition of ATP to a final concentration 4 mM. Aliquots collected at each time point were immediately quenched in stop solution and additional enzyme activity was destroyed by heating to 65 °C for 5 min. Degradation of linear substrate was monitored using 1 % agarose/TAE gel (Fig. 20). Since it is not possible to quantify the density of formed products in this case, as the enzyme cleaves two-site linear substrate randomly producing DNA species of different length which appear on agarose gel as a smear (Janscak et al., 1996), the density of DNA substrate was used for quantification analysis (Fig. 21). WT enzyme showed the highest endonuclease activity and cleaved up to 80 % of two-site linear DNA substrate as was quantified. Mutation of the centrally located K527 residue to alanine as well as double K527A_Y736A mutation lead to a full loss of endonuclease activity, the

linear DNA substrate remains intact through the whole restriction time course. EcoR124I with alanine substitution of Y736 located in the interdomain linker region lead to reduction of endonuclease activity *in vitro*, mutant enzyme cleaves only ~30 % of two-site linear DNA substrate with 2 min delay.

Interestingly, substitution of D796 to alanine in the helical domain, canceling one of two possible H-bonds from K527 in the crystal structure of HsdR, does not lead to a drastic decrease in endonuclease activity *in vitro*, but D796A mutant enzyme cleaves ~75 % of two-site linear DNA substrate. However, the smear pattern for D796A mutant is different from that of WT. As already mentioned above in the paragraph about the ATPase assay, the seemingly contradicting result might be potentially caused by several reasons originating in the different conditions of the assays: the triplex displacement assay was carried out at 23 °C and restriction activity was tested at 37 °C; the distance between recognition sequences on two-site linear DNA pDRM-2R is 455 bp where two enzymes translocate toward each other, while single enzyme travels 1517 bp along DNA duplex to displace TFO. Thus, cleavage of particular DNA substrate requires less translocation effort than triplex displacement. Also all target sites on DNA are ocupied by translocating enzymes in the restriction activity assays and excess of enzyme from solution might bind nonspecifically to DNA anywhere between the two target sites presenting potential partner for collision for specifically bound enzymes. If indeed both activities, ATPase and translocase, would be uncoupled in the D796A mutant and the enzyme would cleave without prior translocation or only with very reduced translocation rate, DNA species of ~700 bp and ~1550 bp would be produced when cleaving two-site DNA close to target site. In comparison, WT enzyme would produce the whole range of DNA pieces from ~700 bp to ~2000 bp in length. The fact that a smear is observed in the mutant, too, excludes the possibility that the enzyme would not translocate prior to cleavage.

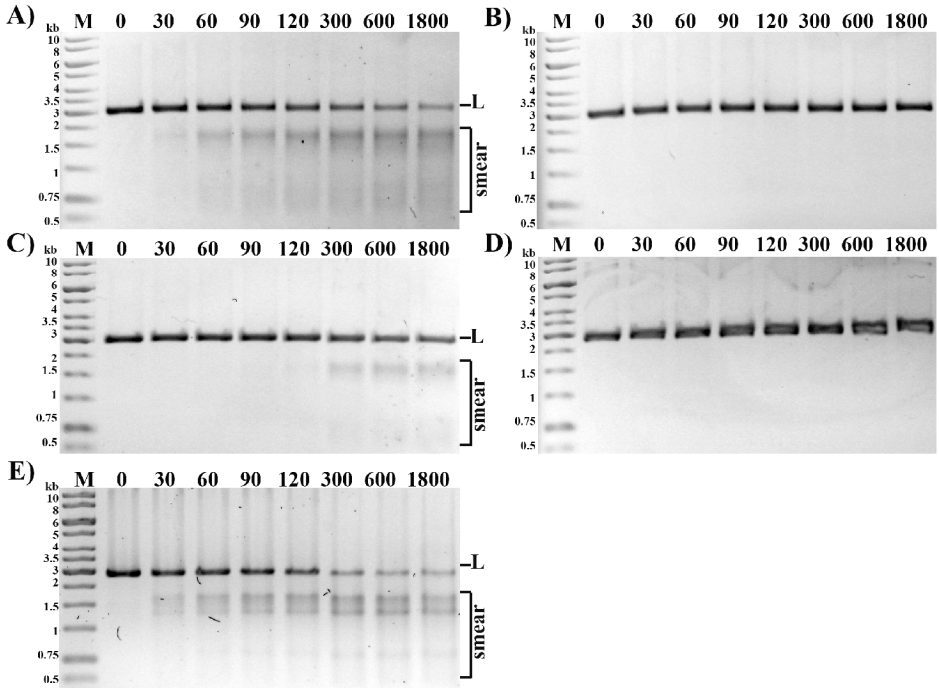



**Figure 20.** Cleavage of two-site linear DNA pDRM-2R. EcoR124I R-M complex was reconstituted *in vitro* using separately purified methyltransferase and WT or mutant HsdR subunits: A) WT; B) K527A; C) Y736A; D) K527A_Y736A; E) D796A. Aliquots were quenched at time points indicated in seconds and resolved by 1 % agarose/TAE gel stained with ethydium bromide. The gel was photographed under UV illumination, image was converted to grey scale and then inverted. The line corresponding to time point zero (before initiation reaction with ATP) corresponds to 100 % of linear substrate. Linear substrate is denoted as L, products of random cleavage are denoted as smear, DNA molecular weight marker with band size in kb is denoted as M.

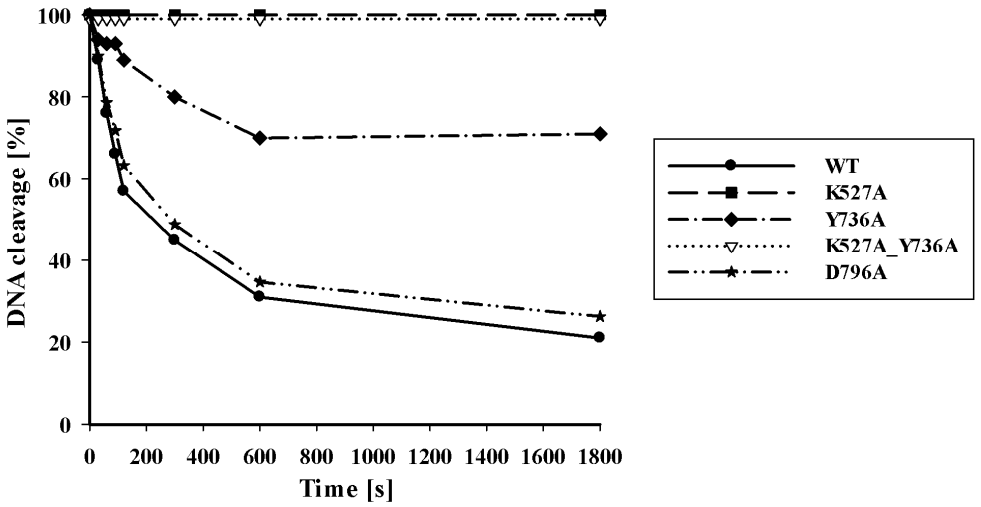



**Figure 21.** Quantitation of cleavage of two-site linear DNA substrate by WT, K527A, Y736A, K527A_Y736A and D796A mutant enzymes from Fig. 20. Decrease of DNA substrate was quantified by scanning the gels using ImageJ software (Abramoff et al., 2004), taking zero time point as 100 % of DNA substrate.

### Cleavage of single-site circular DNA substrate

Effective cleavage of circular DNA substrate with one target site for EcoR124I requires a lower enzyme concentration than the two-site linear DNA substrate, in this case one target site is occupied by one specifically bound enzyme. In comparison with two-site linear DNA substrate, cleavage of single-site circular DNA substrates has different mechanism of stalling of DNA translocation and character of formed products. Endonuclease activity on single-site circular DNA substrate was carried out as for two-site linear DNA substrate. For efficient DNA cleavage 1:1 ration of EcoR124I R-M enzyme to DNA was used as described in (Janscak et al., 1996). Although the cleavage is non-specific and distant from the single recognition sequence, the resulting linearized DNAs are all of the same length. The concentration of EcoR124I R-M enzyme was 15 nM. The EcoR124I pentameric complex was reconstituted *in vitro* using 15 nM methyltransferase and a 6 molar excess of HsdR subunit. Initially EcoR124I introduces a single strand cut producing the open circle (or nicked DNA) as intermediate product which then is degraded after cleavage of the second DNA strand into the linear DNA product.

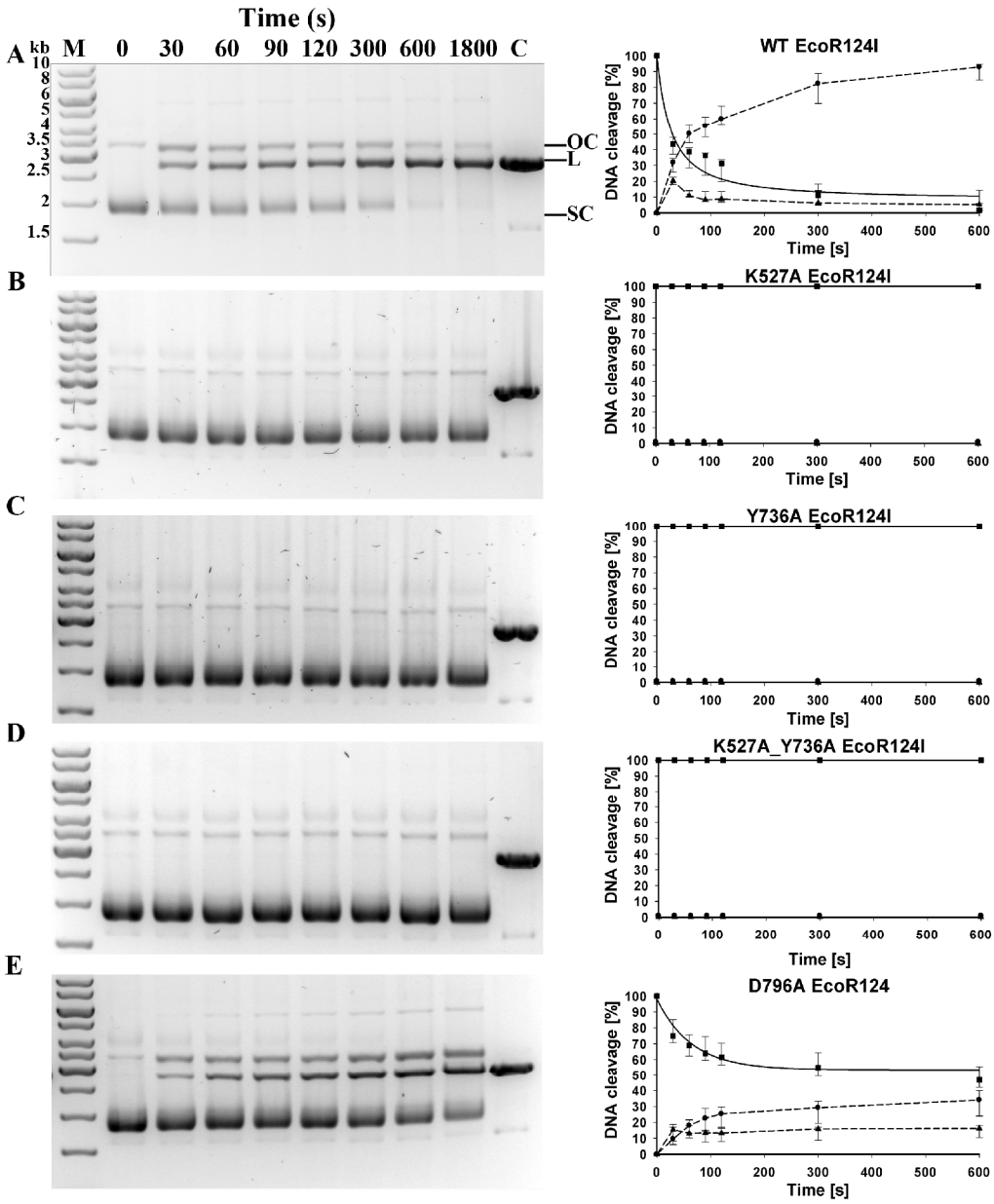



**Figure 22.** Cleavage of supercoiled substrate. **A.** Supercoiled DNA substrate plasmid pRK carrying single EcoR124I recognition site is cleaved by EcoR124I R-M complex *in vitro*. EcoR124I was reconstituted from HsdS1HsdM2 methyltransferase and A) WT HsdR or mutant HsdRs B) K527A, C) Y736A, D) K527A_Y736A and E) D796A. Aliquots were quenched at the time points indicated in seconds and resolved on agarose gels stained with ethidium bromide. On panel A open circle and linear products are denoted as OC and L, respectively, and supercoiled substrate as SC; control (plasmid pRK linearized by HindIII restriction enzyme) is

denoted as C; DNA molecular weight marker (M) with marked band size in kb on panel A. **B.** Quantification. DNA species were quantified using ImageJ software [Abramoff, Magalhaes & Ram, 2004]. Gels were scanned under UV illumination, and the image converted to grey scale and then inverted. The band density in the lane at time zero (before initiation of the reaction with ATP) is taken as 100 % of supercoiled substrate. The three indicated DNA species OC, open circular product (▲); L, linear product (●); SC, supercoiled substrate (■) were quantified individually from densities of bands in the gels on the left panels. Error bars represent standard deviations calculated from at least three independent experiments for each enzyme. Rates for the decrease of supercoiled DNA substrate were derived by fitting an exponential decay function to the data in SigmaPlot (solid lines); dashed lines connect the points only to guide the eye and do not represent fits to the data.

With reconstituted WT enzyme in 30 min of incubation, essentially all the circular substrate with a single EcoR124I recognition site is cleaved on one or both strands, generating approximately 90% linear product and 10% nicked DNA ("open circle" on Fig 2.). Mutants K527A, Y736A, and the double mutant K527A_Y736A are unable to cleave this circular DNA. Mutant D796A exhibits measurable but significantly decreased cleavage activity compared to WT, with ~47 % of the DNA remaining circular, ~35 % cleaved completely, and ~18 % nicked. To quantify the relative cleavage rates of different mutant proteins, rate constants for the disappearance of supercoiled DNA were estimated for WT and mutant D796A by fitting the data on Fig 2 with an exponential function. The best-fit value of the rate constant λ for the reconstituted WT enzyme is 0.0238 $s^{-1}$, whereas the rate constant λ for mutant D796A is approximately two thirds the WT value, 0.0154 $s^{-1}$. The difference between WT and mutant values is in a similar range as found previously for other mutants that display a restriction-deficient phenotype *in vivo* (Csefalvay et al., 2015). Thus the *in vivo* results can be considered a more sensitive indication of enzyme activity than the *in vitro* results.

### 3.2.4. Endonuclease activity *in vivo* reveals restriction phenotype

The ability of mutant HsdR subunits to bind to the MTase core and to form the pentameric EcoR124I R-M complex as well as the restriction proficiency of the formed complex were determined *in vivo* using a qualitative drop test (Colson et al., 1965). Unmodified bacteriophage λvir.0 was used for

simulation of a phage attack. The plasmid DNA pTrcR124 with corresponding introduced mutation(s) was transformed into competent cells JM109(DE3)[pACMS] and JM109(DE3)[pKF650]. JM109(DE3)[pACMS] cells are able to produce HsdM and HsdS subunits of EcoR124II and display the r− phenotype. JM109(DE3)[pKF650] cells expressing all three subunits for EcoR124II R-M system display the $r^+$ phenotype. JM109(DE3)[pACMS] and JM109(DE3)[pKF650] were used for positive and negative complementation tests, respectively. Low level of expression of *hsd* genes from pACMS and pKF650 plasmids without IPTG induction was sufficient for the complementation assay (Sisáková et al. 2008b). In combination, both tests allow to determine the restriction phenotype and the ability of mutant HsdR subunits to bind to methyltransferase.

EcoR124I and EcoR124II are members of the same Type IC family, show a high degree of similarity and can interchange with its HsdR subunits preserving translocation/endonuclease activities (Dreier et al., 1996; Janscak et al., 1999). The complementation tests are based on this feature. In negative complementation test HsdR subunits of EcoR124I expressed from pTrcR124 and HsdR subunits of EcoR124II expressed form pKF650 compete for binding to MTase of EcoR124II. Mutant HsdRs, which can bind to MTase and have affected endonuclease activity, alter the $r^+$ phenotype of the host to r± or r−, displaying a transdominant effect. Mutant HsdRs, which are capable of assembly with MTase and possess intact endonuclease activity would not show any effect in the negative complementation test, however in the positive complementation test they would change an r− phenotype to $r^+$.

WT HsdR complements restriction in the $r^-$ host in the positive complementation test (Table 3, $r^-$ host), as expected (value 0.001, within the range 0.0001–0.01 for the restriction proficient phenotype). All mutants, K527A, Y736A, K527A_Y736A and D796A, fail to complement restriction in the $r^-$ host (all values within the range 0.1–1 for the restriction-deficient phenotype) and are therefore restriction-deficient. In the negative complementation test the reduced restriction activity of the all mutants in the $r^+$ host (Table 3, $r^+$ host) indicates that the mutant subunits are fully competent for assembly with methyltransferase to form the endonuclease complex. The results for mutants K527A, Y736A, and K527A_Y736A were suggested already by the

*in vitro* tests. The substitution of D796 to alanine clearly alters enzymatic activity of the EcoR124I R-M complex, leading to a restriction-deficient (r−) phenotype *in vivo*. This result is in apparent contradiction to the *in vitro* results with this mutant, although the *in vivo* results can be considered a better indication of enzyme activity than the *in vitro* results.

**Table 3.** Effect of mutations on the restriction phenotype of EcoR124I. Values 0.0001-0.01 corresponds to r+ phenotype, 0.01-0.1 – r± and 0.1-1.0 – r−.

| | **restriction**[a] | | **Ability of cleavage** | **Ability of assembly** |
|---|---|---|---|---|
| **HsdR** | **r− host** [b] | **r+ host**[c] | | |
| WT | 0.0012±0.0005[SD] | 0.0011±0.0004 | Yes | Yes |
| K527A | 0.4388±0.3009 | 0.1092±0.0857 | No | Yes |
| Y736A | 0.3197±0.2147 | 0.2307±0.0694 | No | Yes |
| K527A+Y736A | 0.4230±0.1953 | 0.1619±0.0666 | No | Yes |
| D796A | 0.3962±0.0885 | 0.1965±0.1335 | No | Yes |

[a] restriction activity was determined as the efficiency of plating of λvir.0 on tested strains relative to the efficiency of plating of λvir.0 on *E. coli* JM109(DE3) indicator (nonrestricting) strain.
[b] The positive complementation was tested in r− host *E. coli* JM109(DE3)[pACMS] (r−m+).
[c] negative complementation (transdominant effect) in r+ host *E. coli* JM109(DE3)[pKF650] (r+m+).

### 3.2.5. Rotational motion of the helicase 2 domain in HsdR subunit of EcoR124I during translocation cycle

HsdR subunit of EcoR124I belongs to SF2 helicases and 2$^{nd}$ domain of helicases is suggested to rotate during translocation activity having alternating affinity to DNA. Thereby interactions occurring at the helical-helicase 2 domain interface in HsdR subunit appear to play important role in relative rotation of the helicase 2 domain and dsDNA advancement. Several crystal structures of SF2 helicases have been used for modeling of relative position of the helicase 2 domain in HsdR of EcoR124I. In YASARA the excised helicase domain 2 from HsdR was aligned with corresponding domains from other helicase structures and positioned in HsdR in this new orientation. These new positions of the helicase 2 domain in HsdR, modeled in analogy of conformations observed in

crystal structures of Rad54, RecG, NS3, RecQ, UvrB, were ordered into a sequence of motions that the helicase 2 domain in HsdR might undergo during DNA advancement (Fig. 24).

At the initiation stage, which was modeled from Rad54 (1Z63) structure (Dürr et al., 2005), the helicase 2 domain is rotated 180° relative to the helicase 1 domain. The helicase 2 domain is in open conformation and its helicase motifs IVa and VI are moved away from ATP binding pocket as in Rad54. The helicase motifs Ia, Ib and IIa are compactly clustered in the helicase 1 domain and ready for dsDNA binding. Such DNA free open conformation was shown to exist in solution for Rad54 by FRET experiments (Lewis et al., 2008). EcoR124I pentameric complex was suggested to be in relaxed conformation prior DNA binding (Kennaway et al., 2012), and likely 180° relative rotation of the helicase 2 domain is expected.

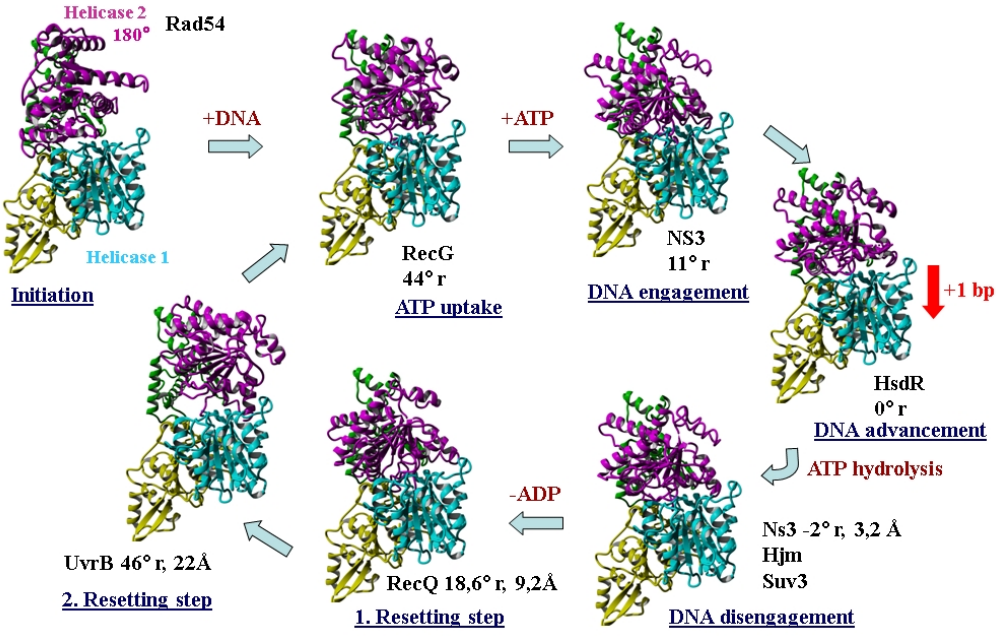



**Figure 23.** Proposed positions of the helicase 2 domain in HsdR subunit of EcoR124I during translocation cycle. Rotational angles were measured through N632 (Cα) – center of mass of the helicase 2 domain – center of mass of the helicase 1 domain – T364 (Cα). Widening of the cleft was measured as distance between N632 (Cα) in modeled conformation and N632 (Cα) in the HsdR (2w00) crystal structure.

The first stage, ATP uptake, would corresponds to the conformation observed in the structure of RecG (pdb 1GM5) (Singleton et al., 2001), with 44° relative rotation of the helicase 2 domain about the same axis as seen for the 180° rotation at the initiation stage. This conformation is favorable for ATP to enter ATP binding pocket. The helicase 1 domain contacts dsDNA via Ia, Ib and IIa motifs in all stages of the translocation cycle. Motif IVa gets closer to the helicase 1 domain and is well oriented to contact DNA.

The second stage, DNA engagement, is represented by the structure with 11° rotation, as observed in NS3 (pdb 3O8D) (Appleby et al., 2011). At this stage the helicase 2 domain is strongly bound to dsDNA what would correspond to a high affinity state. The helicase cleft is in a closed conformation with ATP in the binding site interacting with Q, I and II motifs. The intermediate rotational angle suggests that DNA is already involved in a translational movement and half-way to its advanced position.

The third, DNA advancement, stage is observed in the HsdR structure itself, here we observe 0° rotation and ATP is present in ATP binding pocket. The helicase cleft is collapsed what results in the advancement of DNA by 1 bp in the 3'-5' direction. The characteristic helicase motifs in HsdR (2W00) (Lapkouski et al., 2009) are compactly clustered around the ATP binding pocket in two RecA-like helicase domains and keep the position of ATP. The helicase motifs Ia, Ib and IIa from the helicase 1 domain and IVa from the helicase 2 domain contact dsDNA substrate via 5'-3' strand and 3'-5' strands, respectively. DNA is already advanced by 1 bp. During DNA advancement, a rapid hydrolysis of ATP takes place what leads to a conformational change that cause unbinding of the helicase domain 2 from DNA.

At the fourth stage, DNA disengagement, ADP and Pi are in the ATP binding pocket, -2° rotation is observed, as in NS3 (pdb 3O8C) (Appleby et al., 2011). ATP hydrolysis promotes initial conformational changes associated with DNA unbinding from the helicase 2 domain what is suggested by the small widening of the helicase cleft by 3.2 Å and motifs IVa and VI are shifted away from the ATP binding pocket. This stage corresponds to a state with low affinity for DNA binding as DNA starts to loose interactions with the helicase 2

domain. Release of ADP and Pi promotes large rotation of the helicase domain 2 and wide opening of the helicase cleft what results in a 1st resetting stage.

The 1st resetting step, modeled from the conformation observed in RecQ (pdb 2V1X) (Pike et al., 2009), represents a stage with 18.6° rotation and 9.2 Å helicase cleft widening. ADP and Pi are released, the helicase cleft is widely open and domain 2 is rotated by a quite large angle, thus the helicase 2 domain completely loses its interactions with DNA and can easily return to its initial position without pushing the DNA back.

The 2nd resetting step (modeled from UvrB (pdb 1D9X) (Theis et al., 1999)) with 46° rotation and 22 Å helicase cleft opening returns the helicase 2 domain the completely back to the first stage with binding to DNA substrate upstream, as seen in RecG. Then the translocation cycle can be repeated.

In contrast to the impression one might get from crystal structures, proteins in solution are dynamic. The HsdR crystal structure (2W00) with interdomain interactions involving K527 and Y736 residues represents only a single snapshot the HsdR subunit would sample during translocation cycle. So, H-bonding occuring between the helical and the helicase 2 domain via K527 and Y736 are not constant during translocation cycle. HsdR structures with modeled position of the helicase 2 domain could represent another snapshots the HsdR subunit would undergo during translocation.

So at the initiation stage, K527 residue is moved away from its partners D796 and N794. Orientation of the helicase 2 domain is stabilized by H-bonds occurring via side chain of D796 residue to backbone of T667 residue and side chain of N794 to backbone oxygen of K697. Y736 is in the same orientation as in 2W00. E730 is moved away from Y736 and stabilized by H-bonding with side chains of K731 and K474.

At the firs stage, ATP uptake, H-bond is formed between side chains of K527 and D796, also side chain of N794 interacts with I791 residue. Domain's orientation is stabilized via H-bonding between side chains of E730 and K474 residues.

At the second stage, DNA engagement, side chain of K527 residue forms H-bonds to side chain and backbone oxygen of N794 residue. Additionally, the side chain of N794 is stabilized by interaction with backbone

oxygen of I791. OH-group of Y736 interacts with backbone oxygen of K731 forming H-bond.

At the third stage, DNA advancement, side chain of K527 residue forms bifurcated contact to D796 and N794 residues as was described earlier. Side chain of N794 is additionally stabilized by H-bonding with side chain of N701 and backbone oxygen of I791. Side chain of E730 residue contacting OH-group of Y736 is additionally stabilized by H-bond with backbone NH of K731.

At the fourth stage, DNA disengagement, side chain of K527 preserves H-bond to backbone of N794 and looses its interaction with side chain of D796 residue. Side chain of N794 still contacts I791 residue. OH-group of Y736 interacts with backbone oxygen of K731 residue. Side chain of E730 forms two H-bonds to side chains of K474 and K731 residues.

At 1$^{st}$ and 2$^{nd}$ resetting steps, the H-bonding network formig by K527 and Y736 residues is not present anymore and other residues are involved in stabilizing the relative orientation of the helicase 2 domain in HsdR.

Thus, interactions provided by K527 and Y736 are not constant during translocation cycle. They appear to play important role in DNA advancement. Binding and hydrolysis of ATP with release of ADP and Pi promote shutting of the helicase cleft with consequent rearangment of interactions at the helical-helicase 2 domain interface.

## 3.3. The helicase motif III and its extended region

Translocation activity of EcoR124I R-M complex is housed on the motor subunit in the two helicase domains where the seven characteristic motifs of DEAD-box helicases are clustered (Gorbalenya & Koonin, 1993; Murray, 2000; McClelland & Szczelkun, 2004). Some members of SF2 helicases, for example Rad54 from *S. cerevisiae*, lack strand separation activity like other classical helicases and act as true ATP-dependent translocases (Flaus & Owen-Hughes, 2011). The sequence region in Rad54 from yeast adjacent to the actual helicase motif III originally was thought to be a conserved proliferating cell nuclear antigen (PCNA)-interacting protein (PIP) motif. Recently it was shown that this motif actually plays no role in stabilizing PCNA-interactions, instead, it is critical for ATPase activity and required for proper functioning of the

translocase (Burgess et al., 2013; Zhang et al., 2013). In HsdR of EcoR124I, residues 436 to 448 represent a partially conserved extended region of the helicase motif III, also it is partially conserved throughout the Type I R-M systems family IC, with the critical leucine fully conserved, and the critical glutamine replaced by the related glutamic acid (Fig. 24). The high similarity between glutamine and glutamic acid should conserve the three-dimensional structure and is marked in the alignment on figure 24 with '**:**'. However, sequence alignment suggests the motif III is not extended in other Type I R-M families IA and IB (Fig. 24).


**Figure 24.** Sequence alignment of HsdR of EcoR124I from *E. coli* with Rad54 and other Type I members. A) The sequence of HsdR of EcoR124I has partially conserved extended region of motif III as in Rad54 translocases from Swi/Snf2 protein family (UniProt accession numbers: *S. cerevisiae* Rad54 – P32863, *D. rerio* (Zebrafish) Rad54 - Q7ZV09, *H. sapiens* Rad54 - Q92698). The location of the helicase motif III and its extended region in Rad54 sequence is marked below the alignment. B) The HsdR subunit of EcoR124I from Type IC family is shown to have the extended region of the helicase motif III. In motor subunits of EcoKI (UniProt: P08956) and EcoAI (UniProt: Q07736) from IA and IB families, respectively, it is not present. The conservation of the amino acid sequence around motif III is depicted: '*' indicates positions which have a single, fully conserved residue; '**:**' indicates highly similar residues that should conserve the structure and '**.**' indicates somewhat similar residues.

Mutational analysis within extended region of the motif III in Rad54 from yeasts shows that substitution of two highly conserved aromatic residues Y494 and F495 to alanine resulted in a dramatic reduction in ATPase activity. In contrast to WT Rad54, this double mutant has been shown to be defective in branch migration of a Hollyday junction (Burgess et al., 2013; Zhang et al., 2013). While these two residues do not have an equivalent in EcoR124I,

mutations in the conserved positions 488 and 491, Q488A and L491Q, lead to a drastic reduction in ATPase and D-loop formation activities (Burgess et al., 2013). Docking of a dsDNA filament into the Rad54 structure, modeled from the X-ray crystal structure of Rad54 from zebrafish (pdb id: 1Z3I), suggests that the extended motif III directly contacts DNA. Thus the extended region of motif III is important for dsDNA binding which is likely linked to ATPase activity for Rad54 (Zhang et al., 2013), however mutational analysis within the motif III with G484R substitution suggests that the conserved glycine is important only for DNA-dependent ATPase activity but not for dsDNA binding (Smirnova et al., 2004).

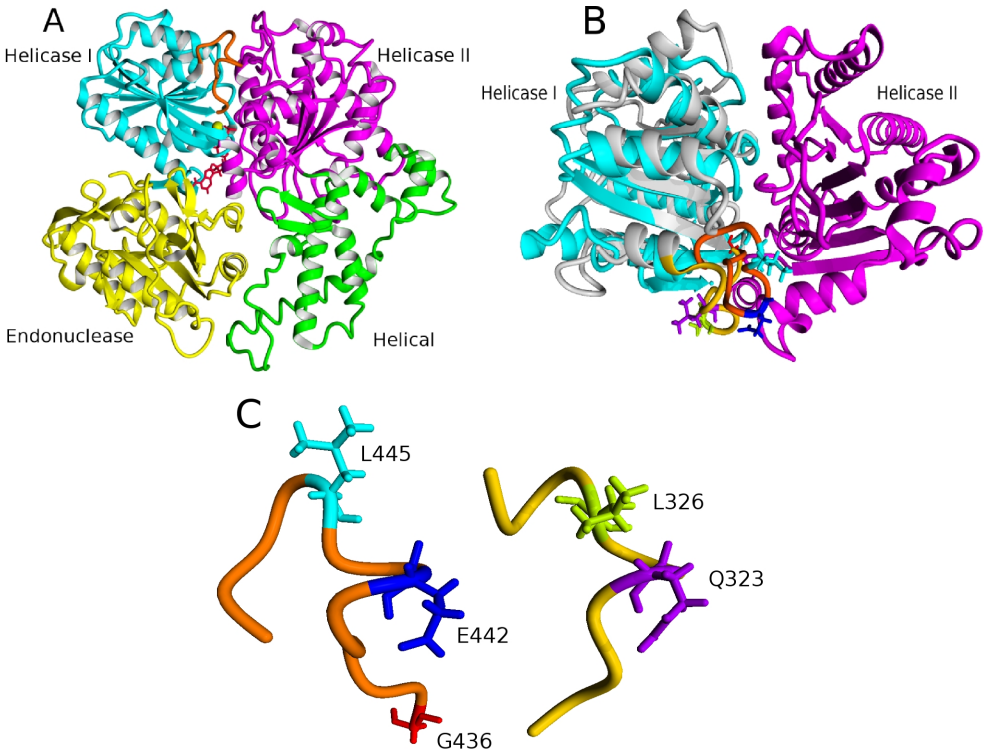



**Figure 25.** 3D-arrangement of the extended region of motif III. A) Structure of HsdR consisting of four domains which are color coded here, Yellow is endonuclease domain while Helicase 1 and Helicase 2 domains are depicted cyan and magenta, respectively. The helical domain is shown in green. ATP is represented in red as a skeletal model and magnesium is indicated by yellow sphere. The extended motif III is in orange. B) Structural alignment of HsdR of EcoR124I from *E. coli* with Rad54 from *D. rerio* (zebrafish). The orientation of the helicase

domins of HsdR of EcoR124I corresponds to top view on panel A. The crystal structure of HsdR of EcoR124I (2w00) and Rad54 from *D. rerio* (zebrafish, pdb id: 1z3i) were used for structural pairwise alignment in YASARA using helicase 1 domains (in cyan) and helicase 2 domain (in magenta) of EcoR124I and helicase 1 domain (in gray) of Rad54. Residue Q448 and L491 in Rad54 from *S. cerevisiae* correspond to Q323 and L326 of Rad54 from *D. rerio*, respectively. HsdR loop representing the extended motif III and containing amino acid residues for mutagenesis is shown in orange and the corresponding loop from Rad54 in dark yellow. Side chains of residues in HsdR are G436 (red), E442 (blue) and L445 (cyan) and in Rad54 are Q323 (light green) and L326 (purple). C) Conformational alignment of loop of HsdR in orange and Rad54 in yellow with the respective side chains of residues in colors as on panel B.

The extended region of motif III from Rad54 and HsdR of EcoR124I bears sequence and structural similarities. As in Rad54, in HsdR subunit the extended region with conserved E442 and L445 residues forms a loop following the β-strand and placed at the interface between two helicase domains outside the helicase cleft (Fig. 25). Side chains of both residues are oriented towards the solvent. Conserved residue G436 within the helicase motif III in the HsdR crystal structure (2W00) is located at the C-terminal end of β-strand *q* facing the ATP binding pocket, however the closest atoms of the ATP molecule is approximately 7 Å far from it. Thus, the location of residues E442 and L445 in the HsdR subunit and corresponding residues in Rad54 is very similar. The small differences in the loop position might be easily explained by the fact that the second helicase domain in Rad54 is in a different conformation (Dürr et al., 2005) compared to EcoR124I, where this domain undergoes large conformational changes during translocation that influence the interdomain interface (Lewis et al., 2008). In order to elucidate the role of extended motif III we prepared EcoR124I mutants G436A, E442A, L445A and L445E (see methods) and tested them *in vitro* and *in vivo*. For more details, see **Paper 4**.

### 3.3.1. ATP hydrolysis *in vitro*

Residue E442 and L445 in the extended region of the helicase motif III in the HsdR subunit corresponding to those critical for ATPase activity in Rad54 from yeast and residue G436 within the core of the helicase motif III were substituted by alanine, L445 additionally was replaced by glutamic acid. Standard malachite green assays (Chan et al., 1986) were used to measure the

rate of ATP hydrolysis of reconstituted mutant enzymes. The ratio of R-subunit to MTase was 6:1 and circular DNA substrate pRK with single EcoR124I target sequence was in 6-fold excess over enzyme concentration. ATP hydrolysis was detected only when all reaction components were present in the reaction mixture. In the absence of DNA substrate, or MTase or HsdR subunit only negligible ATP hydrolysis was observed. To determine kinetic constants ATPase activity for WT and mutant enzymes was measured at ATP concentrations up to 4 mM, higher ATP concentrations lead to decrease in the initial velocity similarly as was reported previously for EcoR124II R-M enzyme (Dreier & Bickle, 1996). Mutations introduced within the extended region of the helicase motif III only slightly altered ATPase activity compared to WT (Fig. 26, Tab. 4). E442A mutant has $K_m$ value 82.4 ± 8.1 μM as WT, while mutant enzymes L445A and L445E are slightly less efficient ATPases with $K_m$ values of 129.8 ± 9.3 μM and 102.4 ± 7.9 μM, respectively. All mutants from the extended region showed a lower $V_{max}$ in comparison to WT but within the same range. However, substitution of the highly conserved residue in an intrinsic part of the helicase motif III, highly conserved among all Type I members and other helicases (Sharples et al., 1994; Webb et al., 1996; Davies et al., 1999) resulted in a ~5-fold decrease in $V_{max}$ and ~4.5-fold increase in Km value in comparison to WT.

**Table 1.** Kinetic constants from ATPase activity assays. SD - standard deviation from 3 repeats.

| | **$K_m$ (μM) (±SD)** | **Vmax ($s^{-1}$) (±SD)[a]** | **Relative ATPase rate[b]** |
|---|---|---|---|
| **WT** | 81 ± 8 | 519 ± 34 | 1 |
| **G436A** | 373 ± 10 | 99 ± 28 | 0.19 |
| **E442A** | 82 ± 8 | 443 ± 36 | 0.86 |
| **L445A** | 129 ± 9 | 461 ± 38 | 0.89 |
| **L445E** | 102 ± 7 | 453 ± 27 | 0.87 |

[a] ATP per HsdR assuming a fully saturated system and 100% occupancy.

[b] Vmax,mutant relative to Vmax,WT.

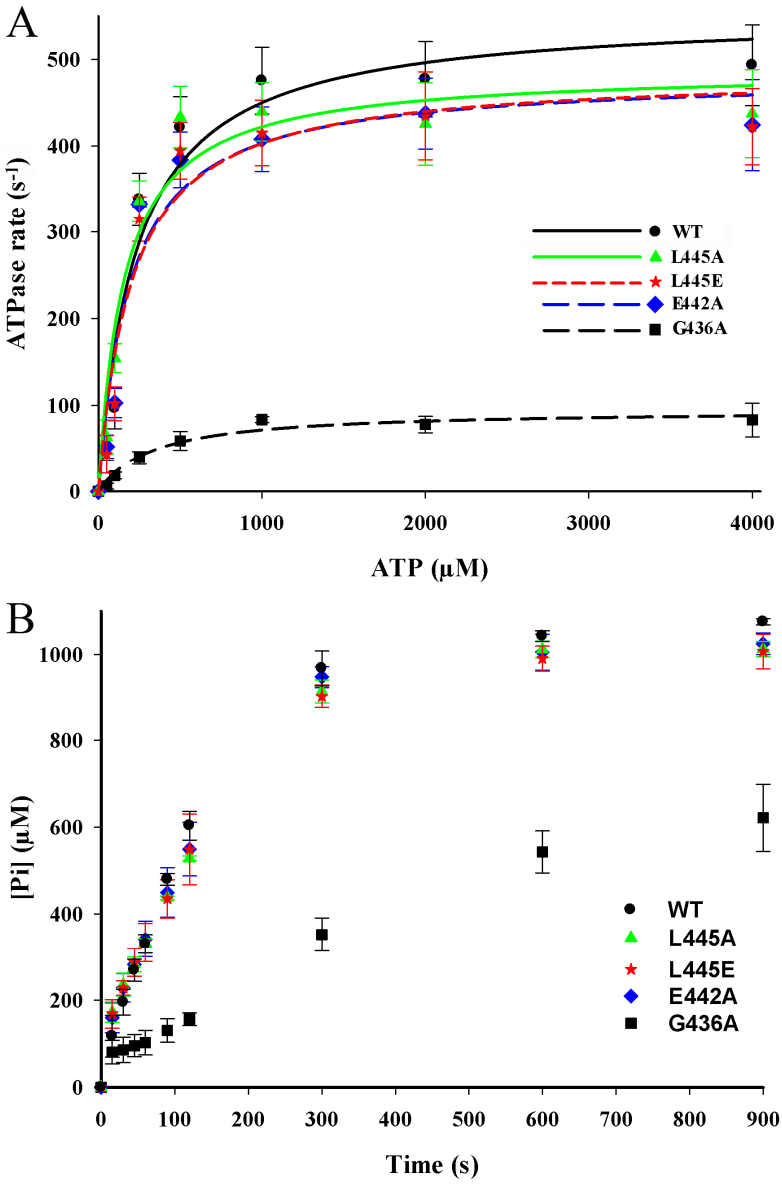



**Figure 26. ATPase activity as a function of ATP concentration.** A) The amount of inorganic phosphate released during DNA-dependent ATPase reaction was monitored using a standard malachite green assay. The DNA-dependent data was fitted to a Michaelis–Menten relationship in SigmaPlot to derive the curves shown and the values in Table 1. Standard deviation is calculated from three independent replicates and plotted as *error bars*. B) ATP hydrolysis rate at 4 mM ATP concentration. The concentration of inorganic phosphate released in the DNA-dependent reaction is plotted as a function of time. The error bars represent one standard deviation calculated from three independent replicates.

G436A mutant enzyme with $K_m$ value of 373.9 ± 10.7 μM is less efficient motor compared to WT. Earlier it was shown that opposite conservative change of alanine to glycine in equivalent position (A619) in EcoKI only slightly reduced ATPase activity *in vitro*, however increase in side-chain length of residue in this position (A619V) lead to complete loss of ATPase activity *in vitro*. These effects were interpreted in the light of a direct contact between ATP and the helicase motif III, and a role of motif III in coordinating the γ-phosphate was assumed (Davies et al., 1998, 1999). Also we attempted to introduce G436V mutation (not mentioned in methods section), however it was not possible to express protein with such substitution even at low temperature. Likely, longer residue side chain in this position provokes misfolding and aggregation of the motor subunit, G436V mutant HsdR was found only in the pellet fraction.

### 3.3.2. Cleavage of supercoiled substrate *in vitro*

Translocation activity powered by ATP is coupled to endonuclease activity in EcoR124I, and the signaling pathway involves two loops, 180s and 220s, and nucleotide competent states determine the rotation of the endonuclease domain leading to possible DNA cleavage (Sinha et al., 2014). Any alteration in ATPase activity disturbing DNA translocation might affect consequent endonuclease activity. The ability of EcoR124I with mutations in the helicase motif III and its extended region to cleave single-site supercoiled DNA substrate was tested as was described previously in section 3.2.3. Decrease of supercoiled DNA substrate pRK (described as pDRM-1R in Janscak et al., 1996) was used to evaluate effectivness of DNA cleavage by mutant enzymes. The amount of remaining supercoiled DNA substrate was quantified for each time point from three independent time course experiments. Average values for remaining DNA substrate at each individual time point were plotted as a function of time, error bars were calculated from three replicates. Additionally, three replicates plotted for each individual time point allow estimating cumulative variability of DNA cleavage assay coming from agarose gel preparation, sample loading and quantification of DNA species. To distinguish slight alterations in the endonuclease activity of EcoR124I on

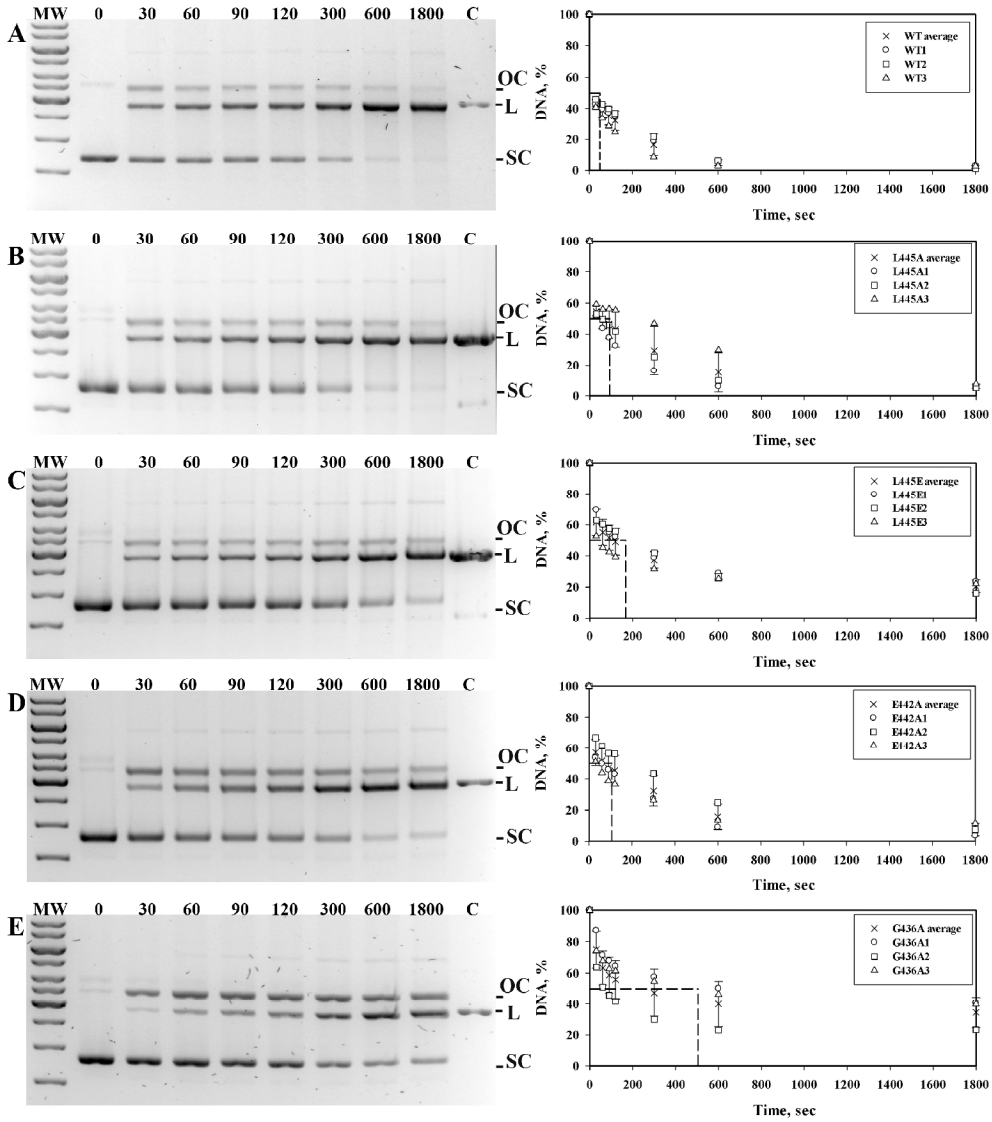



**Figure 27.** Cleavage of supercoiled DNA substrate. Typical gels from *in vitro* clevage of single-site supercoiled DNA substrate are shown (on the left). EcoR124I was reconstituted from methyltransferase and A) WT HsdR or mutant HsdRs B) L445A, C) L445E, D) E442A and E) G436A. Aliquots were quenched at indicated in second time points and resolved by agarose gel stained with ethydium bromide. The gel was scanned under UV illumination, image was converted to gray scale and then inverted. The line corresponding to time point zero (before initiation reaction with ATP) is to contain 100 % of supercoiled DNA substrate. Control (linearized plasmid DNA pRK by HindIII) is denoted as C; DNA molecular weight marker is denoted as MW, OC – open circle, L – linear DNA, SC – supecoiled DNA. Quantification analysis (on the right). Remaining supercoiled DNA substrate in the restriction time course experiment quantified individually using ImageJ software (Abramoff et al., 2004) is plotted as a

function of time. Three replicates and average value with error bars are shown. Dashed perpendicular on X-axis indicates $T_{1/2}$ value, time when half of supercoiled DNA substrate is cleaved (see text).

supercoiled DNA substrate caused by point mutation in the helicase motif III and its extended region, the $T_{1/2}$ value was introduced as the time where 50 % of DNA substrate is cleaved already (Fig. 27, $T_{1/2}$ is derived from an exponential decay fit (not shown) and visualized by dashed lines perpendicular to the X-axis). Mutant enzymes E442A and L445A require approximately twice as long to cleave 50 % of DNA substrate in contrast to WT. The $T_{1/2}$ value for WT corresponds to 48.8 sec, while E442A and L445A mutants reach half DNA substrate cleavage at 106.6 sec and 92.9 sec, respectively. Introducing of a negatively charged glutamic acid instead of L445 leads to ~2.5-fold increase in $T_{1/2}$ value, this mutant enzyme degrades 50 % of DNA substrate in 166.9 sec. While all this values are still close to each other, the substitution of the conserved G436 residue by an alanine resulted in a drastic increase in $T_{1/2}$ value with one order difference: $T_{1/2}$ value for G436A mutant enzyme is 504.9 sec. Our data indicates that point mutations in the extended region of the helicase motif III have only a minor or slight effect on the endonuclease activity of EcoR124I compared to substitution of the conserved residue in motif III itself. Despite conservative change of glycine to alanine in position 436, already the minor increase in residue side chain length affects the endonuclease activity (consequence of affected ATPase activity).

### 3.3.3. Endonuclease activity *in vivo*

The ability of mutant HsdR subunits to bind to the MTase core was additionally confirmed by *in vivo* experiments (for more details see section 3.2.4.). Mutations in the helicase motif III and its extended region do not affect the assembly of EcoR124I in living cells. All mutant enzymes with mutations in the extended region were able to restore the restriction phenotype in positive complementation test and in negative complementation tests the restriction-proficient phenotype remained unchainged (Tab. 5). Mutant EcoR124 enzyme with only nicking activity on single-site supercoiled DNA substrate leads to restriction-deficient phenotype *in vivo* (Sinha et al., 2014). Despite the fact that the G436A mutant enzyme is able to degrade intermediate open circle DNA

into a linear DNA product, the second DNA strand cut occurs with large delay as judged by the $T_{1/2}$ value from *in vitro* DNA cleavage assay. Large $T_{1/2}$ values indicate that G436A mutant enzyme requires more time to translocate the entire DNA circle until translocation is impeded, or it takes a longer time to initiate DNA translocation due to possible structural changes in the ATP binding pocket produced by longer residue side chain. So, the G436A mutant enzyme is very likely to be a slower and less effective motor, and consequently *E. coli* cells expressing G436A mutant enzyme are restriction-deficient and cannot successfully resist phage infection.

**Table 5.** Restriction phenotype of EcoR124I with mutations in the helicase motif III and its extended region.

| | **restriction** [a] | | **Ability of cleavage** | **Ability of assembly** |
|---|---|---|---|---|
| **HsdR** | **r− host**[b] | **r+ host**[c] | | |
| WT | 0.0049±0.0031[SD] | 0.0013±0.0012 | Yes | Yes |
| G436A | 0.10±0.033 | 0.024±0.031 | No | Yes |
| E442A | 0.0089±0.0092 | 0.006±0.0073 | Yes | Yes |
| L445A | 0.0117±0.0037 | 0.004±0.002 | Yes | Yes |
| L445E | 0.0079±0.004 | 0.0035±0.0027 | Yes | Yes |

[a] restriction activity was determined as the efficiency of plating of λvir.0 on tested strains relative to the efficiency of plating of λvir.0 on *E. coli* JM109(DE3) indicator (nonrestricting) strain.

[b] The positive complementation was tested in r− host *E. coli* JM109(DE3)[pACMS] (r−m+).

[c] negative complementation (transdominant effect) in r+ host *E. coli* JM109(DE3)[ pKF650] (r+m+).

SD - The standard deviation

Data from MD simulations and principle components analysis show that oscillation of the helicase 2 domain relative to the helicase 1 domain in G436A mutant HsdR is larger than in WT and other mutant HsdRs. Comparison of mutant HsdR structures before and after simulations revealed that in G436A mutant HsdR displacement of the helicase 2 domain is larger while for other mutant HsdRs it is almost identical to WT, hence the increased interdomain distance measured between centers of mass of both helicase domains in G436A mutant HsdR is suggesting opening of the helicase cleft. Hydrogen bonding plays a crucial role in stabilizing domain-domain interfaces. So, the most probable explanation of the helicase cleft opening is the observed lower number of H-bonds at the interface formed by the helicase domains if compared to WT

and other mutant HsdRs. Structural comparison with other related helicases suggests that G436A mutant HsdR might represent single snapshot from DNA translocation cycle, however from such domain's arangement motor subunit cannot normally initiate translocation and possibly higher number of attempts is required to start it. Detailed analysis of MD simulations is given in **Paper 4**.

### 3.3.4 Crystallization of G436A and data collection

G436A mutant HsdR subunit was purified as was described previously (see section 2.4.) (Janscak et al., 1996). 50 kDa cutoff spin concentrator (Millipore) was used for protein concentration and buffer exchange (to 20 mM phosphate buffer, pH 7.5, 0.1 M KCl). Freshly isolated sample of the G436A mutant HsdR subunit of EcoR124I was used for crystallization trials under the same conditions that were used for the wild type of the protein (PDB ID 2W00) (Lapkouski et al., 2007). The crystals of introduced mutant variant were diffracted to a maximum resolution in 10 Å. In order to improve the crystal quality packing optimization steps were applied in crystallization experiments. The sitting-drop vapor-diffusion method was used for protein crystallization and following protein concentrations were used 27 mg $ml^{-1}$, 22 mg $ml^{-1}$, 15 mg $ml^{-1}$ and 11 mg $ml^{-1}$ supplemented with ATP to a final concentration 5 mM. Precipitant (500 µl reservoir solution (24-well Cryschem Plate with well capacity: 1.5 ml.)) containing 0.2 M $Li_2SO_4$, 8 % PEG 20 000 and increased concentration of PEG550 MME in precipitant to 15 % or 20 % as well as doubled volume of crystallization drop (4 µl of precipitant and 4 µl of protein) allowed obtaining crystals of bigger size with dimensions of about 0.3 x 0.2 x 0.05 mm (Fig. 28 A). A crystal mounted in a nylon loop (Hampton Research) was transferred for a few seconds to 10 µl of cryosolution (25 % glycerol (in water) or 100 % PEG550 MME) and then was flash-frozen in liquid nitrogen prior measurement at a synchrotron facility. Finally, the diffraction data was collected at the MX14.1 beamline operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron storage ring (Berlin-Adlershof, Germany) (Mueller et al., 2012). The crystal diffracted to max resolution up to 3 Å, but with a horrible anisotropy: good diffraction at some angle and almost nothing after rotation to 90° (Fig. 29).

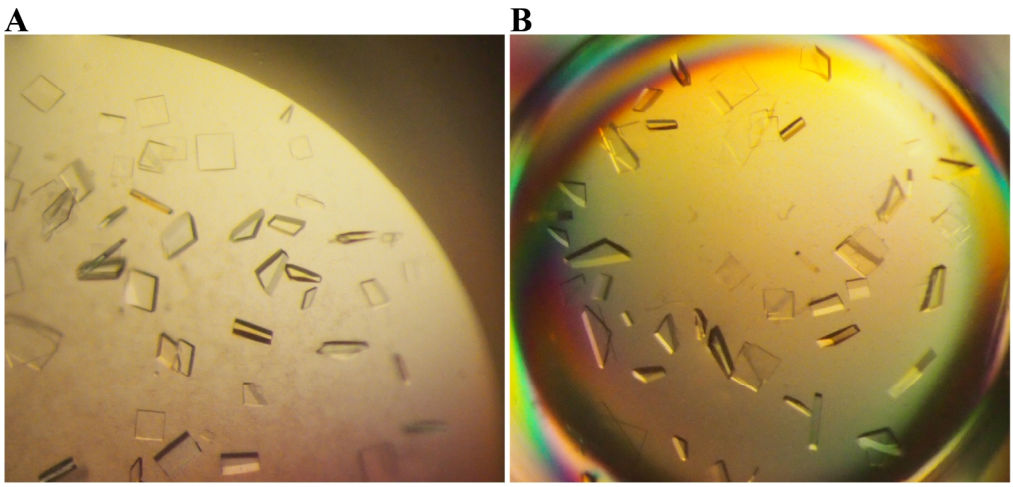



**Figure 28.** Crystallization of G436A mutant HsdR subunit. A) 0.2 M $Li_2SO_4$, 8 % PEG 20 000 and 8 % PEG550 MME (4 µl of precipitant and 4 µl of protein (15 mg $ml^{-1}$)); B) 0.01 M $NiCl_2$, 8 % PEG 20 000 and 8 % PEG550 MME (4.5 µl drops containing 2 µl of precipitant and 2 µl of protein (15 mg $ml^{-1}$) and 0.5 µl of additive).

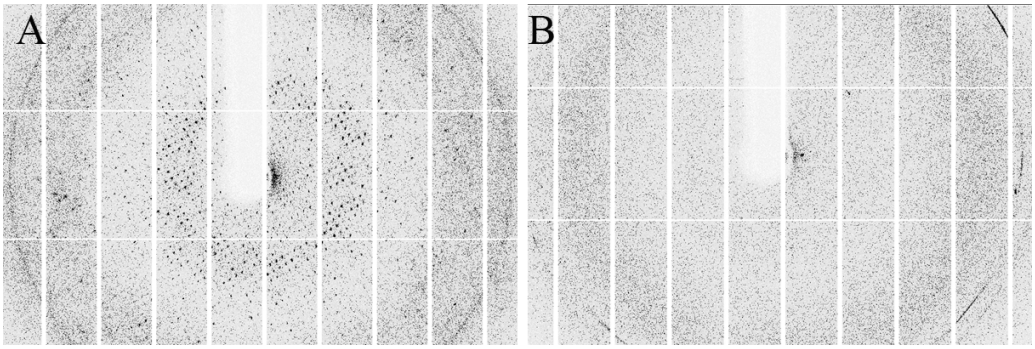



**Figure 29.** Diffraction image of G436A mutant crystal collected at A) 0° and B) 90° rotation angles.

In order to improve the crystal quality, gel filtration (prepacked column HiLoad 16/600 Superdex 200 pg from GE Healthcare) was used after DEAE ion-exchange chromatography for buffer exchange and encrease purity of protein sample. Growth of G436A mutant HsdR crystals was screened in precipitant (90 µl reservoir solution (Jena Bioscience, Compact Clover Plate with 24 partitioned reservoirs, the capacity of each of the four partitioned areas of the reservoir is 250 µl)) containing 8 % PEG 20 000 and 8 % PEG550 MME with Additive screen HT (Hampton Research) in 4.5 µl drops containing 2 µl of precipitant and 2 µl of protein and 0.5 µl of additive. Crystallization condition

A12 containing 0.01 M $NiCl_2$ resulted in lower degree of protein precipitation in crystallization drop and lower number of 2D crystals (Fig. 28 B). 3D crystals in A12 condition were of bigger size and more stable. Preliminary testing of X-Ray diffraction of G436A mutant HsdR crystals (grown in A12 crystallization condition) at room temperature in capillaries (MiTeGen) on Bruker D8 Discover diffractometer revealed diffraction at 0° and 90° rotation angles. The best crystals from A12 conditions were mounted in a nylon loops (Hampton Research), transferred for a few seconds to 10 µl of cryosolution (25 % glycerol (in water) or 100 % PEG550 MME) and then was flash-frozen in liquid nitrogen prior to measuring at a synchrotron facility.

## 3.4. Proteolysis of EcoR124I

Limited proteolysis was performed following Carey (2000). Methyltransferase and WT HsdR were used in a ratio of 1:3 for reconstitution of EcoR124I *in vitro*. Trypsin (Sigma Aldrich) was diluted from freshly prepared stock.

| Protease | pH optimum | Cleavage |
|---|---|---|
| Trypsin **T** | 7 - 9 | lysine, arginine |

The proteolytic reaction was performed in 500 µL Eppendorf tubes on ice. 15 µl of protein mixture contained ~20 pmoles of methyltransferase and ~60 pmoles of WT HsdR in R buffer. Reaction was started by adding of 2 µl of $10^1$, $10^2$, $10^3$ and $10^4$ dilutions of trypsin stock solution and stopped after 30 minutes by adding of 4x SDS-loading buffer and immediate boiling for 5 min. Then proteolytic digestion was analyzed by 10 % SDS-PAGE (Fig. 30). After that reaction conditions for $10^2$ dilution of trypsin stock were chosen as working conditions for further proteolysis of EcoR124I to obtain digestion products in large scale. Protein blotting was used for transfer of digestion products from SDS-PAGE on PVDF membrane. Then bands of interest were cut out from PVDF membrane and sent for Edman degradation protein sequencing (**was done in collaboration with Jannette Carey, Princeton University**). Analyzed bands come from trypsin digestion of HsdR subunit of EcoR124I, all bands are above 60 kDa on SDS-PAGE (MW of HsdM - 58 kDa and HsdS - 46 kDa). 5

samples from 1-7 sent for sequencing (marked on fig. 30) resulted in following signals at N-terminus (data from 5 cycles of Edman degradation):
sample 2: (I,L)(Q)(T,E);
sample 4: (H)((Q),(P))(T)(S)(T);
sample 5: SLETE;
sample 6: (L)(P)(A)(Q)(N);
sample 7: (F)(I,P)(Q)(Q)(K).

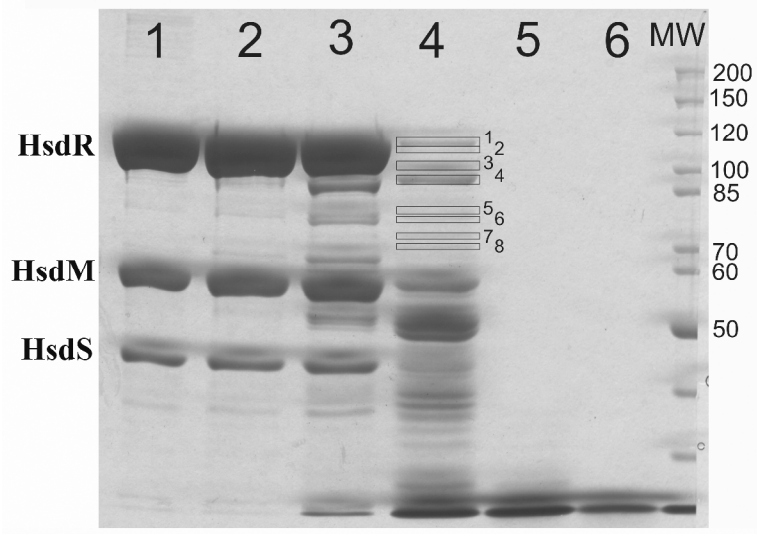



**Figure 30.** Digestion of EcoR124I by trypsin. Line 1 – intact methyltransferase (~20 pmoles) and WT HsdR (~60 pmoles); lines 2 to 5 – the same amount (as on line 1) of methyltransferase and WT HsdR was digested with $10^4$ (line 2), $10^3$ (line 3), $10^2$ (line 4) and $10^1$ (line 5) trypsin dilutions; line 6 – undiluted trypsin; MW – molecular weight marker. On line 4 bands of interest are marked and numbered; this reaction was repeated and used for protein blotting to obtain bands of interest in large scale for protein sequencing.

Identification of cleavage position in the HsdR subunit was not possible for samples 4, 6 and 7. For sample 2 there could be two possibilities: (1) IQE 513-515; in that case the N-terminal would be IQE and in P1 would be Arginine; (2) LQE 560-562; in that case the N-terminal would be LQE and P1 would be Arginine. For sample 5 SLETE 488-492, Lysine in P1 position as expected for trypsin (Fig. 31).

In summary, described above proteolytic experiments were aimed to find appropriate conditions for treatment of EcoR124I with trypsin which then

will be applied also to HsdR subunit alone and EcoR124I in complex with DNA duplex. Also mass spectrometry for protein sequencing will be employed.

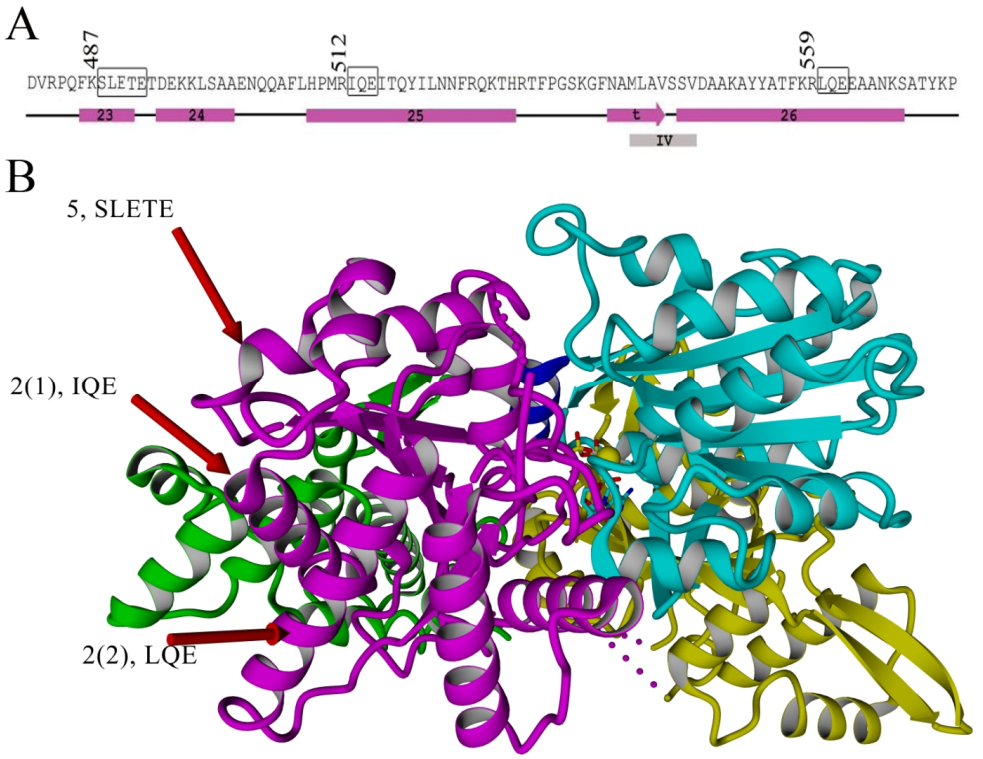



**Figure 31.** Cleavage positions for trypsin on HsdR subunit of Ecor124I. A) Sequence fragment of the helicase 2 domain of HsdR subunit with identified cleavage sites: numbered residues (K487, R512, R559), N-terminal sequences from protein sequencing are in rectangles. B) Top-view of the crystal structure of HsdR subunit of EcoR124I (2W00). Cleavage positions are denoted with arrows (Figure prepared in YASARA).

## 3.5. Interdomain communication in the motor subunit of EcoR124I

Structurally the HsdR motor subunit of EcoR124I R-M enzyme was characterized in 2009 by Lapkouski et al. Already prior to structure determination the DEAD-box motifs in the RecA-like helicase domains, the catalytic residues and the QxxxY motif in the endonuclease domain were identified. The crystal structure of HsdR subunit (pdb id: 2W00) co-crystallyzed with ATP revealed an unusual contact in which the endonuclease

domain contacts the ATP molecule directly via lysine 220. The K220 residue from the so called 220-loop is located ~20 Å from the endonuclease active site and contacting the ATP molecule was proposed to play a role in interdomain communication coupling helicase and nuclease functions in the HsdR subunit of EcoR124I. Mutational analysis of position 220 revealed that the endonuclease activity of mutant enzymes tested *in vitro* is ordered WT ≈ K220R > K220A > K220E accordingly rate constant and crossover times values. This data is in agreement with data from *in vivo* experiments; however, ATPase activity for all tested enzymes was unaltered, indicating that translocation is not affected by the mutations, and thus both activities are uncoupled in the mutants. In the crystal structure of K220R mutant HsdR similarly as in WT contact of the endonuclease domain to ATP is preserved, and in K220A and K220E this contact is lost. The details of this work are described in **Paper 2** (Csefalvay et al., 2015).

The so-called 180-loop is missing in the 2W00 HsdR structure but became available in the crystal structure of K220A mutant HsdR. This allowed building of a full WT model with the 180-loop being present and a consequent study of the $Q_{179}xxxY_{183}$ motif which was assigned to play auxiliary role during DNA cleavage in the endonuclease domain. During the analysis of the behavior of conserved Y183 in MD simulations it was noted that the movement of the side chain of adjacent R182 strongly correlates with the movement of the 220-loop bearing the K220 residue. Also, a loss of the K220-ATP contact during simulation lead to rotation of the endonuclease domain, that is not happening in the R182A mutant HsdR. R182A mutant enzyme preserved WT-like ATPase activity, but was able only to nick supercoiled DNA substrate *in vitro* and resulted in restriction-deficient phenotype *in vivo*. Based on these results the R182A residue was assigned a role in communicating the ATP-ligation state through the endonuclease/motor subunit. The details of this work are described in **Paper 1** (Sinha et al., 2014).

# 4. CONCLUSIONS

Complex multi-subunit organization in the restriction-modification system EcoR124I determines its multi-functionality. EcoR124I is an example of a molecular machinery with consequent coordination of various enzyme functions. Firstly, target recognition occurs that then leads to ATPase/translocation activity that later culminates in DNA cleavage. Two activities, endonuclease as well as translocation activities, are housed on the motor subunit.

The structurally characterized homologous enzymes, which are the SF2 helicases for the translocation activity and the Type II restriction endonucleases for the endonuclease activity, act as isolated proteins and no additional domains or complex formation with other proteins is needed for enzyme function, translocation or restriction activity is thus carried out by a single enzyme bearing one activity only. The RecA-like helicase characteristic motifs are conserved and determine ATPase and translocation activities of SF2 superfamily helicases. *In vitro* purified Rad54 translocase from SWI2/SNF2 family displays branch migration as well as translocation on dsDNA including displacement of of Rad51 activities in an ATP-dependent fashion (Burgess et al., 2013; Zhang et al., 2013). Type II restriction endonucleases of the PD-(E/D)xK superfamily are represented by two separate proteins MTase and restriction endonuclease, each protein has a single domain organization with a particular subdomain for sequence recognition, cleavage of DNA and dimerization. The restriction function requires the formation of an enzyme homodimer. No ATP is needed, there is no translocation activity and cleavage of DNA occurs at or in close proximity to the recognition site (Pingoud et al., 2005; Loenen et al., 2014a). In Type I restriction-modification systems neither the translocase nor the endonuclease are present as isolated proteins as we find them only as part of the array of functionally integrated domains that exhibit mutual contacts in the available crystal structure. The X-ray crystal structure of HsdR subunit of EcoR124I released in 2009 by Lapkouski et al. shed light on some questions and raised new ones. The HsdR subunit represents a symbiotic fusion of the RecA-like helicase fold characteristic for SF2 helicases and the endonuclease unit with an auxiliary helical domain at C-terminal of one polypeptide chain.

Implying that translocation function in EcoR124I is carried out similarly as in Rad54, where helicase 2 domain undergoes a rotational movement during DNA advancement, one would expect the possibility of free rotation for helicase 2 domain. However, we observe interdomain interactions in form of hydrogen bonds at the helical-helicase 2 domain interface which seems to be a contradiction at the first glance.

So why do we observe these contacts in the crystal structure of the HsdR subunit of EcoR124I between the helical and the helicase 2 domains? Type I restriction-modification system needs to be able to turn on and off the individual enzyme functions based on the status of the other enzymatic functions. So unmethylated target sequence provokes initiation of DNA translocation in an ATP-dependent manner. No methylation occurs in that case, and only when translocation is stalled then endonuclease function is switched on. At certain conditions the HsdR subunit participating in DNA cleavage can undergo a turnover and then the next round of DNA translocation and cleavage is possible for this motor subunit (Simons & Szczelkun, 2011). Likely, the observed interdomain interactions at the helical-helicase 2 domain interface described and studied in this work contribute to control the enzymatic function, especially to fine-tune the enzymatic function and to communicate between the "enzymes".

Although it is possible to study all activities of this gigantic pentameric enzyme *in vitro* and restriction and complex assembly even *in vivo* the structural characterization of those processes still presents a difficult task and needs guidance by computational models and consequent molecular dynamics simulations.

Sequence alignment suggests that the main interaction partner in the helicase 2 domain, K527, is conserved throughout Type I R-M families IA, IB, IC, and for this region a similar three-dimensional arangment and function is predicted throughout Type I R-M systems. The low sequence identity with other related translocases from the SF2 superfamily such as Rad54 suggests that the conserved Q/D-KT motif is rather specific for Type I R-M systems supporting the idea that the role of this motif is in contacting the helical domain.

Mutations of the residues participating in helical-helicase 2 domain interactions were designed to probe the potential contribution of these residues

either to the motor activity of the helicase domains, or the stabilization of a particular translocation stage. Mutagenesis of K527 and Y736 residues demonstrated their critical role for all enzyme activities housed on the HsdR subunit, K527A mutant enzyme as well as double K527A_Y736A mutant enzyme were not able to hydrolyse ATP and translocate DNA, consequently no endonuclease activity was detected. The single Y736A muant enzyme displayed a low rate of ATPase activity, no cleavage activity on circular DNA and only residual cleavage activity on two-site linear DNA substrate. The computational results allow to interprete the consequence of the mutation to be a conformational change that changes the position of helicase domain 2 with respect to helicase domain 1. For the D796A mutation computations predict the preservation of the H-bond between K527 and D794 residues, however it is not sufficient for sucsessfull functioning of EcoR124I. The D796A mutant enzyme shows a clear r− phenotype in the *in vivo* restriction test, even though we find some ATPase activity *in vitro* and partial cleavage of circular DNA substrate. This strongly suggests that the H-bond between K527 and D796 is not less important for enzyme activities and cannot be easily substituted by the contact with D794.

Available crystal structures of related SF2 helicase show different conformations with respect of the mutual position of the helicase domains, allowing modeling of various conformations of the motor subunit that can be used to propose a translocation cycle for the motor subunit of EcoR124I and describe the rotational motion of the helicase 2 domain. Our models suggest that the helical-helicase 2 domain interactions observed in the 2W00 HsdR crystal structure are not persistant throughout the whole translocation cycle but are present predominantly at the DNA advancement stage, stabilizing this specific conformation that is not sampled in the absence of these interactions. Binding and hydrolysis of ATP lead to a rearangment of interdomain interactions at the helical-helicase 2 domain interface.

A similar effect is observed when probing a G436A mutant in helicase motif III, where the substitutions of conserved G436 residue by alanine lead to a drastic decrease in ATPase activity and consequent endonuclease activity. *E. coli* cells expressing G436A mutant enzyme were not able to survive in *in vivo* tests revealing restriction-deficient phenotype. Data from MD simulations

shows that the interdomain distance between the helicase domains in G436A mutant HsdR is increased, particularly, arrangement of the helicase 2 domain in G436A mutant HsdR represent more open conformation wich possibly is sampled by motor subunit during translocation. However, in this case the mutation would drastically shift the equilibrium from the translocation stage represented by the conformation observed in the crystal structure towards a more open conformation, making probably sampling of the crystal-like conformation merely impossible and hindering the helicase from performing the whole translocation cycle.

Summarizing all the data from biochemical and structural analysis, molecular dynamics simulations, very specific contacts of the helicase 2 domain to the helical domain (and between both helicase domains) in the HsdR subunit are required to sample all stages of the translocation cycle.

# 5. REFERENCES

Abadjieva, A., Patel, J., Zinkevich, V., Firman, K. (1994). Deletions within the DNA recognition subunit of M.EcoR124I that identify a region involved in protein protein interactions between HsdS and HsdM. *J. Mol. Biol*. 241, 35-43.

Abadjieva, A., Scarlett, G., Janscak, P., Dutta, C.F., Firman, K. (2003) Characterization of an EcoR124I restriction-modification enzyme produced from a deleted form of the DNA-binding subunit, which results in a novel DNA specificity. *Folia Microbiol*. 48(3), 319-28.

Abramoff, M.D., Magalhaes, P.J., Ram, S.J. (2004) Image processing with ImageJ. *Biophoton Int* 11(7), 36–42.

Amadei, A., Linnssen, A.B., Berendsen, H.J. (1993) Essential dynamics of proteins. *Proteins*. 17, 412–25.

Appleby, T.C., Anderson, R., Fedorova, O., Pyle, A.M., Wang, R., Liu, X., Brendza, K.M., Somoza, J.R. (2011) Visualizing ATP-dependent RNA Translocation by the NS3 Helicase from HCV. *J Mol Biol*. 405(5), 1139-53.

Arber, W., and Dussoix D. (1962) Host specificity of DNA produced by *Escherichia coli*. I. Host controlled modification of bacteriophage lambda. *J. Mol. Biol*. 5, 18–36.

Barcus,V.A., Titheradge,A.J. and Murray,N.E. (1995) The diversity of alleles at the hsd locus in natural populations of *Escherichia coli*. *Genetics*. 140, 1187–1197.

Berendsen, H.J.C., van der Spoel, D., van Drunen, R. (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comp Phys Comm*. 91, 43–56.

Bertani, G. and Weigle, J.J. (1953) Host-controlled variation in bacterial viruses. *J. Bacteriol*. 65, 113–121.

Bianco, P.R. and Hurley, E.M. (2005) The type I restriction endonuclease EcoR124I, couples ATP hydrolysis to bidirectional DNA translocation. *J. Mol. Biol*. 352, 837–859.

Blakely, G.W. and Murray, N.E. (2006) Control of the endonuclease activity of type I restriction-modification systems is required to maintain chromosome integrity following homologous recombination. *Mol Microbiol*. 60(4), 883-93.

Bleichert, F. & Baserga, S.J. (2007). The long unwinding road of RNA helicases. *Mol. Cell*. 27, 339–352.

Burckhardt, J., Weisemann, J. & Yuan, R. (1981). Characterisation of the DNA methylase activity of the restriction enzyme from *Escherichia coli* K. *J. Biol. Chem*. 256, 4024–4032.

Burgess, R.C., Sebesta, M., Sisakova, A., Marini, V.P., Lisby, M., Damborsky, J., Klein, H., Rothstein, R., Krejci, L. (2013) The PCNA interaction protein box sequence in Rad54 is an integral part of its ATPase domain and is required for efficient DNA repair and recombination. *PLoS One*. 8(12).

Bussi, G., Donadio, D., Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J Chem Phys*. 126, 014101.

Calisto, B.M., Pich, O.Q., Pinol, J., Fita, I., Querol, E., Carpena, X. (2005). Crystal structure of a putative type I restriction-modification S subunit from mycoplasma genitalium. *J. Mol. Biol*. 351, 749-762.

Carey, J. (2000) A systematic and general proteolytic method for defining structural and functional domains of proteins. *Methods Enzymol.* 328:499-514.

Caruthers, J.M. & McKay, D.B. (2002). Helicase structure and mechanism. *Curr. Opin. Struct. Biol*. 12, 123–133.

Chan, K.M., Delfert, D., Junger, K.D. (1986) A direct calorimetric assay for Ca2+ stimulated ATPase activity. *Anal Biochem*. 157,375–380.

Colson, C., Glover, S.W., Symons, N., Stanley, K.A. (1965) The location of the genes for host-controlled modification and restriction in *Escherichia coli* K-12. *Genetics*. 52, 1043–50.

Cooper, L.P. and Dryden, D.T.F. (1994) The domains of a type I DNA methyltransferase. Interactions and role in recognition of DNA methylation. *J. Mol. Biol*. 236, 1011–1021.

Cordin, O., Banroques, J., Tanner, N.K. & Linder, P. (2006). The DEAD-box protein family of RNA helicases. *Gene*, 367, 17–37.

Cowan, G.M., Gann, A.A. and Murray, N.E. (1989) Conservation of complex DNA recognition domains between families of restriction enzymes. *Cell*, 56, 103–109.

Darden, T., York, D., Pedersen, L., Ewald, P. (1993) An N·log(N) method for Ewald sums in large systems. *J Chem Phys*. 98(12), 10089–92.

Davies, G.P., Powell, L.M., Webb, J.L., Cooper, L.P., Murray, N.E.. (1998) EcoKI with an amino acid substitution in any one of seven DEAD-box motifs has impaired ATPase and endonuclease activities. *Nucleic Acids Res*. 26(21), 4828-36.

Davies, G.P., Kemp, P., Molineux, I.J., and Murray, N.E. (1999) The DNA translocation and ATPase activities of restriction-deficient mutants of EcoKI. *J. Mol. Biol*. 292, 787–796.

Davies, G.P., Martin, I., Sturrock, S.S., Cronshaw, A., Murray, N.E., Dryden, D.T.F. (1999) On the structure and operation of type I DNA restriction enzymes. *J Mol Biol*., 290, 565–579.

Dillingham, M.S., Wigley D.B., Webb, M.R. (2000) Demonstration of unidirectional single-stranded DNA translocation by PcrA helicase: measurement of step size and translocation speed. *Biochemistry*, 39, 205-12.

Doronina, V.A., Murray, N.E. (2001)The proteolytic control of restriction activity in *Escherichia coli* K-12. *Mol Microbiol*. 39(2), 416-28.

Dreier, J., Bickle, T.A. (1996) ATPase activity of the type IC restriction-modification system EcoR124II. *J Mol Biol*. 257(5), 960-9.

Dreier, J., MacWilliams, M.P. and Bickle, T.A. (1996) DNA cleavage by the type IC restriction–modification enzyme EcoR124II. *J. Mol. Biol*., 264, 722–733.

Dryden, D.T.F., Cooper, L.P. and Murray, N.E. (1993) Purification and characterization of the methyltransferase from the type 1 restriction and modification system of *Escherichia coli* K12. *J. Biol. Chem*., 268, 13228–13236.

Dryden, D.T.F., Cooper, L.P. and Murray, N.E. (1997b) Assembly of the multifunctional EcoKI DNA restriction enzyme *in vitro*. *Tech. Prot. Chem*., 8, 593–601.

Dryden, D.T.F., Cooper, L.P., Thorpe, P.H. and Byron, O. (1997a) The in vitro assembly of the EcoKI type I DNA restriction/ modification enzyme and its *in vivo* implications. *Biochemistry*, 36, 1065–1076.

Dryden, D.T.F., Murray, N.E., & Rao, D.N. (2001). Nucleoside triphosphate-dependent restriction enzymes. *Nucleic Acids Research, 29*(18), 3728-3741.

Dürr, H., Körner, C., Müller, M., Hickmann, V., and Hopfner, K.P. (2005) X-ray structures of the Sulfolobus solfataricus SWI2/SNF2 ATPase core and its complex with DNA. *Cell*. 121, 363–373.

Eskin, B. and Linn, S. (1972) The deoxyribonucleic acid modification and restriction enzymes of *Escherichia coli* B. III. Studies of the restriction adenosine triphosphatase. *J. Biol. Chem*., 247, 6192–6196.

Firman, K. & Szczelkun, M. D. (2000). Measuring motion on DNA by the type I restriction endonuclease EcoR124I using triplex displacement. *EMBO J*. 19, 2094–2102.

Firman, K., Dutta, C., Weiserova, M. and Janscak, P. (2000) The role of subunit assembly in the functional control of Type I restriction-modification enzymes. *Mol. Biol. Today*, 1, 35–41.

Firman, K., Szczelkun, M.D. (2000). Measuring motion on DNA by the type I restriction endonuclease EcoR1241 using triplex displacement. *EMBO J*. 19, 2094-2102.

Flaus, A. & Owen-Hughes, T. (2011) Mechanisms for ATP-dependent chromatin remodelling: the means to the end. *FEBS J*. 278(19), 3579-95.

Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Pople, J.A. et al. (2004) GAUSSIAN 03 (revision C.02). Gaussian, Inc., Wallingford.

Fuller-Pace, F.V., Bullas, L.R., Delius, H. and Murray, N.E. (1984) Genetic recombination can generate altered restriction specificity. *Proc. Natl Acad. Sci. USA*, 81, 6095–6099.

García, L.R. and Molineux, I.J. (1999) Translocation and specific cleavage of bacteriophage T7 DNA *in vivo* by EcoKI. *Proc Natl Acad Sci U S A*. 96, 12430–12435.

Gill, S.C. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem*. 182, 319-326.

Gorbalenya, A.E. & Koonin, E.V. (1991) Endonuclease (R) subunits of type I and type III restriction enzymes contain a helicase-like domain. *FEBS Letters*. 291, 277-281.

Gorbalenya, A.E. and Koonin, E.V. (1993) Helicases: amino acid sequence comparisons and structure-function relationships. *Curr Opin Struct Biol*. 3, 419–29.

Gubler, M., Bickle, T.A. (1991) Increased protein flexibility leads to promiscuous protein-DNA interactions in type IC restriction-modification systems. *EMBO J*. 10(4), 951-7.

Hadi, S.M., Bickle, T.A. & Yuan, R. (1975). The role of *S*-adenosyl methionine in the cleavage of deoxyribonucleic acid by the restriction endonuclease from *Escherichia coli* K. *J. Biol. Chem.* 250, 4159–4164.

Hall, M.C. & Matson, S.W. (1999) Helicase motifs: the engine that powers DNA unwinding. *Mol. Microbiol*, 34, 867-877.

Hedges, R. W.,and Datta, N. (1972) R124, an *fi*$^{+}$ R Factor of a New Compatibility Class. *J. Gen. Microbiol*., 71, 403-405.

Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M. (1997) LINCS: a linear constraint solver for molecular simulations. *J Comp Chem*. 18(12), 1463–1472.

Holubova, I., Vejsadová, Š., Firman, K., Weiserova, M. (2004) Cellular localization of type I restriction-modification enzymes is family dependent. *Biochem Biophys Res Commun*, 319, 375–380.

Holubova, I., Vejsadova, S., Weiserova, M., and Firman, K. (2000) Localization of the Type I restriction–modification enzyme EcoKI in the bacterial cell. *Biochem. Biophys. Res. Commun*. 270, 46–51.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C. (2006) Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins*. 65, 712–725.

Jacob, F., and Wollman, E.L. (1954) Etude génétique d'un bactériophage tempéré d'Escherichia coli. III. Effet du rayonnement ultraviolet sur la recombinaison génétique. *Ann Inst Pasteur*. 87, 653-673.

Janscak P, Sandmeier U, Bickle TA (1999) Single amino acid substitutions in the HsdR subunit of the type IB restriction enzyme EcoAI uncouple the DNA translocation and DNA cleavage activities of the enzyme. *Nucleic Acids Res*., 27, 2638–2643.

Janscak, P., Abadjieva, A. and Firman, K. (1996) The type I restriction endonuclease R.EcoR124I: over-production and biochemical properties *J. Mol. Biol*., 257, 977–991.

Janscak, P., and Bickle, T.A. (2000) DNA supercoiling during ATP-dependent DNA translocation by the type I restriction enzyme EcoAI. *J. Mol. Biol*. 295, 1089-1099.

Janscak, P., Bickle, T.A. (1998). The DNA recognition subunit of the Type IB restriction-modification enzyme EcoAI tolerates circular permutations of its polypeptide chain. *J. Mol. Biol*. 284, 937-948.

Janscak, P., Dryden, D.T.F. and Firman, K. (1998) Analysis of the subunit assembly of the type IC restriction–modification enzyme *Eco*R124I. *Nucleic Acids Res*., 26, 4439–4445.

Janscak, P., MacWilliams, M.P., Sandmeier, U., Nagaraja, V., and Bickle, T.A. (1999) DNA translocation blockage, a general mechanism of cleavage site selection by type I restriction enzymes. *EMBO J*., 18, 2638-2647.

Janscak, P., Sandmeier, U., and Bickle, T.A. (1999) Single amino acid substitutions in the HsdR subunit of the type IB restriction enzyme EcoAI uncouple the DNA translocation and DNA cleavage activities of the enzyme. *Nucleic Acids Res*. 27(13), 2638-2643.

Janscak, P., Sandmeier, U., Szczelkun, M.D., Bickle, T.A. (2001) Subunit assembly and mode of DNA cleavage of the type III restriction endonucleases EcoP1I and EcoP15I. *J Mol Biol*. 306(3), 417-31.

Janscak, P., Weiserova, M., Hubacek, J., Holubova, I., Dutta, C.F., Firman, K. (2000) Two temperature-sensitive mutations in the DNA binding subunit of EcoKI with differing properties. *FEMS Microbiol Lett*. 182(1), 99-104.

Jindrova, E., Schmid-Nuoffer, S., Hamburger, F., Janscak, P., Bickle, T.A. (2005) On the DNA cleavage mechanism of Type I restriction enzymes. *Nucleic Acids Res*. 33, 1760-1766.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 79, 926.

Kennaway, C.K., Taylor, J.E., Song, C.F., Potrzebowski, W., Nicholson, W., White, J.H., ... Dryden, D.T.F. (2012). Structure and operation of the DNA-translocating type I DNA restriction enzymes. *Genes & Development*, *26*(1), 92-104.

Kim, J.S., DeGiovanni, A., Jancarik, J., Adams, P.D., Yokota, H., Kim, R., Kim, S.H. (2005). Crystal structure of DNA sequence specificity subunit of a type I restrictionmodification enzyme and its functional implications. *PNAS*. 102, 3248-3253.

Kneale, G.G. (1994). A symmetrical model for the domain structure of type I DNA methyltransferases. *Mol Biol*. 243, 1-5.

Konagurthu, A.S., Whisstock, J.C., Stuckey, P.J., Lesk, A.M. (2006) MUSTANG: a multiple structural alignment algorithm. *Proteins*. 64, 559–574.

Krieger, E., Koraimann, G., Vriend, G. (2002) Increasing the precision of comparative models with YASARA NOVA; a self-parameterizing force field. *Proteins*. 47, 393–402.

Kulik, E.M., Bickle, T.A. (1996) Regulation of the activity of the type IC EcoR124I restriction enzyme. *J Mol Biol*. 264(5), 891-906.

Lapkouski, M., Panjikar, S., Janscak, P., Smatanova, I.K., Carey, J., Ettrich, R., Csefalvay, E. (2009). Structure of the motor subunit of type I restriction-modification complex EcoR124I. *Nature Struct. Mol. Biol*. 16, 94-99.

Lapkouski, M., Panjikar, S., Smatanova, I.K, Csefalvay, E. (2007). Purification, crystallization and preliminary X-ray analysis of the HsdR subunit of the EcoR124I endonuclease from *Escherichia coli*. *Acta Crystallogr*. 63, 582-585.

Lee, M.S., Marians, K.J. (1990) Differential ATP requirements distinguish the DNA translocation and DNA unwinding activities of the Escherichia coli PRI A protein. *J Biol Chem*., 265, 17078-83.

Lewis, R., Dürr, H., Hopfner, K.P., Michaelis, J. (2008) Conformational changes of a Swi2/Snf2 ATPase during its mechano-chemical cycle. *Nucleic Acids Res*. 36(6), 1881-90.

Linn, S. M. and Roberts, R. J. (Editors) (1982) Nucleases, vol.14.

Linn, S., and Arber, W. (1968) Host specificity of DNA produced by Escherichia coli. X. *In vitro* restriction of phage fd replicative form. *Proc. Natl. Acad. Sci. USA* 59, 1300–1306.

Loenen, W.A. M., Dryden, D.T.F., Raleigh, E.A., Wilson, G.G., & Murray, N.E. (2014a). Highlights of the DNA cutters: a short history of the restriction enzymes. *Nucl. Acids Res.* 42, (1) 3–19.

Loenen, W.A. M., Dryden, D.T.F., Raleigh, E.A., & Wilson, G.G. (2014b). Type I restriction enzymes and their relatives. *Nucleic Acids Research*. 42 (1), 20-44.

Lohman, T.M., Tomko, E.J., Wu, C.G. (2008) Non-hexameric DNA helicases and translocases: mechanisms and regulation. *Nat Rev Mol Cell Biol*. 9(5), 391-401.

Makovets, S., Powell, L.M., Titheradge, A.J.B., Blakely, G.W. and Murray, N.E. (2004) Is modification sufficient to protect a bacterial chromosome from a resident restriction endonuclease? *Mol. Microbiol*., 51, 135–147.

Marini, V. & Krejci, L. (2012) Unwinding of synthetic replication and recombination substrates by Srs2. *DNA Repair*. 11, 789–798.

McClelland, S.E. & Szczelkun, M.D. (2004) The Type I and III restriction endonucleases: structural elements in molecular motors that process DNA. In *Nucleic Acids and Molecular Biology – Restriction Enzymes* (Pingound, A., ed.), Springer, Germany, vol 14, pp. 111–135.

McClelland, S.E., Dryden, D.T.F., and Szczelkun, M.D. (2005) Continuous assays for DNA translocation using fluorescent triplex dissociation: application to type I restriction endonucleases. *J. Mol. Biol*. 348, 895–915.

Mernagh, D. R., Janscak, P., Firman, K. and Kneale, G. G. (1998) Protei-protein and protein-DNA interactions in the type I restriction endonuclease R.EcoR124I. *Biol. Chem*. 379, 497-503.

Meselson, M., and R. Yuan. 1968. DNA restriction enzyme from E. coli. Nature 217, 1110–1114.

Meselson, M., Yuan, R. and Heywood, J. (1972) Restriction and modification of DNA. *Annu. Rev. Biochem*. 41, 447–462.

Mueller, U., Darowski, N., Fuchs, M.R., Förster, R., Hellmig, M., Paithankar, K.S., Pühringer, S., Steffien, M., Zocher, G., Weiss, M.S. (2012) Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin. *J Synchrotron Radiat.* 19, 442-449.

Murray, N.E. (2000) Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol Mol Biol Rev*. 64(2), 412-34.

Murray, N.E., Batten, P.L. and Murray, K. (1973) Restriction of bacteriophage l by *Escherichia coli* K. *J. Mol. Biol*., 81, 395–407.

Nagaraja, V., Shepherd, J.C. and Bickle, T.A. (1985) A hybrid recognition sequence in a recombinant restriction enzyme and the evolution of DNA sequence specificity. *Nature*, 316, 371–372.

Niv, M.Y., Ripoll, D.R., Vila, J.A., Liwo, A., Vanamee, E.S., Aggarwal, A.K., Weinstein, H., Scheraga, H.A. (2007) Topology of type II REases revisited; structural classes and the common conserved core. *Nucleic Acids Res.,* 35: 2227–2237.

Obarska, A., Blundell, A., Feder, M., Vejsadova, S., Sisáková, E., Weiserova, M., Bujnicki, J.M., Firman, K. (2006) Structural model for the multisubunit Type IC restriction-modification DNA methyltransferase M.EcoR124I in complex with DNA. Nucleic Acids Res. 34(7), 1992-2005.

Obarska-Kosinska, A., Taylor, J.E., Callow, P., Orlowski, J., Bujnicki, J.M., Kneale, G.G. (2008) HsdR subunit of the type I restriction-modification enzyme EcoR124I: biophysical characterisation and structural modelling. *J Mol Biol*. 376(2), 438-52.

O'Neill, M., Dryden, D.T., Murray, N.E. (1998) Localization of a protein-DNA interface by random mutagenesis. *EMBO J*. 17(23), 7118-27.

Parrinello, M., Rahman, A. (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys*. 52, 7182.

Patel, J., Taylor, I., Dutta, C.F., Kneale, G.G., Firman, K. (1992). High-level expression of the cloned genes encoding the subunits of and the intact DNA methyltransferase, M.EcoR124. *Gene* 112, 21-27.

Piekarowicz, A., Goguen, J.D., Skrzypek, E. (1985) The EcoDXX1 restriction and modification system of Escherichia coli ET7. Purification, subunit structure and properties of the restriction endonuclease. *Eur J Biochem*. 152(2), 387-93.

Pike, A.C., Shrestha, B., Popuri, V., Burgess-Brown, N., Muzzolini, L., Costantini, S., Vindigni, A., Gileadi, O. (2009) Structure of the human RECQ1 helicase reveals a putative strand-separation pin. *Proc Natl Acad Sci U S A.* 106(4), 1039-44.

Pingoud, A. and Jeltsch, A. (1997) Recognition and cleavage of DNA by type-II restriction endonucleases. *Eur. J. Biochem*., 246, 1–22.

Pingoud, A., Fuxreiter, M., Pingoud, V., Wende, W. (2005) Type II restriction endonucleases: structure and mechanism. *Cell Mol Life Sci*. 62(6), 685-707.

Powell, L.M., Dryden, D.T.F., Willcock, D.F., Pain, R.H. and Murray, N.E. (1993) DNA Recognition by the EcoK methyltransferase. *J. Mol. Biol*. 234, 60–71.

Price, C., Pripfl, T. & Bickle, T. A. (1987a). EcoR124 and EcoR124/3: the first members of a new family of type I restriction and modification systems. *Eur. J. Biochem*. 167, 111–115.

Price, C., Shepherd, J.C.W., Bickle, T.A. (1987b). DNA recognition by a new family of type I restriction enzymes: a unique relationship between two different DNA specificities. *Eur. Mol. Biol. Organ. J*. 6, 1493-1498.

Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M.R., Smith, J.C., Kasson, P.M., van der Spoel, D., Hess, B., Lindahl, E. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*. 29(7), 845–854.

Randerath, K., Randerath, E. (1964) Ion-exchange chromatography of nucleotides on poly-(ethyleneimine)-celluslose thin layers. *J Chromatogr*. 16, 111–125.

Redaschi, N., Bickle, T.A. DNA restriction and modification systems. In: Neidhardt FC, Curtiss R III, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE, editors. Escherichia coli and Salmonella: cellular and molecular biology. 2nd ed. Washington, D.C.: *ASM Press*; 1996. pp. 773–781.

Richmond, E., Peterson, C.L. (1996) Functional analysis of the DNA-stimulated ATPase domain of yeast SWI2/SNF2. *Nucleic Acids Res*. 24(19), 3685-92.

Roberts, R.J., Belfort, M., Bestor, T., Bhagwat, A.S., Bickle, T.A., Bitinaite, J. et al. (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucl. Acids Res*. 31, 1805–1812.

Roberts, R.J., Vincze, T., Posfai, J., Macelis, D. (2010) REBASE – a database for DNA restriction and modification: enzymes, genes and genomes. *Nucl. Acids Res*. 38, D234–D236.

Roman, L.J., Kowalczykowski, S.C. (1989) Characterization of the helicase activity of the *Escherichia coli* RecBCD enzyme using a novel helicase assay. Biochemistry; 28, 2863-73.

Rosamond, J., Endlich, B. and Linn, S. (1979) Electron microscopic studies of the mechanism of action of the restriction enzyme of *Escherichia coli* B. *J. Mol. Biol*., 129, 619–635.

Seidel, R., Bloom, J.G., Dekker, C., and Szczelkun, M.D. (2008) Motor step size and ATP coupling efficiency of the dsDNA translocase EcoR124I. *EMBO J*. 27(9), 1388-1398.

Seidel, R., Bloom, J.G., van Noort, J., Dutta, C.F., Dekker, N.H.,Firman, K., Szczelkun, M.D. and Dekker, C. (2005) Dynamics of initiation, termination and reinitiation of DNA translocation by the motor protein EcoR124I. *EMBO J*., 24, 4188–4197.

Seidel, R., van Noort, J., van der Scheer, C., Bloom, J.G., Dekker, N.H., Dutta, C.F., Blundell, A., Robinson, T., Firman, K., Dekker, C. (2004) Real-time observation of DNA translocation by the type I restriction modification enzyme EcoR124I. *Nat Struct Mol Biol*. 11, 838–843.

Seidel,R., Bloom,J.G., van Noort,J., Dutta,C.F., Dekker,N.H.,Firman,K., Szczelkun,M.D. and Dekker,C. (2005) Dynamics of initiation, termination and reinitiation of DNA translocation by the motor protein EcoR124I. *EMBO J*., 24, 4188–4197.

Sharples GJ, Whitby MC, Ryder L, Lloyd RG. (1994) A mutation in helicase motif III of *E. coli* RecG protein abolishes branch migration of Holliday junctions. *Nucleic Acids Res.* 22(3), 308-13.

Simons, M. and Szczelkun, M.D. (2011) Recycling of protein subunits during DNA translocation and cleavage by Type I restriction-modification enzymes. *Nucleic Acids Research*, 39, 7656–7666.

Simons, M., Diffin, F.M., Szczelkun, M.D. (2014) ClpXP protease targets long-lived DNA translocation states of a helicase-like motor to cause restriction alleviation. *Nucleic Acids Res*. 42(19), 12082-91.

Singleton, M.R., Scaife, S., Wigley, D.B (2001) Structural analysis of DNA replication fork reversal by RecG. *Cell*. 107(1), 79-89.

Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Ann. Rev. Biochem*. 76, 23-50.

Sinha, D., Shamayeva, K., Ramasubramani, V., Reha, D., Bialevich, V., Khabiri, M., et al. (2014) Interdomain communication in the endonuclease/motor subunit of Type I restrictionmodification enzyme EcoR124I. *J Mol Model*. 20(7), 2334.

Sisáková, E., Stanley, L.K., Weiserova, M., Szczelkun, M.D. (2008a) A RecB-family nuclease motif in the Type I restriction endonuclease EcoR124I. *Nucleic Acids Res*. 36(12), 3939-49.

Sisáková, E., Weiserova, M., Dekker, C., Seidel, R., Szczelkun, M.D. (2008b) The interrelationship of helicase and nuclease domains during DNA translocation by the molecular motor EcoR124I. *J Mol Biol*. 384(5), 1273-86.

Skrzypek, E., Piekarowicz, A. (1989) The EcoDXX1 restriction and modification system: cloning the genes and homology to type I restriction and modification systems. *Plasmid*. 21(3), 195-204.

Smirnova, M., Van Komen, S., Sung, P., Hannah, L.K. (2004) Effects of tumor associated mutations on Rad54 functions. *J Biol Chem*. 279(23), 24081-8.

Smith, J.D., Arber, W. and Kuhnlein, U. (1972) Host specificity of DNA produced by *Escherichia coli*. XIV. The role of nucleotide methylation in *in vivo* B-specific modification. *J. Mol. Biol*. 63, 1–8.

Stanley, L.K., Seidel, R., van der Scheer, C., Dekker, N.H., and Szczelkun, M.D., and Dekker, C. (2006) When a helicase is not a helicase: dsDNA tracking by the motor protein EcoR124I. *EMBO J.* 25(10), 2230-2239.

Strick, T.R., Allemand, J.F., Bensimon, D., Croquette, V. (1998) Behavior of supercoiled DNA. *Biophys J.*, 74, 2016–2028.

Studier, F.W. and Bandyopadhyay, P.K. (1988) Model for how type I restriction enzymes select cleavage sites in DNA. *Proc. Natl Acad. Sci. USA*, 85, 4677–4681.

Suri, B. & Bickle, T.A. (1985). EcoA: the first member of a new family of type I resriction modification systems. Gene organization and enzymatic activities. *J. Mol. Biol*. 186, 77–85.

Suri. B., Shepherd, J. C. W. and Bickle, T.A. (1984) The EcoA restriction and modification system of Escherichia coli 15T-: enzyme structure and DNA recognition sequence. *Cold Spring Harbor Symp. Quant. Biol*., 43, 1217-1221.

Szczelkun, M.D., Dillingham, M.S., Janscak, P., Firman, K., Halford, S.E. (1996). Repercussions of DNA tracking by the type IC restriction endonuclease EcoR124I on linear, circular and catenated substrates. *EMBO J.* 15, 6335-6347.

Szczelkun, M.D., Janscak, P., Firman, K. and Halford, S.E. (1997) Selection of non-specific DNA cleavage sites by the type IC restriction endonuclease EcoR124I. *J. Mol. Biol*., 271, 112–123.

Tanner, N.K. & Linder, P. (2001). DExD/H box RNA helicases: from generic motors to specific dissociation functions. *Mol. Cell*, 8, 251–262.

Taylor, I. A., Davis, K. G., Watts, D., Kneale, G. G. (1994). DNA binding induces a major structural transition in a type I methyltransferase. *Eur. Mol. Biol. Org. J*. 13, 5772-5778.

Taylor, I., Patel, J., Firman, K., Kneale, G.G. (1992). Purification and biochemical characterisation of the EcoR124 modification methylase. *Nucleic Acids Res.* 20, 179-186.

Taylor, I., Watts, D., Kneale, G. (1993). Substrate recognition and selectivity in the type IC DNA modification methylase M.EcoR124I. *Nucleic Acids Res.* 21, 4929-4935.

Taylor, J.E., Callow, P., Swiderska, A., Kneale, G.G. (2010) Structural and functional analysis of the engineered type I DNA methyltransferase EcoR124I(NT). *J Mol Biol*. 398(3), 391-9.

Theis, K., Chen, P.J., Skorvaga, M., Van Houten, B., Kisker, C. (1999) Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. *EMBO J*. 18(24), 6899-907.

Tock, M.R. and Dryden, D.T. (2005) The biology of restriction and anti-restriction. *Curr. Opin. Microbiol*. 8, 466–472.

Tyndall,C., Meister,J. and Bickle,T.A. (1994) The *Escherichia coli* prr region encodes a functional type IC DNA restriction system closely integrated with an anticodon nuclease gene. *J. Mol. Biol*., 237, 266–274.

van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., Berendsen, H.J.C. (2005) GROMACS: fast, flexible, and free. *J Comput Chem*. 26, 1701–1718.

van Noort, J., van der Heijden, T., Dutta, C.F., Firman, K., Dekker, C. (2004). Initiation of translocation by Type I restriction-modification enzymes is associated with a short DNA extrusion. *Nucleic Acids Res*. 32, 6540-6547.

Vovis, G.F., Horiuchi, K., Zinder, N.D. (1974) Kinetics of methylation of DNA by a restriction endonuclease from Escherichia coli B. *Proc Natl Acad Sci U S A*. 71(10), 3810-3.

Waldron, D.E., Lindsay, J.A. (2006) Sau1: A novel lineage-specific type I restriction-modification system that blocks horizontal gene transfer into Staphylococcus aureus and between S. aureus isolates of different lineages. *J. Bacteriol*. 188: 5578–5585.

Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A. (2004) Development and testing of a general AMBER force field. *J Comput Chem*. 25, 1157–1174.

Webb, J.L., King, G., Ternent, D., Titheradge, A.J.B., Murray, N.E. (1996) Restriction by EcoKI is enhanced by cooperative interactions between target sequences and is dependent on DEAD-box motifs. *EMBO J*. 15(8), 2003-9.

Webb, M., Taylor, I.A., Firman, K., and Kneale, G.G. (1995) Probing the domain structure of the type IC DNA methyltransferase M.EcoR124I by limited proteolysis. *J. Mol. Biol*. 250(2), 181-190.

Weiserova, M., Dutta, C.F., Firman, K. (2000). A novel mutant of the type I restrictionmodification enzyme EcoR124I is altered at a key stage in the subunit assembly pathway. *J. Mol. Biol*. 304, 301-310.

Weiserova, M., Firman, K. (1998). Isolation of a non-classical mutant of the DNA recognition subunit of the type I restriction endonuclease R.EcoR124I. *Biol. Chem.* 379, 585-589.

Willcock, D.F., Dryden, D.T.F., and Murray, N.E. (1994) A mutational analysis of the two motifs common to adenine methyltransferases. *EMBO J*. 13, 3902–3908.

Yanisch-Perron, C., Vieira, J., Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*. 33, 103–119.

Youell, J., Firman, K. (2008) EcoR124I: from plasmid-encoded restriction-modification system to nanodevice. *Microbiol Mol Biol Rev*. 72(2), 365-77.

Yuan, R., Hamilton, D.L., Burckhardt, J. (1980) DNA translocation by the restriction enzyme from *Escherichia coli* K. *Cell*, 20, 237–244.

Zhang, X.P., Janke, R., Kingsley, J., Luo, J., Fasching, C., Ehmsen, K.T., Heyer, W.D. (2013) A conserved sequence extending motif III of the motor domain in the Snf2-family DNA translocase Rad54 is critical for ATPase activity. *PLoS One*. 8(12).

Zinkevich, V., Heslop, P., Glover, S.W., Weiserova, M., Hubacek, J. and Firman, K. (1992) Mutation in the specificity polypeptide of the type I restriction endonuclease R.EcoK that affects subunit assembly. *J. Mol. Biol.* 227, 597-601.

Zinkevich, V., Popova, L., Kryukov, V., Abadjieva, A., Bogdarina, I., Janscak, P., and Firman, K. (1997) The HsdR subunit of R.EcoR124II: cloning and over-expression of the gene and unexpected properties of the subunit. *Nucleic Acids Res.* 25(3), 503-511.

# SUPPLEMENTS

# 1.

## Interdomain communication in the endonuclease/motor subunit of Type I restriction-modification enzyme EcoR124I

D. Sinha*, K. Shamayeva*, V. Ramasubramani, D. Reha, **V. Bialevich**, M. Khabiri, N. Milbar, A. Guzanova, M. Weiserova, E. Csefalvay, J. Carey, R. Ettrich. (* - contributed equally)

*Journal of Molecular Modeling* 20 (7):2334; (2014)

**Abstract**: Restriction-modification systems protect bacteria from foreign DNA. Type I restriction-modification enzymes are multifunctional heteromeric complexes with DNA cleavage and ATP-dependent DNA translocation activities located on endonuclease/motor subunit HsdR. The recent structure of the first intact motor subunit of the type I restriction enzyme from plasmid EcoR124I suggested a mechanism by which stalled translocation triggers DNA cleavage via a lysine residue on the endonuclease domain that contacts ATP bound between the two helicase domains. In the present work, molecular dynamics simulations are used to explore this proposal. Molecular dynamics simulations suggest that the Lys–ATP contact alternates with a contact with a nearby loop housing the conserved QxxxY motif that had been implicated in DNA cleavage. This model is tested here using in vivo and in vitro experiments. The results indicate how local interactions are transduced to domain motions within the endonuclease/motor subunit.

**Keywords:** DNA restriction enzymes, Molecularmodeling, QM/MMcalculations, Principal components analysis, *E. coli*, Multisubunit enzyme complex, Correlated loop motions

# 2.

## Functional coupling of duplex translocation to DNA cleavage in a Type I restriction enzyme

E. Csefalvay, M. Lapkouski, A. Guzanova, L. Csefalvay, T. Baikova, I. Shevelev, **V. Bialevich**, K. Shamayeva, P. Janscak, I. Kuta Smatanova, S. Panjikar, J. Carey, M. Weiserova, R. Ettrich.

*PLoS ONE* 10 (6):e0128700; (2015)

**Abstract**: Type I restriction-modification enzymes are multifunctional heteromeric complexes with DNA cleavage and ATP-dependent DNA translocation activities located on motor subunit HsdR. Functional coupling of DNA cleavage and translocation is a hallmark of the Type I restriction systems that is consistent with their proposed role in horizontal gene transfer. DNA cleavage occurs at nonspecific sites distant from the cognate recognition sequence, apparently triggered by stalled translocation. The X-ray crystal structure of the complete HsdR subunit from E. coli plasmid R124 suggested that the triggering mechanism involves interdomain contacts mediated by ATP. In the present work, in vivo and in vitro activity assays and crystal structures of three mutants of EcoR124I HsdR designed to probe this mechanism are reported. The results indicate that interdomain engagement via ATP is indeed responsible for signal transmission between the endonuclease and helicase domains of the motor subunit. A previously identified sequence motif that is shared by the RecB nucleases and some Type I endonucleases is implicated in signaling.

# 3.

## The helical domain of the EcoR124I motor subunit participates in ATPase activity and dsDNA translocation

**V. Bialevich**[#], D. Sinha[#], K. Shamayeva, A. Guzanova, D. Řeha, E. Csefalvay, J. Carey, M. Weiserova, R. Ettrich ([#] *these authors contributed equally and should be considered as first authors*)

*PeerJ* (accepted 8.12.2016)

**Abstract**: Type I restriction-modification enzymes are multisubunit, multifunctional molecular machines that recognize specific DNA target sequences, and their multisubunit organization underlies their multifunctionality. EcoR124I is the archetype of Type I restriction-modification family IC and is composed of three subunit types, HsdS, HsdM, and HsdR. DNA cleavage and ATP-dependent DNA translocation activities are housed in the distinct domains of the endonuclease/motor subunit HsdR. Because the multiple functions are integrated in this large subunit of 1038 residues, a large number of interdomain contacts might be expected. The crystal structure of EcoR124I HsdR reveals a surprisingly sparse number of contacts between helicase domain 2 and the C-terminal helical domain that is thought to be involved in assembly with HsdM. Only two potential hydrogen-bonding contacts are found in a very small contact region. In the present work, the relevance of these two potential hydrogen-bonding interactions for the multiple activities of EcoR124I is evaluated by analysing mutant enzymes using *in vivo* and *in vitro* experiments. Molecular dynamics simulations are employed to provide structural interpretation of the functional data. The results indicate that the helical C-terminal domain is involved in the DNA translocation, cleavage, and ATPase activities of HsdR, and a role in controlling those activities is suggested.

# 4.

## The role of motif III and its extended region in positioning the two helicase domains in the motor subunit of the restriction-modification system EcoR124I

Sinha D*, **Bialevich V***, Shamayeva K, Řeha D, Guzanova A, Sisakova A, Csefalvay E, Krejci L, Carey J, Weiserova M and Ettrich R (* - contributed equally)

*Manuscript*

**Abstract**:: The type I restriction-modification enzymes differ significantly from the type II enzymes commonly used as molecular biology reagents. On hemi-methylated DNAs type I enzymes, such as EcoR124I restriction-modification complex, act as conventional adenine methylases at their specific target sequences, but unmethylated targets induce them to translocate thousands of base pairs through the enzyme before cleaving distant sites nonspecifically. EcoR124I belongs to the same superfamily as yeast dsDNA translocase Rad54 (Superfamily 2 helicases) and two RecA-like helicase domains are involved in translocation and contain the seven motifs characteristic for DEAD box helicases family. In Rad54 region adjacent to motif III, so called an extended region, is involved in ATPase activity. In EcoR124I this extended region, although it bears sequence and structural similarities with Rad54, does not influence ATPase activity nor restriction activity. However, mutagenesis of the conserved glycine residue in motif III in Ecor124I leads to altered ATPase activity and cleavage activity. The full model of HsdR of EcoR124I based on the crystal structure of WT and several mutant structures that were used to complete unresolved loop regions allowed the interpretation of experimentally studied functional effects of several mutants in the extended motif III as well as in the motif III itself through the lens of molecular dynamics. Computational analysis suggests that this conserved glycine residue in motif III plays a role in the positioning of the two helicase domains towards each other.