University of South Bohemia, České Budějovice
Faculty of Science

Involvement of the endonuclease domain of the EcoR124I restriction-modification complex in interdomain communication

Ph.D. Thesis

Katsiaryna Shamayeva, MSc.

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

České Budějovice 2015

This thesis should be cited as:

Shamayeva K., 2015: Involvement of the endonuclease domain of the EcoR124I restriction-modification complex in interdomain communication. Ph.D. Thesis. University of South Bohemia, Faculty of Science, České Budějovice, Czech Republic, 112 pp.

Annotation

Type I restriction-modification (R-M) enzymes recognize specific sequences on foreign DNA invading the bacterial cell. At a first sight, also host DNA containing the specific target site for Type I R-M enzyme would cleavage, too, but this doesn't happen as the enzymes are able to distinguish between hostband foreign DNA. Normally the specific sequence on host DNA is either fully methylated or the Type I R-M enzyme recognizes the hemi-methylated state of the DNA, switches to the modification mode and methylates the second strand of the hemi-methylated DNA. Recognition of unmethylated foreign DNA invading the bacterial cell from the outer environment leads to a switching to the restriction mode, initiating endonuclease activity. The R-M complex tightly bound to the recognition sequence on foreign DNA then starts to translocate dsDNA in an ATP-dependent manner towards the stationary enzyme over up to several thousand base pairs. When DNA translocation is finally stalled, the enzyme complex introduces a double strand break, seemingly in a random site distant from the recognition sequence. Multi-subunit structure determines complex behavior of Type I R-M enzymes. The fully assembled Type I R-M enzyme consists of five subunits which are encoded by *hsd* genes (host specificity for DNA): one copy of HsdS subunit together with two copies of HsdM subunit form the trimeric $HsdS_1$-$HsdM_2$ methyltransferase complex which recognizes and binds to a specific DNA sequence and bears the methylation function. The fully assembled $HsdS_1$-$HsdM_2$-$HsdR_2$ complex possesses ATP-dependent DNA translocation and endonuclease activities located on its HsdR subunits. The X-ray crystal structure of HsdR of EcoR124I with bound ATP gives a first insight of structural/functional correlation in the HsdR subunit. In this work the involvement of the endonuclease domain in interdomain communication within the HsdR motor subunit of EcoR124I is probed experimentally, confronted with computational predictions and discussed in the light of the fully functional pentameric complex.

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, 21.09.2015 .....

Katsiaryna Shamayeva

This thesis originated from a partnership of the Faculty of Science, University of South Bohemia, and the Institute of Nanobiology and Structural Biology, Global Change Research Center, 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 and MW), http://www.gacr.cz and the grand agency of the University of South Bohemia via the doctoral team grants №170/2010/P and №141/2013/P.

List of papers and author's contribution

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

1. D. Sinha*, **K. Shamayeva***, V. Ramasubramani, D. Reha, 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)

   ***KS** performed site-directed mutagenesis, expressed and purified WT and mutant proteins, tested purified enzymes in vitro, performed in vivo experiments, analyzed experimental data, participated in writing of the manuscript.*

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)

   ***KS** expressed and purified WT and mutant proteins, performed restriction activity assays in vitro.*

Publication not included in the thesis:

3. V. Bialevich, D. Sinha, **K. Shamayeva**, A. Guzanova, A. Sisakova, E. Csfalvay, L. Krejci, J. Carey, M. Weiserova, R. Ettrich (2015) 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, submitted to *BMC Biophysics*. (IF=2.89)

   ***KS** performed in vivo experiments and determined restriction phenotype.*

4. D. Sinha, V. Bialevich, **K. Shamayeva,** A. Guzanova, E. Csefalvay, J. Carey, M. Weiserova, R. Ettrich (2015) The helical domain of the motor subunit of EcoR124I participates in ATPase activity and dsDNA translocation, prepared for submission to *NAR.* (IF=9.112)

***KS** performed in vivo experiments and determined restriction phenotype.*

Prof. RNDr. Rüdiger H. Ettrich, Ph.D.

I dedicate this work to my family

ACKNOWLEGMENTS

I would like to acknowledge everyone who has contributed to my research work in the **Institute of Nanobiology and Structural Biology** at the **Department of Structure and Function of Proteins.**

Primarily, I would like to thank my supervisor **Prof. Rudiger Ettrich** who gave me a chance to work on this project, helped and supported me during my PhD study and special thanks for involving me in many interesting scientific projects in Nove Hrady.

My thanks belong to **Prof. Jannette Carey** for her great and creative ideas, for excellent advices and for her enthusiasm related to my research.

I also wish to acknowledge my colleagues from Prague: **Dr. Marie Weiserova, Alena Guzanova** and **Anicka Stahova**. Thanks to them for helping me with *in vivo* experiments.

I'm grateful to **Dr. Jost Ludwig** for giving me the opportunity to visit his laboratory in Bonn for international practice. As well, thanks a lot for professional recommendations during my work and friendly atmosphere in our lab in Nove Hrady.

I would like to express my thanks to my colleagues from the computational group, especially to **Dr. David Reha** for his professionalism and interesting ideas during preparation of manuscripts.

Many thanks to **Dr. Eva Csefalvay** for introducing me in the field of molecular biology and **Dr. Ivana Kuta Smatanova** for help in the beginning of my PhD study. And thanks a lot to our fantastic lab technician **Kamila Blafkova** for her sense of humor and good time during my work in the lab.

Thanks to the **technical** and **administrative** stuff of the Castle for their kindness, and many thanks to **Marcela Kotrcova** for her wonderful ideas and good time during the FEBS course and Summer School projects preparation.

I am very grateful to my parents **Aleksandr** and **Elena**, my grandmothers **Mariya** and **Anna**, and my grandfathers **Stanislav** and **Michail** for their love and care.

And of course special thanks belong to my dear colleague, my beloved husband **Vitali Bialevich**. I cannot imagine my work without his support.

TABLE OF CONTENTS

LIST OF ABREVIATIONS

Å	Angstrom
AdoMet	S-Adenosyl methionine
AFM	Atomic Force Microscopy
ATP	Adenosine-5'-triphosphate
bp	base pair
DEAE	Diethylaminoethyl
DNA	Deoxyribonucleic acid Acid
dsDNA	double-stranded Deoxyribonucleic Acid
DTT	Dithiothreitol
e.o.p.	Efficienty of plating
EDTA	Ethylenediaminetetraacetic Acid
HJ	Holliday Junction
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kd	dissociation constant
kDa	kilodalton
L	Liter
LB	Luria-Bertani
M	Molar
MD	Molecular Dynamics
OD	Optical Density
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
QM/MM	Quantum Mechanics/ Molecular Mechanics
REnase	Restriction Endonuclease
R-M	Restriction-Modification
RNA	Ribonucleic Acid
SAD	Single-wavelength Anomalous Dispersion
SANS	Small Angle Neutron Scattering
SAXS	Small Angle X-ray scattering
SeMet	L-Selenomethionine
ssDNA	single stranded Deoxyribonucleic Acid
TBS	Triplex Binding Site
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 was described for the first time by Bertani and Weige in 1953, and is attributed to restriction enzymes [Bertani and Weige, 1953]. Restriction–modification enzymes are enzymatic complexes, which are found only within prokaryotes [Bickle, 1993], and certain viruses, all bacteria and archaea seem to code for them [Redaschi and Bickle, 1996]. These systems protect bacterial cells against invading DNA mostly from bacteriophages. Distinctions in structural organization, cofactor requirements, specificity and manner of DNA cleavage were used for classification of restriction enzymes [Roberts *et al*., 2010]. There are four (I-IV) Types of restriction-modification systems [Loenen *et al*., 2014].

The EcoR124I restriction-modification complex from *E. coli* belongs to Type I restriction-modification systems. In comparison with other Types of restriction-modification systems, enzymes from Type I are more complex and notable functionally for translocating DNA prior to restriction. EcoR124I is comprised of three different subunits each bearing specific enzymatic functions. The HsdS subunit determines specificity to DNA; the HsdM subunit fulfills the DNA modification function; and the HsdR motor subunit carries out DNA translocation and cleavage (endonuclease) functions. This variety of enzyme functions is consequently coordinated depending on the status of individual enzymatic functions. If the target sequence for EcoR124I on DNA is hemi-methylated, EcoR124I recognizes this sequence as its own, switches to the modification mode and methylates the second DNA strand of the recognition sequence. Fully modified DNA is protected from restriction by EcoR124I [Makovets *et al*., 2004]. Unmethylated specific sequences recognized by EcoR124I are considered as foreign and lead to activation of the restriction mode in EcoR124I. The motor subunits start to translocate DNA in an ATP-dependent fashion towards the stationary enzyme itself anchored to the specific sequence and thus forming long DNA loops [van Noort *et al.*, 2004]. Stalling of DNA translocation triggers endonuclease activity and the enzyme cleaves DNA even thousands of base pairs away from the specific target sequence which produces large pieces of DNA [Janscak *et al*., 1996]. This hints strongly to a role of EcoR124I (and type Is in general) in horizontal gene transfer [Waldron

et al., 2006]. The first X-ray crystal structure of a Type I motor subunit was reported for HsdR (2W00) of EcoR124I in 2009 [Lapkouski *et al.*, 2009]. The structural information allowed to characterize the HsdR subunit as a fusion of four functional domains with a square-planar arrangement. The N-terminal endonuclease domain contains a core common for the Type II restriction endonucleases of PD-(E/D)xK superfamily including three catalytic residues in the active site [Obarska-Kosinska *et al.*, 2008] and a $Q_{179}xxxY_{183}$ motif characteristic for RecB-like nucleases [Sisakova et al., 2008b]. Helicase/translocase characteristic motifs conserved throughout Type I and The SF2 superfamily of DEAD-box helicases are clustered in two helicase/translocase domains and determine ATPase and translocation activities. The C-terminal helical is hypothized to be required for assembling with methyltransferase [Sisakova *et al.*, 2008a].

In comparison with Type I, Type II restriction endonucleases of the PD-(E/D)xK superfamily have a single domain organization [Niv *et al.*, 2007]. Dimerization, DNA recognition and DNA cleavage functions are carried out by via particular subdomain. Classical Type II restriction endonucleases cleave DNA within or at fixed position close to recognition sequence without translocation activity and not consuming ATP. Type II enzymes also have modification function which is carried out by separate protein methyltransferase with its own recognition subdomain [Pingoud *et al.*, 2005; Loenen *et al.*, 2014]. Members of the SF2 superfamily helicases from SWI2/SNF2 family as well as Type II restriction endonuclease can act without additional domains or complex formation other proteins. For example, isolated Rad54 translocase from yeast carries out branch migration and displacement of Rad51 from DNA translocating with the helicase domains in ATP-dependent manner [Burgess *et al.*, 2013; Zhang *et al.*, 2013].

The ATP molecule bound between the two helicase domains in the crystal structure of HsdR subunit (2W00) was found unexpectedly also to contact Lys220 residue from the endonuclease domain suggesting a potential involvement of this residue in communication of translocation and endonuclease functions. However, the static nature of the structure cannot give a clear answer how a signal could be transferred through the endonuclease domain to the catalytic site as the whole distance from Lys220-ATP to the

catalytic residues in the endonuclease domain is more than ~20 Å. The 180-loop, unresolved in the first crystal structure probably due to its high flexibility and revealed only in the crystal structure of a specific Lys220Ala HsdR mutant (4BEC)), is bearing a QxxxY motif lying directly above the endonuclease active site. This lead to our work hypothesis that this loop might participate in a signal transfer across the motor subunit, connecting the 220s loop and the catalytic site.

In this work we thus focus on the potential involvement of Lys220 and the QxxxY motif on the 180s loop in coupling of translocation and endonuclease functions. This involves also the probing of various other interdomain contacts of the endonuclease domain to the helical and the helicase 1 domain to determine a potential co-contributiuon.

These aims required:

1. Site-directed mutagenesis of Lys220 residue and residues around it, as well as residues in QxxxY motif and at the endonuclease/helical and endonuclease/helicase 1 domain interdomain interfaces;
2. Determination of expression conditions for mutant HsdR subunits;
3. Isolation of WT and mutant HsdR subunits;
4. Isolation of methyltransferase;
5. Examining of restriction phenotype of host cells expressing mutant HsdRs (*in vivo* restriction assay);
6. Examining of ATPase activity of mutant enzymes *in vitro*;
7. Characterization of the endonuclease activity of mutant enzymes *in vitro*.
8. Computational modeling and molecular dynamics to interpret the gained experimental data.

1. Introduction

1.1. Restriction-modification systems. General information

Restriction–modification enzymes (R-M systems) are enzymatic complexes, which were found only within prokaryotes, certain viruses, all bacteria and archaea seem to code for them. These systems protect bacterial cells against bacteriophage attacks, maintaining integrity of the cellular genome and it is considered as their main role in nature. During phage attack foreign DNA, which is unmethylated at specific sites, is injected into the bacterial cell. Unmodified DNA switches the endonuclease protective mode on and foreign DNA is cleaved. Hemi-methylated DNA is recognized as host DNA and then fully modified (methylated) by the enzyme. Methylated DNA is fully protected from cleavage and no restriction action of the endonuclease is observed [Abadjieva *et al.*, 1994 and 2003]. For the first time this effect was described by Bertani and Weige in 1953. The temperate phages P2 and λ were propagated in *E. coli* strain B and *E. coli* strain C and then used for *E. coli* strain K-12 infection which resulted in the activation of protective properties if unlabeled DNA was transferred to the cell [Bertani and Weige, 1953]. Defined bases of foreign phage DNA can be modified by chance at specific sites by the host modification enzyme. After such modification phage DNA is protected against degradation by endonucleases in the cell and recognized as host own DNA. This property gives to a phage the opportunity to survive and propagate [Arber and Dissoix, 1962; Smith *et al.*, 1972]. Later it was shown *in vitro* that this phenomena has an enzymatic origin. EcoKI from *E. coli* K-12 was the first isolated Type I R-M for which protective properties against foreign DNA were demonstrated [Meselson and Yuan, 1968; Linn and Arber, 1968].

Currently, all restriction enzymes are divided into four (I-IV) types, based on the number and character of subunits, enzyme activities, cofactor requirements, position of cleavage site and sequence specificity [Roberts *et al.*, 2003]. Key characteristics of distinct types of R-M systems are summarized in Table 1.

Table 1. Classification of R-M systems (adapted from Sistla and Rao, 2004, Loenen et al., 2014)

Characteristic	Type I	Type II	Type III	Type IV
Subunits	Three different	Two identical	Two different	Two different
Enzyme activity	Endonuclease, methylation, ATPase	Endonuclease or methylation	Endonuclease, methylation, ATPase	Endonucleaase
DNA translocation	Yes	No	Yes	Yes
Cofactor for DNA cleavage	ATP, Mg^{2+} (AdoMet)	Mg^{2+}	ATP, Mg^{2+}, AdoMet	GTPase, Mg^{2+}
Cleavege site	Random, remote from recognition site	At or near recognition site	25-27 bp from recognition site	Between methylated bases at multiple positions
Recognition sequence	Assymmetric, bipartite	Symmetric	Assymmetric	Bipartite, methylated
Methylation	Mg^{2+}, AdoMet	AdoMet	Mg^{2+}, AdoMet	No
REBASE	104 enzymes, 47 genes cloned, 34 genes sequenced, 5140 putatives	3938 enzymes, 633 genes cloned, 597 sequenced, 9632 putatives	21 enzymes, 19 genes cloned and sequenced, 1889 putatives	18 enzymes and genes cloned,15 sequenced, 4822 putatives

1.1.1. Type I

Structural genes for many Type I R-M systems from *E. coli*, *Salmonella*, and *Citrobacter freundii* have been cloned, sequenced and genetically analyzed [Sistla and Rao, 2004, Loenen *et al.*, 2014]. These enzymes are multi-subunit and multifunctional. The complex includes three different subunits: HsdM (M), HsdR (R) and HsdS (S) encoded by the ***hsdM***, ***hsdR*** and ***hsdS*** genes, respectively, where hsd is Host Specificity for DNA. HsdS subunit comprises two target recognition domains (TRDs), which impart target sequence specificity to both, restriction and modification activities of the complex. The HsdM subunit has an AdoMet binding site, which is required for the methylation activity. Two HsdM subunits and one HsdS subunit form the methyltransferase complex (MTase), which can exist in solution independently

and is sufficient for DNA modification. The HsdR subunit includes an ATP binding pocket, specific helicase motifs essential for DNA translocation and the endonuclease active site [Murray**,** 2000]. Altogether 91 R-M enzymes are assigned as Type I currently [Roberts *et al*., 2010].

Type Is subdivide into four different families based on cross-hybridization of genes and antibodies cross-reactivity. Sequences of R subunits from different type I families share only 20-30 % of identity [Titheradge *et al*., 1996]. Type I includes IA, IB, IC and ID families. Enzymes from AI or K-family (EcoKI, EcoBI) [Bickle, 1987, Wilson, 1991] recognize two constituent sequences, which are separated by a non-specific spacer ($AACN_6GTGC$ for EcoKI, $TGAN_8TGCT$ for EcoBI). A part of EcoKI is localized in the periplasm [Garcia and Molineux, 1999]. The EcoAI enzyme belongs to the IB family and the DNA recognition sequence is $GAGN_6GTCA$ [Suri *et al*., 1984]. The IC family is represented by EcoR124I, EcoR124II, EcoDXXI and EcoprrI, which recognize $GAAn_6RTCG$, $GAAn_7RTCG$ [Price *et al.,* 1987], $TCAn_7ATTC$ [[Piekarowicz *et al.*, 1986] and $CCAn_7RTGC$ [Tyndall *et al.*, 1994], respectively, where n is the spacer from six to seven variable nucleotides and R designates purine. EcoR124I enzyme is localized on the periplasmatic side of the cytoplasmic membrane [Garcia and Molineux, 1999].

KpnAI was found in *Klebsiella pneumoniae*. Based on the high level of identity of *hsdS, hsdM* and *hsdR* genes of EcoR9I and KpnAI [Lee *et al*., 1997] and complementation assays analysis, KpnAI enzymes are assigned to the ID family [Titherage *et al.*, 2001]. The Type ID family is represented by StySb1I as well [Titherage *et al.*, 1996]. KpnBI from *Klebsiella pneumoniae*, not showing a high percentage of similarity with other type I enzymes, was placed into a new family, IE [Chin *et al.*, 2004].

1.1.2. Type II

Type II is the largest group of restriction enzymes. According to REBASE statistics there are found 3762 enzymes in nature. These enzymes are small (in comparison to Type I R-M enzymes), much simpler, recognize and cleave dsDNA at a certain position [Yourell and Firman, 2008]. Restriction and modification activities of Type II R-M systems usually are provided by an

independent methyltransferase and endonuclease. Due to the ability to cleave DNA at predictable positions, Type II enzymes are widely used in molecular biology and commercially available. This huge group is divided into families such as IIS (FokI enzyme) [Szybalski *et al.*, 1991], IIE (NaeI enzyme) [Kruger *et al.*, 1995], IIM (DpnI enzyme) [Lacks and Greenberg, 1975].

1.1.3. Type III

Type III R-M systems cleave DNA 20-30 bp away from the recognition site [Dryden *et al.*, 2001] and this property is regarded as their special family feature. Mod- and Res-subunits of Type III enzymes are encoded by *mod* and *res* genes, respectively. While the Mod-subunit is responsible for modification, the Res-subunit recognizes a specific sequence on DNA and is responsible for restriction [Sistla and Rao, 2004]. The subunit stoichiometry for Type III enzyme is Mod_2Res_2 [Pingoud *et al.*, 2004]. 11 enzymes are currently assigned to type IIIs, accordingly REBASE statistics.

1.1.4. Type IV

These enzymes recognize and cleavage methylated DNA substrates. McrBC enzyme is the most studied in this group. It consists of two different subunits, MrcB, which perform the recognition function, and MrcC, which is responsible for DNA cleavage [Dryden *et al.*, 2001]. Eco57I is closely related to Type III R-M system (AdoMet, Mg^{2+} as cofactor for DNA cleavage), but ATP is not required for enzyme activities [Janulaitis *et al.*, 1992], and thus it was suggested as a representative of a novel Type of restriction enzymes and proposed to be classified as type IV [Petrusyte *et al.*, 1988].

1.2. Structural organization of Type I R-M system

Type I R-M systems are large, oligomeric enzymes that exhibit endonuclease (REase) and/or DNA modification (MTase) activities. [Obarska-Kosinska *et al.*, 2008] Type I restriction endonucleases are multimeric proteins

comprised of three different subunits encoded by *hsd* genes [Bickle, 1987, Wilson, 1991]. The *hsdR* gene encodes for the HsdR subunit which is responsible for restriction of foreign unmethylated DNA and is transcribed from its own promoter (PRES) [Bickle and Kruger, 1993]; while *hsdM* and *hsdS* genes are capable to produce an active DNA methyltransferase or MTase (independent complex with M_2S_1 stoichiometry) [Patel *et al.*, 1992] and transcribed from a separate promoter (PMOD) [Bickle and Kruger, 1993]. The MTases from different types of R-M enzymes vary in their size. The methyltransferase (162 kDa) of EcoR124I has been shown to consist of two copies of the HsdM subunit (each 58 kDa) and one HsdS subunit (46 kDa) [Taylor *et al.*, 1994]. The S-subunit of EcoR124I alone is insoluble and it gains solubility only being assembled with two M-subunits forming the MTase. In comparison, the S-subunit of EcoAI is soluble alone in solution [Mernagh *et al.*, 1997]. The HsdS subunit is responsible for binding to the recognition site [Glover, 1970; Hubacek and Glover, 1970] and HsdM is required for DNA modification [Willcock *et al.,* 1994]. The endonuclease activity demands products of all three genes as well as AdoMet, ATP and Mg^{2+} as cofactors. The EcoR124I multifunctional enzyme complex is capable of both restriction and methylation activities on DNA, which are dependent upon the methylation state of the recognition sequence [Abadjieva *el al.,* 1993].

The plasmid R124, which encodes EcoR124I, was described for the first time in the work of Meynell and Datta in 1966. Originally this plasmid was isolated from *Salmonella enterica* and the enzyme should have been named StyR124 [Youell and Firman., 2008]. The first genetic biochemical characterization of type I R-M enzymes was done by Bannister in 1968. The ability of R-M systems to prevent the growth of bacteriophages T1, λ, and P1 was observed in these experiments and thus they were named as host specifies system I (hspI).

1.2.1. HsdS subunit

A first identification of the secondary structure of the DNA specific subunits of Type I R-M enzyme was provided by Argos *et al.,* [1983]. The *hsdS* gene was cloned into the plasmid pTZ19R [Patel *et al.,* 1992]. The HsdS

subunit plays an important role in recognizing specific unmethylated DNA and binding to foreign DNA. HsdS has two domains: the N-terminal domain of 150 amino acids has been shown to interact with the tri-nucleotide component of the target sequence [Sistla and Rao, 2004] and the C-terminal domain recognizes a tetra-pentanucleotide component [Kannan *et al.*, 1989]. Amino acid sequences of both domains are strongly conserved. Between the two domains is a central conserved region, outside of which the amino acid composition varies. These variable regions were named TRD1 and TRD2 (target recognition domain) and each recognizes part of the bipartite target sequence on DNA [Obarska *et al.*, 2006]. The *hsdS* gene of EcoR124II, in comparison to the *hsdS* gene of EcoR124I, has a 12 bp extra repeat within the central region (two repeats for EcoR124I and three for EcoR124II) [Price *et al.*, 1987]. The HsdS subunit of EcoR124I is insoluble. The solubility can be gained only when complexed with two HsdM subunits [Patel *et al.*, 1992] or as GST fusion protein [Kusiak *et al.*, 1992]. The first structure of HsdS, solved at 2.4 Å, came from *Methanocaldococcus jannaschii* (MjaXIP) [Kim *et al.*, 2005]. The second HsdS structure, solved at 2.3 Å, from *Mycoplasma genitalium* (MgeORF438P) showed the same fold and secondary structure as the first one [Calisto *et al.*, 2005]. Based on these two structures and bioinformatics (fold recognition analysis), Obarska *et al.* proposed a structural model for HsdS of EcoR124I in 2006 (Fig. 1).

Mutational analysis of the HsdS subunit within the conserved region has been performed *in vivo* (Fig. 2) [Weiserova and Firman, 1998, Weiserova *et al.*, 2000]. A phenotype r^-m^- was expected, but, surprisingly, it was r^-m^+. The series of experiments was repeated *in vitro* and again confirmed the complicated behavior of DNA binding to the EcoR124I R-M complex [Weiserova *et al.*, 2000]. Additional mutations were introduced close to the central conserved region and tested *in vivo*. In most cases a r^+m^+ phenotype was detected, what was expected, only Ser154Pro, Arg163Gln, Glu200Gly showed the same phenotype as WT because they are localized at either end of the long coiled-coil which connect both TRDs. Probably, the segments connecting the coiled-coil structure and TRDs could play the role of flexible hinges which give a certain conformation mobility to the TRDs during DNA binding (Fig. 2) [Obarska *et al.*, 2006]. A Leu175Pro mutation, introduced in the center of the coiled-coil

structure, was supposed to have a strong effect on conformational changes. However, the mutation showed only a slight effect on DNA cleavage [Taylor *et al*., 1994]. Trp212 is situated in a flexible “elbow” and plays an important role in binding of HsdS to HsdM. Arg mutant showed lower ability of MTase to be assembled [Weiserova *et al*., 2000].

1.2.2. HsdM subunit and Methyltranspherase

The HsdM subunit in the MTase complex participates in DNA modification, using S-adenosylmethionine (AdoMet) as the donor of methyl groups. Assembly of the HsdR subunit with the Mtase is predicted to happen via the helical domain of the HsdM subunit. Two structures of HsdM are known: M.BthVORF4518P from *Bacteroides thetaiotaomicron* solved at 2.2 Å and HsdM of EcoKI from *E. coli* solved at 2.8 Å. The HsdM subunit catalyzes the transfer of a methyl group from AdoMet to the N_6 position of specific adenine, when the target base is rotated 180° out of the DNA helix towards the catalytic site of the enzyme [Mernagh *et al.,* 1998]. The sequence identity of the M-subunit of EcoR124I with M-subunits from other families is 25-30% [Sharp *et al*., 1992]. The HsdM subunit includes two domains: the HsdMc catalytic domain, which is responsible for binding of the cofactor for DNA methylation, and the HsdMh helical domain, which contacts helical domain of R-subunit. Together with the HsdS subunit two HsdM subunits form the core complex of the methyltransferase or MTase with a M_2S_1 stoichiometry [Taylor *et al*., 1992]. Firstly the ***hsdM*** gene was singly cloned into the vector pUC119. The HsdM subunit of EcoR124I was expressed by Patel from the *lac* operon of the pYC119 plasmid in quantities which were enough for protein purification. The expressed HsdM subunit has a size of 58 kDa and exists in soluble fraction as a monomer. The M subunit alone has no enzymatic activity, but in the MTase complex, using AdoMet as a donor of methyl groups, it is able to modify DNA. Proteolysis analysis shows that HsdM is quite resistant to proteolysis and trypsin removes just a small tail from the S subunit having a globular arrangement [Patel *et al.*, 1992]. The HsdM subunit contains

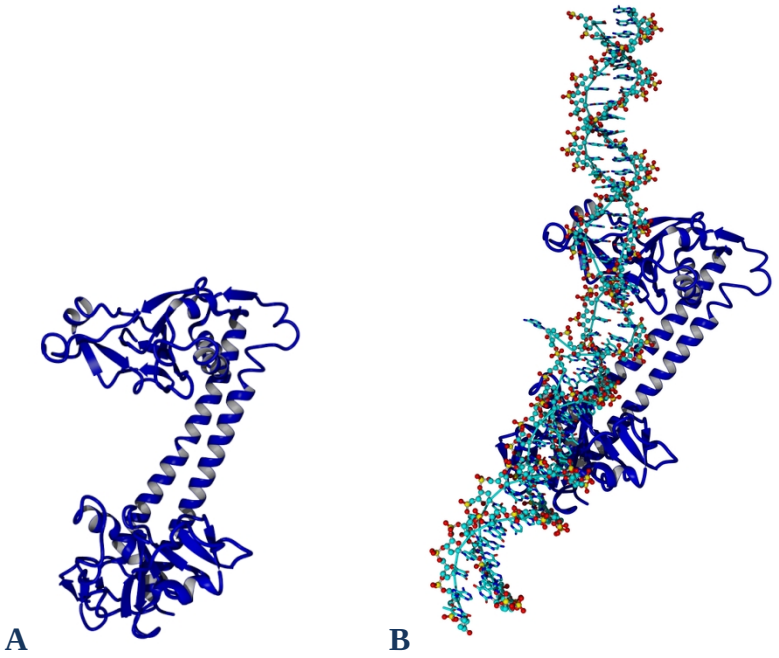


Figure 1. Predicted 3D structure of HsdS from EcoR124I without DNA (A) and with DNA (B) (M.S.EcoR124I.DNA.pdb, ftp://genesilico.pl/iamb/models/MTases/M.EcoR124I, Obarska *et al.*, 2006). Figure prepared in YASARA.

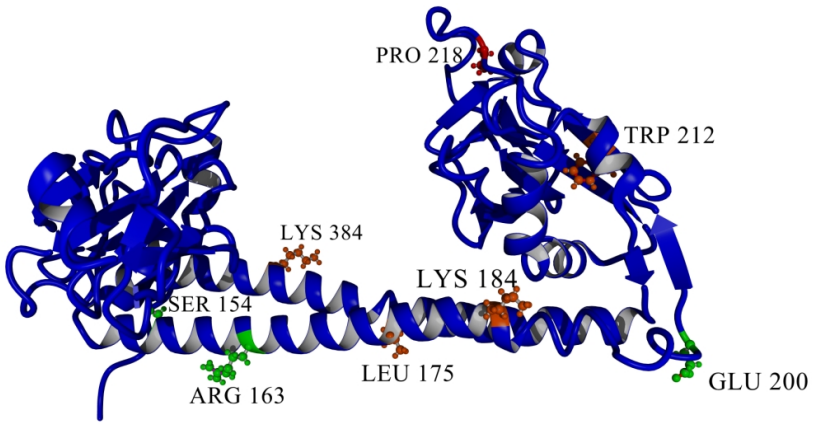


Figure 2. Mutations in S-subunit (blue) which lead to distinct phenotypes (r^+m^+ amino acids residues are green; r^-m^+ are orange; r^-m^- are red) (M.S.EcoR124I.DNA.pdb, ftp://genesilico.pl/iamb/models/MTases/M.EcoR124I, Obarska *et al.*, 2006). Figure prepared in YASARA.

characteristic motifs typical for all adenine methyltransferases [Loenen *et al.*, 1987]: motif I (D/E/SXFXGXG), which is responsible for AdoMet binding; motif IV (N/DPPF/Y/W), which is involved in catalysis reaction [Willcock *et al.*, 1994]. Only the fully assembled complex with M_2S_1 stoichiometry is able to modify DNA. M_1S_1 is a heterodimer, which was identified just for EcoKI [Cooper and Dryden, 1994], and can only bind to DNA [Powell *et al.*, 1998]. Enzymes from different families have a distinct methylation activity depending on the target DNA sequence. EcoAI from IB family shows no affinity for hemi-methylated DNA as most other methyltransferases. Footprinting experiments demonstrated that M.EcoR124I forms contacts in the major groove of the DNA helix [Mernagh *et al.*, 1996]. In comparison the enzymes from IA (EcoKI) and IC (EcoR124I) families strongly prefer hemi-methylatehed substrates [Murray, 2000]. Janscak *et al.* (1998) report a molar extinction coefficient for metyltransferase of EcoR124I of 160400 M^{-1} cm^{-1}. A predicted structural model of M.EcoR124I methyltransferase in complex with DNA was reported by Obarska et al., in 2006 (Fig. 3).

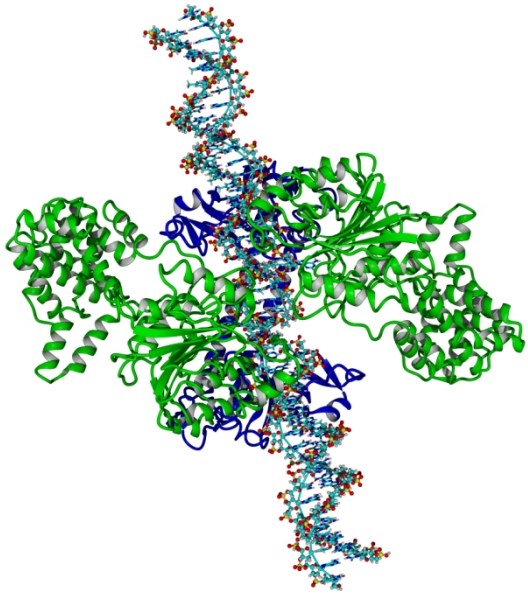


Figure 3. Predicted structure of methyltransferase of EcoR124I with DNA. HsdS subunit is in blue, HsdMs subunits are in green (M.S.EcoR124I.DNA.pdb, ftp://genesilico.pl/iamb/models/MTases/M.EcoR124I, Obarska *et al.*, 2006). Figure prepared in YASARA.

The structure of MTase complex was determined at low-resolution using SANS (small angle neutron-scattering) [Callow *et al.*, 2007]. The binding affinity of MTase of EcoR124I for its cognate DNA recognition sequence is 10^8 M^{-1} [Taylor *et al.*, 1993]. It was suggested by X-ray solution scattering studies that the shape of M.EcoR124I is more compact when the complex interacts with specific DNA sequence and no influence on compactization is observed in the presence of AdoMet. Large increase of circular dichroism confirms that not only MTase undergoes large changes upon DNA binding, but that the conformation of DNA is changed too [Taylor *et al.*, 1994]. Proteolysis experiments with HsdS and HsdM subunits of M.EcoKI (Type IA) have shown that both subunits can be cleaved. In comparison, only the S-subunit of M.EcoR124I is digested by proteases, while the M-subunit remains intact. The S-subunit of EcoR124I, in the MTase complex bound to the cognate sequence, hidden under DNA gains additional protection and the HsdS surface becomes inaccessible for proteases. MTase, after treatment by trypsin, still maintains three subunits assembled together where two M-subunits are uncleaved and the S-subunit is partially digested. Such proteolysed MTase binds to target sequence worse due to a significantly decreased ability of the S-subunit for binding. The DNA-MTase complex, probably, is more resistant for trypsitation as result of the presence of bound DNA and the more compact character of such complex [Cooper and Dryden, 1994].

Chemical modification analysis of the surface accessible Lysine residues of M.EcoR124I shows that they are crucial for interactions with DNA in the free M.EcoR124I enzyme and DNA-protein complex. 19 Lysine residues of HsdS subunit and 11 Lysine residues of each M subunit from 109 total Lysine residues of M.EcoR124I have susceptibility to modification and Lysine residues located in S subunit are protected for modification better than those from the M subunit [Obarska *et al.*, 2006].

Footprint experiments display that adenine is a target for methylation by M.EcoR124I. M.EcoR124I binds better when a base in the target site is replaced by target adenines. During binding of the enzyme a strong conformational change in the sugar-phosphate skeleton of DNA around methylation region is observed [Mernagh *et al.*, 1996].

1.2.3. HsdR subunit

The HsdR subunit is the biggest part of whole complex and is responsible for restriction. PcrA from *Bacillus stearothermophilus* was the first crystal structure from SF1 helicases described. PcrA is comprised of 4 distinct domains: 1A and 2A contain seven motifs characteristic for all DEAD-box helicases, which are important for interactions with nucleic acid and for hydrolysis of ATP; 1B and 2B, presence of which varies from helicase to helicase [Velenkar *et al.*, 1999]. Domains 1A and 2A are also present in the recombination enzyme RecA from *E. coli*, two domains containing seven characteristic motifs form the NTP-binding site located between them. The structure of PcrA in complex with DNA reveals that a ssDNA strand of the substrate binds to the top of 1A and 2A domains and the double-stranded portion of the substrate is bound along the domain 2B [Singleton and Wigley, 2002]. Based on the structural organization of the HsdR subunit of Type I R-M enzymes all of them can be classified as SF2 helicases which translocate dsDNA in an ATP-dependent manner. EcoR124I translocates dsDNA in the 3'-5' direction without duplex unwinding. [Stanley *et al.*, 2006].

A model of the R-subunit was proposed by Obarska-Kosinska *et al.* in 2008. It is possible to reconstitute the holoenzyme complex in a $R_2M_2S_1$ stoichiometric ratio *in vitro* simply by mixing R-subunits and MTase as was shown for EcoR124I [Janscak *et al.*, 1998], EcoR124II [Dreier *et al.*, 1996], EcoKI [Dryden *et al.*, 1997]. The EcoBI R-M enzyme can have several stoichiometric forms (e.g. $R_1M_2S_1$, $R_2M_2S_1$ and $R_4M_2S_1$) [Janscak *et al.*, 1996]. Only the whole R-M complex of EcoR124I with a stoichiometry of $R_2M_2S_1$ displays both ATPase/translocation and endonuclease activities [Janscak *et al.*, 1998]. The structure of the 120 kDa HsdR subunit of HsdR which is transcribed in *E. coli* from the P_{res} promoter [Zinkevich *et al.*, 1997] in complex with ATP was determined at 2.6 Å resolution using Single Wavelength Anomalous Dispersion (SAD) phasing using SeMet [Lapkouski *et al.*, 2007 and 2009]. The R-subunit is arranged in four approximately globular domains: a N-terminal endonuclease domain (has endonuclease active site), two RecA-like helicase domains (take part in ATP-dependent DNA translocation) and a C-terminal helical terminal domain (proposed complex assembly through interaction with

MTase) (Fig. 4). The endonuclease domain has two parts: a central αβα core characteristic for Type II endonucleases and the subdomain II with an unknown function [Niv *et al.*, 2007]. The persistent structural core present in the region X can be classified as part of PD-(E/D)xK family. This includes R-M System types I-III, as well as corrective and recombination nucleases [Sisakova *et al.*, 2008]. According to Stanley et al., [2006] helicases are divided into two classes based on which type of nucleic acids they bindnd and how they interact with them. SF1 helicases have a binding pocket with hydrophobic interactions formed by aromatic residues [Soultanas and Wigley, 2001]. SF2 enzymes (including Type I R-M enzymes) contact the nucleic acid substrate via the backbone [Kim *et al.*, 1998]. However, considering all helicase characteristic motifs all of them are divided into four helicase superfamilies SF1, SF2, SF3 and SF4 [Hall and Matson, 1999]. Helicase domain 1 and 2 in HsdR, or two RecA-like domains, contain the seven helicase characteristic motifs including DEAD-box motifs conserved in all helicases. Both domains have a similar structural organization forming nearly a homodimer and can bind one molecule of ATP [Lapkouski *et al.*, 2009]. The helicase domains of HsdR subunit include motifs Ia, Ib, Ic and II clustered in domain 1 and motif IVa in domain 2 as in other related SF2 helicases [Durr *et al.*, 2005]. The N-terminal domain (13-260 residues in 2W00) has a high degree of similarity with the endonuclease domains from other Type I R-M enzymes. Both helicase domains (261-461 and 470-731 residues) are organized in similar way as in RecA. Fourth, the C-terminal helical domain, includes 732-892 residues [Lapkouski *et al.*, 2009].

Sequence analysis of R-subunits of EcoR124I, EcoAI and EcoKI (Fig. 5) shows that Asp151, Glu165 and Lys167 residues are conserved and form the endonuclease active site. *In vivo* restriction phenotype and restriction analysis of these mutants confirmed their role for restriction activity of EcoR124I killing endonuclease activity in all three cases and additionally affecting translocation [Sisakova *et al.*, 2008].

According to gel filtration analysis [Zinkevich *et al.*, 1997], DLS, Sedimentation velocity experiments and SANS [Obarska-Kosinska *et al.*, 2008] the R-subunit in solution is present in its monomeric form. The molar extinction coefficient of HsdR subunit derived from contribution of all Tryptophan and Tyrosine residues is 98225 M^{-1} cm^{-1} [Janscak *et al.*, 1998].

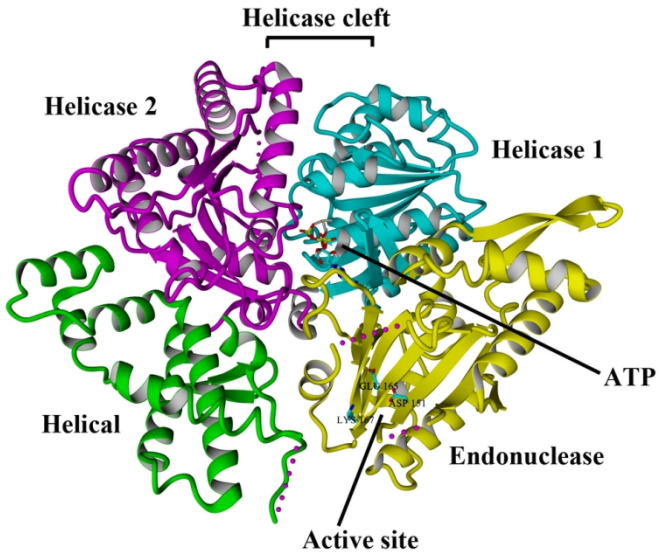
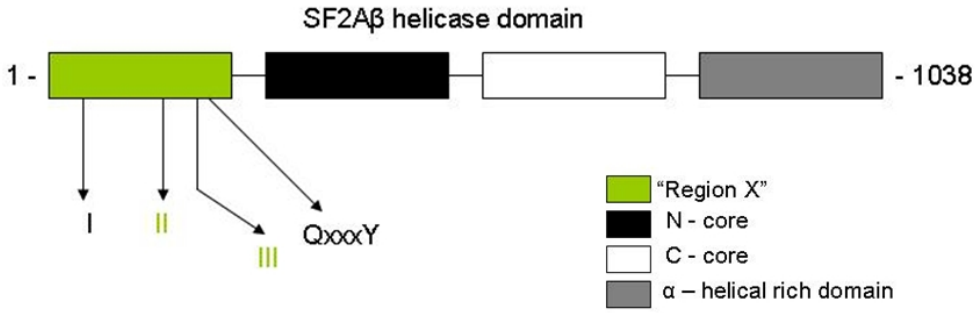


Figure 4. Structure of HsdR subunit with ATP (PDB ID: 2W00, Lapkouski *et al.*, 2009). Figure prepared in YASARA.



EcoR124I	**E**$_{37}$	Y**D**$_{151}$	VTILV-NGLPLVQI**E**$_{165}$	L**K**$_{167}$	AFN**Q**$_{179}$	IHR**Y**$_{183}$
EcoAI	**E**$_{11}$	A**D**$_{61}$	IVLIHKPGIPLAVI**E**$_{76}$	A**K**$_{78}$	GMQ**Q**$_{90}$	GIE**Y**$_{94}$
EcoKI	**E**$_{242}$	A**D**$_{298}$	YVLFV-GLKPIAVV**E**$_{312}$	A**K**$_{314}$	RLN**E**$_{326}$	SYR**Y**$_{330}$

Figure 5. Domain structural organization and nuclease motifs in type I R-M systems of EcoR124I, EcoAI and EcoKI. Sequence alignment of "Region X" (endonuclease domain) and important amino acid residues from II and III motifs are shown. QxxxY motif could not be identified in EcoKI while being present in EcoAI and EcoR124I [adapted from Sisakova *et al.*, 2008].

1.3. EcoR124I R-M complex

EcoR124I is a pentameric enzymatic complex composed of three different subunits (HsdS, HsdM and HsdR) encoded by three *hsd* (*hsdS*, *hdsM* and *hsdR*) genes. They expressed from two promoters: initiation of the hsdR gene transcription occurs from the P_{Res} promoter, P_{Mod} is required for *hsdM* and *hsdS* transcription. For the first time restriction activity of R-M system encoded by the conjugative plasmid R124 was noticed by Bannister and Glover (1968). They demonstrated the negative influence of this system on growth of bacteriophages T1, *Φ*80, P1, and λ. The enzyme isolated from R124 plasmid was named EcoR124, but originally, *S. enterica* was the source of R124 plasmid and the enzyme should have been named StyR124 [Youell and Firman, 2008]. The HsdS subunit with two HsdM subunits forms an independent methyltransferase complex as described earlier. The HsdS subunit of EcoR124I requires the assembly with HsdM subunits to be soluble and can be purified only as part of MTase complex with a M_2S_1 stoichiometry [Taylor *et al*., 1992]. The M_1S_1 complex does not exist in the case of EcoR124I while in case of EcoKI the MTase with a stoichiometry M_2S_1 can dissociate into the inactive M_1S_1 complex and a free M-subunit [Meselson *et al.*, 1972; Dryden *et al*., 1993 and 1997]). MTase and R-subunits can be reconstituted in pentameric EcoR124I R-M complex *in vitro*.

Translocation and restriction activities require the full prentameric EcoR124I R-M complex with the stoichiometry of $R_2M_2S_1$ to be assembled. Affinity of binding of the first R-subunit to MTase is higher than for the second R-subunit resulting in formation of a EcoR124I complex with a stoichiometry of $R_1M_2S_1$ (Fig. 6). In contrast to the fully assembled EcoR124I with stoichiometry $R_2M_2S_1$ this complex is not able to cleave the DNA substrate. The fully assembled EcoR124I R-M complex with stoichiometry of $R_2M_2S_1$ can dissociate into $R_1M_2S_1$ complex and free HsdR subunit with a K_d of ~2.4×10^{-7} M [Janscak *et al*., 1998]. When the R-subunit binds to the MTase core its interaction leads to changes in DNA conformation [Halford *et al.*, 2004]. In the EcoKI $R_2M_2S_1$ complex each R-subunit has a similar binding affinity to the MTase [Dryden *et al*., 1997]. The EcoAI restriction endonuclease is a weaker complex and easily dissociates into methyltransferase and R-subunits. Only the

reconstitution of the holoenzyme pentameric complex *in vitro* with appropriate HsdR subunit excess can produce an active endonuclease suitable for testing of biochemical properties [Suri *et al.*, 1984]. Restriction enzyme EcoBI is present in different stoichiometric forms such as $R_4M_2S_1$, $R_2M_2S_1$, $R_1M_2S_1$ and $R_1M_1S_1$ [Eskin and Linn, 1972]. The binding affinity of each R-subunit of EcoR124I is different and the ratio of MTase to the HsdR subunit is greater than is required for the $R_2M_2S_1$ complex to be formed [Janscak *et al.*, 1998]. Termination of dsDNA translocation by EcoR124I leads to dissociation of the R-subunits from the MTase core, however they remain bound to DNA, thus, the MTase-HsdR complex can be considered as complex with a short life-time, in comparison to the DNA-MTase complex where MTase is bound to DNA after the cleavage event [Seidel *et al.*, 2005].

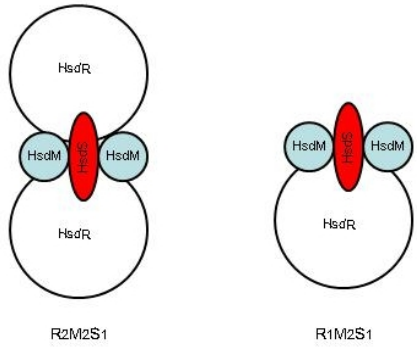


Figure 6. Subunit assembly in EcoR124I R-M complex.

Control of the endonuclease activity *in vivo* is based on subunit assembly [Janscak *et al.*, 1998]. The first stable complex formed in living cell is the methyltransferase. After the ***hsd*** genes are introduced into the host, the restriction proficient $R_2M_2S_1$ will be formed only if there is no free MTase left for the R_1 complex formation, as the high binding affinity of the first R-subunit to methylase shifts the equilibrium to the formation of $R_1M_2S_1$ complex in the beginning. This gives enough time for the MTase and/or the R_1 complex for modification of host DNA and thus the cell genetic material is protected against self-degradation by the R_2 complex [Janscak *et al.*, 1996]. EcoKI has a similar mechanism of subunit assembly as EcoR124I [Dryden *et al.*, 1997]. Restriction activity of EcoR124I can be turned off by dissociation of the R_2 complex with

the release of one HsdR into solution. Such mechanism of control of the endonuclease and modification activities might be essential for transmissibility of the plasmid R124 [Janscak and Bickle, 2000]. No restriction alleviation was observed for EcoR124I in comparison to EcoKI, this mechanism of enzyme control is family depended and depends on holoenzyme localization in the host cell [Holubova *et al*., 2000 and 2004].

SAXS, SANS and electron microscopy experiments done by Kennaway *et al*. 2012 resulted in two atomic models for EcoR124I with and without DNA. The Enzyme+DNA models show that the MTase core is in a compact, closed conformation after accommodation of DNA and EcoR124I without DNA prefers a very loose and open conformation in solution with respect to the interface between the S and M subunits.

1.4. Translocation activity

Translocation of enzymes along the DNA plays an important role for many existential processes as reparation, replication, recombination, transcription, translation and restriction [Seidel *et al.*, 2004]. Protein moves along DNA during ATP-dependent translocation and the phenomena of DNA metabolism in the living cell is observed [Stanley *et al.*, 2006]. Translocation of DNA includes such steps as initiation, termination and reinitiation [Seidel *et al.*, 2005]. Initiation of DNA translocation requires a specific sequence such as a promoter or a specific nucleic acid structure like a dsDNA break [Farah and Smith, 1997]. Enzyme movement along DNA occurs with different mechanisms depending on the particular enzyme. DNA polymerase forms a clamp with its β subunit that slides loosely along DNA [Kong *et al.*, 1992]. RNA polymerases and DNA helicases can move along the minor groove on right-handed DNA double helices [Droge, 1992]. R-M enzymes recognize specific non-palindromic DNA sequences (GAGnnnnnnnGTCA where n is any nucleotide) [Janscak *et al.*, 1999] and translocate DNA towards itself in both directions in ATP-dependent fashion on their HsdR subunits [Studier and Bandyopadhyay, 1988]. Translocation of DNA moves the possible cleavage site for Type I R-M enzymes up to several thousand base pairs away from their recognition site [Eskin and Linn, 1972].

The EcoR124I R-M complex, tightly anchored on the DNA recognition sequence, translocates dsDNA toward itself forming long DNA loops. It is a ATP-dependent process, advancement of 1 base pair requires one ATP molecule to be hydrolyzed [Seidel *et al.*, 2008]. Approximately 60 molecules of EcoKI R-M enzyme are present at the same time in the cell [Kelleher and Raleigh, 1994], and they can translocate with a speed of 1000 bp/s using 1 molecule ATP per 1 bp advanced [Endlich and Linn, 1985]. During translocation ATP is hydrolyzed to ADP and Pi is released in solution. A malachite green assay [Janscak *et al.*, 1996] can be used to monitor the rate of ATP hydrolyses *in vitro*. This method allows the measurement of Pi concentration based on spectrophotometrical quantification of the phosphomolibdate-malachite green complex. Using this technique we can judge indirectly about dsDNA translocation in terms of ATP hydrolysis. The malachite green assay is a very simple technique and it can be used for preliminary routine testing.

As described previously, translocation starts with binding of the Type I R-M enzyme to the recognition sequence [Seidel *et al.*, 2005]. DNA binding of EcoR124I was reported by Janscak *et al.*, (1996) using surface plasmom resonance (SPR) analysis. To determine of the influence of cofactor AdoMet on binding to DNA experiments reaction mixtures with or without AdoMet were analyzed. Experimental data suggests that enzyme dissociates much faster in the absence of AdoMet and more slowly if AdoMet is present in solution. Binding of MTase of EcoR124I to DNA in the absence of AdoMet cofactor is associated with large changes in the quaternary structure of MTase [Taylor *et al.*, 1994].

A single molecule magnetic tweezers technique was developed for studies of translocation activity (Fig. 7) [Strick *et al.*, 1998]. The triplex displacement method as an alternative method for studies of translocation activity of EcoR124I is using a special triplex forming oligonucleotide (TFO) which binds to a specific triplex binding site (TBS) located distantly from the enzyme binding site on DNA (Fig. 8) [Firman and Szczelkun, 2000].

These methods allow the observation of a detailed picture of translocation including translocation rates (Fig. 9). The initiation of translocation activity begins with binding of MTase to the recognition site on DNA (magnetic-tweezer assay [Seidel *et al.*, 2005]) and binding of HsdR subunit and can be expressed as following scheme:

$$\text{EcoR124I (HsdR+MTase)} \xrightarrow{k_{\text{bind, HsdR}}} R_0 \xrightarrow{k_{\text{ini}}} R_1^{\text{trans}}$$

where R_0 is non-translocating R-subunit; R_1^{trans} is the translocating R-subunit; $k_{\text{bind, HsdR}}$ is the binding constant of R-subunit to MTase core; k_{ini} is the initiation of translocation constant including HsdR and ATP binding and loop formation events.

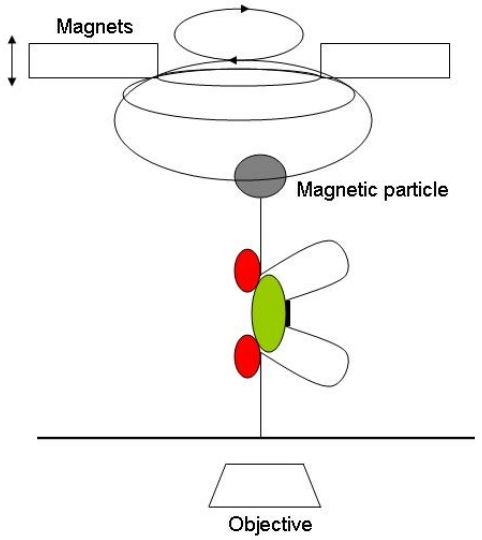


Figure 7. Schematic representation of magnetic tweezers. Linear DNA with one EcoR124I recognition site is fixed on a glass slide and magnetic particle is bound to the second end of DNA molecule. Magnets keep the DNA stretched. Video microscopy is used for measuring end-to-end distance which is decreasing during translocation by EcoR124I. Red ovals are R-subunits and green oval is MTase. [adapted from Seidel *et al.*, 2004].

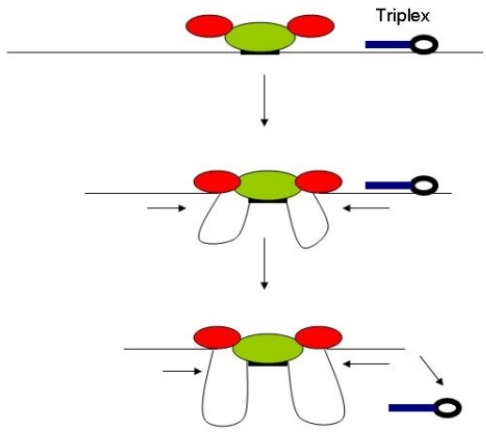


Figure 8. Triplex displacement method. EcoR124I binds to single-site linear DNA with TFO bound remotely from EcoR124I recognition sequence. EcoR124I R-M enzyme anchored tightly on DNA translocates DNA toward itself forming DNA loops. DNA translocation leads to displacement of bound TFO form DNA into solution and change of fluorescence is monitored by fixed a stopped-flow machine in real time [adapted from Seidel *et al*., 2005].

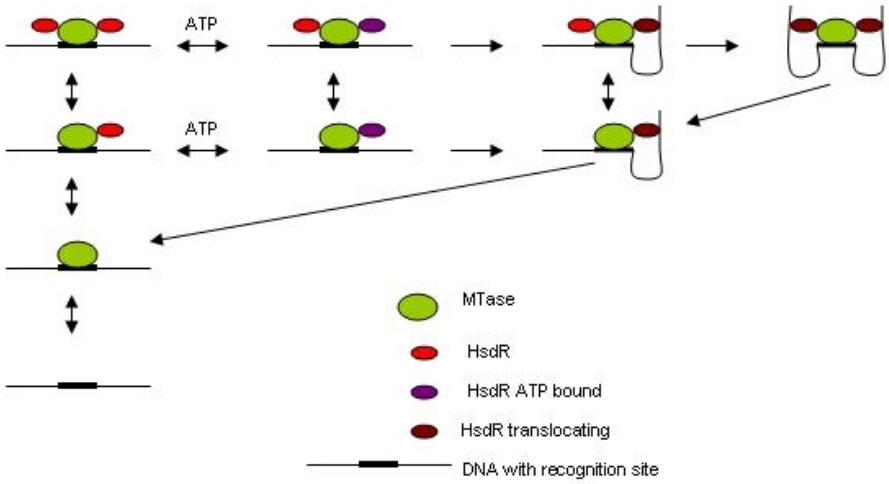


Figure 9. Schematic representation of initiation and termination of DNA translocation by EcoR124I [adapted from Seidel *et al*., 2005].

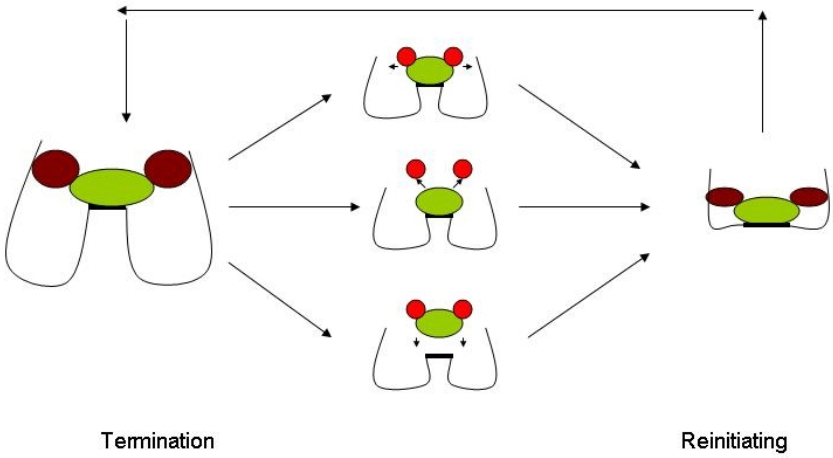


Figure 10. Termination and reinitiation of translocation by Type I R-M systems. Three possible mechanisms for termination of translocation by translocating R-M enzyme: 1 - translocated DNA can disengage from R-subunit while enzyme complex remains bound to target sequence; 2 – dissociation of R-subunits from MTase core; 3 - enzyme complex dissociates from DNA completely. Reinitiation of translocation can occur after termination of translocation depending on the termination mechanism [adapted from Seidel *et al*., 2005].

Using atomic force microscopy (AFM) it was demonstrated that binding of the holoenzyme to DNA and initiation of translocation lead to large conformational changes on DNA in the region of the target sequence [van Noort *et al.*, 2004]. After MTase binding DNA bents about 50° and in the end of initiation step the length of DNA is decreased by 8 nm allowing loop formation [Seidel *et al*., 2005]. Initiation of translocation occurs in an ATP-dependent manner and a DNA loop initially appears in the R_1-complex during ATP-saturation. Termination of the translocation occurs when HsdR dissociates from MTase [Siedel *et al.*, 2004]. There are three possible ways of termination of translocation (Fig. 10). The second R-subunit in the R_2-complex translocates independently from the first R-subunit, so two rates of DNA translocation for R_1 and R_2 complex can be observed. Termination of translocation happens when one of two translocating HsdR subunits dissociate [Seidel *et al.*, 2005].

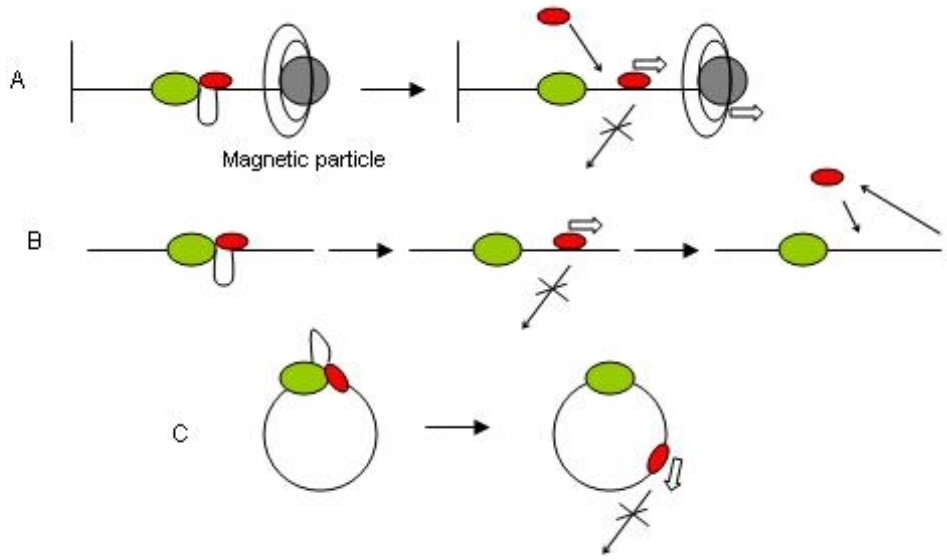


Figure 11. Turnover of HsdR subunit on different DNA substrates A) Magnetic tweezers experiment; B) Linear DNA substrate; C) Supercoiled DNA substrate [adapted from Simons and Szczelkun, 2011].

Translocation activity of EcoR124I R-M complex occurs without unwinding of the DNA duplex as was shown for other helicases [Lia *et al.*, 2006; Stanley *et al.*, 2006]. The enzyme moves the minor groove of DNA helix in 3'-5' direction, interactions with the 5'-3' strand is weaker but very important for stabilization of the translocating complex on DNA.

Turnover of the HsdR subunit of EcoR124I after translocation termination and dissociation of HsdR from MTase can occur only from free DNA ends [Simons and Szczelkun, 2011]. Translocation is terminated after the HsdR subunit dissociates [Seidel *et al.*, 2004]. Turnover of HsdR subunit is not possible in the magnetic-tweezers experiment as there is no free end on DNA. If the R-subunit is free in solution it can assemble with MTase bound to the target site and start translocation. No turnover of the HsdR subunit on circular DNA substrate is observed, only new HsdR from solution after binding to MTase can initiate translocation (Fig. 11).

1.5. Restriction activity

First studies of restriction and modification activities of type I R-M systems were done by Studier and Bandyopadyay back in 1988. Type I R-M enzymes specifically bind to a target sequence on DNA and cleave DNA substrate randomly and remotely from this specific sequence due to DNA translocation [Janscak *et al.*, 1999]. EcoR124I translocates DNA in both directions with a rate of a few hundred bp per second [McClelland *et al.*, 2005] forming two DNA loops [Yuan *et al.*, 1980] and introducing double strand cuts up to several thousand bp away from the recognition sequence after further translocation is impeded [Studier and Bandyopadyay, 1988]. Type I enzymes can dissociate into a R_1 complex under *in vitro* conditions, releasing one HsdR subunit [Suri *et al.*, 1984]. Affinity of binding for the second HsdR subunit to MTase in EcoR124I in contrast to other type I R-M systems is weaker and the formation of $R_1M_2S_1$ and $R_2M_2S_1$ is determined by the ratio between HsdR and MTase. The R_1 complex is able to modify DNA but is restriction deficient, while the R_2 complex possesses both activities [Seidel *et al.*, 2004].

The cleavage mechanism for Type I R-M enzymes is not fully understood yet. The mode of enzyme action depends on the methylation status of the DNA substrate. On unmodified DNA substrate DNA translocation precedes restriction activity. A blocking mechanism of DNA translocation for DNA cleavage was proposed by Janscak *et al.* (1999). It was shown that not only protein-protein interaction but also DNA with a Holliday junction (HJ) present a steric clash for DNA translocation or even the degree of supercoiling on DNA substrate or other steric blockage leading to impeding of DNA translocation triggers endonuclease activity of Type I R-M enzymes (Fig. 12). It was shown that positively and negatively supercoiling of DNA substrates lowers the rate of cleavage by EcoAI and EcoR124I [Janscak *et al.*, 1996], alterations of the rate of DNA translocation are not excluded [Janscak *et al.*, 1999]. EcoR124I cuts from either side of the Holliday junction on circular DNA. The complex might not attain the Holliday junction due to the higher degree of DNA substrate supercoiling.

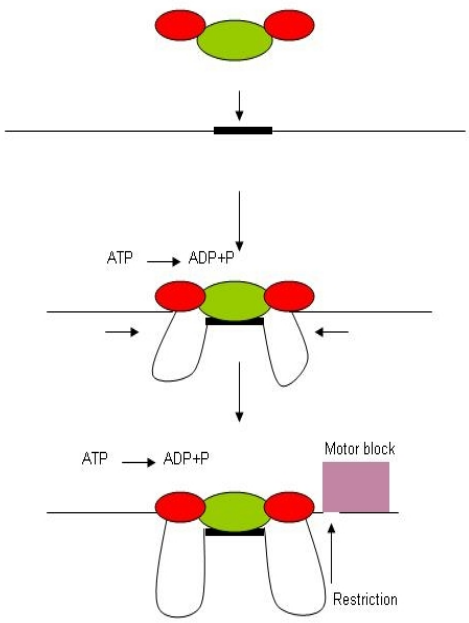


Figure 12. DNA cleavage model by Type I R-M enzymes, collision model. R-M enzyme binds to target sequence and starts to translocate DNA in ATP-dependent fashion producing DNA loops. When RM complex collides with a barrier, for example with another translocating Type I enzyme, or other physical block stalls translocation, enzyme introduces double strand break on DNA This model can be also applied for uni-direction translocation [adapted from Janscak *et al.*, 1999].

Biochemical analysis shows that the continuity of the 3'-5' DNA strand is more important for DNA translocation and interactions with the 5'-3' DNA strand stabilize the enzyme complex on DNA substrate [Stanley et al., 2006]. Experiments with physical blocks such as Holliday junctions or positive supercoiling which can impede DNA translocation also suggest that nicking (single strand break) of DNA can also pause translocation [Dreier et al., 1996].

Different DNA substrates (linear DNA and supercoiled plasmid DNA with one or two recognition sites) were used for studies of restriction activity of Type I R-M enzymes. Single-site linear DNA substrates support DNA cleavage, however such linear DNA substrates require a very high concentration of the enzyme and could not be fully cleaved [Murray *et al.*, 1973]. Collision of translocating R-M enzyme with another one unspecifically bound was proposed to lead to DNA cleavage [Studier and Bandyopadhyay, 1988].

Type I R-M enzymes cleave circular DNA substrate with one recognition site efficiently generating two types of products such as intermediate nicked DNA product and fully cleaved linear DNA product (Fig. 13) [Rosamond *et al*., 1979].

Type I R-M enzymes cleave linear DNA substrates with two or more target sites between them, products appear on agarose gel as a smear (many species of DNA of different length as result of random cleavage) [Studier and Bandyopadhyay, 1988] and change of orientation of asymmetric target sequences does not affect the restriction [Dreier *et al*., 1996].

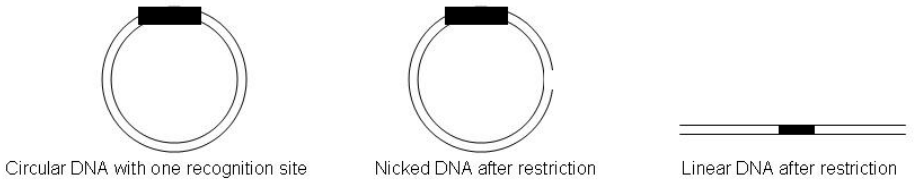


Figure 13. Circular DNA substrate and its restriction DNA products.

When the enzyme moves along a circular DNA substrate, the degree of supercoiling is increasing. In case of linear DNA substrate, both free ends of the DNA can freely move and rotate thus preventing tension. Translocation of circular DNA by the HsdR subunits of the enzymatic complex tightly anchored on recognition sequence leads to negative supercoiling of DNA in loops and a high degree of positive supercoiling of untranslocated DNA resulting in a decrease of the translocation rate. Translocation is impeded only when the supercoiling on a short stretch of DNA is too high presenting a thermodynamic barrier for further translocation [Szczelkun *et al.*, 1996].

Janscak *et al*. in 1996 studied the cleavage of supercoiled DNA substrate with one and two target sequences for EcoR124I. Analysis of cleavage kinetics of both DNA substrates shows that single-site circular substrate is converted efficiently into a linear product by EcoR124I at enzyme-to-DNA ratio of 4 and further increase in enzyme concentration doesn’t lead to increase of linear product. In contrast, enzyme-to-DNA ratio greater than 4 leads to further degradation of two-site circular DNA substrate, and the generated linear DNA

product is cleaved between the two EcoR124I recognition sequences more efficiently.

Interactions of two type I enzymes, belonging to different families, in cleavage of linear substrate were studied by Janscak et al. (1999). The main idea of these experiments was to demonstrate that only collision of two translocating Type I enzymes is sufficient for DNA cleavage and no protein-protein interactions are required. Thus, Janscak's model suggests that stalling of DNA translocation is the only reason leading to DNA cleavage by Type I R-M enzyme.

Jindrova *et al.,* (2005) characterized the DNA ends formed by EcoR124I, EcoAI and EcoKI after cleavage of single-site circular DNA. DNA cleavage by all enzymes resulted in the formation of both 5'- and 3'-overhangs of different length and no blunt ends were observed. EcoR124I showed a preference for 5'-overhangs of 3-5 nt; EcoAI showed a preference for 3'-overhangs of 2-3 nt and EcoKI displayed a preference for formation of 5'-overhangs of 6-7 nt. The double-strand break on DNA after cleavage by Type I R-M enzymes is the result of the interaction between the translocating R-subunit and a second R-subunit; which is hypothesized to be either is free in solution or being part of another assembled enzyme complex. However, there are no clear evidences yet for either possibility. The enzyme cuts at non-specific DNA sites and local DNA sequences can have a different rate of cleavage. One R-subunit performs the double-stand break faster than other. Together two HsdR subunits can cleave DNA resulting in a double-strand break.

Restriction activity assay *in vitro* doesn't reflect fully the processes occurring in nature, as the in vivo situation in the cell is more complex and various additional factors might influence the enzyme activity, and that is why *in vivo* restriction activities assays were employed, too. This method provides information about restriction and modification functions, as well as complex assembly under *in vivo* conditions. A r^- phenotype can be examined by the so-called complementation test, which reveals the ability of mutated HsdR for complex assembly [Sisakova *et al.*, 2008]. Not only studies on HsdR but also *in vivo* studies of mutations in HsdS and HsdM subunits were reported previously [Weiserova *et al.*, 1998].

2. Materials and Methods

2.1. Preparation of competent cells

Competent cells from stock were inoculated in 5 ml of LB media and grown at 37 °C overnight with shaking. Next morning 1 ml overnight culture was inoculated in 100 ml of LB media and grown at 180 RPM shaking until OD_{600} (optical density at 600 nm) was ~0.4. Then cell were cooled on ice and harvested in by centrifugation at 4000 RPM for 15 min at 4 °C. Harvested cells were resuspended in 50 ml of 100 mM $CaCl_2$ (pH=8.0) and kept on ice for 20 min. Cells were harvested again at 4000 RPM for 10 min at 4 °C, resuspended in 1.75 ml of 100 mM $CaCl_2$ with 30 % glycerol solution (pH=8.0) and hold on ice for 30 min. Cells were stored in 50-100 μl aliquots at -70 °C.

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

2.2. Site-directed mutagenesis

Plasmid pTrcR124 carrying *hsdR* gene [Janscak *et al.*, 1996] and oligonucleotide primers containing desired mutations for each mutant were used for site-directed mutagenesis. Sequences of forward primes are given and reverse primer is its reverse complement, mutated codons are underlined and mutated nucleotides are in bold:

5’-CACGCTAAC**GC**TTATGACG-3’ for Arg149Ala;
5'-CCAGATACAT**GCT**TACAGTAAAGAGAG-3' for Arg182Ala;
5’- TGCCAACACGACA**GC**GCGCGATAAAAACAG-3’ for Lys217Ala;
5’- GCCAACACGACAA**C**GCGCGATAAAAACAG-3’ for Lys217Thr;
5’- GCCAACACGACAAAG**GC**CGATAAAAACAG-3’ for Arg218Ala;
5’-CACGACAAAGCG**C**GCTAAAAACAGTTTTG-3’ for Asp219Ala;
5’- CGCGAT**GC**AAACAGTTTTG -3’ for Lys220Ala;
5’- CGCGATA**G**AAACAGTTTTG -3’ for Lys220Arg;
5’- CGCGAT**G**AAAACAGTTTTG -3’ for Lys220Glu;

5’-AACAGTTTT**GC**CTTCACCATGAACTGGG-3’ for Asp224Ala;
5’-GCGTTTTTGACAG**GC**TCAGACACTACTGG-3’ for Ser266Ala;
5’-CCAACCGCATT**GC**TGACGCCACTAAAACC-3’ for Tyr693Ala;
5’-CCAACCGCATTTATG**CC**GCCACTAAAACC-3’ for Asp694Ala;
5'-TTGAGGTC**GCT**TTGCTGAAGTCTCAGG-3' for Asp881Ala;
5'-TTGAGGTC**AAA**TTGCTGAAGTCTCAGG-3' for Asp881Lys;
5’-GAGGTCGATTTGCTG**GC**GTCTCAGG-3’ for Lys884Ala;
5’- GAGGTCGATTTGCTG**GA**GTCTCAGG-3’ for Lys884Glu;
5’- GATATTCGC**GCC**TGGCAGC- 3’ for Asp855Ala;
5’- ACTGGCAG**GCC**CGTGAGAAAG-3’ for Arg858Ala;
5’- GCAGCGC**GCT**GAGAAAGAAG-3’ for Arg859Ala;
5’-GAAGCTGA**GGC**AAAAGAGAAATC-3’ for Lys865Ala;
5’-GCTGAGAAA**GCA**GAGAAATCAACC-3’ for Lys866Ala.

Double mutant Arg218Ala+Lys884Glu was obtained from pTrcR124 [Janscak *et al*., 1996] already containing either single mutation. Double mutants Glu165His+Asp855Ala and Glu165His+Arg858Ala were obtained from pACR124 plasmid [Chang and Cohen, 1978; Zinkevich *et al*., 1997] carrying single mutation Glu165His [Sisakova *et al.*, 2008]. Double mutants Arg858Ala+Arg859Ala and Lys865Ala+Lys866Ala were obtained from pTrcR124 containing single mutation using oligonucleotide primers with double mutation 5’-GCAG**GCC**GC**T**GAGAAAGAAG-3’ and 5’-GAAGCTGAG**GCA**G**C**AGAGAAATC-3’, respectively.

Mutant Phe140Gly…Arg149Gly was prepared by deletion of 8 amino acids, Phe140 and Arg149 were replaced by Glycine creating Glycine’s bridge using following primer (reverse primer is is reverse complement) 5’-CAGATTATCCAGCAG**GGTGGT**TATGACGTAACTATCCTGG-3’.

Polymerase chain reaction was performed in Eppendorf Mastercycler® gradient. 200 μl tube contained 20 μl of PCR reaction mixture: 1.5 U of Expand Long Range polymerase (Roche); 5X Expand Long Range buffer with 12.5 mM $MgCl_2$; 100 ng of plasmid DNA pTrcR124 or pACR124; 0.3 μM of Forward and Reverse primers; 500 μM PCR Nucleotide Mix (dATP, dCTP, dGTP, dTTP at 10 mM each); DMSO 3-12% and H_2O. PCR conditions for pTrcR124 plasmid were following: 2 min - 96 °C 1X; (30 sec - 96 °C; 1 min 30 sec - 55

°C, 8 min - 68 °C) 30X; 7 min - 68 °C 1X. PCR conditions for pACR124 plasmid were following: 2 min – 96 °C 1X; (10 sec – 96 °C; 1 min – 59 °C; 8 min 30 sec – 59 °C) 20X; 7 min – 72 °C (Fig.14 (1 to 3)). After PCR Type II restriction enzyme DpnI was used to remove parental methylated plasmid, 20 µl of PCR mixture were incubated for 1.5-2 hours at 37 °C and then inactivated for 20 min at 80 °C. Then PCR mixture was transformed into Dh5α *E. coli* competent cells using standard transformation protocol and cells were grown overnight at 37 °C on LB/agar plates containing 100 µg ml^{-1} of ampicillin for pTrcR124 plasmid and 50 µg ml^{-1} of chloramphenicol for pACR124 plasmid (Fig.14 (4)). Then colonies formed on LB/agar plates were transferred in LB medium supplemented with ampicillin or chloramphenicol and grown overnight at 37 °C. Next morning cells were harvested and DNA was isolated using Zyppy™ Plasmid Miniprep Kit. All mutants were fully sequenced and DNA with successfully introduced mutation was used for protein production.

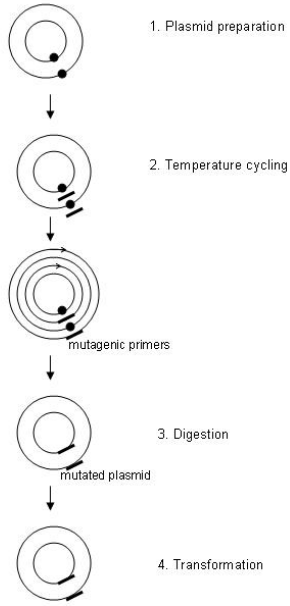


Figure 14. Site-directed mutagenesis (adapted from Stratagene QuikChange XL Site-Directed Mutagenesis Kit Instruction Manual, 2008)

2.3. Expression of HsdR and mutated HsdRs

Plasmid DNA pTrcR124 carrying unmutated or mutated *hsdR* gene was transformed into *E. coli* BL21(DE3)Gold competent cells [Janscak *et al*., 1996] and grown on LB/agar plates containing 100 μg ml^{-1} of ampicillin at 37 °C, or plasmid DNA pACR124 with mutated *hsdR* gene was transformed in *E. coli* JM109(DE3) competent cells and grown on LB/agar plates containing 50 μg ml^{-1} of chloramphenicol at 37 °C. Single colony was transferred in LB medium with corresponding antibiotic to prepare overnight culture. Then overnight culture was diluted 1:100 in 0.5 L of LB in 3 L flask. LB medium supplemented with corresponding antibiotic (100 μg ml^{-1} of ampicillin or 50 μg ml^{-1} of chloramphenicol) and grown with shaking (180 RPM) at 37 °C until OD_{600} was ~0.5–0.6. Protein expression was induced by addition of 1 mM IPTG (isopropyl β-D-thiogalactopyranoside) and the culture was incubated additional 4 hours at 37°C for Arg149Ala, Phe140Gly…Arg149Gly, Arg182Ala, Lys217Ala, Lys217Thr, Arg218Ala, Asp219Ala, Lys220Ala, Lys220Arg, Lys220Glu, Asp224Ala, Ser266Ala, Arg218Ala+Lys884Glu, Glu165His+Asp855Ala and Glu165His+Arg858Ala mutant HsdR subunits at 180 RPM or at 30°C for Tyr693Ala, Asp694Ala, Asp881Ala, Lys884Ala, Lys884Glu, Asp855Ala, Arg858Ala, Arg859Ala, Lys865Ala, Lys866Ala, Arg858Ala+Arg859Ala, Lys865Ala+Lys866Ala. The culture with Asp881Lys mutant HsdR expression was incubated at 15°C overnight after induction with 1 mM IPTG. Then cells were harvested by centrifugation at 4000 RPM for 20 min at 4 °C, washed in STE buffer (pH 8.0) and cell pellet was stored in 50 ml tubes at -20 °C until purification. STE buffer: 0.1 M NaCl; 10 mM This-HCl (pH=8.0); 1 mM EDTA (pH=8.0).

2.4. Purification of WT and mutant HsdR subunits

Cell pellet (obtained from 0.5 L of the culture) was defrosted for 30 min on ice and resuspended in 7 ml of buffer A supplemented with 100 mg ml^{-1} of lysozyme. After 30 min of incubation on ice cells were disrupted using French press (15000 psi). The cell lysate was clarified by ultracentrifugation (30000 RPM for 1.5 hours at 4 °C). Anion-exchange chromatography was used for

HsdR purification. Supernatant was applied on 25 ml DEAE Sepharose Fast Flow (Sigma-Aldrich) column pre-equilibrated with 10 column volumes of buffer A. Bound proteins were eluted with a linear gradient of NaCl (0.05–1 M; 250 ml), WT or mutated HsdRs were eluted as a single peak. Purification was performed on AKTA Purifier. Elution of protein was observed by monitoring of UV absorption at wavelength 280 nm. Collected fractions were analyzed by 10% SDS–PAGE [Maniatis *et al*., 1982]. Fractions contained protein of interest were pooled together and concentrated using 50 kDa cut off spin concentrators (Millipore) and washed 2 times in buffer A. Concentrated protein was stored at -20 °C with 50 % of glycerol. Protein concentration was determined from the absorbance at 280 nm and molar extinction coefficient derived from the amino-acid sequence (98 225 M^{-1} cm^{-1}).

Buffer **A**: 20 mM Tris pH = 8.0; 50mM NaCl; 0.1 mM EDTA; 1 mM DTT

Buffer **B** : 20 mM Tris pH=8.0; 1M NaCl; 0.1 mM EDTA; 1 mM DTT

2.5. Expression of MTase

Plasmids DNA pJS491 (HsdS subunit of EcoR124I) [Patel *et al*., 1992] and pAC15M (HsdM subunit of EcoR124I) [Holubova *et al*., 2004] were used for expression of MTase of EcoR124I in *E. coli* JM109(DE3). 20 ml of LB medium supplemented with 100 μg ml^{-1} of ampicillin and 12.5 μg ml^{-1} of tetracycline was inoculated with 200 μl of *E. coli* JM109(DE3)[pAC15M, pJS491] cell stock and grown overnight. Next morning 0.5 L of 2xYT media supplemented with 100 μg ml^{-1} of ampicillin and 12.5 μg ml^{-1} of tetracycline in 1 L flask was inoculated with 5 ml (1:100) of overnight culture and incubated at 200 RPM at 37 °C until OD_{600} reached ~0.4. Then the culture was induced by adding of 0.5 mM of IPTG and incubated overnight at 37 °C at 200 RPM. Next morning the culture was cooled on ice for 15-30 minutes, cells were harvested and washed with ice-cold STE buffer 2 times and weight of cell paste was determined. Efficiency of induction was analyzed by 10 % SDS-PAGE [Maniatis *et al*., 1982].

2.6. Sample preparation and purification of MTase

1.13 g of cell paste was resuspended in Lysis buffer (buffer L: 50 mM Tris-HCl; pH 8;0; 25 % w/v sucrose; 5 mM EDTA, pH 8.0; 3 mM DTT) (10 ml per 1 g) with Complete protease inhibitor cocktail tablet from Roche). Cells were broken by sonication on ice for 5X 2 min bursts with 1 min cooling in between. Cell lysate was clarified by centrifugation at 18000 RPM for 20 min at 4 °C. $(NH_4)_2SO_4$ precipitation was performed for 4 h at 4 °C and precipitated proteins were harvested by centrifugation at 18000 RPM for 30 min at 4 °C. Protein pellet was resuspended in 15 ml of buffer L and dialyzed against 4 L of buffer L overnight. Dialyzed sample was clarified by ultracentrifugation at 47000 RPM for 1h at +4 °C.

Then supernatant was applied on DEAE Sepharose FF column (20 ml, pre-packed, GE Healthcare) pre-equilibrated with 10 column volumes of buffer L. Bound proteins were eluted with a linear gradient of NaCl (0.05 – 0.5 M; 100 ml) using AKTA Purifier System. Eluted fractions were visualized by 10 % SDS-PAGE. Fractions, contained MTase were pooled together and diluted with buffer Zero to 25 ml final volume to decrease NaCl concentration to 50 mM. Diluted sample was applied directly on 5 ml Heparin column (GE Healthcare) pre-equilibrated with 10 column volumes of buffer L. Bound proteins were eluted with a linear gradient (0.05 – 1 M NaCl, 50 ml). Collected fractions were analyzed by 10% SDS–PAGE [Maniatis ***et al.***, 1982]. Fractions contained protein of interest were pooled together and concentrated using 50 kDa cut off spin concentrators (Millipore). Concentrated protein with 50 % of glycerol was stored at -20 °C. Protein concentration was determined from the absorbance at 280 nm and 320 nm using a molar extinction coefficient (160 400 M^{-1} cm^{-1}):

$$C= (A280 – A320)/ E_{Mtase}$$

Buffer **L** (Low salt solution):
10 mM Tris-HCl, pH 8.0; 50 mM NaCl; 0.1 mM EDTA, pH 8.0; 1 mM DTT;
Buffer **H** (High salt solution):
10 mM Tris-HCl, pH 8.0; 2 M NaCl; 0.1 mM EDTA, pH 8.0; 1 mM DTT;
Buffer **Zero** (Buffer without salt):
10 mM Tris-HCl, pH 8.0; 0.1 mM EDTA, pH 8.0; 1 mM DTT

2.7. Restriction and modification activities assays *in vivo*

Determination of restriction phenotype (r). Complementation analysis was used for determination of restriction phenotype [Jacob and Wollman, 1954] in a plate assay measuring the ability of virulent λ phage [Colson *et al*., 1965] to lysate E. coli. This test is based on the ability of enzymes from the same Type I family to interchange with its HsdR subunits, thus HsdR subunits of two enzymes produced in the same cell compete with each other to assemble with MTase. Plasmid DNA pTrcR124 containing ***hsdR*** gene with desired mutation was transformed in JM109(DE3)[pKF650] and JM109(DE3)[pACM+SR124II] competent cells with chloramphenicol resistance for negative and positive complementation tests, respectively. Then single colonies were isolated and later used for preparation of overnight culture. 1 ml of overnight culture was inoculated in 10 ml LB with appropriate antibiotics and grown at 37 °C until OD_{600} was ~0.6 and then 0.5 ml of culture was transferred in 3 ml of soft agar pre-heated to 45 °C, mixed gently and poured onto the surface of agar plate with appropriate antibiotics. After soft agar solidified phage λ0 [Jacob and Wollman, 1954] (phage stock 10^9-10^{10} PFU/ml was used) was dropped (30 μl) on plates (phage dilutions: 10^6, 10^5, 10^4, 10^3, 10^2), and dried at room temperature [Colson *et al*., 1965]. After overnight incubation at 37 °C number of plagues in drops was scored. The efficiency of plating was determined as a fraction of bacteriophage surviving on tested strain compared to that growing on the non-restricting control strain at the same temperature and at the same concentration of phage. Numbers 0.0001-0.01 correspond to r+ phenotype, 0.01-0.1 to the r± and 0.1-1 to the r– phenotype [Yanisch-Perron *et al*., 1985].

Determination of modification phenotype (m). The phage lysate *λvir* grown on a tested strain was used. A plaque of λ grown on tested strain was stabbed with a sterile straight platinum wire and phage particles were resuspended in 1 ml of phage buffer to give approximately 10^5 PFU/ml. The suspension was then appropriately diluted and the dilutions were spotted onto soft agar overlays of each indicator strain C122 (r^-_{R124I} m^-_{R124I}) and C122/R124I (r^+_{R124I} m^+_{R124I}), and grown overnight at 37 °C. The efficiency of modification of specificity EcoR124I was scored as a ratio of titer of a phage, grown on the strain containing plasmid with a given specificity, and a titer of a phage, grown

on non-modifying strain C122. Numbers 1.0-0.5 correspond to the phenotype m^+, 0.5-0.09 to the $m^\pm$ and below 0.09 correspond to the phenotype m^-.

Phage buffer: 22 mM KH_2PO_4; 49.1 mM Na_2HPO_4; 85.6 mM NaCl, pH adjusted to 7.2 before autoclaving. 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.0001% gelatine was added after autoclaving.

2.8. Restriction activity *in vitro*

One-site circular plasmid DNA pDRM-1R [Taylor *et al.*, 1992] was used as substrate for DNA cleavage *in vitro*. Endonuclease activity was assayed at 37 °C in 40 μl reaction mixture contained buffer R (50 mM Tris-HCl (pH 8.0), 1 mM DTT, 10 mM $MgCl_2$ and 50 mM NaCl), 15 nM of pDRM-1R plasmid, 15 nM MTase and 90 nM HsdR. After 1 min pre-incubation reaction was started by addition of ATP and AdoMetto 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 (30, 60, 90, 120, 300, 600, 1800 sec). Enzyme activity was destroyed by addition of 13 μl of stop solution (3% SDS; 0.15 M EDTA; 10% glycerol; 0.1% bromophenol blue) and heating at 65 °C for 5 minutes. Samples were loaded into 1 % agarose/TAE gel and electrophoresed at 5 V/cm (100 V) for 130 min. The gels were stained in 2 μg/ml ethidium bromide solution during 20 minutes and photographed. The percentage of DNA in each band was evaluated by densitometry (ImageJ software) of the ethidium bromide fluorescence.

Two-site linear DNA pDRM-2R [Janscak *et al.*, 1996] linearized by Type II restriction enzyme ScaI was used as substrate for DNA cleavage *in vitro*. Endonuclease activity was assayed at 37 °C in 40 μl reaction mixture contained buffer R (50 mM Tris-HCl (pH 8.0), 1 mM DTT, 10 mM $MgCl_2$ and 50 mM NaCl), 5 nM of pDRM-2R, 50 nM MTase and 300 nM HsdR. After 1 min pre-incubation reaction was started by addition of ATP and AdoMet 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 (30, 60, 90, 120, 300, 600, 1800 sec). Enzyme activity was destroyed by addition of 13 μl of stop reagent solution (3% SDS; 0.15 M EDTA; 10% glycerol; 0.1% bromophenol blue) and heating at 65°C for 5 minutes. Samples were loaded

into 1 % agarose/TAE gel and electrophoresed at 1,25 V/cm (25 V) for 5 h. The gels were stained in 2 μg/ml ethidium bromide solution during 20 minutes and photographed. The DNA substrate and products were quantified by densitometry (ImageJ software).

2.9. ATPase activity assay

2.9.1. Malachite green method

ATPase activity assay was determined in buffer containing 10 mM $MgCl_2$, 50 mM Tris-HCl (pH 8.0), 50 mM NaCl and 1 mM DDT. 15 nM MTase and 90 nM HsdR subunit and 90 nM pDRM-1R DNA was used. The time course with time points 15, 30, 45, 60, 120, 300, 600 and 900 sec was initiated by addition 4 mM ATP and carried out incubated at 37 °C. The reaction was stopped by addition 20 μl 0.1 M EDTA (pH 8.0). The concentration of inorganic phosphate (Pi) released from 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 and H_2O). The plates were examined after 5 min at 630 nm in the microplate ELISA reader. The concentration of Pi was than determined from a calibration curve derived from solutions of known Pi concentration (KH_2PO_4).

2.9.2. λ-P^{32}-ATP method

Additionally, the λ-P^{32}-ATP method that should be slightly more robust was used in our group. First, cellulose plates were prepared. Positions for sample drops on cellulose plates were marked with a pencil: 1.5 cm from the bottom of the plate and 1 cm in between samples. 50 ml of liquid (running buffer (0.4 M LiCl, 1 M Formic acid) or water) was carefully poured in the glass chamber, avoiding drops on the walls. Cellulose plates were then prerun 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 for 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 (pDRM-1R) and NEB2 reaction buffer. Reaction was initiated by addition of ATP mixture containing 0.16 µCi (0.0013 mM) (5 µl 100 mM ATP + 0.2 µl [γ-P^{32}]ATP. Final concentration of ATP was 2 mM. Reaction was carried out at 37 °C and 4 µl aliquots were withdrawn at 0, 15, 30, 60, 90, 120, 180, 300 and 900 sec. Reaction was stopped transferring aliquots at indicated time points immediately into tubes containing 10 µl of 1% SDS. Sample was loaded on the cellulose plate at the marked positions and then dried at room temperature. Then the plate was run in running buffer up to the top and dried. Dried plate was wrapped with film and exposed for 10 min. Then screen was scanned and data was processed. The hydrolyzed [γ-P^{32}] was separated from [γ-P^{32}]ATP by TLC (stationary phase - celulose, mobile phase – 0,4 M $LiCl_2$ and 1M formic acid) and portion of [γ-P^{32}]/[γ-P32]ATP was visualized by Fujitsu 9000 and evaluated. [Randerath E. and Randerath K, 1964; Marini and Krejci, 2012]. NEB2 buffer: 10 mM Tris-Cl (pH 7.9); 10 mM mg $MgCl_2$; 50 mM NaCl; 1 mM DTT.

2.10. DNA-binding assays

The DNA binding assays was performed and analyzed by our colleagues in the Institute of Microbiology of the ASCR, v.v.i. in Prague. The 30-mer duplex was obtained by annealing equimolar concentration of complementary oligonucleotides and 5´-end-labelled with [γ^{32}P]-ATP using T4 polynucleotide kinase in kinase buffer and incubated for 30 min at 37°C. The unincorporated ATP was removed with QIAquick Nucleotide removal Kit (QIAGEN). DNA binding reactions were performed in volume of 10 µl in a buffer (50 mM Tris (pH 8.0), 25 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, and 10% (v/v) glycerol). 5 nM end-labeled DNA duplex was incubated with 40 nM MTase, and MTase was additionally assembled with WT or selected mutant HsdRs *in vitro*. MTase/HsdR mixtures at a range of molar ratios were incubated at room temperature for 10 minutes and additionally for 10 min after adding duplex DNA. Bound and unbound DNA was separated in 6 % polyacrylamide non-denaturing TAE gels run at 4 °C and 100V. Gels were dried under vacuum for

30 minutes at 80°C (Model 583 Gel dryer, Bio-Rad) and visualized using a Molecular Dynamics PhosphorImager (Model BAS 5000, Fuji).

2.11. DNA translocation assay

DNA translocation assays were performed for WT, Lys220Ala, Lys220Glu and Lys220Arg mutants and analyzed by our colleagues in Institute of Microbiology of the ASCR, v.v.i. in Prague. Translocation analysis was performed by a combination of two previously described triplex oligonucleotide displacement assays [Firman and Szczelkun, 2000; Stanley and Szczelkun; 2006] with some modifications. Plasmid pLKS5 carrying a triplex binding site 2093 bp downstream of the recognition sequence for EcoR124I enzyme and the triplex-forming-oligonucleotide TFO14 (5' TTCTTTTCTTTCTTCTTTCTTT 3') were used [Stanley and Szczelkun; 2006]. Covalently closed pLKS5 DNA and ^{32}P-labelled TFO14 were mixed in equimolar concentration (50 nM) in MM buffer and incubated at 20°C overnight. The resulting triplex (10 nM) was pre-incubated with MTase (100 nM) and HsdR (250 nM) at 20°C in R buffer for 5 min. Reaction was started by addition of ATP to a final concentration of 4 mM. The reaction was stopped after 1 h incubation at 20 °C by addition of GSMB buffer (15% (w/v) glucose, 3% (w/v) SDS, 250 mM MOPS, pH 5.5, 0.4 mg/ml bromophenol blue) and the reaction products were analysed by electrophoresis in 1.5% (w/v) agarose gels (40 mM Tris–acetate, pH 5.5, 5 mM sodium acetate and 1 mM $MgCl_{2)}$ at 10 V/cm for 3 h at 4 °C. Wet gels were exposed on an imaging plate (BAS-IP MS 2025, Fuji film) and visualized in a Molecular Imager FX Pro Plus system (Bio Rad).

Buffer MM: 25 mM MES pH 5.5; 12.5 mM $MgCl_2$

Buffer R: 50 mM Tris–HCl, pH 8.0; 10 mM $MgCl_2$; 1 mM DTT

2.12. Molecular dynamics and QM/MM calculations

Molecular dynamics simulations and QM/MM calculations were carried out within our group by my colleques Dhiraj Sinha and David Reha. YASARA [Krieger *et al.*, 2002; Konagurthu *et al.*, 2006] GROMACS version 4.55 [Berendsen *et al.*, 1995; Van Der Spoel *et al.*, 2005; Pronk *et al.*, 2013] and

AMBER99SB force field [Hornak *et al.*, 2006] were used to prepare the system and perform MD simulation. The standard RESP procedure using Antechamber [Wang *et al.*, 2004] was applied for ATP parameterization.

To calculate long-range electrostatic interactions the particle-mesh Ewald method [Darden *et al.*, 1993] was applied. The LINCS algorithm of fourth-order expansion was used to constrain bond lengths [Hess *et al.*, 1997].

DynDom [Hayward *et al.*, 1997; Hayward and Berendsen, 1998] was used to carry out protein domain motion analysis and principal-components analysis [Amadei *et al.*, 1993] was applied to identify the global motions of the systems.

QM/MM calculations were performed using QSite from the Schrödinger package [Schrödinger LLC, 2011]. The QM energies were calculated using Gaussian03 [Frisch *et al.*, 2004] by the density functional theory method with a B3LYP functional including additive dispersion (treated by DFTD3 [Grimme *et al.*, 2010]), using the 6–31G* basis set. The basis set superposition error was treated by the counterpoise correction method [Boys and Bernardi, 1970]. Contributions to the QM/MM coupling energy from van der Waals interactions were determined in the MM program Impact [Schrödinger LLC, 2011] using the OPLS2005 force field [Jorgensen and Tirado-Rives, 2005]. All molecular structures were analyzed using VMD or YASARA [Humphrey *et al.*, 1996].

3. Results

3.1. Transformation, expression and purification of HsdR subunits

The preparation of plasmid pTrc99 (or pTrcR124) carrying *hsdR* gene [Janscak *et al*., 1996] encoded WT HsdR subunit or with desired mutations was checked on agarose gel electrophoresis (Fig.15**)** and then plasmid DNA was used for transformation experiment.

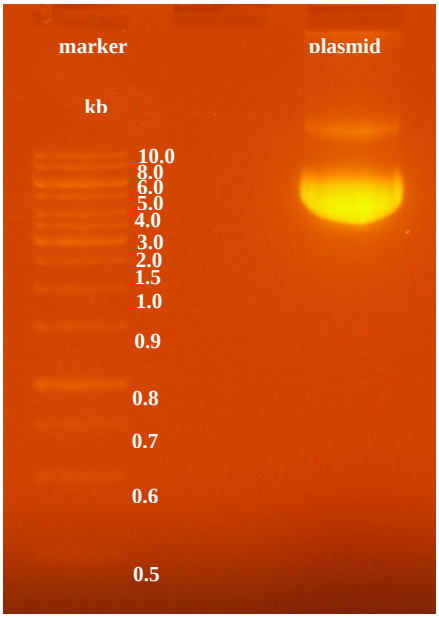


Figure 15. Agarose gel electrophoresis. 5 µl (0.25 µg) of 1kb DNA Ladder and 1 µl of pTrcr124 plasmid preparation were loaded per line.

Usually 10pg to 100ng of pTrcR124 plasmid DNA was mixed with 50 µl of BL21(DE3)Gold *Escherichia coli* competent cells. Transformation was done using a standard protocol [Bergmans *et al*., 1981]; cell suspension was spread on agar plate supplemented with 100 µg ml^{-1} of ampicillin and incubated at +37°C overnight. The next day one single colony was used for preparation of overnight culture: one colony was transferred in 5 ml of LB media supplemented with 100 µg ml^{-1} of ampicillin and grown 15 hours. WT or mutant HsdR subunit was expressed from the P*trs* promoter following Janscak et al., (1996) (Fig. 16).

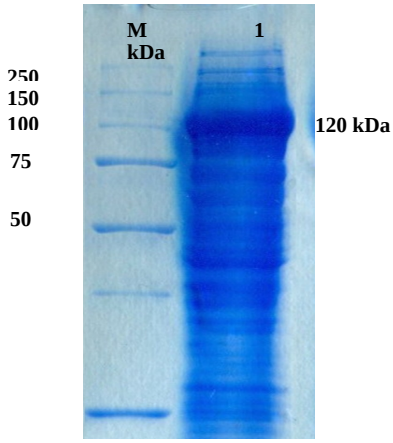


Figure 16. Expression test for WT HsdR of EcoR124I subunit. Line **M** is molecular weight marker; line **1** is expressed WT.

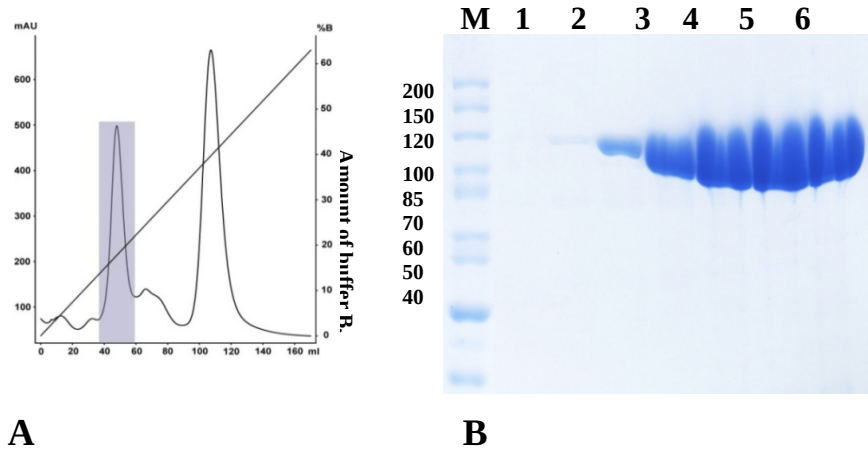


Figure 17. Purification of WT HsdR. A) Elution profile (280nm subscript) of purification of WT HsdR subunit of EcoR124I endonuclease on DEAE Sepharose Fast Flow. System: ÄKTA purifier. B) 10% SDS-PAGE visualization of fractions eluted from DEAE Fast Flow column which correspond to the peak in grey in elution profile. **M** is molecular weight marker with bands size in kDA on the left and **1-7** are collected fractions.

WT HsdR subunit was purified as described previously [Janscak *et al.*, 1996]. Anion-exchange chromatography was used, eluted fractions from the

DEAE Sepharose Fast Flow column were analyzed on 10% SDS-PAGE [Maniatis *et al*., 1982] (Fig. 17). Fractions containing the protein of interest were pooled together and concentrated using 50 kDa cut off spin concentrators (Millipore). Concentrated protein then was stored at -20 °C as 50 % mixture with glycerol. Usually concentration of WT and mutant HsdR subunits varied and was up to 300 μM (Fig. 20; line 2). Wild-type HsdR and mutated HsdRs were used for analysis reported in **Paper 1** (Sinha *et al*., 2014; JMM) and **Paper 2** (Csefalvay *et al.,* 2015; PLoS One).

3.2. Transformation, expression and purification of Methyltransferase

Plasmid pAC15M (with tetracycline resistance) with the ***hsdM*** gene for methyltransferase of EcoR124I derived from the pACYC184 vector was used for expression of the HsdM subunit from the P_{T7g10} promoter [Holubova *et al*., 2004]. Plasmid pJS49 (with ampicillin resistance) carrying the ***hsdS*** gene for HsdS subunit of EcoR124I derived from pTZ19R vector was used for expression of the HsdS subunit from the $T7_{g10}$ promoter [Patel *et al*., 1992]. Both plasmids (Fig. 18) were used for double-transformation as the HsdS subunit of EcoR124I is insoluble alone.

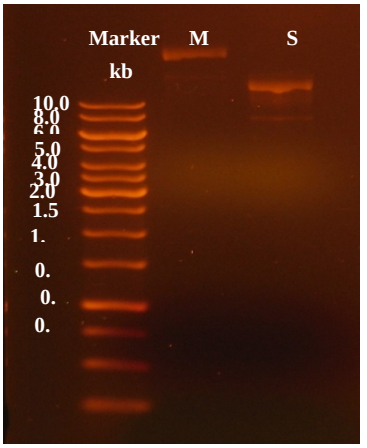


Figure 18. Agarose-gel electrophoresis. Marker is a 1kb DNA Ladder loaded 5 μl (0.25 μg) per line, line **M** is 1 μl of pAC15M plasmid preparation; line **S** is 1 μl of pJS491 plasmid preparation.

10pg to 100ng of each plasmid DNA were mixed with 50 µl of JM109(DE3) *Escherichia coli* competent cells and spread on agar plate supplemented with 12.5 µg ml^{-1} of tetracycline and 100 µg ml^{-1} of ampicillin. MTase was expressed and purified as described in Taylor *et al.*, 1992 (Fig. 19). The purity of the protein after the first purification (DEAE Sepharose Fast Flow column) was more than 90% and more than 95% after second purification (Heparin column). Concentrated protein as 50 % mixture with glycerol was stored at -20 °C. The concentration of methyltransferase usually was 40-96 µM (Fig. 20 line 1). Methylase was used for biochemical assays as described in **Paper 1** (Sinha *et al.*, 2014; JMM) and **Paper 2** (Csefalvay *et al.*, 2015; PLoS One).

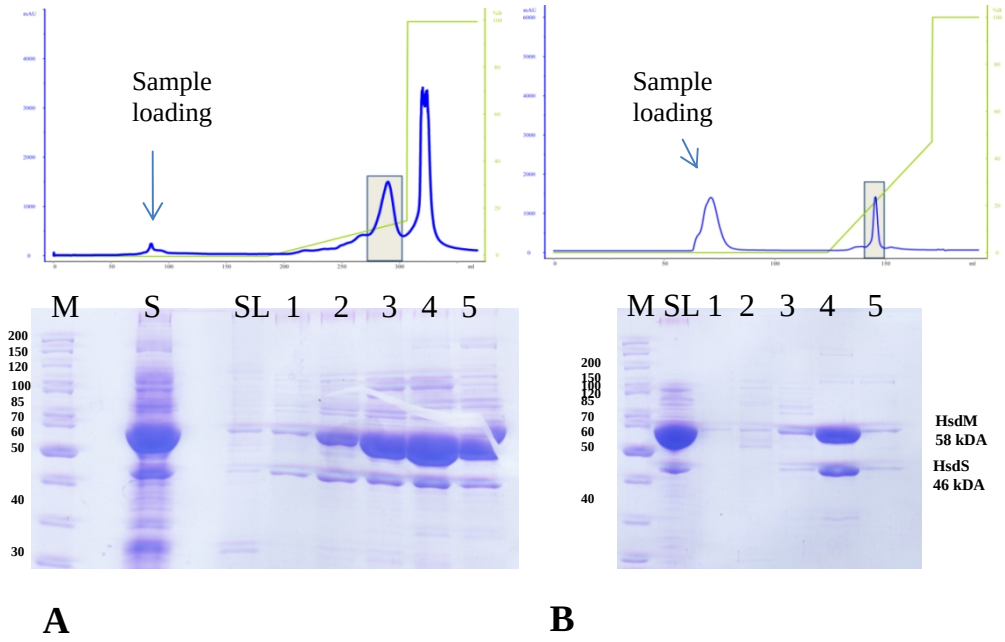


Figure 19. MTase purification. A) Elution profile (280nm subscript) of purification of MTase of EcoR124I on DEAE Sepharose Fast Flow. System: ÄKTA purifier. The peak in grey corresponds to collected fractions on 10% SDS-PAGE below (lines 1-4) containing MTase. Line **M** is molecular weight marker; line **S** is supernatant; line **SL** is sample loading; lines **1-4** are eluted fractions. B) Elution profile (280nm subscript) of purification of MTase of EcoR124I on Heparin column. System: ÄKTA purifier. The peak in grey represents corresponds to collected fractions on 10% SDS-PAGE below (lines 3-4) containing MTase. Line **M** is molecular weight marker; line **SL** is sample loading; lines **1-4** are eluted fractions. Only one fraction (line 4) was used for concentration.

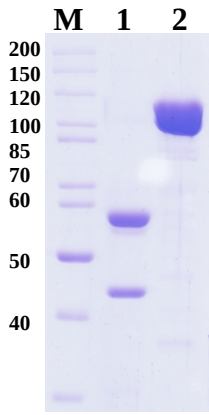


Figure 20. Concentrated enzymes. Line **M** is molecular weight marker; line **1** is MTase, HsdM (58 kDa) and HsdS (46 kDa) subunit are seen; line **2** is WT HsdR.

3.3. Arg182 residue located within QxxxY motif in the endonuclease domain

Type I R-M enzymes display very low sequence identities and even similarities (less than 20% between the different families) and hence structural predictions were merely impossible for HsdR of EcoR124I before its structure was determined. The primary HsdR structure allows only to distinguish the DECHR-variant of a DEAD-box motif and other helicase characteristic motifs clustered in two RecA-like helicase domains [Gorbalenya and Koonin, 1991], as well as three catalytic residues in the endonuclease active site [Obarska-Kosinska *et al*., 2008] and a QxxxY motif characteristic for RecB-like helicases [Sisakova *et al*., 2008b]. Luckily, the crystal structure of HsdR with ATP bound between two helicase domains was reported by Lapkouski *et al*. in 2009 (PDB entry 2W00), however it lacks some residues at the N and C terminal ends and several loops: It starts from 13th residue, four loops in the structure remain unresolved (residues 142-147, 182-189, 585-590, 859-869) as well as large (893-1038) C-terminal part that was however confirmed to be present in the crystal. Surprisingly, ATP bound between helicase domains also contacts endonuclease domain, specifically the Lys220 residue lying ~20 Å from endonuclease active site on a short helical segment not comprising the core fold. Such a contact was not detected for other structures of helicses and translocases suggesting strongly that this contact might potentially be involved in coupling helicase and endonuclease activities. All details of this work are given in the attached **Paper 1** (Sinha *et al*., 2014; JMM).

Segment 182-189, where Arg182 and Tyr183 are from QxxxY motif, remains unresolved in the HsdR crystal structure and however it seems to be positioned close to the endonuclease active site. Preceding the HsdR crystal structure determination, the QxxxY motif was suggested to play an auxiliary role during DNA cleavage [Sisakova *et al*., 2008b]. Thus, the mutational analysis of the Tyr183 residue from motif QxxxY, which is conserved among HsdR homologs [Davies *et al*., 1999], revealed more than a two-fold decrease in nicking of the first DNA strand leading to severely impaired cleavage of the second DNA strand in Tyr183A HsdR mutant, however DNA-dependent

ATPase activity of the mutant was comparable to the WT enzyme assuming that translocation activity was unaffected. In order to aid the interpretation of the role of the QxxxY motif in enzyme function MD simulations were performed in our group and referred here as Sinha *et al.*, 2014. For that, based on the HsdR crystal structure (2W00), a new model of WT HsdR was created completing the above mentioned missing segments by loop modeling in YASARA and borrowing the 182-189 segment form the crystal structure of a specific mutant, Lys220Ala (PDB entry 4BEC) (Fig. 21), where this segment is fully resolved. In our new HsdR model this loop is placed just above the endonuclease active site with the QxxxY motif pointing toward the 220s loop bearing Lys220 which also contacts ATP molecule [Csefalvay *et al.*, 2015].

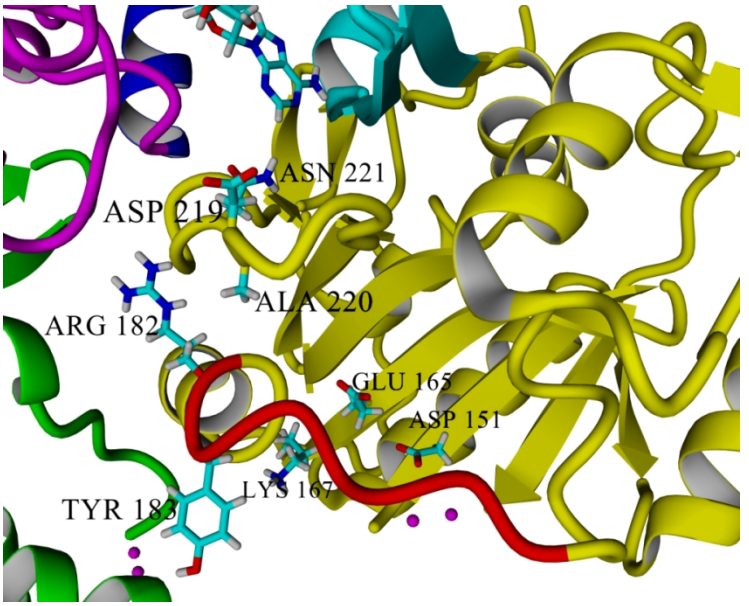


Figure. 21. 180s loop bearing Arg182 and Tyr183 residues from QxxxY motif in the Lys220Ala HsdR crystal structure (PDB id: 4BEC, Csefalvay *et al.*, 2015). 180s loop is lying above endonuclease active site, residue Arg182 is pointing toward 220s loop with Lys220Ala residue. 180s loop is in red, indicated residues side chains are in element color, HsdR domains colors as on figure 4. Figure prepared in YASARA.

While analyzing the Tyr183 behavior in the MD simulations, the focus switched soon to nearby Arg182, as its behavior seemed to correlate with the behavior of the loop bearing Lys220 residue. Thus, detailed analysis was

carried out on Arg182 and virtual mutant Arg182Ala in MD simulations, as well as Arg182Ala mutant was tested *in vitro* and *in vivo*. Hereby, the computational studies were used for an atomistic explanation of the observed experimental results. Thus we were able to suggest that the Arg182 residue plays a key role in communicating the ATP-ligation state through the endonuclease/motor subunit. It was shown that there is an alternative conformation to the Lys220-ATP interaction that is sampled in case this interaction is lost. Lys220 that is not in contact with the ATP moves as well as the whole 220s loop does away from the ATP, and the entire endonuclease domain rotates relative to the other three domains; additionally, after this transition, Lys220 has a persistent interaction with adjacent Asp219 and the 180s and 220s loops interact via Arg182 and Asn221 residues. Contrary, in MD simulations of the Arg182Ala mutant no correlation in 180s and 220s loops movements was observed, the 220s loop gained more degrees of freedom in the absence of the Arg182-Asn221 interaction and the rotation degree of the entire endonuclease domain was less and not stable in time. Thus, a stabilized rotated conformation was observed only in WT simulations but not in Arg182Ala suggesting that the Agr182 residue plays a stabilizing role in the rotational movement of the endonuclease domain required for cleavage function. Indeed the Arg182Ala mutation affects HsdR of EcoR124 functions *in vitro* and *in vivo*.

3.3.1. Determination of the restriction phenotype of Arg182Ala HsdR mutant *in vivo*

Firstly, the effect of introduced Arg182Ala mutation was studied *in vivo* and its restriction phenotype was determined. Plasmid pTrcR124 [Janscak *et al.*, 1996] was used and the restriction phenotype of the Arg182Ala mutant was determined *in vivo* by testing the ability of cells expressing mutated or unmutated HsdR subunits from pTrcR124 plasmid to cleave the growth of unmodified bacteriophage λvir0. Positive and negative complementation tests described previously [Sisakova *et al.*, 2008] were used. For positive complementation test restriction-deficient host cells with expression system only for MTase of EcoR124II were used, *E.coli* JM109(DE3) with pACMS

plasmid, and this test reveals whether tested HsdR is restriction-deficient. In negative complementation tests, the plasmid pKF650 in *E.coli* JM109(DE3) host cells is the source of MTase and HsdR subunit of EcoR124II, so this host is restriction-proficient, noted as r^+ host, and it is able to survive phage infection. When a restriction-deficient HsdR subunit is tested on a r^+ host it is possible to determine whether such HsdR can assemble with the MTase. Restriction-deficient HsdR subunits compete with WT HsdR subunits expressed from pKF650 plasmid for complex formation with the Mtase and reduce thus the ability of the cells to survive phage attacks displaying a so-called trans-dominant effect. Efficiency of plating (e.o.p.) is used to distinguish restriction phenotypes, and it was determined as a fraction of bacteriophage surviving on the tested strain compared to that growing on the non-restricting control strain at the same temperature and at the same concentration of phage lysate. Numbers 0.0001-0.01 correspond to the r^+ phenotype, 0.01-0.1 to the $r^\pm$ and 0.1-1 to the r^- phenotype.

The results of both complementation tests for WT and Arg182Ala mutant are summarized in Table 2. The positive complementation test in the r^- host showed that Arg182Ala mutant fails to complement restriction (value 0.2; within range 0.1-1 for the restriction-deficient phenotype), displays a 100-fold difference in e.o.p. in comparison to WT, and is therefore restriction-deficient. The negative complementation test in the r^+ host revealed that the endonuclease activity *in vivo* for Arg182Ala mutant is 500-fold reduced relative to WT, as well this test confirms that Arg182Ala mutant HsdR subunits preserve the ability to assemble with MTase to form EcoR124I complex. Taking together, the *in vivo* data indicates that Arg182Ala mutation in HsdR of EcoR124I has an effect on the DNA cleavage activity *in vivo* similar to that observed with Tyr183Ala mutation [Sisakova et al., 2008].

Table 2. Complementation test for WT HsdR and mutant Arg182Ala

	restriction [a]			
HsdR	**r^- host** [b]	**r^+ host**[c]	**Ability of cleavage**	**Ability of assembly**
WT	0.002	0.0001	Yes	Yes
Arg182Ala	0.2	0.05	No	Yes, but less than WT

[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^+).

3.3.2. Effect of the Arg182Ala mutation on endonuclease activity *in vitro*

Mutated Arg182Ala HsdR was expressed and purified under the same conditions as unmutated WT HsdR subunit. Endonuclease activity of Arg182Ala HsdR mutant was tested on two substrates, covalently closed circular DNA pDRM-1R with one target site for EcoR124I and linear DNA with two recognition sites for EcoR124I derived from plasmid pDRM-2R [Janscak *et al.*, 1996] by using Type II restriction enzyme ScaI. Both plasmids were isolated from *E. coli* Dh5α competent cells. The DNA concentration was determined spectrophotometrically. The endonuclease activity was assayed using EcoR124I reconstituted from separately purified methyltransferase and either purified WT HsdR or Arg182Ala HsdR mutant enzyme. Following Janscak *et al.*, 1996, the HsdR subunit was in excess in the reaction mixture, the molar ratio of MTase to HsdR subunit was 1:6; DNA:enzyme ratios of 1:1 for circular substrate and 1:10 for linear substrate were used. The assay was carried out as the time course for WT and mutated HsdR. Restriction activity was initiated after an initial incubation time with 4 mM ATP, reaction aliquots were

quenched at indicated time points, DNA substrate and cleavage products were visualized using agarose gel electrophoresis. Agarose gels were scanned under UV illumination and analysed in ImageJ software taking the time point “0” sec as 100% of substrate to be present.

Circular substrate with a single recognition site is believed to be cleaved when translocation of such substrate is no longer possible due to steric hindrances on supercoiled substrate or full length of circular substrate was processed by the complex [Janscak ***et al.***, 1999]. The cleavage of circular substrate results in two specific products: open-circle DNA or nicked DNA is formed after one strand of dsDNA is cut and cleavage of the second DNA strand gives linear product. Results from the endonuclease activity assayed on supercoiled DNA substrate show that WT cleaves ~80% of substrate within 30 min, ~70% of substrate converted into linear product and ~10% is nicked (Fig. 22). In contrast, the EcoR124I complex with Arg182 mutant HsdR displays drastically impaired restriction activity: such complex does not result in a linear product within 30 min, but instead is only able to cleave one strand of dsDNA in 60 sec after reaction initiation and form open circle (nicked) DNA product from fewer than 20% of supercoiled substrate. The result obtained for Arg182Ala mutant enzyme is very similar to that earlier observed for Tyr183Ala mutant enzyme [Sisakova ***et al.***, 2008b] and is in agreement with our data for Arg182Ala mutant from ***in vivo*** complementation tests.

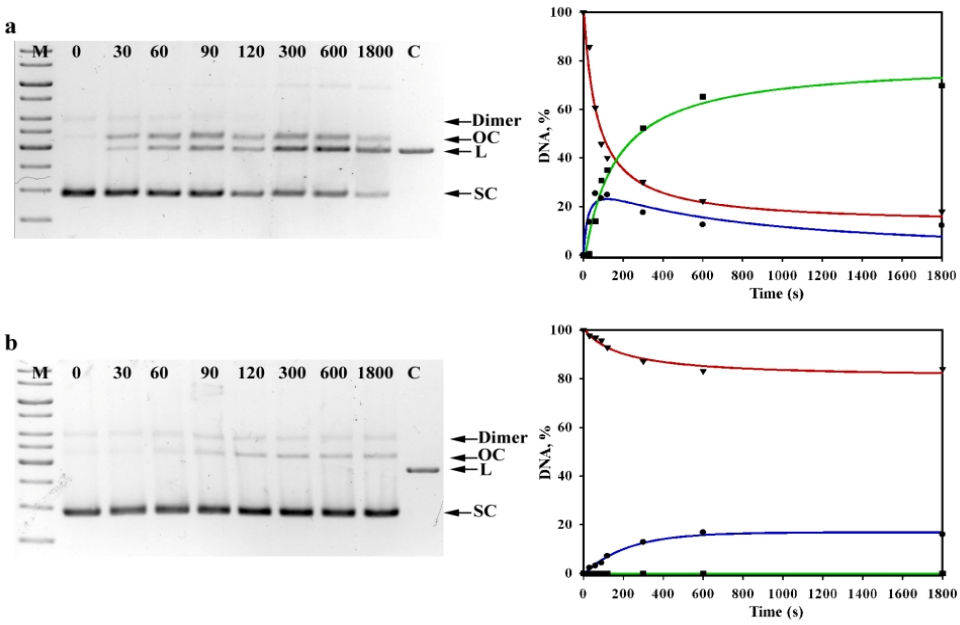


Figure 22. Cleavage of supercoiled DNA. Supercoiled substrate with one EcoR124I recognition site was cleaved by reconstituted enzyme as described in the text. Aliquots were removed at the times indicated (in seconds) and resolved on agarose gels stained with ethidium bromide (left). M – linear DNA molecular weight markers. C – supercoiled plasmid DNA pDRM-1R linearized with Type II restriction enzyme HindIII was used as a marker for the linear product of EcoR124I cleavage. The identity of the DNA species is indicated on the right of each gel. Dimer double-size plasmid present as a trace constituent of the substrate. OC – open circular DNA resulting from the nicking of one strand of the duplex. L – linear DNA resulting from cleavage of both DNA strands. SC – supercoiled substrate DNA. The extent of reaction was quantified (right) by scanning the gels under UV illumination. The indicated DNA species were quantified individually by scanning the gels using ImageJ software [Abramoff *et al*., 2004], taking the time point "0" as 100 %, and best-fit curves were fitted to the data in SigmaPlot 12.0. **a)** EcoR124I reconstituted with unmutated HsdR. OC open circular product (blue, log normal fit). L linear product (green, rectangular hyperbolic fit). SC supercoiled substrate (red, hyperbolic decay fit). **b)** EcoR124I reconstituted with Arg182Ala mutant HsdR. SC supercoiled substrate (red, fitted with a modified single-hyperbolic decay). OC open circular product (blue, fitted with a double-exponential rise to maximum). Reprinted with permission from Springer from **Paper 1** (2014, Fig. 5a-b).

The mechanism of DNA cleavage of linear substrate with two specific sites for EcoR124I differs from circular substrate, the enzyme complex introduces double strand break in random position located between enzyme recognition sequences after the translocating enzyme complex collides with

another Type I enzyme specifically bound on DNA or meets a barrier which makes further translocation impossible [Studier and Bandyopadhyay, 1988]. The cleavage products of such randomly cleaved linear substrate appear on agarose gels as smears. Unmutated WT enzyme is able to degrade ~80% of two-sites linear substrate (Fig. 23), while Arg182Ala mutant enzyme displays no endonuclease activity on such linear substrate leaving 100% of intact linear substrate within 30 min time course. So our data suggests that the Arg182Ala substitution in HsdR subunit makes the EcoR124I enzyme complex even more defective than the substitution of the conserved Tyr to Ala in 183 position reported earlier [Sisakova *et al.*, 2008b].

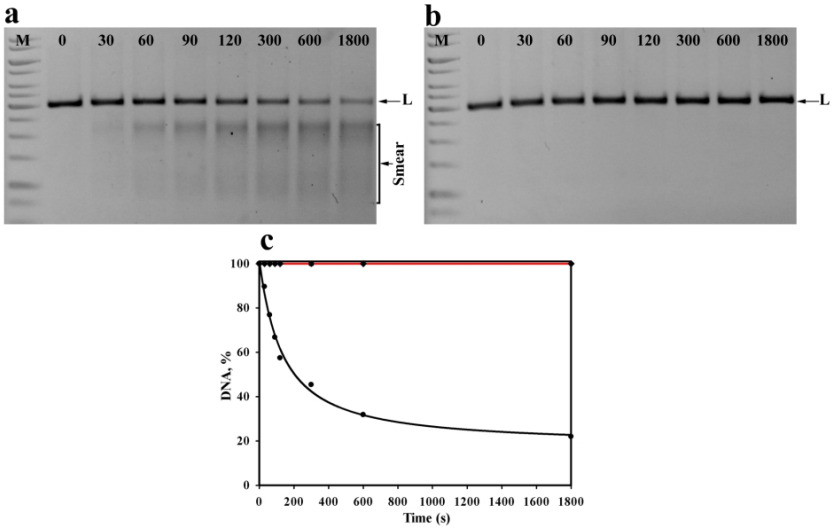


Figure 23. Cleavage of linear DNA. Linear substrate with two EcoR124I recognition sites was cleaved by reconstituted enzyme and analyzed as described in Fig. 22. The parenthesis on the right side of panel **a** marks the region of the smeared products that result from cleavage at random locations between the two recognition sites. **a)** EcoR124I reconstituted with unmutated HsdR. **b)** EcoR124I reconstituted with Arg182Ala mutant HsdR. **c)** Quantitation of the linear substrate DNA for unmutated enzyme (black points, black line) and for Arg182Ala mutant enzyme (black points, red line). The decrease in the amount of DNA substrate was quantified by scanning the gels using ImageJ software [Abramoff *et al.,* 2004], taking the time point "0" as 100 %. A single-exponential curve (black line) was fitted to the data for the unmutated enzyme in SigmaPlot 12.0, taking the zero time point as 100 %. Reprinted with permition from Springer from **Paper 1** (2014, Fig. 6a-c)

3.3.3. Whether Arg182Ala substitution in QxxxY motif affects ATPase activity *in vitro*

The rate of inorganic phosphate accumulation during ATP hydrolysis might serve for indirect evaluation of DNA translocation [Janscak *et al*., 1996]. EcoR124I hydrolyses ~1 ATP molecule per base pair advanced during translocation process [Seidel *et al*., 2008]. Here we used standard malachite green method to determine inorganic phosphate released during motor activity of enzyme [Chan *et al*., 1986]. ATPase activity of WT and Arg182Ala mutant enzyme was measured using reconstituted EcoR124I, as in endonuclease activity assays *in vitro*, and pDRM-1R plasmid DNA with single recognition site for EcoR124I. At least one HsdR subunit is required to be assembled with MTase for DNA-dependent ATPase activity [Janscak *et al.*, 1998]. Control tests showed no ATPase activity in the absence of either HsdR subunits, or cognate DNA or MTase (data not shown) implying that ATPase activity reports on translocation activity and ATP degradation rate was negligible. ATPase activity assay for WT and Arg182Ala mutant enzymes is shown in Fig. 24, three independent experiments of three replicates each were performed over time period of nine months. Here Arg182Ala mutant enzyme displays slightly higher ATPase activity in comparison to WT enzyme, however it can be considered as that observed for WT enzyme taking into account relatively large variability in this manual method due to unavoidable reaction initiation and aliquots quenching time variance. So, unaltered ATPase activity implies that DNA translocation for Arg182Ala mutant enzyme is similar as in WT enzyme and to Tyr183Ala mutant enzyme reported earlier [Sisakova *et al*., 2008b].

3.3.4. The role of Arg182 in the QxxxY motif

Summarising data from biochemical analysis and MD simulations, 220s loop and 180s loop are attributed a role in switching between two conformations with alternative ATP engagement states in HsdR subunit of EcoR124I. So relative rotation of endonuclease domain occurs when both loops interact and after loss of ATP-Lys220 interaction, such rotation might be linked to signalling across endonuclease domain about ATP-ligation status and further

enzyme action. Despite the fact that R182 residue is less conserved than Tyr183, our data suggests this residue to play a key role in the switch as it can form persistent interaction at some distance with its long charged side chain. Analysis of QxxxY sequence across other HsdR subunits reveals either Lys or Arg might be in equivalent to Arg182 position, so the function proposed here for Arg182 residue might be similarly fulfilled in other motor subunits possessing QxxxY endonuclease motif. Concerning Tyr183, which was proposed to engage with DNA [Sisakova *et al.*, 2008b] in RecB-family helicases and HsdRs, the solvent-exposed orientation observed here supports this view. Thus, our results suggest additional function for QxxxY motif via Arg182 residue in communicating a signal across the enzyme about its ATP-ligation status.

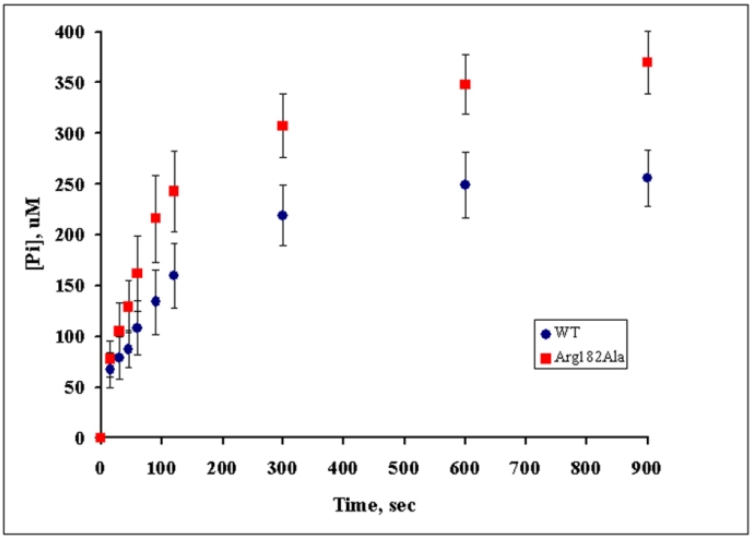


Figure 24. ATPase activity. The concentration of inorganic phosphate released in the DNA-dependent reaction is plotted as a function of time for enzyme reconstituted with unmutated (blue) or Arg182Ala mutant (red) HsdR. The error bars represent one standard deviation calculated from three independent replicates, each with triplicate determinations. Reprinted with permission from Springer from **Paper 1** (2014, Fig. 7).

3.4. Arg182 and Asp855 residues in Glu165His mutant HsdR

HsdR subunit represents a fusion of domains with distinct functions activated only within EcoR124I R-M complex. The whole structural/functional organization of EcoR124I R-M enzyme is primarily aimed to stop phage infection of the cell, thus motor subunit activities such as ATPase and translocation which are preceding endonuclease activity serve for the type I R-M complex to cleave foreign DNA. However, the fact that cleavage occurs far from the recognition site and large pieces of foreign DNA are obtained as a results hints to the major role in horizontal gene transfer. The endonuclease domain of the HsdR subunit partially resembles a RecB-like nuclease from the PD-(E/D)xK superfamily and contains motifs I, II, III and QxxxY characteristic for RecB-like nucleases, which as well were found in other repair and recombination nucleases [Sisakova *et al*., 2008, 2008b]. Residues Asp151, Glu165 and Lys167 forming endonuclease active site are conserved among type I R-M enzymes and play a key role in dsDNA cleavage in HsdR of EcoR124I. Mutational analysis of those residues showed not only the expected loss of endonuclease activity in mutant enzymes, but also a profound effect on the motor activity of EcoR124I. Thus, substitution of residues in the endonuclease domain of the HsdR subunit can have multiple effects on the distant helicase domains affecting duplex translocation suggesting that domains of the same polypeptide have a close interrelationship [SIsakova *et al*., 2008]. Unfortunately biochemical characterization of those mutant enzymes tells nothing about structural implications which led to the impaired enzyme behavior. All three residues from the endonuclease active site are present in the crystal structure of HsdR subunit (pdb id: 2W00) [Lapkouski *et al*., 2009] while the 182-189 segment lying above those residues is missing, however this segment was later resolved in the Lys220Ala mutant HsdR [pdb id: 4BEC] [Csefalvay *et al*., 2015]. In this work it was shown already that the 180s loop can have two alternative conformations and residue Arg182 from the QxxxY motif does either contact the 220s loop or does it contact residue Asp881 [Sinha *et al*., 2014]. In the crystal structure of the Glu165His (pdb id: not available yet) mutant HsdR subunit the 180s loop is trapped in a new conformation where Arg182 residue contacts Asp855 residue from the helical domain, forming a

salt-bridge (Fig. 25). This crystal structure was used to make an attempt to explain previously reported biochemical analysis of endonuclease active site mutants in terms of structural dynamics. In MD simulations both, Arg182 and Asp855 keep a stable distance, even in long simulations. However, this distance is not due to a direct contact between them, but the 182 guanidine group makes aromatic stacking with Arg858, and this interaction is very persistent throughout the simulation. MD simulations of a *in silico* chimera of WT with 180s loop conformation as in the crystal structure Glu165His show that the 182-855 contact is not persistent in this construct but is quickly lost after a few nanoseconds, and the loop gets back to conformations that was observed in the WT simulations [Sinha *et al*., 2014], means oscillating between contacting loop 220 and residue Asp881. Thus it seems that the new conformation of this loop is a feature exclusively of the mutant enzyme that traps the loop in this state. To test whether salt-bridge interaction between Arg182 and Asp855 in the Glu165His mutant is the cause of reduction in ATPase/translocation activity reported earlier [Sisakova *et al*., 2008], we introduced a second mutation Asp855Ala in Glu165His mutant to nullify the possibility of a salt-bridge, as well we additionally substituted Arg858 with Alanine in Glu165His mutant to remove the stacking interaction stabilizing the Arg182-Asp855 salt bridge. Also single mutations Asp855Ala and Arg858Ala were introduced. EcoR124I R-M enzyme reconstituted with MTase and separately purified mutant HsdR subunits were tested *in vitro* and *in vivo*.

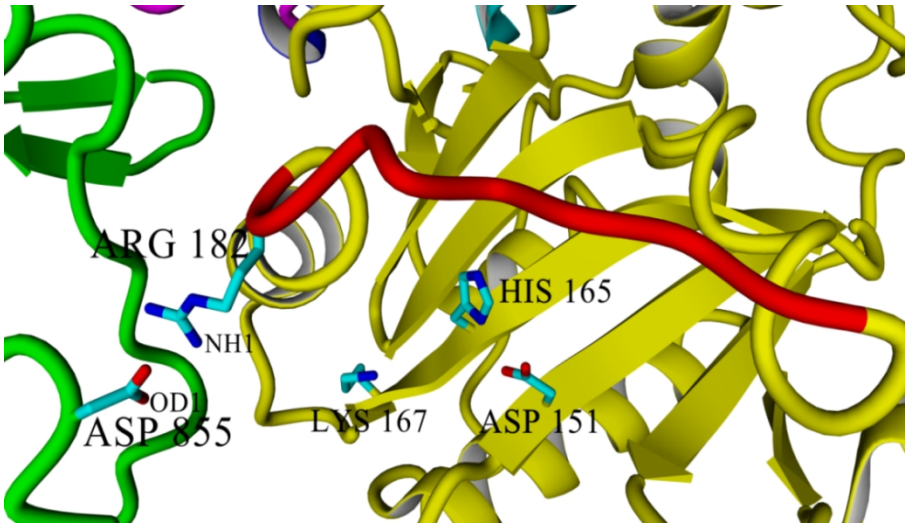


Figure 25. Arg182 residue from 180s loop in the crystal structure of Glu165His mutant HsdR (pdb id: not available yet). The distance of 3.169 Å between NH1 of Arg182 and OD1 of Asp855 allows formation of salt bridge. Side chains of Asp855 residue and Arg182 residue from motif QxxxY in 180s loop as well as residues from the endonuclease active site are shown. 180s loop is in red, residue's side chains are in element colors, the endonuclease domain is in yellow and the helical domain is in green. Figure prepared in YASARA.

3.4.1. Mutant enzymes and *in vivo* restriction phenotype

All mutant HsdR subunits were tested *in vivo* as was described previously in 1954 by Jakob and Wollman. Not surprisingly, both double mutants Glu165His+Asp855Ala and Glu165His+Arg858Ala showed the same results as single mutant Glu165His reported by Sisakova in 2008, they are able to compete for binding to MTase complex but cells with these mutant HsdR subunits are not able to survive phage infection. Interestingly, Asp855Ala mutation resulted in r^- phenotype (value 0.278 within the range of 0.1-1) if tested on restriction-deficient host in positive complementation test (Tab. 3). Competition for binding to methyltransferase of Asp855Ala mutant HsdRs with WT HsdRs in the negative complementation test revealed $r^{\pm}$ phenotype (with value 0.014 within the range 0.01-0.1 but close to r^- range), indicating successful competition and thus the ability to assemble with the MTase. Both tests indicate that Asp855Ala mutant HsdR is restriction-deficient with retained

ability for MTase binding. The Arg858Ala single mutant was able to restore restriction in the r^- host but only partially, host cells expressing this mutant HsdRs displayed a $r^{\pm}$ phenotype (value of 0.028 within the range 0.01-0.1) in the positive complementation test.

Table 3. Complementation tests of the WT HsdR and mutant HsdRs

	restriction [a]		**Ability of cleavage**	**Ability of assembly**
HsdR	**r^- host** [b]	**r^+ host**[c]		
WT	0.003533	0.000852	Yes	Yes
Asp855Ala	0.278413	0.014609	No	Yes
Arg858Ala	0.028258	0.007843	Yes	Yes
Glu165His+Asp855Ala	0.220625	0.179583	No	Yes
Glu165His+Arg858Ala	0.300694	0.284742	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^+).

Thus, *in vivo* experiments demonstrate that both double mutant of HsdRs lose their endonuclease activity due to the Glu165His mutation, as expected, however bind well to methylase competing successfully with wild-type HsdRs. In the single mutants Asp855Ala and Arg858Ala, restriction here might be affected as a consequence of these residues being involved in translocation or restriction, since their binding to MTase is not affected; the Arg858Ala mutation seems to affect restriction *in vivo* less than Asp855Ala. This excludes the possibility that misfolding of the R-subunit leading to unsuccessful complex assembly is the cause of the deficiency. Initially, the fact that both, Asp855Ala and Arg858Ala mutant HsdRs for *in vitro* assays were expressed from pTrcR124 vector at lowered temperature (30 °C) to keep

proteins in soluble fraction and double mutants were normally expressed at 37 °C either from pTrcR124 or pACMS vectors leads to assumption that Asr855Ala and Arg858Ala mutations might cause changes in the helical domain packing making such motor subunits inaccessible for MTase binding and competition with WT HsdRs at the conditions of induced *in vivo* experiment.

3.4.2. Cleavage of supercoiled DNA substrate *in vitro*

Endonuclease activity of enzymes reconstituted with both single mutant HsdR subunits, either Asp855Ala or Arg858Ala, assayed *in vitro* on single-site circular DNA substrate, are similar to that observed for WT enzyme (Fig. 26). Induced expression of both mutant subunits was done at 30 °C, while standard 37 °C caused formation of inclusion bodies. Restriction activity *in vitro* ion a first glance opposes the *in vivo* experiments, as it seems that Asp855Ala and Arg858Ala mutations do not affect endonuclease activity in EcoR124I, they preserve the same activity as it is in WT.

However, the *in vivo* situation is much more complex than the in vitro situation, and the *in vivo* system might be more sensitive to subtle changes than the *in vitro* system. New MD simulations for Asp855Ala and Arg858Ala mutant HsdRs as well as double mutants are needed to better understand the subtle behavior of single mutants at the helical-endonuclease domains interface in terms of structural dynamics, as the explanation doesn't seem as simple as just a loss of the salt-bridge, that would, if sampled regularly during normal enzyme action, lead to alterations in the *in vitro* tests, too. However, the fact that the new conformation is rarely sampled in the MD simulations of the WT enzyme and is trapped in the Glu165His mutant, might be an indication that this conformation is only sampled (and important) under very specific circumstances that occur *in vivo*, but not *in vitro*.

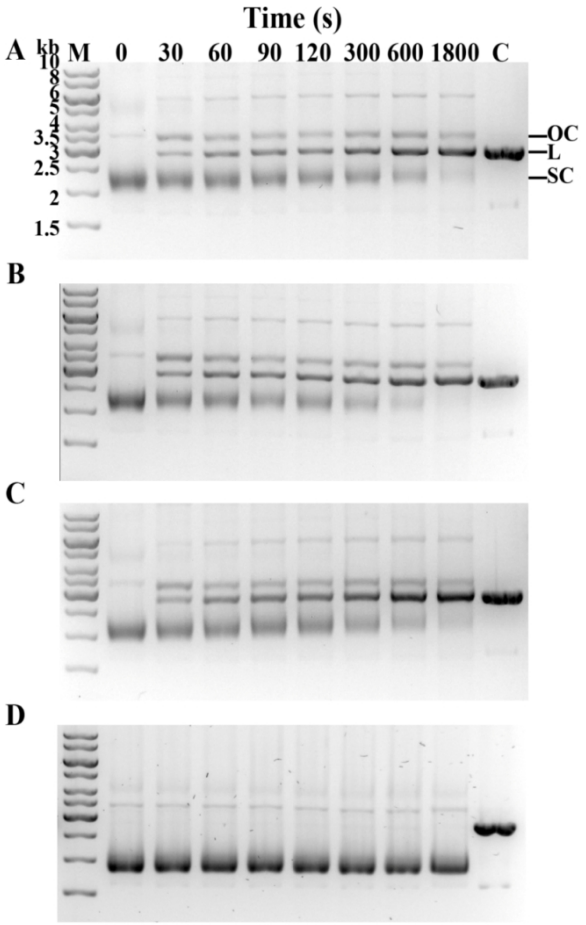


Fig. 26. Cleavage of supercoiled substrate. Supercoiled DNA substrate pDRM-1R carrying single EcoR124I recognition site is treated with EcoR124I R-M complex *in vitro*. EcoR124I was reconstituted from $HsdS_1HsdM_2$ methyltransferase and A) WT HsdR or mutant HsdRs B) Asp855Ala, C) Arg858Ala, D) Glu165His+Asp855Ala. 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 "0" (before initiation reaction with ATP) is to contain 100 % of supercoiled substrate. On panel A open circle and linear products are denoted as OC and L, respectively, and supercoiled substrate as SC; control (linearized plasmid DNA pDRM-1R after cleavage by HindIII restriction enzyme) is denoted as C; DNA molecular weight marker (M) with marked band size in kb on panel A is the same on all panels.

3.4.3. The effect of mutants on the ATPase activity *in vitro*

ATPase activity was tested on circular DNA substrates with a single target site in the presence of 4 mM ATP as described earlier. Glu165His, Asp855Ala and Glu165His+Asp855Ala mutant enzymes displayed ATPase activity slightly higher than that for WT enzyme. Both Arg858Ala and Glu165His+Arg858Ala mutant enzymes showed ATPase rate similar to WT enzyme. Considering the relatively large variability in this manual method much of which is attributed to unavoidable time variance in initiating the reactions results from ATPase assay for the first group of mutants are not substantially different from the second group.

Taking together all experimental data and computational analysis of Glu165His-Asp855 salt-bridge interaction, it can be concluded that Asp855Ala and Arg858Ala mutations preserve WT-like activities *in vitro* and translocation and complex assembly *in vivo*, too, but are restriction deficient *in vivo*.

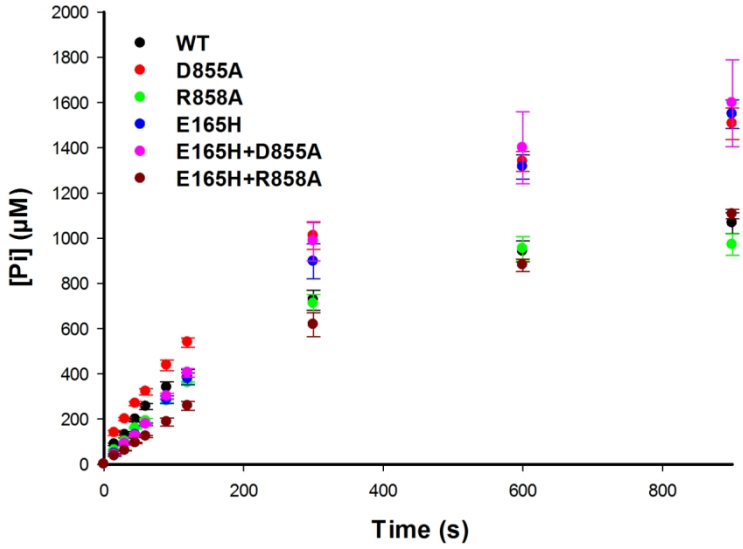


Figure 27. ATPase activity. The amount of inorganic phosphate released during DNA-dependent ATP hydrolysis was monitored using standard malachite green method. Amount of inorganic phosphate is plotted as a function of time. Standard deviation is calculated from three independent replicates and *error bars* are plotted in both directions.

3.5. Lys220 residue as a potential key residue in functional coupling of duplex translocation and DNA cleavage

The crystal structure of WT HsdR containing an ATP molecule bound between the two helicase domains determined by Lapkouski *et al*., (2009) was found unexpectedly also to contact Lys220 residue from the endonuclease domain (Fig. 28). The Lys–ATP contact suggests a potential means for coupling helicase and endonuclease activities during enzyme activation. The crystal structure of Lys220Ala HsdR mutant (pdb id: 4BEC) [Csefalvay *et al*., 2015] containing resolved 180s loop segment allowed a detailed structural analysis of the broader Lys residue region as well as monitoring the behavior of the QxxxY motif lying directly above the endonuclease active site using MD simulations. As reported earlier in this work, MD simulations showed that there is correlation in 220s and 180s loops motions after Lys–ATP contact is lost what results in establishment of persistent contact between them via Arg182 residue [Sinha *et al*., 2014]. So, global motion movement observed in MD simulations of WT HsdR after Lys–ATP contact was lost together with analysis of Arg182Ala mutant HsdR and restrained WT HsdR simulations allowed to hypothesize that signal transmission between the endonuclease and helicase domains of the motor subunit occurs by means of interdomain engagement via ATP. Thus, a role of Lys220 residue was probed employing crystal structures of three mutants of HsdR of EcoR124I as well as *in vivo* and *in vitro* assays. In the crystal structure of a HsdR mutant with Arg substitution of Lys220 hydrogen bonding of 220 residue to ATP was preserved as well, while Ala and Glu mutants in 220 position lacked this interaction.

Specially designed primers were used for site-directed mutagenesis (Methods) to replace positively charged Lys with: Alanine with neutral charge; Arginine with positive charge and Glutamic acid with negative charge. Three mutants were expressed and purified at the same conditions as WT HsdR [Janscak *et al*., 1996]. Endonuclease activity of WT HsdR and mutants HsdRs was assayed *in vitro* on circular DNA substrate with single target sequence for EcoR124I (for more details see endonuclease activity *in vitro* previously described Arg182Ala mutant HsdR). The results of *in vitro* restriction activity of WT and HsdR mutant enzymes led to the general conclusion that all mutated

HsdR subunits are capable of assembling with MTase and form the EcoR124I complex and all mutant enzymes can cleave circular DNA substrate, however to a different extent and with altered kinetics (Fig. 29).

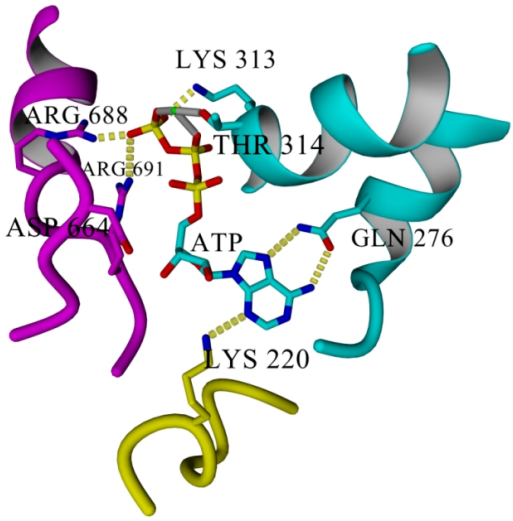


Figure 28. ATP contacts in the WT HsdR crystal structure (PDB id: 2W00, Lapkouski *et al.*, 2009). Side chains of residues from the helicase 1 and 2 domains and Lys220 residue from the endonuclease domain, which form H-bonds to ATP or positioned close enough to permit bonding interactions, are shown. Domain segments (ribbons) and selected residues (stick models) are color-coded as was described previously, N- and O-atoms of residue side chains in blue and in red, respectively, and with Mg ion shown as a green sphere. Figure prepared in YASARA.

A more detailed analysis of agarose gels allows to distinguish between the endonuclease activity of mutant enzymes and a quantification analysis was carried out (Fig. 30). Intensities of DNA bands were individually quantified using ImageJ software 1.45s. The time point, where remaining supercoiled substrate concentration equals to the linearized DNA product concentration, is denoted hereafter as crossover time and is used here as an inverse measure of enzyme activity. Thus, for 6 molar excess in HsdR subunit over methyltransferase used to yield maximum enzyme activity resulted in ~50 seconds crossover time for WT EcoR124I; additional HsdR does not increase activity. At the same reaction conditions nearly WT activity with a crossover time of ~40-50 sec was observed for Lys220Arg mutant enzyme. While Lys220Ala and Lys220Glu mutant enzymes displayed approximately half and one-third of WT activity with crossover times of ~90-100 and ~150 sec, respectively.

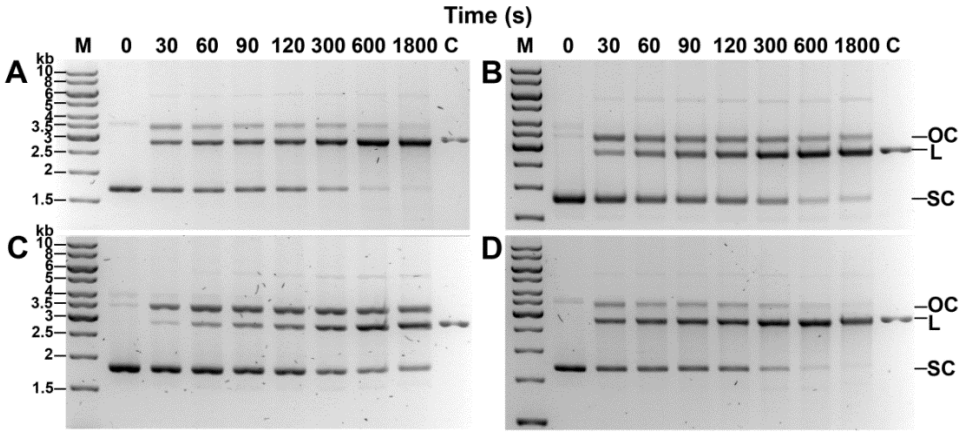


Figure 29. Cleavage of circular DNA. Circular plasmid DNA bearing one EcoR124I recognition site was reacted with enzymes reconstituted from HsdS1-HsdM2 methylase and WT (A), Lys220Ala (B), Lys220Glu (C) and Lys220Arg (D). Reactions stopped at the indicated time points were applied to 1.2% agarose gels and visualized by ethidium bromide staining. M is the marker of the indicated numbers of base pairs shown at the left side of panel A and C. C is the linearized plasmid DNA as a control; OC is open circular product; L is linear product; and SC is supercoiled substrate indicated at the right side of panel B and D. (Csefalvay E, Lapkouski M, Guzanova A, Csefalvay L, Baikova T, Shevelev I, et al. (2015) Functional Coupling of Duplex Translocation to DNA Cleavage in a Type I Restriction Enzyme. PLoS ONE 10(6): e0128700. doi:10.1371/journal.pone.0128700).

Then bands densities plotted for linearized DNA product were fitted with an exponential rise to maximum (function f = a*(1-exp(-λ*x)) using SigmaPlot Version 12.0), and rate constants λ were derived from those fits. The rate constant λ for WT was 0.0122 s^{-1} and almost the same for Lys220Arg mutant enzyme with value of 0.0127 s^{-1}, while Lys220Ala and Lys220Glu displayed 0.0081 s^{-1} and 0.0048 s^{-1} rate constant values. Rate constants and the crossover times were used for ranking endonuclease activity of mutant enzymes. These rate constant values combining with crossover time show that *in vitro* endonuclease activity of Lys220Arg mutant enzyme is preserved as for WT, while the lowest enzyme activity was observed for Lys220Glu mutant enzyme. Thus, accordingly rate constant and crossover times values tested enzymes can be ordered WT≈Lys220Arg>Lys220Ala>Lys220Glu in their *in vitro* endonuclease activity.

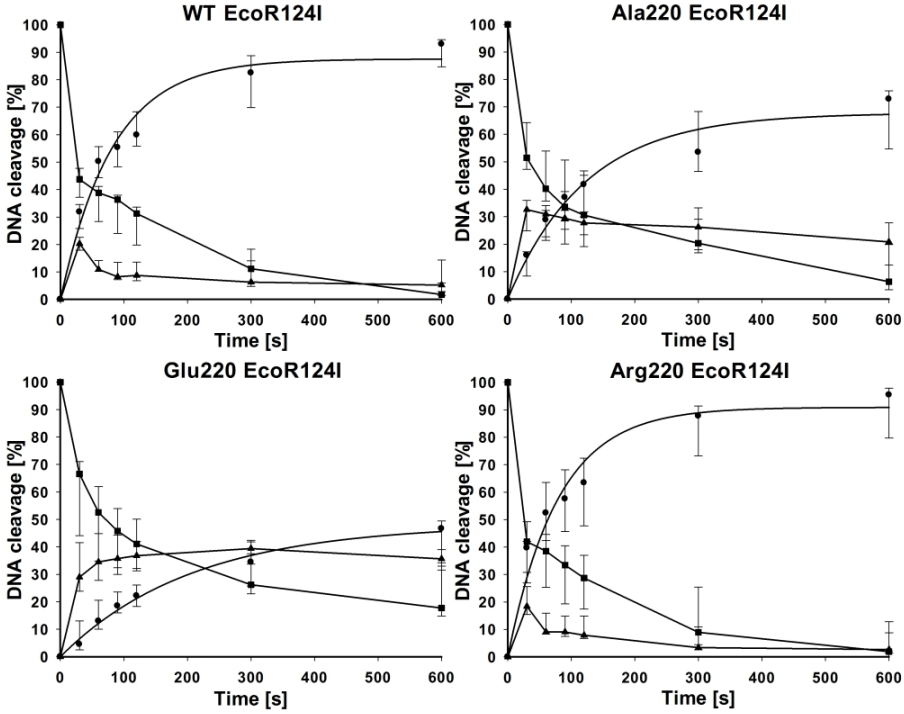


Figure 30. Quantification analysis of WT and mutants. OC, open circular product (▲); L, linear product (●); SC, supercoiled substrate (■). The three indicated DNA species were quantified individually. Plots for the increase of linear DNA product were derived by fitting an exponential rise to maximum function in SigmaPlot 12.0. The points are given for quantification of the gels (Fig.29), and standard deviations are given from the mean of six repetitions. (Csefalvay E, Lapkouski M, Guzanova A, Csefalvay L, Baikova T, Shevelev I, et al. (2015) Functional Coupling of Duplex Translocation to DNA Cleavage in a Type I Restriction Enzyme. PLoS ONE 10(6): e0128700. doi:10.1371/journal.pone.0128700).

Summarizing all functional tests, no defects in complex assembly or DNA binding was revealed for all enzymes. As well ATPase activity assay and DNA translocation assay show that all mutants translocate duplex DNA and hydrolyze ATP. All enzymes are able to cleave circular DNA with single target site for EcoR124I. Results from *in vitro* endonuclease activity are consistent with data obtained from *in vivo* experiments, where WT and Lys220Arg are r^+ restriction phenotype, Lys220Ala is ~10% of r^+ and Lys220Glu is r^-. Thus, mutational analysis of Lys220 residue, however it is situated ~20 Å away from the endonuclease active site, results in specifically impaired only nuclease activity. All details are described in **Paper 2** (Csefalvay *et al.*, 2015; PloS One).

3.6. Interactions of residues from 220-loop with other domains

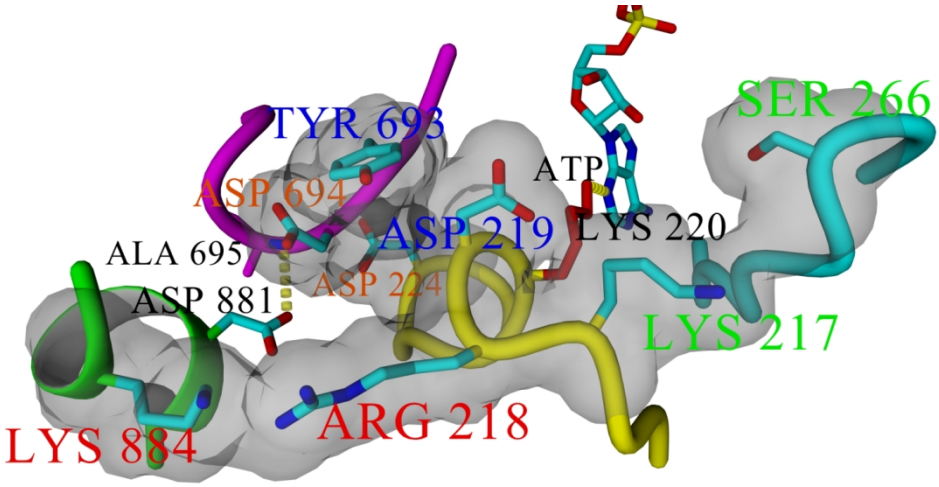


Figure 31. Residues from 220-loop interacting with other domains in the HsdR crystal structure (PDB id: 2W00, Lapkouski *et al.*, 2009). Gray clouds represent van der Waals interactions of residues and interacting residue pairs are labeled with the same colors. Domain fragments in ribbon representation are colored as was described previously. Figure prepared in YASARA.

EcoR124I is a member of SF2 helicases and has a collection of characteristic helicases motifs clustered in two RecA-like helicase domains [Gorbalenya and Koonin, 1993]. This could imply for HsdR of EcoR124I a similar mechanism of DNA translocation as in related double-strand translocases from SF2 [Durr *et al.*, 2005]. But still there is no clear answer where dsDNA passes over the square-planar array of functionally integrated domains once it leaves the helicase cleft and somehow has to contact the endonuclease active site prior to restriction. Currently two models propose different paths for DNA and each of them has its right to exist since we lack detailed information related to this issue. The model coming from our group and proposed in Lapkouski *et al.*, 2009 does not explain the restriction deficiency of the $R_1M_2S_1$ complex while the model proposed by Kennaway *et al.*, 2012 provides a logical explanation of this feature but does not reflect the location of DNA after it leaves the helicase cleft. However $R_2M_2S_1$ complex restriction-proficiency might have another still undiscovered mechanism. Here

we examine amino acid residues at the extent between the ATP binding pocket and the endonuclease active site, particularly in the 220-loop, whether they could be involved in ATPase/translocation activity of EcoR124I complex through interactions with dsDNA or forming interdomain communication network. Four residues with different charges situated around Lys220 (Fig. 31): two amino acids with positive charge (Lys217 and Arg218) and two with negative charge (Asp219 and Asp224) were replaced with neutral Alanine. In the crystal structure of HsdR those four residues have partners situated in other domains. Lys217 has partner in the helicase 1 domain – Ser266, distance between side chain of Lys217 and backbone oxygen of Ser266 is ~3.2 Å that theoretically might be sufficient for hydrogen bonding in the dynamic structure. Both Arg218 and Lys884 from the helical domain are turned with their side chains towards the groove where dsDNA was proposed to pass, and the close position to the endonuclease active site predicts they might be involved in stabilizing the DNA substrate as it comes from the helicase cleft. Both Asp in positions 219 and 224 provide van der Waals interactions to the helicase 2 domain via Tyr693 and Asp694, respectively. Each above mentioned residue was substituted individually with Alanine, and Lys884 was replaced by Glu. Removal of any residues Asp219, Asp224, Tyr693 or Asp694 from negatively charged network bellow DNA path across motor subunit would bring new local structural arrangement decreasing repulsion interactions with DNA and stabilizing it on the track. If the DNA is bent as it leaves the helicase cleft and goes across domains of motor subunit near the endonuclease active site then amino acid residues located in the interdomain groove could be involved in the ATPase/translocation process. Prominent positively charged residues Arg218 and Lys884 which are close to each other in the crystal structure of HsdR would be apart in the presence of DNA and would attract negatively charged substrate to the surface of the motor subunit. The absence of either of them would lead to a at least minor destabilization of the directional movement of DNA during translocation. This is still ongoing work and the results will be used for future publication.

3.6.1. Site-directed mutagenesis, expression and purification

All mentioned above residues were substituted with neutral Alanine, Lys884 with Glutamic acid and Lys217 with Threonine. Interestingly, all mutant HsdRs with substitution in the endonuclease domain were well expressed at the same conditions as WT [Janscak *et al.*, 1996]. Mutations introduced in other domains required lower expression temperature 30 °C (Fig. 32). Lower temperature would suggest that mutations cause local structural changes impacting protein packing what results in aggregation of protein at 37 °C during induced expression. Then mutant HsdR subunits were purified.

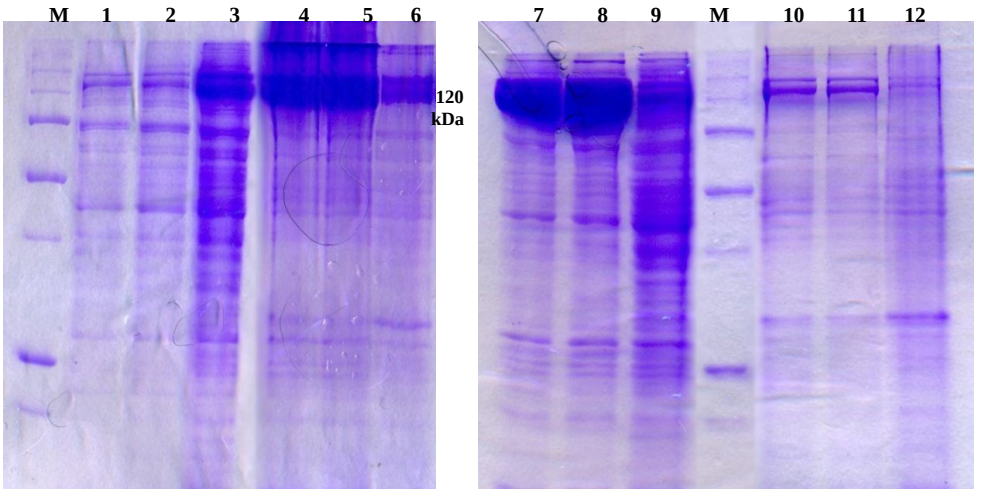


Figure 32. Expression test for **Tyr693Ala** mutated HsdR. **M** is molecular weight marker; **1)** soluble fraction, OD = 0.4; 37 °C; 1 mM IPTG; **2)** soluble fraction, OD = 0.6; 37 °C; 1 mM IPTG; **3)** soluble fraction, 37 °C; no IPTG; **4)** insoluble fraction, OD = 0.4; 37 °C; 1 mM IPTG; **5)** insoluble fraction, OD = 0.6; 37 °C; 1 mM IPTG; **6)** insoluble fraction, 37 °C; no IPTG); **7)** soluble fraction, OD = 0.4; 30 °C; 1 mM IPTG; **8)** soluble fraction, OD = 0.6; 30 °C; 1 mM IPTG; **9)** soluble fraction, 30 °C; no IPTG; **10)** insoluble fraction, OD = 0.4; 30 °C; 1 mM IPTG; **11)** insoluble fraction, OD = 0.6; 30 °C; 1 mM IPTG; **12)** insoluble fraction, 30 °C; no IPTG.

3.6.2. Effect of introduces mutations on endonuclease activity *in vitro*

All purified mutant HsdR subunits together with WT MTase were reconstituted and tested *in vitro* as was described earlier. Quite interesting effect

for endonuclease activity was observed when Asp224 and Tyr693 were substituted by neutral Alanine. The Asp224 mutant enzyme showed a bit better endonuclease activity than the WT enzyme, rapidly accumulating OC and then converting this intermediate product to more than 80 % of linear product within 30 min time course. Tyr693 mutation resulted in slower accumulation and degradation of OC product reaching ~60 % of linear product in the end of the time course comparing to ~70 % of WT yield. It seems that Asp224Ala mutation situated in the next turn of α-helix might influence motional activity of entire 220-loop altering its movement or removal of negative charge close to ATP binding pocket might alter ATP uptake by motor subunit and results in increased ATPase/translocation activity consequently cleaving DNA substrate better. Concerning the Tyr693 substitution to Alanine, substitution of such large side chain might result in a local rearrangement due to the loss of its potential partner for H-bonding resulting in impaired endonuclease activity *in vitro*. Thus, standard 37 °C temperature leads to aggregation of Tyr693Ala mutant motor subunit during protein expression.

Mutagenesis of Arg218 to Alanine in the interdomain groove does not alter endonuclease activity significantly as its partner Lys884 does changed by Glutamic acid or when both mutations are present. Interestingly, single Lys884Glu mutant enzymes degrades ~50 % of supercoiled substrate, while double mutant Arg218Ala+Lys884Glu cleaves single-site circular DNA substrate even better degrading ~60 % of substrate indicating that Lys884Glu substitution has major impact on DNA cleavage in double mutant. Located distantly from ATP binding pocket and being prominent on the surface of the motor subunit these two residues, at least Lys884 residue, potentially could direct dsDNA substrate from the helicase cleft to the endonuclease active site.

Substitution of Asp in positions 219 and 694 and Ser266 for Alanine or Lys217 for Ala/Thr displays the same restriction activity as WT HsdR.

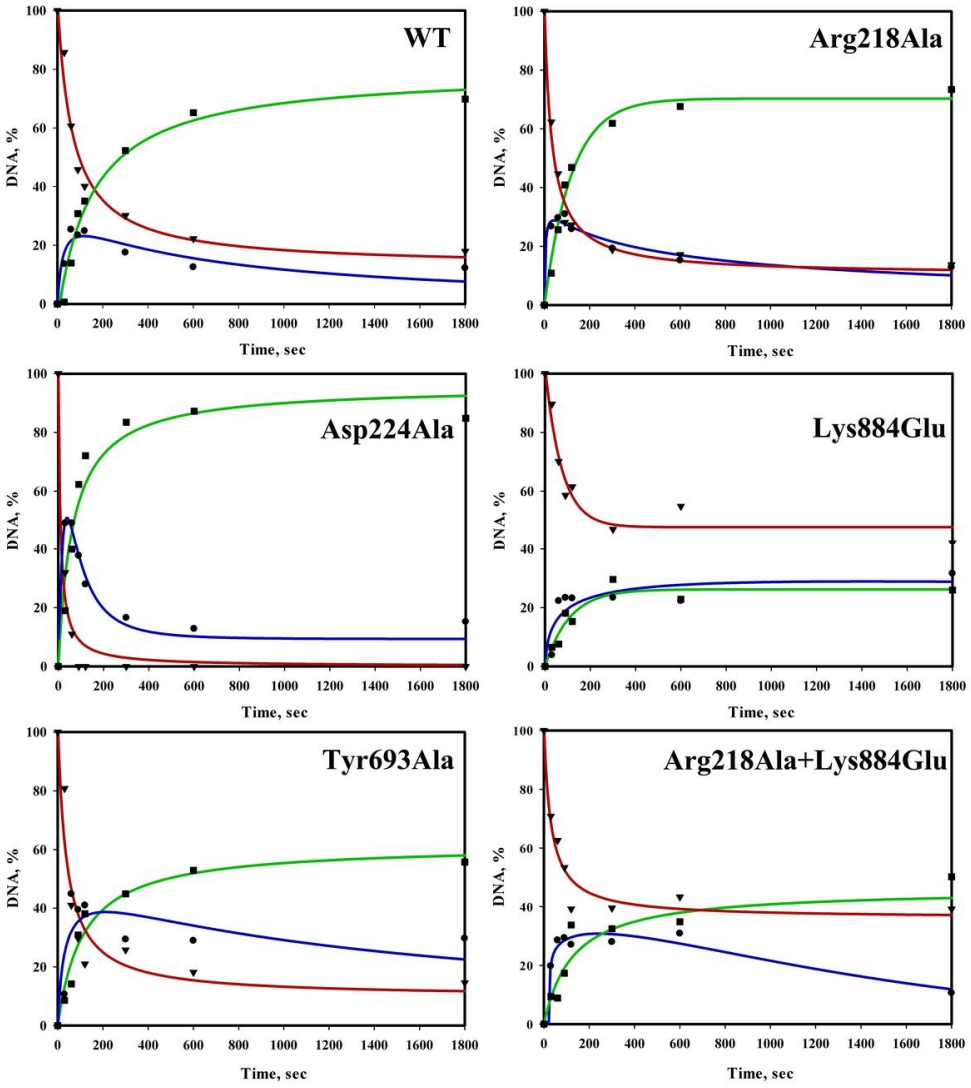


Figure 33. Quantification analysis of single-site DNA cleavage by mutant enzymes with mutations introduced around Lys220 residue. The indicated DNA species were quantified individually by scanning the gels (not shown) using ImageJ software [Abramoff *et al*., 2004], taking the time point "0" as 100 %, and best-fit curves were fitted to the data in SigmaPlot 12.0 (Supercoiled DNA substrate – red and ▼, open circle product – blue and ●, litnear product – green and ■).

3.6.3. Hydrolysis of ATP *in vitro*

In vitro restriction activity as the secondary activity of the holoenzyme EcoR124I after ATPase and translocation [Janscak *et al.*, 1999] represents only ability of the complex to cleave substrate while restriction deficiency in HsdR subunit could have different causes not associated with negative influences of introduced mutations on catalytic residues in the endonuclease active site. To investigate whether the above mentioned point mutations lead to affected restriction through impaired DNA translocation ATPase activity assays were employed. DNA advancement step in 1 bp length by HsdR motor subunit is coupled to consumption of 1 ATP molecule [Seidel *et al.*, 2008] and, thus the rate of accumulation of inorganic phosphate released during ATP hydrolysis could indirectly indicate weather translocation activity is altered. Malachite green method [Chan *et al.*, 1986] was used to compare ATP hydrolysis rates of HsdR mutants and WT.

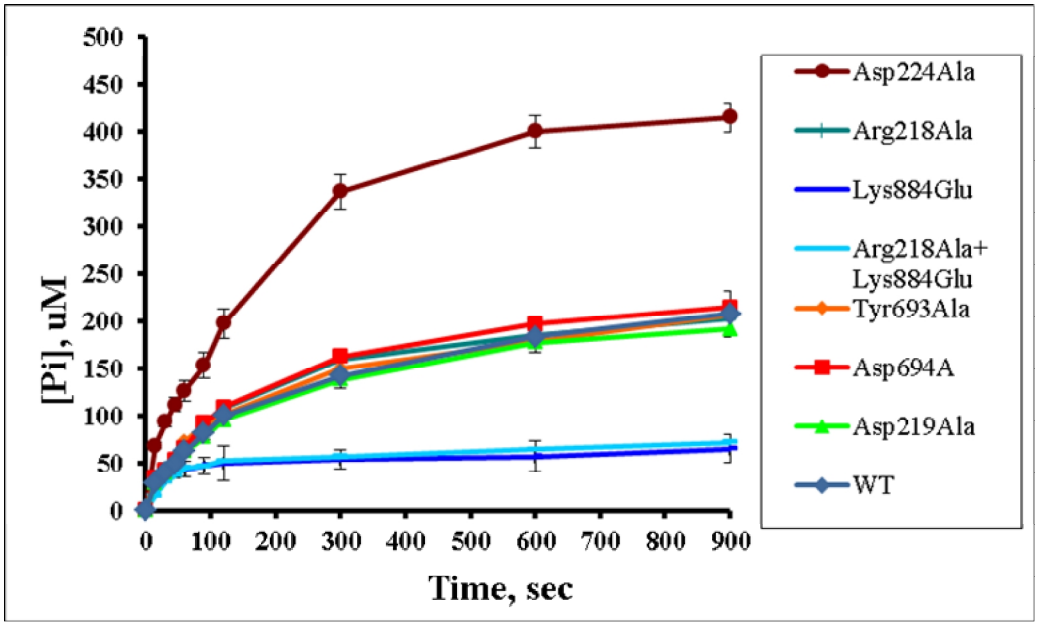


Figure 34. Influence of mutations introduced around Lys220 on ATPase activity EcoR124I R-M complex *in vitro*. For better visualization data points are connected with colored lines and error bars are plotted only for WT, Asp224Ala and Lys884Glu.

Asp224Ala is the most interesting mutant from the group located near the ATP binding pocket, it displays ~2-folds higher ATPase rate in contrast to WT, while Asp694Ala, D219Ala and Tyr693Ala behave as WT (Fig. 34). The high rate of translocation and ATP hydrolysis [Seidel *et al.*, 2008] by the motor subunit implies fast entry of ATP into the ATP binding pocket with consequent fast release of products of its hydrolysis what raises the question how ATP gets in there. If negatively charged DNA goes from the helicase cleft to the endonuclease active site across motor subunit covering a passage which could be the potential entrance for ATP to its binding site from the front then removal of negatively charged Asp224 behind would not lead to any effects. But it seems that 2-folds higher ATPase activity of Asp224Ala mutant might suggest better capturing of negatively charged ATP from opposite site of HsdR. Other mutants from this group due to their location and weak interactions between them have no such effect.

Groove mutations Lys884Glu and double Arg218Ala+Lys884Glu result in about ~3-folds lower ATPase rate in comparison to WT, while no change is detected if Arg218Ala is present alone. Translocation of DNA implies keeping of the double strand close to the motor subunit through interactions with residues located on the surface. Thus many residues might be involved in such stabilization of DNA and abolition of one-two residues would not be sufficient enough to knock out ATPase activity, otherwise the absence of any particular residue could be compensated by a neighboring one. Lys884Ala, Lys217Ala/Thr and Ser266Ala mutations do not alter the ATPase activity (not shown).

3.6.4. *In vivo* restriction activity

HsdR mutants which displayed distinct from WT behavior *in vitro* were characterized *in vivo*. *In vivo* experiments display that *in vitro* reduced endonuclease activity and ATPase activity for the Arg218+Lys884Glu mutant enzyme where the secondary cleavage event is present but less efficient is still sufficient to protect host cell from virus infection. Asp224Ala with a slightly better endonuclease activity *in vitro* and higher ATPase activity *in vitro* in

comparison to WT enzyme shows a better ability to resist phage infection *in vivo* (Tab.4).

Table 4. Complementation tests of the WT HsdR and mutants

	restriction [a]		**Ability of cleavage**	**Ability of assembly**
HsdR	**r^- host** [b]	**r^+ host** [c]		
WT	0.002	0.0001	Yes	Yes
Arg218Ala+Lys884Glu	0.003	0.0003	Yes	Yes
Asp224Ala	0.0008	0.0004	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^+).

3.7. How a D881 residue substitution influences EcoR124I activities

Asp881 residue from the helical domain attracts special interest; this residue faces both the helicase 2 domain and the endonuclease domain close to 180- and 220-loop. The side chain of this residue interacts with the helicase 2 domain via the backbone amino group of Ala695 residue (Fig. 31) in the HsdR crystal structure (2W00) and in the dynamical HsdR structural model we demonstrate that it can interact with the endonuclease domain via the side chain of Arg182 (Fig. 35). MD simulation predicted that the Arg182 residue alternates between Asp881 residue and 220-loop in WT HsdR [Sinha *et al*., 2014]. So, we mutated Asp881 residue to neutral Alanine and positive Lys to prevent H-bonding and studied enzyme activities *in vivo* and *in vitro*.

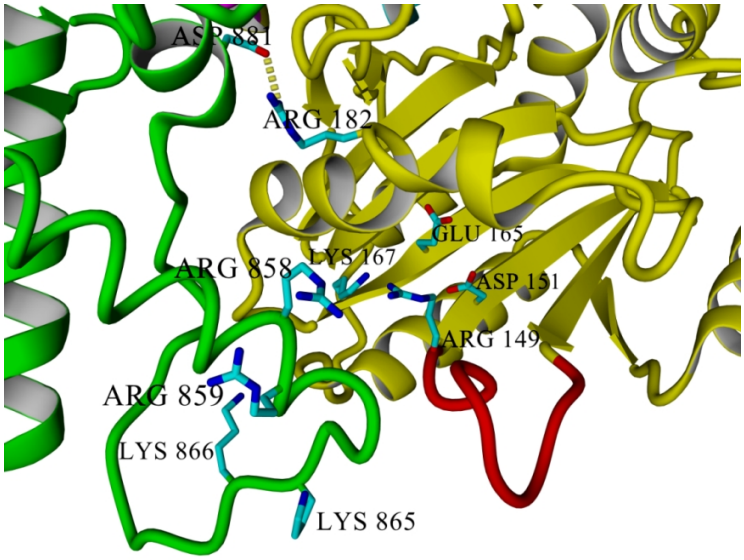


Figure 35. Modeled segments in the WT HsdR of EcoR124I. Residues 142-147, 182-189 and 859-869 unresolved in the HsdR crystal structure were modeled in YASARA. Residues from modeled regions Arg (858, 859) and Lys (865, 866) in the helical domain and Arg149 residue from the endonuclease domain were substituted with Ala, also 140-149 segment (in red) was deleted introducing Gly in 140 and 149 positions. Asp881 residue side chain contacting R182 residue side chain here in modeled structure or backbone amino group of Ala 695 in 2W00 structure forming H-bond in both cases (Fig. 31). The helical and the endonuclease domains are in green and yellow, respectively. Side chains of residues are in element color.

Asp881Ala mutant HsdR was expressed at 30 °C, otherwise the protein appeared in pellet fraction at standard 37 °C temperature. Surprisingly, Lys substitution at 881 position required even lower expression temperature and longer incubation time, Asp881Lys mutant HsdR was expressed at 15 °C overnight. Observed effects for both Asp881 mutants might indicate that this centrally located residue is essential for protein packing, any change either loss of potential H-bonding or introducing of larger residue side chain might bring to improper protein folding which in the living cell would result in elimination of such HsdR subunits. Quite high expression level induced with IPTG could cause this solubility problem, however expression of WT HsdR from the same plasmid at standard conditions was not causing any problems for protein solubility.

Both mutated HsdRs were purified and analyzed *in vitro* and *in vivo.* Restriction activity *in vitro* was assayed using pDRM-1R circular DNA substrate, Ala and Lys mutants were able fully degrade circular substrate. Despite disruption of the hydrogen bond between Asp881 and Ala695 as in the HsdR crystal structure these mutant enzymes preserve WT-like endonuclease activity *in vitro*.

A drastic increase in ATPase activity ~2.5- and ~5.3-fold in comparison to WT was detected for Asp881Ala and Asp881Lys mutant enzymes (Fig. 36), respectively. The absence of negatively charged Asp881 residue might have a positive impact on Lys884 located in the next turn of the same helix which would face dsDNA during translocation and attract duplex to the endonuclease active site. Additionally molecular dynamic simulation predicts formation of hydrogen bond between Asp881Lys mutant and Asp694 residue from the helicase 2 domain which could introduce more stability in rotational motion of the helicase 2 domain (how other helicases do [Durr *et al.*, 2005; Lewis *et at.*, 2008]) during the translocation cycle.

Thus Asp881Ala mutant was questioned whether it preserves this increased *in vitro* endonuclease activity also in living cells. At first sight it seems that the absence of Asp881-Ala695 hydrogen bond leads to crucial conformational changes in the motor subunit which would be incompatible with cell survival during phage attack at 37 °C due to inaccessibility of insoluble motor subunit, but data from *in vivo* studies suggests the opposite. This mutant

HsdR binds well to methyltransferase of EcoR124II and displays a more efficient resistance to phage attack in conditions of *in vivo* experiment carried out at 37 °C (Table 5). Taking into account standard endonuclease and higher ATPase activities *in vitro* of Asp881Ala mutant enzyme it could be concluded that the introduced mutation (but only at high level of induced expression) makes the protein inaccessible for protein purification at standard expression conditions. However, this appears not to happen in living cells where the expression level is significantly lower. Indeed, the increased enzymatic activities of Asp881Ala mutant HsdR emphasize the important role of Asp881 residue; it could improve survival of *E. coli* strains during intensive phage attack.

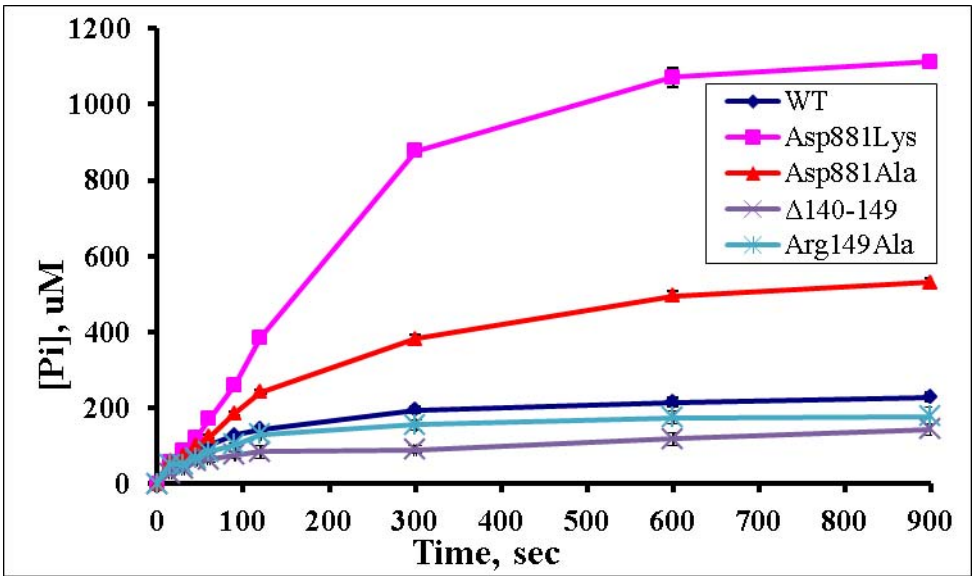


Figure 36. ATPase activity of EcoR124I mutant enzymes *in vitro*. For better visualization data points are connected with colored lines. Change of charge from negative to positive in 881 position leads to drastic increase in ATPase activity. Point mutation Arg149Ala has insignificant effect on ATPase activity of EcoR124I *in vitro*, however deletion of 140-149 segment close to the endonuclease active site resulted in slightly lower ATPase activity than Arg149Ala substitution.

Table 5. Complementation tests of the WT HsdR and mutants

	restriction [a]		**Ability of cleavage**	**Ability of assembly**
HsdR	**r^- host** [b]	**r^+ host**[c]		
WT	0.002	0.0001	Yes	Yes
Asp881Ala	0.00008	0.0007	Yes, superactive	Yes
Arg149Ala	0.003	0.0004	Yes	Yes
Phe140Gly…Arg149Gly	0.02	0.02	Yes, but less than others	Yes, but less than others

[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^+).

Surprisingly, the substitution of Asp881 with positively charged Lys leading to extremely high ATPase activity *in vitro* had a different effect in *in vivo* experiments. Despite the fact that the same plasmid was used in both complementation tests it was not possible to grow any overnight culture suitable for plating in soft agar layer. Thus, this example of a "synthetic" Asp881Lys mutant HsdR with drastically increased ATPase activity *in vitro* reminds once more that conditions within the living cell are more complex and many factors influence not only protein packing but its enzymatic activity as well.

3.8. Arg149 mutant and 140-149 loop deletion

Additionally, our interest in the endonuclease domain was focused on the Arg149 residue situated close to the Asp151 catalytic residue (Fig. 35). As the catalytic residues in the endonuclease active site are required for successful dsDNA cleavage [Sisakova *et al.,* 2008], it implies that DNA substrate is passing above those residues and long positively charged side chain of Arg149 residue might be involved in attracting of dsDNA toward the active site residue prior/or within the cleavage event. So, Arg149 was replaced by an Alanine and analyzed the same way as previous mutant HsdR subunits. Also we prepared mutant HsdR with loop deletion, Phe140 and Arg149 were substituted by Glycines and 8 amino acids between them were removed. Unfortunately in the 2W00 crystal structure residues 142-147 are unresolved. Relying on the modeled 140-149 loop in the WT structure and close location to the endonuclease active site residues, its direct involvement in dsDNA cleavage can be predicted.

Both Arg149Ala and a 140-149 deletion mutant HsdRs were purified and tested *in vivo* and *in vitro* as was described earlier for other HsdRs. Proteins were obtained under the same conditions as WT [Janscak *et al.*, 1996]. ATPase and endonuclease activities of both mutant enzymes were studied using single-site circular DNA pDRM-1R, two-site linear DNA substrate pDRM-2R additionally was used for cleavage assays. Interestingly, substitution of Arg149Ala neighboring to catalytic Asp151 residue leads to accumulation of more than ~55 % of nicked product within 5 min after initiation of reaction and about 45 % of this intermediate product is left after 30 min. The only substantial linear product is visible starting from 5 min in the time course, or at least after 2 min as no time points were taken in this time interval; Arg149Ala mutant enzymes accumulates about 30 % of linear product within 30 min time course (Fig. 37). As for two-site linear DNA, about 60 % of this substrate was cleaved by Arg149Ala mutant enzyme, only after 90 sec of reaction incubation visible cleavage products appear as a smear on the agarose gel. So, such restriction behavior could suggest that this positively charged residue might be involved in interaction with DNA substrate keeping it on the right track near

catalytic residues and/or substitution of Arg149 residue adjacent to the e-β-sheet leads to local structural changes affecting functionality of Asp151 catalytic residue where the alanine mutation knocks out endonuclease activity [Sisakova *et al*., 2008]. Concerning the more serious structural change as 140-149 segment deletion in the endonuclease domain, it is not surprisingly that such mutant enzyme is not able to cleave any of both tested DNA substrates (not shown).

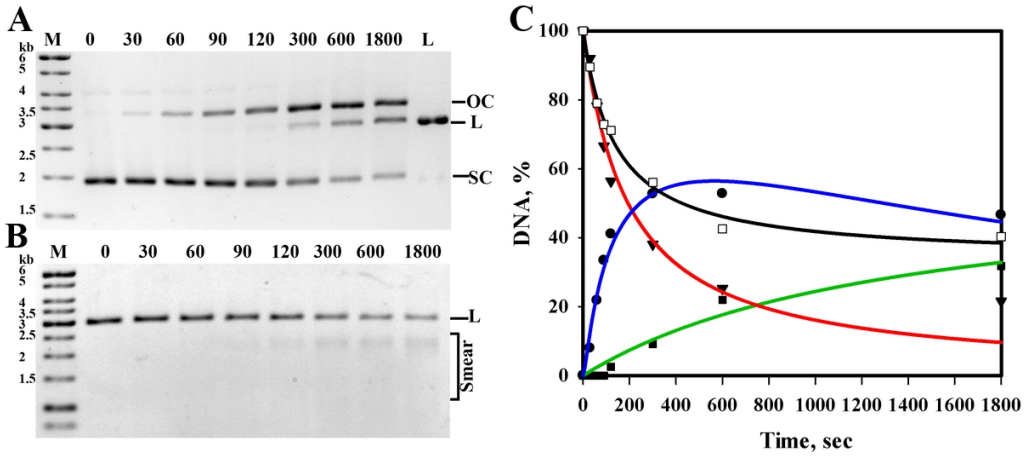


Figure 37. Cleavage of supercoiled and linear DNA by Arg149Ala mutant enzyme. Single-site supercoiled DNA substrate (**A**) and linear DNA substrate (**B**) with two EcoR124I recognition sites were cleaved by reconstituted enzyme and analyzed as described in Fig. 22. On panel **A** OC open circular product, L linear product, SC supercoiled substrate. The parenthesis on the right side of panel **B** marks the region of the smeared products that result from cleavage at random locations between the two recognition sites. **C)** Quantitation of the gel on panel **A:** OC open circular product (blue, ●, log normal fit). L linear product (green, ■, rectangular hyperbolic fit), SC supercoiled substrate (red, ▼, hyperbolic decay fit); and of the gel on panel **B**: linear DNA substrate (black , □, hyperbolic decay fit).

A slight decrease in ATPase activity was detected in both cases, the 140-149 segment deletion affects ATP hydrolysis more than Arg149Ala point mutation but still close to WT considering variations in this manual method. Nevertheless, it seems that in both cases some interactions with the DNA might be lost and this might facilitate the detachment of duplex from the protein surface in that region of the endonuclease domain altering ATPase/DNA translocation activities.

However, the Arg149Ala mutant HsdR with partially affected endonuclease activity *in vitro* behaves as WT in *in vivo* conditions (Table 5) implying that only second DNA strand cut is significant for successful resistance to phage infection if compare to Arg182Ala mutation described earlier. *In vivo* restriction assays confirms that the HsdR subunit with 140-149 loop deletion is able to assemble with methyltransferase. Surprisingly, despite failed *in vitro* endonuclease activity, host cells expressing this mutant display a $r^{\pm}$ phenotype in both complementation tests. This significant partial activity in the HsdR construct lacking this segment deserves a more detailed structure-functional analysis in the future. The results of this work will be improved and used for future publication.

3.9. Mutational analysis of residues in 859-869 unresolved segment

HsdR subunit of EcoR124I is suggested to bind methyltransferase with its helical domain which includes residues from 732 to 892 [Lapkouski *et al*., 2009]. Unfortunately the large C-terminal part starting from 885 residue is unresolved in 2W00 WT HsdR crystal structure as well as 859-869 segment in the helical domain. The situation was not improved after the crystal structures of Lys220 mutant motor subunits were determined: only in the Lys220Glu mutant HsdR structure side chains of Arg858 and Arg859 are present. So, 859-869 segment was modeled in YASARA. This segment in modeled structure is located close to the active site of the endonuclease domain (Fig. 35). Considering the close orientation of this loop-segment rich with positively charged Arg (858, 859) and Lys (865, 866) to the endonuclease domain and proposed DNA path through the groove formed by the helical and the endonuclease domains after it leaves the helicase cleft, it might be assigned a role of an exit clamp for helical DNA which would keep the duplex going close to the endonuclease active site residues as EcoR124I translocates. So, we tried to find out whether those positively charged Arg and Lys are important for EcoR124I R-M complex activities. Arg858, Arg859, Lys865 and Lys866 were replaced by Alanine, and two double mutants (Arg858Ala+Arg859Ala and

Lys865Ala+Lys866Ala) were prepared as well. All mutant motor subunits were expressed, purified and analyzed in reconstituted EcoR124I R-M complex *in vitro* and *in vivo*. All ATPase and endonuclease activities *in vitro* as well as *in vivo* complementation tests showed no difference from data obtained for WT enzyme. One possible explanation is that introducing single or double mutations is not enough to alter DNA-protein interactions on the protein surface in such huge multi-subunit complex, or still unresolved C-terminal fragment is involved in accommodation of DNA duplex near the endonuclease active site. The only distinction between WT and above mentioned mutant HsdRs was that all of them were well expressed at 30 °C and at 37 °C were found in the pellet fraction. We tried to create triple mutants as well by introducing a third Alanine substitution in either Arg or Lys double mutant and as result *E. coli* cells were not able even survive when pTrcR124 plasmid with triple mutation was transformed. Thus, mapping of DNA path across HsdR subunit with point mutations is a difficult task and might be not suitable enough taking into account variety of functions accumulated in this motor subunit and complex organization and interrelationship of subunits in pentameric EcoR124I R-M complex. Alternative techniques allowing studies of protein-DNA interactions of such restriction machinery without introducing any structural changes will be explored in the future.

4. Conclusions

Studies of structure-functional interrelationship in the motor subunit of the EcoR124I R-M complex present a complex task as only the fully assembled EcoR124I R-M complex exhibits the individual enzymatic function, the ones on the methyltransferase as well as the variety of functions accommodated in the 120 kDa HsdR subunit, such as ATPase/translocation and endonuclease which are assigned to helicase 1 and 2 domains and the endonuclease domain, respectively. The endonuclease domain of HsdR of EcoR124I partially resembles a RecB-like nuclease of the PD-(E/D)xK superfamily and contains motifs I, II, III and QxxxY characteristic for RecB-like nucleases [Sisakova *et al.*, 2008b], which as well were found in other repair and recombination nucleases. Helicase characteristic motifs clustered in the helicase domains of the motor subunit are conserved throughout Type I and the SF2 superfamily of DEAD-box helicases and are involved in DNA translocation and ATP binding. The HsdR subunit is predicted to interact with the methyltransferase through the C-terminal helical domain to form EcoR124I restriction-modification complex. Thus, the HsdR subunit is a fusion of functional domains with multiple functions. Such multifunctionality of the enzyme has to be controlled, activation of each function has to be regulated based on the status of the other enzymatic functions. So, the endonuclease activity is switched on as the result of stalled DNA translocation and enzyme cuts dsDNA distantly from recognition sequence [Janscak *et al.*, 1999]. While EcoR124I translocates DNA in ATP-dependent manner the endonuclease activity is switched off until translocation is stopped. Modification activity of EcoR124I precedes DNA translocation and the endonuclease activities (Janscak *et al.*, 1999).

In comparison with Type I, Type II restriction endonucleases of the PD-(E/D)xK superfamily act as dimers each represented by a single domain with subdomain organization. Dimerization, DNA recognition and DNA cleavage functions are determined by particular subdomain [Pingoud *et al.*, 2005; Loenen *et al.*, 2014]. Classical Type II restriction endonucleases function without ATP-dependent DNA translocation activity and cleave DNA within or at fixed position close to recognition site [Niv *et al.*, 2007]. Methylation function of Type II restriction endonucleases is housed in separate protein MTase, so each enzyme either endonuclease or MTase has own target recognition subdomain and additional domains or complex formation with other proteins are not

needed [Pingoud *et al.*, 2005; Loenen *et al.*, 2014]. Members of the SF2 superfamily helicases from SWI2/SNF2 family lacking strand separation activity act as true translocases. Isolated protein, for example Rad54 from yeast, is able to carry out branch migration and to remove Rad51 from DNA translocating with the helicase domains in ATP-dependent manner [Burgess *et al.*, 2013; Zhang *et al.*, 2013].

The X-ray crystal structure of the HsdR subunit (2W00) became available back in 2009. However the HsdR structure has unresolved segments, it allows undertaking more detailed and directed structure-functional analysis. As well new other crystal structures of mutant HsdRs help to fill gaps in 2W00 HsdR crystal structure. The crystal structure of HsdR containing ATP molecule bound between the two helicase domains was found unexpectedly also to contact Lys220 residue lying in the endonuclease domain on a short helical segment not comprising the core fold indicating a direct communication between translocase and endonuclease via ATP and that was never detected for other structures of helicases and translocases. But the static nature of the X-ray structure still left open the question how a signal is communicated over the whole distance of more than ~20 Å from the ATP to the catalytic center of the endonuclease domain. The crystal structure of the Lys220Ala HsdR mutant (4BEC) containing the resolved 180s loop segment allowed structural analysis and interpretation of the region around Lys220 residue as well as QxxxY motif lying directly above the endonuclease active site using MD simulations with a hybrid construct in which the loop structure was taken into the WT structure from the specific mutant. In addition to the contact with ATP, computational modeling predicts various contacts of the endonuclease either with the adjacent helical domain or DNA, and the involvement of these in either interdomain communication between translocase and endonuclease or in activation of endonuclease activity has not yet been described or discussed.

Our efforts were focused on studies of structural elements of the endonuclease domain in interdomain communication in the HsdR motor subunit of EcoR124I. Using site-directed mutagenesis with combination of *in vitro* and *in vivo* characterization of tested mutant enzymes we assigned additional function for $Q_{179}xxxY_{183}$ motif via Arg182 residue in communicating a signal across the enzyme about its ATP-ligation status. We showed that substitution of

Arg182 side chain with a shorter side chain leads to an interruption in the communication of the translocation and the endonuclease functions, and such mutant enzyme is not able to cleave two-site linear DNA *in vitro* assays and the substrate remains uncut the second DNA strand in the *in vitro* test with circular DNA. However, the translocation function is unaltered. Consequently broken coordination of the endonuclease function in the motor subunit results in successful phage infection on *E. coli* cells with such defective mutant enzyme.

Data from MD simulations and crystal structures of Lys220 mutant HsdRs suggested that functional coupling of duplex translocation to DNA cleavage in the motor subunit occurs by means of interdomain engagement via ATP. Lys220 contacting ATP molecule in 2W00 HsdR crystal structure and located ~20 Å away from the endonuclease active site results in specifically impaired only nuclease activity if substituted with Arg, Ala or Glu. Both *in vitro* and *in vivo* studies allowed to order those mutant enzymes WT≈Lys220Arg>Lys220Ala>Lys220Glu accordingly their endonuclease activity. So, the Lys220-ATP interaction as well as motions of 220-loop and Arg182 from 180-loop represent key elements in switching mechanism of the endonuclease function in HsdR of EcoR124I.

Probing of interactions occurring through residues in proximity Lys220 to other domains did not reveal any substantial alterations in enzyme activities, and interactions of these residues are rather random and not persistent nor specific. However, deficient endonuclease activity observed *in vitro* not always corresponds to restriction-deficient phenotype *in vivo* as it was shown for the Phe140Gly…Arg149Gly deletion mutant.

Analysis of all studied mutations in the motor subunit described in this work shows that the helical domain is sensitive for any structural changes and introducing of point mutations, for example in the potential groove region that is predicted to bind to DNA, might lead to problematic protein packing at standard expression conditions, nevertheless when playing with the conditions these constructs can be expressed and folded. However, particular substitutions in the helical domain such as Asp855 and Arg858 for Alanine lead to more serious alterations in the protein packing than for most of other mutant. These mutant enzymes expressed at 30 °C are fully functional *in vitro* however these mutations develop partial restriction deficiency for cell phenotype implying that

C-terminal domain might be packed wrongly under *in vivo* conditions. Interestingly, while the single mutant shows a temperature-sensitive behavior, a double mutant with the additional Glu165His mutation well expressed at standard 37 °C temperature.

We could show that translocation and endonuclease functions in the motor subunit of EcoR124I communicate via ATP. The signal about ATP ligation status represents cascade of interactions across the endonuclease domain of the motor subunit. The loss of the ATP-Lys220 interaction leads to movement of side chain of Lys220 and 220-loop itself. The signal from 220-loop is transferred via Arg182 on 180-loop. Entire endonuclease domain rotates relative to the other three domains and the signal reaches the catalytic residues. After that the enzyme cleaves the DNA duplex.

Structure-functional studies of such large complex enzyme as EcoR124I require comprehensive analysis. More detailed analysis employing MD simulations and development of new techniques in future will allow better functional understanding EcoR124I R-M enzyme, as well as facilitate characterization of R-M enzymes form other Type I families.

5. Publications

5.1.

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

5.2.

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)

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.

6. References

Abadjieva, A., Patel, J., Webb, M., Zinkevich, V., Firman, K. (1993) A deletion mutant of the type IC restriction endonuclease EcoR1241 expressing a novel DNA specificity. *Nucleic Acids Research*, 21(19): 4435-4443.

Abadjieva, A., Webb, M., 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(1): 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,*. 3: 319–328.

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.

Aravind, L., Makarova, K.S., Koonin, E.V. (2000) Holiday junction resolvases and related nucleases: identification of new families, phyletic distribution and evoluationary trajectories. *Nucleic Acids Res*., 28: 3417-3432.

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.

Argos, P., Hanei, M., Wilson, J.M., Kelley, W.N. (1983) A possible nucleotide-binding domain in the tertiary fold of phosphoribosyltransferases. *J. Biol. Chem.,* 258 (10): 6450-7.

Bannister, D. and Glover, S.W. (1968) Restriction and modification of bacteriophages by r^+ strains of *Escherichia coli* K-12. *Biochem. Biophys. Res. Com.,* 30(6): 735-738.

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.

Bergmans, H.E., van Die, I.M., Hoekstra, W.P. (1981) Transformation in *Escherichia coli*: Stages in the process. *J. Bacteriol.,* 146(2): 564-70.

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(4): 837-859.

Bianco, P.R., Xu, C., Chi, M. (2009) Type I restriction endonucleases are true catalytic enzymes. *Nucleic Acids Res.,* 37(10): 3377–3390.

Bickle, T.A. and Kruder, D.H. (1993) Biology of DNA restriction. *Microbiol. Reviews,* 57(2): 434-450.

Bickle, T.A. (1987) DNA restriction and modification systems. In *Escherichia coli* and *Salmonella typhimurium. Cell. Mol. Biol.,* (Ingraham, J.L., Low, K.B., Magasanik, B., Neidhardt, F.C., Schaechter, M., Umbarger, H.E., Eds.), 692-696.

Bickle, T.A. (1993) The ATP-dependent restriction enzymes. In Nucleases (Linn, S.M., Lloyd, R.S. Roberts, R.J., eds), 2nd edit., pp. 35±88, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Boys, S.F. and Bernardi, F. (1970) Calculation of small molecular interactions by differences of separate total energies—some procedures with reduced errors. *Molec. Phys.,* 19: 553–556.

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).

Calisto, R.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.

Callow, P., Sukhodub, A., Taylor, J.E., Kneale, G.G. (2007) Shape and subunit organisation of the DNA methyltransferase M.AhdI by small-angle neutron scattering, *J. Mol. Biol.,* 369: 177–185.

Chan, K.M., Delfert, D., Junger, K.D. (1986) A direct calorimetric assay for Ca^{2+}-stimulated ATPase activity. *Anal. Biochem.* 157: 375–380.

Chang, A.C. and Cohen, S.N. (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.,* 134: 1141–1156.

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: 1043-1050.

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.

Csefalvay, E., Lapkouski, M., Guzanova, A., Csefalvay, L., Baikova, T., Shevelev, I., Bialevich, V., et al. (2015) Functional coupling of duplex

translocation to DNA cleavage in a type I restriction enzyme. *PLOS ONE*, 10(6).

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): 10089–10092.

Datta, N., Lawn, A.M., Meynell, E. (1966) The relationship of F type piliation and F phage sensitivity to drug resistance transfer in R+F- *Escherichia coli* K 12. *J. Gen. Microbiol.,* 45(2): 365-76.

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.

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

Dröge, P. (1992) Recombination of nicked DNA knots by gamma delta resolvase suggests a variant model for the mechanism of strand exchange. *Nucleic Acids Res.,* 20(23): 6159-66.

Dryden, D.T.F., Cooper, L.P., Murray, N.E. (1993) Purification and characterization of the methyltransferase from the type 01 restriction and modification system of *Escherichia coli* K12. *J. Biol. Chem.,* 268: 13228-13236.

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

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

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

Endlich, B. and Linn, S. (1985) The DNA restriction endonuclease of *Escherichia coli* B.I. Studies of the DNA translocation and the ATPase activities. *J. Biol. Chem.,* 260: 5720–5728, (1985).

Eskin B. and Linn, S. (1972) Deoxyribonucleic acid modification and restriction enzymes of *Escherichia coli* B.2. Purification, subunit structure, and catalytic properties of the restriction endonuclease. *J. Biol. Chem.,* 247: 6183–6191.

Farah, J.A. and Smith, G.R. (1997) The RecBCD enzyme initiation complex for DNA unwinding: enzyme positioning and DNA opening. *J. Mol. Biol.,* 272: 699–715.

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

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.

Garcı´a, L.R. and Molineux, I.J. (1999) Translocation and specific cleavage of bacteriophage T7 DNA in vivo by EcoKI. *Proc. Natl. Acad. Sci. USA*, 96: 12430–12435.

Glover, S.W. (1970) Functional analysis of host-specificity mutants in *Escherichia coli. Genet. Res.,* 15(2): 237-50.

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

Grimme, S., Antony, J., Ehrlich, S., Krieg, H. (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.,* 132: 154104–154123.

Halford, S.E., Welsh, A.J., Szczelkun, M.D. (2004) Enzyme-mediated DNA looping. *Annu. Rev. Biophys. Biomol. Struct.,* 33: 1–24.

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

Hayward, S., Kitao, A., Berendsen, H.J.C. (1997) Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme. *Proteins,* 27: 425–43.

Hayward, S. and Berendsen, H.J.C. (1998) Systematic analysis of domain motions in proteins from conformational change; new results on citrate synthase and T4 lysozyme. *Proteins,* 30: 144–154.

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

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

Holubova, I., Vejsadová, S., Firman, K., Weiserová, M. (2004) Cellular localization of Type I restriction-modification enzymes is family dependent. *Biochem. Biophys. Res. Commun.,* 319(2): 375-380.

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.

Hubacek, J. and Glover, S.W. (1970) Complementation analysis of temperature-sensitive host specificity mutations in *Escherichia coli*. *J. Mol. Biol.,* 50(1): 111-27.

Humphrey, W., Dalke, A., Schulten, K. (1996) VMD—visual molecular dynamics. *J. Mol. Graph.,* 14: 33–38.

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., Abadjieva, A., Firman, K. (1996) The type I restriction endonuclease R.EcoR124I: over-production and biochemical properties. *J. Mol. Biol.,* 257(5): 977-991.

Janscak, P., Dryden, D.T.F. Firman, K. (1998) Analysis of the subunit assembly of the type IC restriction-modification enzyme EcoR124I. *Nucl. Acids. Res.,* 26: 4439–4445.

Janscak, P., MacWilliams, M.P., Sandmeier, U., Nagaraja, V., 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. 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.

Janulaitis, A., Vaisvila, R., Timinskas, A., Klimasauskas, S., Butkus, V. (1992) Cloning and sequence analysis of the genes coding for Eco57I type IV restriction-modification enzymes. *Nucleic Acids Res.,* 20 (22): 6051-6.

Jindrova, E., Schmid-Nuoffer, S., Hamburger, F., Janscak, P., Bickle, T.A. On the DNA cleavage mechanism of type I restriction enzymes. *Nucl. Acids Res.,* 33: 1760-1766.

Jorgensen, W.L. and Tirado-Rives, J. (2005) Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc Natl Acad Sci USA.,* 102: 6665–6670.

Kannan, P., Cowan, G.M., Daniel, A.S., Gann, A.A.F., Murray, N.E. (1989) Conservation of organisation in the specificity polypeptides of two families of type I restriction enzymes. *J. Mol. Biol.,* 209: 335-344.

Kennaway, C.K., Taylor, J.E., Song, C.F., Potrzebowski, W., Nicholson, W., White, J.H., Swiderska, A., et al. (2012) Structure and operation of the DNA-translocating type I DNA restriction enzymes. *Genes Dev,.* 26: 92–104.

Kim, J.L., Morgenstern, K.A., Griffith, J.P., Dwyer, M.D., Thomson, J.A., Murcko, M.A., Lin, C., et al. (1998) Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure,* 6: 89–100.

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 restriction-modification enzyme and its functional implications. *Proc. Natl. Acad. Sci. USA.,* 102: 3248-3253.

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

Kong, X.P., Onrust, R., O'Donnell, M., Kuriyan, J. (1992) Three-dimensional structure of the beta subunit of E. coli DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell,* 69(3): 425-37.

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.

Krüger, D.H, Kupper, D., Meisel, A., Tierlich, M., Reuter, M., Schroeder, C. (1995) Restriction endonucleases functionally interacting with two DNA sites. *Gene,* 157(1-2): 165.

Kusiak, M., Price, C., Rice, D., Hornby, D.P. (1992) The HsdS polypeptide of the type IC restriction enzyme EcoR124 is a sequence-specific DNA-binding protein. *Mol. Microbiol.,* 21: 3251-3256.

Lapkouski, M., Panjikar, S., Kuta Smatanova, I., Csefalvay, E. (2007) Purification, crystallization and preliminary X-ray analysis of the HsdR subunit of the EcoR124I endonuclease from *Escherichia coli*. Acta Crystallogr. *Sect. F Struct. Biol. Cryst. Commun.,* F63: 582-585.

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 Structural & Molecular Biology,* 16: 94-105.

Lee, N.S., Rutebuka, O., Arakawa, T., Bickle, T.A., Ryu, J. (1997) KpnAI a new type I restriction-modification system in *Klebsiella pneumonie*. *J. Mol. Biol.,* 271: 342-348.

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.

Lia, G., Praly, E., Ferreira, H., Stockdale, C., Tse-Dinh, Y.C., Dunlap, D., Croquette, V., Bensimon, D., Owen-Hughes, T. (2006) Direct observation of DNA distortion by the RSC complex. *Mol. Cell*, 21: 417–425.

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., Daniel, A.S., Braymer, H.D., Murray N.E. (1989) Organization and sequence of the hsd genes of *Escherichia coli* K-12. *J. Mol. Biol.,* 198: 159-170.

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

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.

Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. New York.

Marini, V. and Krejci, L. (2012) Unwinding of synthetic replication and recombination substrates by Srs2. *DNA Repair (Amst).,* 11(10): 789-98.

McClelland, S.E. and Szczelkun, M.D. (2004) The Type I and III Restriction Endonuclease: Structural Elements in Molecular Motors that Process DNA. In: Pingound A, editor. Restriction Enzymes, Nucleic Acids and Molecular Biology. Springer Verlag, Berlin, Germany. 14: 111–135.

McClelland, S.E., Dryden, D.T.F., 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. and Kneale, G.G. (1996) High resolution footprinting of a type I methyltransferase reveals a large structural distortion within the DNA recognition site. *Nucleic Acids Res.*, 24(24): 4853-4858.

Mernagh, D.R., Reynolds, L.A., Kneale, G.G. (1997) DNA binding and subunit interactions in the type I methyltransferase M.EcoR124I. *Nucleic Acids Res.,* 25(5): 987-991.

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

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

Meselson, M, Yuan, R, Heywood, J. (1972) Restriction and modification of DNA. *Annu. Rev. Biochem.,* 41: 447-66.

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

Murray, N.E., Daniel, A.S., Cowan, G.M., Sharp, P.M. (1993) Conservation of motifs within the unusually variable polypeptide sequences of type I restriction and modification enzymes. *Mol. Microbiol.,* 9: 133–143.

Murray, N.E. (2000) Type I restriction systems: Sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol. *Mol. Biol. Rev.,* 64: 412–434.

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., Sisakova, 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. *Nucl. Acids Res.,* 34: 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): 438-452.

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

Petrusyte, M., Bitinaite, J., Menkevicius, S., Klimasauskas, S., Butkus, V., Janulaitis, A. (1988) Restriction endonucleases of a new type. *Gene,* 74(1): 89-91.

Piekarowicz, A. and Goguen, J.D. (1986) The DNA sequence recognized by the EcoDXX1 restriction endonuclease. *Eur. J. Biochem.,* 154(2): 295-8.

Pingoud, V., Sudina, A., Geyer, H., Bujnicki, J.M., Lurz, R., Lüder, G., Morgan, R., et al. (2004) Specificity changes in the evolution of type II restriction endonucleases: a biochemical and bioinformatic analysis of restriction enzymes that recognize unrelated sequences. *J. Biol. Chem.,* 280(6): 4289-98.

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., Murray, N.E. (1998) Sequence-specific DNA binding by EcoKI, a type Ia DNA restriction enzyme. *J. Mol. Biol.,* 283: 963–976.

Price, C., Shepherd, J.C., Bickle, T.A. (1987) DNA recognition by a new family of type I restriction enzymes: a unique relationship between two different DNA specificities. *EMBO J.,* 6(5): 1493-1497.

Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, et al. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics,* 29(7): 845–854.

Randerath, E. and Randerath, K. (1964) Resolution of complex nucleotide mixtures by two-dimensional anion-exchange thin-layer chromatography. *J. Chromatogr.,* 16: 126-9.

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.

Roberts, G.A., Cooper, L.P., White, J.H., Su, T.J., Zipprich, J.T., Geary, P., Kennedy, C., et al. (2011) An investigation of the structural requirements for ATP hydrolysis and DNA cleavage by the EcoKI Type I DNA restriction and modification enzyme. *Nucleic Acids Research,* 39(17): 7667–7676.

Roberts, R.J., Belfort, M., Bestor, T., Bhagwat, A.S., Bickle, T.A., Bitinaite, J., Blumenthal, R.M., et al. (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Research,* 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. *Nucleic Acids Research,* 38: 234-236.

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

Schrödinger LLC (2011) QSite 5.7. Schrödinger LLC, New York.

Schrödinger LLC. (2011) Impact 5.7. Schrödinger LLC, New York.

Seidel, R., van Noort, J., van der Scheer, C., Bloom, J.G., Dekker, N.H., Dutta, C.F., Blundell, et al. (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., Dekker, C. (2005) Dynamics of initiation, termination and reinitiation of DNA translocation by the motor protein EcoR124I. *EMBO J.,* 24(23): 4188-4197.

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

Sharp, P.M., Kelleher, J.E., Daniel, A.S., Cowan, G.M., Murray, N.E. (1992) Roles of selection and recombination in the evolution of type I restriction-modification systems in enterobacteria. *Proc. Natl. Acad. Sci. USA.,* 89(20): 9836-40.

Simons, M., Szczelkun, D. (2011) Recycling of protein subunits during DNA translocation and cleavage by Type I restriction-modification enzymes. *Nucleic Acids Research,* 1-11.

Singleton, M.R., and Wigley, D.B. (2002) Modularity and specialization in superfamily 1 and 2 helicases. J. Bacteriol. 184, 1819–1826.

Sisakova, E., Weiserova, M., Dekker, C, Seidel, R, Szczelkun, M.D. (2008) The interrelationship of helicase and nuclease domains during DNA translocation by the molecular motor EcoR124I. *J. Mol. Biol.,* 384: 1273-1286.

Sisakova, E, Stanley, L.K, Weiserova, M, Szczelkun M.D. (2008b) A RecB-family nuclease motif in type I restriction endonuclease EcoR124I. *Nucleic Acids Res.,* 36: 3939–3949.

Sistla, S. and Rao, D.N. (2004) S-Adenosyl-L-methionine-dependent restriction enzymes. *Crit. Rev. Biochem. Mol. Biol.,* 39(1): 1-19.

Smidt, J.D., Kühnlein, U., Werner, A. (1972) Host specificity of DNA produced by *Escherichia coli:* XV. The role of nucleotide methylation in in vitro B-specific modification. *J. Mol. Biol.,* 63: 9-19.

Soultanas, P. and Wigley, D.B. (2001) Unwinding the 'Gordian knot' of helicase action. *Trends. Biochem. Sci*., 26: 47–54.

Stanley, L.K and Szczelkun, M.D. (2006) Direct and random routing of a molecular motor protein at a DNA junction. *Nucleic Acids Res.,* 34(16): 4387-4394.

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

Szybalski, W., Kim, S.C., Hasan, N., Podhajska, A.J. (1991) Class-IIS restriction enzymes - a review. *Gene,* 109(1): 169.

Taylor, I., Patel, J., Firman, K., Kneale, G.G. (1992) Purification and biochemical characterization of the EcoR124 type I 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(21): 4929-4935.

Taylor, I.A., Davis, K.G., Watts, D., Kneale, G.G. (1994) DNA-binding induces a major structural transition in a type I methyltransferase. *EMBO J.,* 13(23): 5772-5778.

Titheradge, A.J.B., Ternent, D., Murray, N.E. (1996) A third family of allelic hsd genes in Salmonella enterica: sequence comparisons with related proteins identify conserved regions implicated in restriction of DNA. *Mol. Microbiol.,* 22: 437-447.

Titherage, A.J.B., King, J., Ryu, J., Murray, N.E. (2001) Families of restriction enzymes: an analysis prompted by molecular and genetic data for ID restriction and modification systems. *Nucleic Acids Research,* 29(20): 4195-4205.

Tyndall, C., Mister, J., Bickle, T.A. (1994) The Escherichia coli prr region encodes a functional region encodes a functional type IC DNA restriction systems 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(22): 6540-6547.

Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S., Wigley, D.B. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell,* 97: 75–84.

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.

Weiserova, M. and 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(4-5): 585-589.

Weiserova, M., Dutta, C.F., Firman, K.A. (2000) Novel mutant of the type I restriction-modification enzyme EcoR124I is altered at a key stage of the subunit assembly pathway. *J. Mol. Biol.,* 304(3): 301-310.

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

Wilson, G.G. (1991) Organization of restriction-modification systems. *Nucleic Acids Research,* 19(10): 2539-2566.

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. and 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 E. coli K. *Cell.,* 20(1): 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., Popova, L., Kryukov, V., Abadjieva, A., Bogdarina, I., Janscak, P., 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.