

Use of protein engineering to study DNA recognition by the *E. coli* DNA methyltransferase (EcoDam)

by

Hany Amin Elsayy

A thesis submitted in partial fulfillment
of the requirements for the degree of

**Doctor of Philosophy
in Biochemistry**

Approved, Thesis Committee

Prof. Dr. Albert Jeltsch

Prof. Dr. Matthias Ullrich

Prof. Dr. Marcelo Fernandez-Lahore

Dr. Antal Kiss

Date of Defence: November 24, 2009

School of Engineering and Science

Declaration

The work described in this thesis is my own work, unless otherwise stated or mentioned in the references. The thesis was written by myself and nobody else.

Hany Amin Elsayy Mostafa

Bremen, september 2009.

List of publications and author contributions

Elsawy, H., Podobinski, S., Chahar, S. and Jeltsch, A. (2009) Transition from EcoDam to T4Dam recognition mechanism without loss of activity and specificity. *ChemBioChem*. 10, 2488-2493.

Site directed mutagenesis, protein purification, enzyme kinetics and electrophoretic mobility shift assay experiments were performed by H. E.

H. E. was involved in data analysis and writing of the manuscript.

Chahar, S., **Elsawy, H.**, Ragozin, S. and Jeltsch, A. (2009) Changing of DNA recognition specificity of the EcoDam DNA-(adenine N6)-methyltransferase by directed evolution. *J. Mol. Biol.* in press.

Screening step (2nd – 4th generations), protein purification, some enzyme kinetics experiments were performed by H. E.

H. E. was involved in some data analysis.

Table of contents

List of publications and authors contribution

Table of contents

Achnowledgements

Abstract

I. Introduction	1
I. 1. DNA methylation	1
I. 2. Chemistry of DNA methylation	2
I. 3. Types of prokaryotic DNA methyltransferases	2
I. 4. The role of DNA methylation in prokaryotes	3
I. 4. 1. Coordination of gene expression and cell cycle	3
I. 4. 2. DNA repair	4
I. 4. 3. Gene expression and phase variation	4
I. 4. 4. Control of DNA uptake	5
I. 4. 5. Bacterial virulence	5
I. 5. The structure of prokaryotic DNA methyltransferases	5
I. 5. 1. The structure of the <i>E.coli</i> DNA adenine methyltransferase	7
I. 5. 2. DNA recognition of the <i>E.coli</i> DNA adenine methyltransferase	7
I. 6. Protein engineering	8
I. 6. 1. Rational design	8
I. 6. 2. Directed evolution	9
I. 6. 2. 1. Biochemical strategies for the randomization of protein encoding DNA sequences	12
I. 6. 2. 2. Selection or screening for desired property	13
II. Aim of this study	14
III. Results and discussion	15
III. A. Changing of DNA recognition specificity of the EcoDam DNA-(adenine N6)-methyltransferase by directed evolution	15
III. B. Transition from EcoDam to T4Dam DNA recognition mechanism without loss of activity and specificity	16
III. C. Discrimination of unmethylated and hemimethylated DNA by <i>E.coli</i> DNA methyltransferase enzyme (EcoDam)	16
IV. Conclusions	23
V. Materials and Methods	28
V. A. Materials	28
V. A. 1. Bacterial strains	28
V. A. 2. Buffers and solutions	28
V. A. 3. Chemicals and kits	29
V. A. 3. 1. Chemicals	29
V. A. 3. 2. Kits	31
V. A. 4. Enzymes	31
V. A. 5. Media and antibiotics	31
V. A. 5. 1. Media	31
V. A. 5. 2. Antibiotics	31
V. A. 6. Molecular weight markers	32

V. A. 6. 1. DNA molecular weight marker	32
V. A. 6. 2. Protein molecular weight marker	32
V. A. 7. Primers for EcoDam gene amplification	33
V. A. 8. Dam substrates	33
V. A. 9. Vectors	34
V. B. Methods	36
V. B. 1. Microbiological methods	36
V. B. 1. 1. Culture and storage of <i>Escherichia coli</i> strains	36
V. B. 1. 2. Growth curve	36
V. B. 1. 3. Preparation of log-phase <i>E.coli</i> cells and protein induction	36
V. B. 1. 4. Transformation	37
V. B. 1. 4. 1. Preparation of CaCl ₂ treated competent cells for the	37
heat-shock transformation	
V. B. 1. 4. 1. 1. Heat-shock transformation of the chemically	37
competent cells	
V. B. 1. 4. 2. Preparation of the electrocompetent cells	38
V. B. 1. 4. 2. 1. Transformation of the electrocompetent <i>E.coli</i>	38
cells (Electroporation)	
V. B. 2. Molecular biology methods	39
V. B. 2. 1. DNA concentration determination	39
V. B. 2. 2. DNA purification using silica membrane spin columns	40
V. B. 2. 3. PCR-product purification	40
V. B. 2. 4. Extraction of DNA from agarose gel	40
V. B. 2. 5. DNA digestion with restriction enzymes	41
V. B. 2. 6. Agarose gel electrophoresis of DNA	41
V. B. 2. 7. Polymerase chain reaction (PCR)	42
V. B. 2. 7. 1. Composition of the PCR mixture	42
V. B. 2. 7. 2. Temperature cycling	43
V. B. 2. 8. Site directed mutagenesis	45
V. B. 2. 9. Error prone PCR	46
V. B. 2. 10. Megaprimer PCR of whole plasmid (MEGAWHOP)	46
V. B. 2. 11. DNA shuffling	48
V. B. 2. 12. DNA sequencing	50
V. B. 2. 13. Electrophoretic Mobility Shift Assay (EMSA)	50
V. B. 3. Protein biochemical methods	51
V. B. 3. 1. Determination of protein concentration by UV absorption	51
V. B. 3. 2. Protein expression and purification	51
V. B. 3. 2. A. IPTG induction for Protein expression	51
V. B. 3. 2. B. Soluble protein fraction from host cell	51
V. B. 3. 2. C. Immobilized metal-ion affinity chromatography (IMAC)	52
V. B. 3. 3. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)	53
V. B. 3. 4. Enzymatic activity assay	54
VI. References	56

Curriculum Vitae

Acknowledgments

I would like to thank many people in this endeavor. First of all I would like to thank God for allowing me to have this amazing opportunity. I would like to express my sincere gratitude to Prof. Dr. Albert Jeltsch, my supervisor, for welcoming me into his research group. During my days in his lab, I was deeply impressed by his endless curiosity in science and his systematic approach to solve problems with deep knowledge and strong enthusiasm. Without his timely advice and guidance throughout my PhD program, most of my projects would not have been successful.

I also would like to acknowledge other committee members, Prof. Dr. Matthias Ullrich and Prof. Dr. Marcelo Fernandez-Lahore. They also served as my candidacy committee and provided me with many comments and much help since then. I appreciate their time and efforts to review my thesis.

I want to express my appreciation to Prof. Dr. Antal Kiss (Institute of Biochemistry, Biological Research Center of the Hungarian Academy of Sciences, Szeged, Hungary) for being a co-referee of my thesis. My thanks go to him for his valuable comments and suggestions.

I am grateful to all former and current Jeltsch laboratory members. I was very lucky to work with such wonderful colleagues. In particular, Sanjay Chahar helped me a lot to settle down when I joined Jeltsch laboratory. He knew everything I needed and was always willing to assist me.

I gratefully acknowledge a stipend from the Egyptian Government and giving me the chance to perform my PhD work at Jacobs University Bremen, Germany.

I would like to thank DFG for research funding.

A special thank is also addressed to all staff members and colleagues in the chemistry department, Faculty of Science, Tanta University, Egypt.

Finally, I wish to thank my wife Azza Sadek and my daughters Hweda and Salma for their continual support and encouragement throughout my PhD program. My son Mostafa, who was born in Germany, has added great pleasure to my life.

Abstract

DNA methylation carried out by DNA methyltransferases (MTases) is widespread in prokaryotes and eukaryotes. In this study, we have changed the target specificity of EcoDam from GATC to GATT by directed evolution, combining different random mutagenesis methods with restriction protection at GATT sites for selection and screening. By co-evolution of an enzyme library and a substrate library, we identified GATT as the best non-GATC site and discovered a double mutation, R124S/P134S, as the first step to increase enzyme activity at GATT sites. After four generations of mutagenesis and selection, we obtained enzyme variants with new specificity for GATT. While the wild-type EcoDam shows no detectable activity at GATT sites in *E.coli* cells, some variants prefer methylation at GATT over GATC sites by about 10-fold in cells. In vitro DNA methylation kinetics carried out under single-turnover conditions using a hemimethylated GATC and a GATT oligonucleotide substrate confirmed that the evolved proteins prefer methylation of GATT sites to a similar degree. They show up to 1600-fold change in specificity in vitro and methylate the new GATT target site with 20% of the rate of GATC methylation by the wild-type enzyme, indicating good activity. We conclude that the new methyltransferases are fully functional in vivo and in vitro.

The EcoDam and T4Dam DNA-(adenine N6)-methyltransferases both methylate the adenine residue in GATC sites. They are highly related in amino acid sequence and structure but deviate in their contact to the first base pair of the target sequence. EcoDam contacts GATC with K9 (which corresponds to T4Dam A6), while this contact is mediated by R130 in T4Dam (which corresponds to EcoDam Y138). We have transplanted the T4Dam DNA recognition module into EcoDam and show that the EcoDam K9A/Y138R double mutant is highly active and specific. The evolutionary intermediates of this transition were studied as well: the EcoDam K9A variant showed low activity and loss of recognition of the first base pair, whereas; the EcoDam Y138R variant is fully active and specific. This result indicates that there exists a smooth evolutionary pathway changing the EcoDam DNA recognition mode to the T4Dam mode without loss of activity.

Finally, in an attempt to enhance the preference of the L122A EcoDam for hemimethylated DNA to generate an artificial epigenetic system that can propagate patterns of GATC site methylation potentially useful for gene regulation, the L122, P134 and V133 residues were replaced with other amino acids using site directed mutagenesis, and the catalytic activity of

all variants on unmethylated and hemimethylated substrates was studied. Our results showed that, in addition to L122A, the L122A/V133L EcoDam variant was able to sense the methylation status of the GATC target recognition site and methylated only hemimethylated DNA.

I. Introduction

I. 1. DNA methylation

DNA from various sources contains the methylated bases N⁶-methyladenine, 5-methylcytosine, and N⁴-methylcytosine in addition to the four standard nucleobases (Figure I. 1). These methylated bases are natural components of DNA which distinguishes them from a large variety of chemically modified bases that can be formed by alkylation or oxidative damage of the DNA.

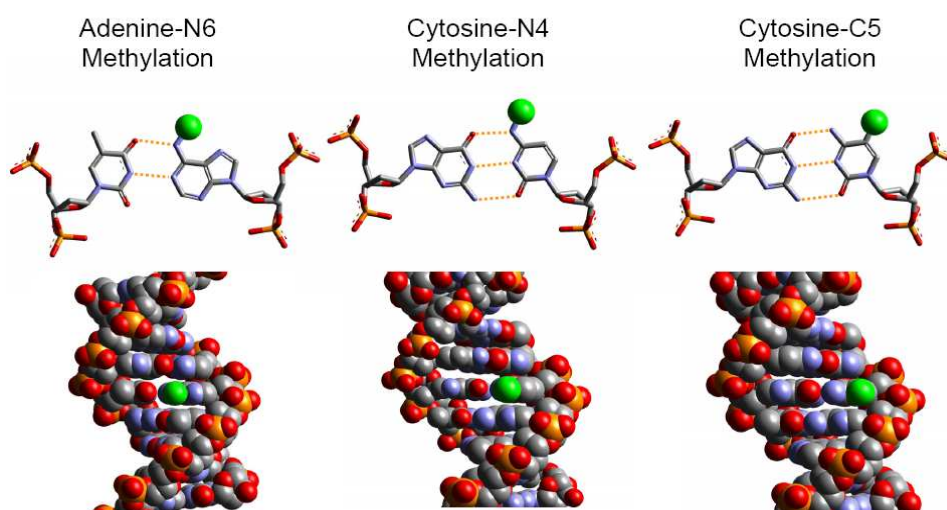


Figure I. 1. **Types of methylated bases observed in DNA.** In the upper part, TA and GC base pairs are shown in stick representation. The positions of methyl groups in N6-methyladenine, N4-methylcytosine and C5-methylcytosine are indicated by green balls. In the lower part, B-DNA structures are shown in space fill representation. The methyl groups positioned in the major groove of the DNA are highlighted in green (Jeltsch & Jurkowska 2009).

DNA methylation is introduced enzymatically by DNA methyltransferases (MTases) after DNA replication. These enzymes use S-adenosyl-L-methionine (AdoMet) as the donor of an activated methyl group and modify the DNA in a sequence-specific manner, usually at palindromic sites, producing S-adenosyl-L-homocysteine (AdoHcy) and methylated DNA (Jeltsch 2002). The methylation does not interfere with the Watson/Crick pairing properties of adenine and cytosine but the methyl group is positioned in the major groove of the DNA, where it can easily be detected by proteins interacting with the DNA. Thereby, methylation adds extra information to the DNA that is not encoded in the sequence, and the methylated bases can be considered the 5th, 6th, and 7th letters of the genetic alphabet.

I. 2. Chemistry of DNA methylation

C-MTases, which form a C-C bond (cytosine-C5 MTases) can be distinguished from N-MTases which form a C-N bond (adenine-N⁶ and cytosine-N⁴ MTases). Different C5-cytosine MTases have been characterized in prokaryotes and eukaryotes. The methylation of C5-cytosine is the most prevalent DNA modification, notably in eukaryotic genomes. So far, C5-methylation has been found in invertebrates, fungi, protozoa, and all of the higher plants and vertebrates studied. However, methylation is not a ubiquitous feature of the eukaryotic genomes because some organisms, including *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, or *Caenorhabditis elegans*, lack detectable methylation (Colot and Rossignol 1999). The N-modified bases (N⁴-methylcytosine and N⁶-methyladenine) are present in prokaryotes and some lower eukaryotes (Jeltsch & Gumpert 2004). It has been known that m6A is present in the DNA of several unicellular eukaryotes, including members of the genera *Chlamydomonas*, *Chlorella*, *Oxytricha*, *Paramecium* and *Tetrahymena* (Ratel *et al.*, 2006).

I. 3. Types of prokaryotic DNA methyltransferases

Most prokaryotic DNA MTases are parts of restriction/modification (RM) systems, which are widely distributed in the bacterial and archeal kingdoms. These systems comprise two enzymes, a restriction endonuclease that specifically recognizes and cleaves DNA within short, often palindromic sequences. The cellular DNA is protected from cleavage by a corresponding MTase, because it modifies the DNA within the same sequence and prevents endonuclease action. Thereby, the methylation pattern imprints a bar code on the DNA that allows the bacteria to distinguish between foreign and own DNA.

A second group of prokaryotic MTases is not accompanied by a restriction enzyme (solitary MTases, also called orphan DNA MTases) as exemplified by the *E.coli* DNA adenine methyltransferase (EcoDam) enzyme that recognizes the GATC palindromic sequence and methylates adenine at N⁶-position and the CcrM (cell cycle-regulated DNA MTase) enzyme, originally identified in *Caulobacter crescentus*, which methylates the adenosine in the sequence 5'-GANTC-3' (where "N" is any nucleotide) (Zweiger *et al.*, 1994).

I. 4. The role of DNA methylation in prokaryotes

In prokaryotes, DNA methylation has five major biological roles (Jeltsch *et al.*, 2007) (Figure I. 2).

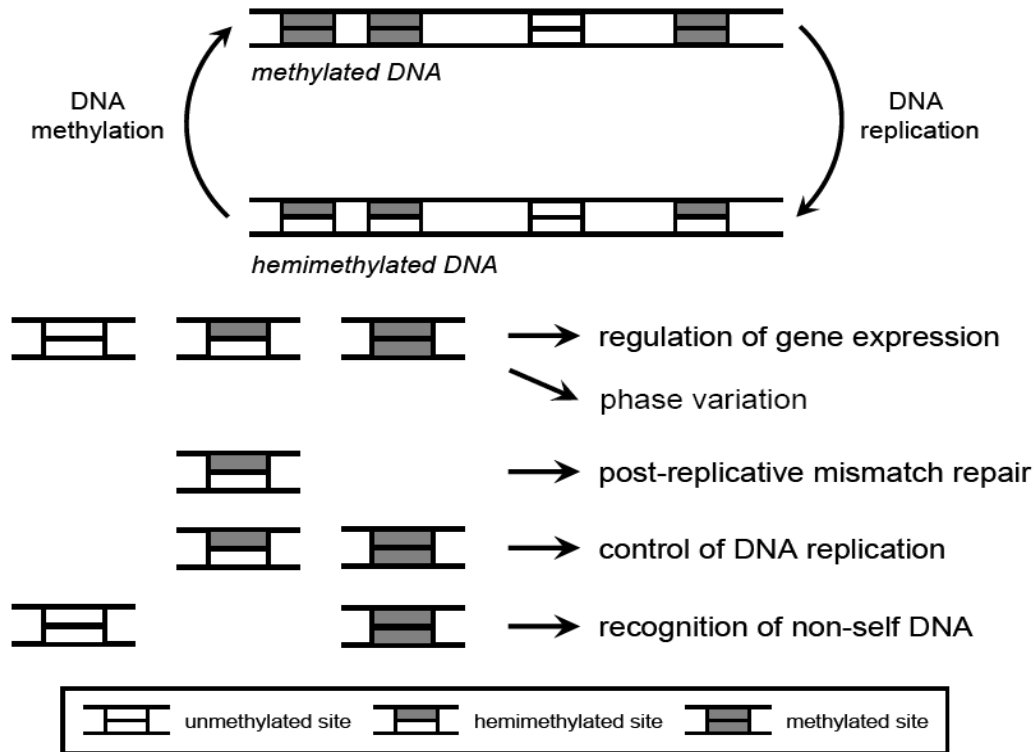


Figure I. 2. **Dynamics and roles of DNA methylation in prokaryotes** (Jeltsch & Jurkowska 2009).

I. 4. 1. Coordination of gene expression and cell cycle

After DNA replication, the transient phase of hemimethylated DNA can be used for specific regulation of the gene expression. In *Escherichia coli* and related bacteria, DNA methylation takes place at the adenine residue within the GATC sites by the action of the DNA adenine methylation (Dam) enzyme. As this site is palindromic, the methyl mark resides on both strands of the DNA. After DNA replication, the DNA is hemimethylated and usually gets rapidly converted into the fully methylated state. However, the origin of replication of the *E. coli* chromosome, which is enriched in GATC sites, stays hemimethylated for 15-20 min because the binding of the SeqA protein prevents remethylation. As hemimethylated origins are not active, this prevents reinitiation of DNA replication before the cell cycle is completed (Lobner-Olesen *et al.*, 2005).

I. 4. 2. DNA repair

The brief period during which methylation sites are hemimethylated after DNA replication allows to distinguish the parental and daughter strand of DNA synthesis. In *E. coli*, this property enables the directed repair of the replication errors by the MutHLS mismatch repair system. In this process, MutS recognizes mismatches and it recruits MutL and MutH, which then move to a GATC site. MutH cleaves the unmethylated strand, and the daughter strand is hydrolyzed by an exonuclease. The repair of the mismatch is finalized by the synthesis of a new DNA strand (Joseph *et al.*, 2006).

I. 4. 3. Gene expression and phase variation

Methylation of MTase target sites can influence the binding of bacterial transcription factors to the DNA and thereby regulate gene expression. In the case of the *E. coli* Dam system, there are some promoters that contain sets of Dam sites, whose methylation is variable. Mutual exclusive methylation of Dam sites at the promoter of P-pili, for example, regulates the expression of these genes and influences pathogenicity of *E. coli* (Figure I. 3). Involvement of DNA methylation in bacterial phase variation and pathogenicity has been demonstrated in other cases as well (Hernady *et al.*, 2002 and Heusipp *et al.*, 2007).

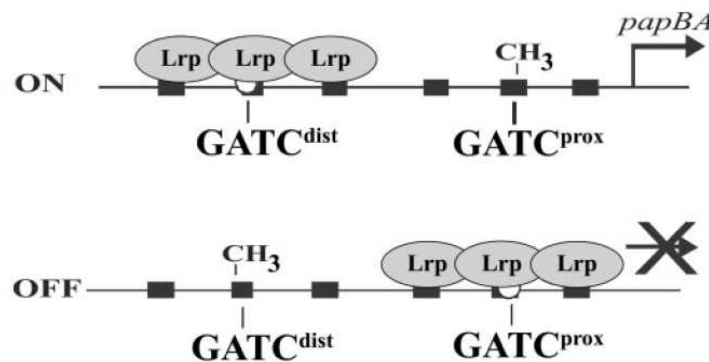


Figure I. 3. **DNA methylation states of phase ON and phase OFF cells.** Binding of Lrp at sites 4-6 in phase ON cells and sites 1-3 in phase OFF cells controls the DNA methylation pattern by blocking methylation of the bound GATC site (Hernady *et al.*, 2002).

I. 4. 4. *Control of DNA uptake*

The majority of bacterial DNA MTases are parts of the restriction-modification systems (Pingoud *et al.*, 2005 & Jeltsch 2002). In such systems, the MTase acts in concert with a restriction enzyme of a matching specificity. The restriction enzyme cuts incoming DNA, whereas genomic DNA is protected against cleavage by methylation at specific sites. This mechanism is employed to protect the bacteria against phage infection, as well as against the uptake of DNA from the environment. These systems thereby constitute a barrier to horizontal gene transfer that might have an important role in bacterial speciation (Jeltsch 2003).

I. 4. 5. *Bacterial virulence*

DNA adenine methyltransferase (Dam) has been reported to control the expression of a number of virulence genes (Mahan *et al.*, 2000 and Mahan & Low 2001). In *Salmonella*, *Haemophilus*, and certain strains of *Yersinia pseudotuberculosis*, lack of Dam methylation causes attenuation of virulence in model animals (Taylor *et al.*, 2005 and Watson *et al.*, 2004). Deletion of Dam erases DNA methylation pattern, which could alter the binding of regulatory proteins to a number of regions on the bacterial chromosome. In the absence of Dam, overexpression of genes could occur if GATC methylation blocked binding of an activator or enhanced the binding of a repressor. Conversely, underexpression of a gene would occur in the absence of Dam if GATC methylation blocked binding of a repressor or enhanced binding of an activator. However, in both *Yersinia pseudotuberculosis* and *Vibrio cholerae*, virulence attenuation is observed if Dam methylase is overproduced (Chen *et al.*, 2003 and Julio *et al.*, 2002). Albeit widespread, the involvement of Dam methylation in bacterial virulence is not universal; for instance, Dam⁻ mutants of *Shigella flexneri* are not attenuated (Honma *et al.*, 2004).

I. 5. The structure of prokaryotic DNA methyltransferases

Both C-MTases and N-MTases are two-domain proteins comprising one large and one small domain with the DNA binding cleft being located at the domain interface. The large domain contains a set of up to ten conserved amino acid motifs, which differ

between C- and N-MTases (Figure I. 4) (Wilson 1992; Kumar *et al.*, 1993 & Malone *et al.*, 1995).

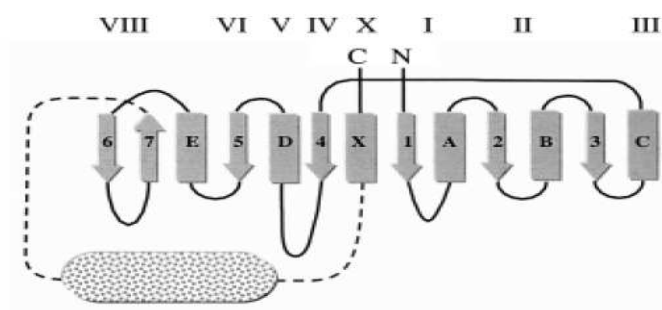


Figure I. 4. **The topology and location of conserved motifs in DNA MTases** (Jeltsch 2002).

Exocyclic amino MTases are subdivided further into six groups (α , β , γ , ζ , δ and ϵ) according to the possible linear arrangement of three domains, the AdoMet-binding domain (FXGXG), the TRD (target-recognition domain) and the catalytic domain (DPPY) (Fig. I. 5). The majority of exocyclic amino MTases fall into the α , β , and γ subgroups. M.BssHI is the only DNA MTase for which the ζ architecture has been confirmed (Bujnicki 2002). EcoDam is classified in the α group, whereas CcrM is an example from β group. The C-MTases are most similar to the γ class of N-MTases. Most cytosine- N^4 MTases are members of the β group of N-MTases, but some examples of cytosine- N^4 MTases are also found in the α and γ groups.

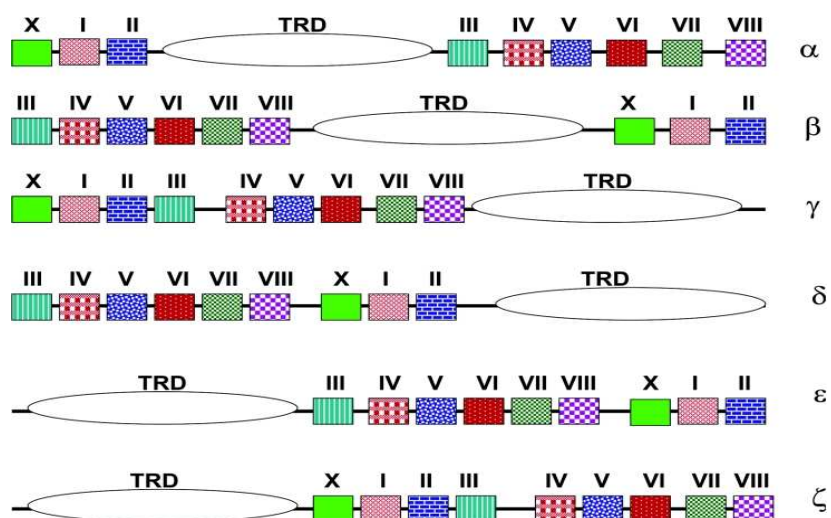


Figure I. 5. **Arrangement of conserved motifs in the primary structure of exocyclic DNA MTases** (Bujnicki 2002) .

I. 5. 1. The structure of the E.coli DNA adenine methyltransferase

The crystal structure of EcoDam (Horton *et al.*, 2006) shows two distinct domains: a seven-stranded catalytic domain (residue 1-56 and 145-270) harboring the binding site for AdoHcy and a DNA binding domain (residues 57-144) consisting of a five-helix bundle and a β -hairpin loop (residues 118-139) that is conserved in the family of GATC-related MTase orthologs (Yang *et al.*, 2003) (Figure I. 6). The DNA interactions are made by a β -hairpin loop and an N-terminal extension of the catalytic domain (residues 7 -10).



Figure I. 6. **The crystal structure of EcoDam complexed with DNA** (Jurkowski *et al.*, 2007).

I. 5. 2. DNA recognition of the E.coli DNA adenine methyltransferase

Key DNA recognition (Horton *et al.*, 2006) contacts between EcoDam and the GATC site are hydrogen bonds between K9 and Y138 and Gua1, contacts of L122 and P134 to the third base pair and an interaction of R124 to the Gua in the fourth base pair. Exchange of any of these residues led to major changes in the DNA recognition specificity of EcoDam (Horton *et al.*, 2005).

I. 6. Protein engineering

Protein engineering aims at creating new or improved enzymes or proteins by altering an enzyme or a protein with inadequate properties. There are two general strategies for protein engineering: rational redesign (Bornscheuer and Pohl 2001; Bornscheuer 2002) and directed evolution (Farinas, Schwaneberg *et al.*, 2000) (Figure I. 7).

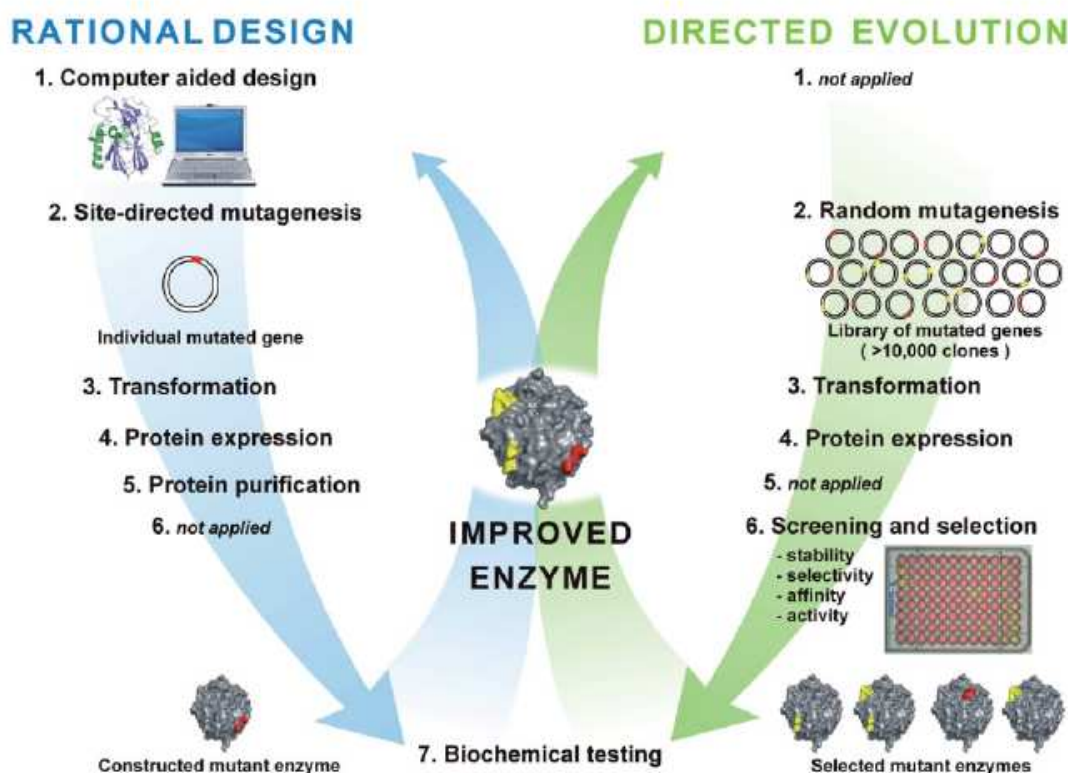


Figure I. 7. **Rational design and directed evolution represent two conceptually different approaches used for engineering of proteins** (Damborsky 2007).

I. 6. 1. Rational design

In rational design, precise changes in amino acid sequence are preconceived based on a detailed knowledge of protein structure, function and mechanism, and are then introduced using site directed mutagenesis (Chen 1999). This technology holds strong promise for optimizing desired properties for commercial applications. It also greatly enhances our understanding of enzyme binding and catalytic mechanisms, thus increasing the success

of future enzyme engineering efforts and laying the foundation for functional prediction of new protein sequences in databases (Chen 2001).

The power of rational design has been demonstrated by the generation of a faster superoxide dismutase, already one of the fastest known enzymes in nature (Getzoff *et al.*, 1992) and complete inversion of coenzyme specificities for both isocitrate and isopropylmalate dehydrogenases (Chen *et al.*, 1996; Hurley *et al.*, 1996 ; Chen 1999).

However, rational design is greatly hindered in practice by the complexity of protein function and the still limited information of structures and functions of proteins. Generally, simple mutations (to be introduced by site-directed mutagenesis, for example) are not expected to have as drastic effects as altering an enzyme's substrate recognition pattern, as many amino acid residues in the enzyme (often not close to one another in the primary structure of the protein) affect the binding pocket of the substrate in the enzyme. Obviously, several amino acid residues may need to be altered simultaneously to achieve the goal of altering substrate specificities. However, as any amino acid residue may be altered into 19 other ones, the number of amino acid combinations that can be made if mutations are introduced at various residues simultaneously can become very large (Glick and Pasternak 2003).

I. 6. 2. Directed evolution

All species have evolved primarily by consecutive rounds of diversification, selection, and amplification. This is based on the principle that the fittest organisms (e.g., those exhibiting the highest growth rate) under a given set of conditions survive and outcompete others in a population. Recent advances in technology in the field of molecular biology have allowed to mimic this process at the molecular level, and to evolve the functions of molecules. Proteins, RNA, and DNA are the major targets of molecular evolution (Wilson *et al.*, 1999; Joyce 1994; Matsuura *et al.*, 2004; Waldo 2003; Aharoni *et al.*, 2005 & Joyce 2004). These molecules are all polymers composed of combinations of different monomer units: 20 different amino acids for proteins and 4 deoxyribonucleotides for DNA or ribonucleotides for RNA.

Directed evolution has become a very popular strategy for improving or altering the biophysical properties of proteins, or even for generating proteins with novel functions.

The alteration of the properties of proteins are of great interest, as some proteins are useful for industrial applications (Ogino and Ishikawa, 2001), medical applications (Sidhu, 2000), and also for understanding the biophysical properties of proteins (Vogit *et al.*, 2000). The molecular evolution of proteins consists of two steps: diversification and selection (Figure I. 8).

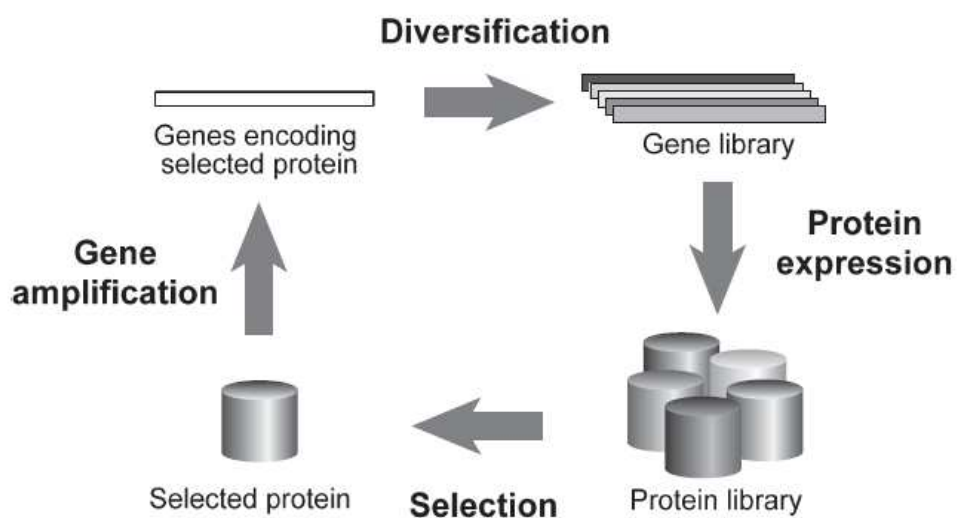


Figure I. 8. **Schematic of molecular evolution of proteins** (Matsuura & Yomo 2006).

The diversification of molecules is carried out by randomly introducing mutations, such as point mutations, insertions or deletions into genes encoding proteins. Thereby, a library of protein variant genes is generated. While this step can be carried out easily, it defines the quality of the gene library that may determine the outcomes of the experiments. Selection steps are not easy when working in a high-throughput format, particularly for proteins; thus, specialized systems are required for this purpose. The evolution of proteins differs from that of RNA or DNA in that proteins can not be amplified themselves, and thus selection based on their properties must simultaneously select the genes encoding them. Hence, genotype and phenotype must be linked physically in some way (Matsuura and Yomo, 2006).

DNA MTases are particularly prone to in vitro evolution approaches as these enzymes modify DNA at specific sites. Thus, the result characterizing the enzyme activity and specificity can be detected on the DNA coding for the protein, giving a unique coupling

of genotype (DNA sequence) and phenotype (enzymatic properties) on the individual DNA molecule (Jeltsch *et al.*, 1996). This facilitates directed evolution experiments enormously, because screening can be performed with a library of DNA molecules coding for different MTases variants and, consequently, carrying different methylation patterns.

The concept of directed evolution was introduced as early as 1967 (Mills *et al.*, 1967). In the past several years, directed evolution has emerged as an alternative approach to rational design, enabling the improvement of structural and functional properties of enzymes, such as stability and performance under different conditions (e.g., at extreme temperatures, pH, and in the presence of organic solvents), or changes in the reaction and substrate specificity (Tao & Cornish 2002). Directed evolution is particularly well suited for tuning enzyme function, that is, improving an activity that already exists at some albeit low level. It was also shown to be useful for combining properties not necessarily found together in any naturally occurring enzyme (Schmidt-Dannert & Arnold 1999). Directed evolution implements an iterative Darwinian optimization process, whereby, the fittest variants are selected from an ensemble of random mutations (Roodveldt *et al.*, 2005).

Directed evolution differs from natural evolution (Table I. 1.) in two key aspects, (1) natural evolution occurs under multiple and variable selection pressures, whereas directed evolution is accomplished under controlled selection pressure for predetermined functions. (2) in directed evolution, ‘non-natural’ functions of practical use can be obtained through the design of appropriate selection schemes, whereas natural evolution favours functions advantageous to the survival of the organism (Schmidt-Dannert 2001 & Williams *et al.*, 2004).

Table I. 1. **Comparison of mutational and selective mechanisms for enzyme improvement between natural and directed evolution.**

	Natural evolution	Directed evolution
Point mutation	Moderate	Frequent
Deletion	Frequent	Rare
Insertion	Frequent	Rare
Inversion	Frequent	Rare
Duplication	Frequent	Rare
Fusion	Frequent	Rare
Recombination	Sexual and somatic	Sexual PCR
Selection	Natural selection (fitness)	Selection/high throuput screen

I. 6. 2. 1. Biochemical strategies for the randomization of protein encoding DNA sequences: Library construction methods for directed evolution.

Methods for the creation of protein-encoding DNA libraries may broadly be divided into three categories (Figure I. 9).

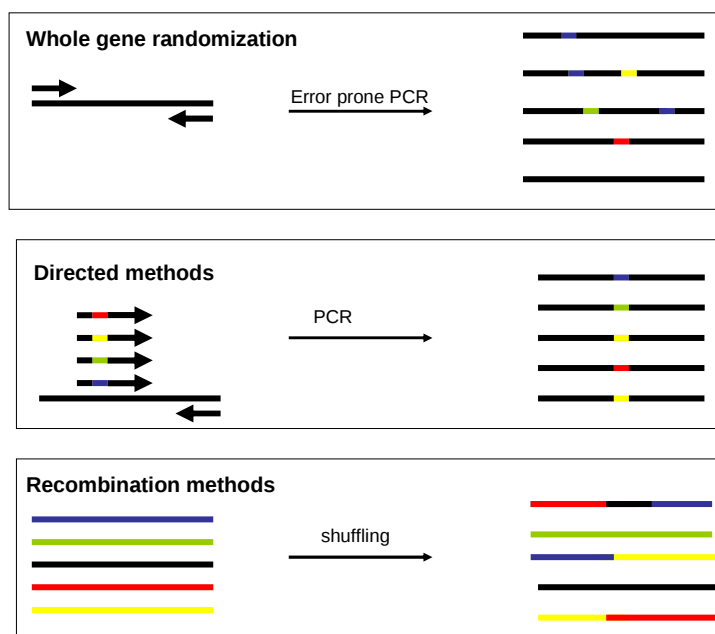


Figure I. 9 . **Overview of methods for the randomization of DNA sequences.** Random methods introduce changes at positions throughout the gene sequence. Directed methods will randomize only a specific segment. Recombination methods bring existing sequence diversity, either from point mutants or from different parental DNA sequences, together in novel combinations.

The first two categories encompass techniques that directly generate sequence diversity in the form of point mutations, insertions, or deletions. These can be divided into methods where changes are made at random within the whole gene and methods that involve randomization at specific positions within a gene sequence.

The first category, randomly targeted methods, encompasses most techniques in which the copying of a DNA sequence is deliberately disturbed. These methods, which include the use of physical and chemical mutagens, mutator strains, and some forms of insertion and deletions mutagenesis, as well as the various forms of error-prone polymerase chain reaction (epPCR), generate diversity at random positions within the DNA being copied. A statistical analysis showed that these gene mutagenesis methods are limited and that not every amino acid can be exchanged into each other (Wong *et al.*, 2004).

The second category of methods targets a controlled level of randomization to specific positions within the DNA sequence. Most work using rational design focuses on mutations close to the active site often by using site directed mutagenesis with primers containing randomized regions, whereas directed evolution experiments often find mutations far from the active site (Sen *et al.*, 2007).

The third category of techniques for library construction are those that do not directly create new sequence diversity but combine existing diversity in new ways. These are the recombination techniques, such as DNA shuffling (Stemmer 1994) and Staggered Extension Process (Zhao *et al.*, 1998) that take portions of existing sequences and mix them in novel combinations. These techniques make it possible to bring together advantageous mutations while removing deleterious mutations in a manner analogous to sexual recombination. It should be noted that while these techniques do not in principle produce new point mutations they are generally dependent on a PCR reconstruction process that can be error prone, and new point mutations are usually produced as a by-product of these techniques.

I. 6. 2. 2. Selection or screening for desired property

The sorting step is the most critical part of any evolution experiment: useful variants must be efficiently recognized and isolated from a large pool of mutants. Evolution requires a link between genotype (a nucleic acid that can be replicated) and phenotype (a functional

trait such as binding of a transition state analogue or catalytic activity). This linkage can be generated by physical coupling of the gene to the encoded protein variant through protein display on phages, viruses, bacteria or eukaryotic cells or by using *in vitro* methods like ribosome display (Becker *et al.*, 2004). Finding the enzyme with the desired properties can be accomplished by genetic selection for a specific property, by phage display and selection for the binding property or direct screening all the candidates for the desired property.

Screening, i.e. the one by one analysis of the catalytic properties of each member of the repertoire, allows typically between 10^3 and 10^6 distinct proteins to be considered (Lin and Cornish, 2002). Selections, which are different from screening, allow the simultaneous analysis of protein properties for sets up to about 10^{8-10} different proteins. Recently, new high-throughput screening/selection methods are emerging, e.g. cell surface display and sorting, compartment translation, *in vitro* ribosome or mRNA translation. *In vitro* restriction protection assay is the screening method used in this study.

II. Aim of this study

The objectives of this study were to

- generate a variant of EcoDam-(adenine N6)-methyltransferase that efficiently methylates a novel target site by using a directed evolution approach.
- transplant the T4Dam recognition module into EcoDam by using a rational design approach.
- enhance the preference of L122A variant (Horton *et al.*, 2006) for hemimethylated DNA by using a rational design approach.

III. Results and discussion

III. A. Changing the DNA recognition specificity of the EcoDam DNA-(adenine N6)-methyltransferase by directed evolution

This part of the thesis describes directed evolution experiments of the EcoDam-(adenine N6)-methyltransferase. The objective of this project was to generate an EcoDam variant that efficiently methylates a novel target site. To achieve this goal, we started the design of EcoDam variants towards new specificity by site saturation mutagenesis at the R124 and P134 residues that are directly involved in the DNA recognition of the TC part of the GATC site combined with a GANN substrate library. After selection, many clones were screened for BspEI protection. After sequencing and data analysis we identified that, GATT is the best non GATC recognition site. Among the clones methylating GATT sites, there is strong enrichment for the R124S and P134S exchanges. Then, we applied site saturation mutagenesis at R124 and P134 residues combined with fixed GATT target site. After screening, sequencing and data analysis we did not observe a better protection and also enrichment for the R124S and P134S exchanges. So, better enzymes could be only obtained after randomization of additional parts of the enzyme. We prepared a second generation library by whole gene randomization of the R124S/P134S variant on GATT substrate. After selection, several clones were screened and 6 clones showed an improved methylation at GATT sites (69 % protection at GATT site). Sequencing showed that two clones were identified twice. Overall, the 4 clones contained 18 different mutations in addition to the original R124S/P134S mutations. Then, we investigated the combinatorial effect of the mutations identified so far and prepared a third generation library by DNA shuffling. After selection, many clones were screened and 2 out of 11 showed improved methylation at GATT sites (78% protection at GATT site).

To investigate the methylation preferences of various EcoDam variants in *E.coli*, we expressed the mutant enzymes in dam⁻negative cells. Under these conditions, a second BspEI site on the plasmid that overlaps with a GATC site became available for cleavage. In dam⁺ positive cells, this site is always protected by the endogenous EcoDam. In dam⁻negative host, the EcoDam variants methylate either GATC or GATT sites depending on their intrinsic preferences. The specificity of methylation was analyzed by a double digestion experiment. So far, we evolved enzyme variants towards methylation of GATT

sites. Here, we evolve proteins against methylation of GATC sites while keeping the activity at GATT sites. To this end, mixture of genes of four mutants obtained from DNA shuffling library were used as a template for generating the fourth generation library using error-prone PCR. Proteins were expressed in dam negative cells, plasmid were isolated and digested using double digestion experiment. After screening, one variant showed an increased specificity as evidenced by the improvement of the ratio of the 4.2 and 5 kb bands (9.3%).

For *in vitro* studies, wild type enzyme, R124S/P134S and best variants from the third and fourth generations were purified and their DNA methylation kinetics using two oligonucleotide substrates containing one GATC or GATT site in an otherwise identical sequence context were determined. Detailed results are described in supplement 1.

III. B. Transition from EcoDam to T4Dam DNA recognition mechanism without loss of activity and specificity

In the second part of this thesis, we performed rational design of EcoDam in order to transplant the T4Dam DNA recognition module into EcoDam. These enzymes are highly related in sequence and structure and they methylate the adenine residue in GATC sites. However, they deviate in their contact to the first base pair of the target sequence. In EcoDam, Gua1 is contacted by K9 (A6 in T4Dam) whereas this contact is mediated by R130 (Y138 in EcoDam) in T4Dam. Our results show that the EcoDam K9A/Y138R double mutant as well as the evolutionary intermediate EcoDam Y138R are highly active and specific. The results illustrate that there exists a smooth evolutionary pathway changing the EcoDam DNA recognition mode to T4Dam mode without loss of activity and without generation of evolutionary intermediates with reduced activity. Detailed results are described in supplement 2.

III. C. Discrimination of unmethylated and hemimethylated DNA by *E.coli* DNA methyltransferase enzyme (EcoDam)

The main goal of the third part of this thesis was to generate an artificial epigenetic system that can propagate patterns of GATC site methylation potentially useful for gene

regulation. To achieve this goal, we replaced L122, P134 and V133 residues with other amino acids using site directed mutagenesis and the catalytic activity of all variants on the unmethylated and hemimethylated substrates as studied on a work was based on the observation that, the L122A EcoDam variant methylates only the hemimethylated substrate (Horton *et al.*, 2006).

Our results showed that the L122A and to some extent the L122A/V133L EcoDam variants are able to sense the methylation status of the GATC target recognition site and methylate only the hemimethylated DNA.

Horton *et al.*, 2006 have shown that the rate of methyl transfer by wild-type EcoDam with the unmethylated substrate is roughly twice as fast as with the hemimethylated substrate. Three, not mutually exclusive, lines of arguments might explain this finding: (1) The unmethylated substrate has twice the number of target sites as the hemimethylated one. (2) If the initial EcoDam binding is random with respect to the two strands, 50% of the binding events with the hemimethylated substrate will be unproductive, because EcoDam will be positioned such that the methylated adenine would be at the target position. (3) The methyl transfer step might be faster with unmethylated DNA. However, surprisingly, the L122A EcoDam variant was almost inactive on unmethylated DNA, while it modified the hemimethylated substrate with a rate similar to that of wild-type EcoDam. The mechanism of this pronounced change in the catalytic properties of L122A is not known.

With this property, the L122A variant may act as a GATC “maintenance enzyme”, that could propagate patterns of GATC site methylation, similarly as the Dnmt1 with its preference for hemimethylated CpG sites. Dnmt1 propagates CpG methylation patterns in mammals (Jeltsch 2008). However, previous studies showed that the preference of L122A for hemimethylated DNA is not sufficient to avoid methylation of unmethylated GATC sites in *E.coli* (S. Chahar, unpublished data).

The objective of this chapter was to enhance the preference of L122A for hemimethylated DNA while maintaining the activity on hemimethylated substrate. Such enzyme could be used to construct an artificial epigenetic system that could be useful for gene regulation. To achieve this goal, the L122 residue was substituted by other amino acids using site directed mutagenesis. In addition, the P134 and V133 residues were also

targets for site directed mutagenesis, because P134 is central for the discrimination of the TA base-pair (Horton *et al.*, 2005) and the side chain of V133 residue interacts with that of the L122 residue. Table (1) summaries all EcoDam variants used in this study.

Target position	Mutation	Combination
L122	L122V	
	L122T	
	L122I	
	L122S	
	L122G	
P134	P134A	L122A/P134A
V133	V133A	
	V133I	
	V133S	
	V133L	L122A/V133L

Table 1. EcoDam variants.

All mutants were generated by site directed mutagenesis. Successful mutagenesis and the absence of additional mutations were confirmed by sequencing of the complete EcoDam genes. The His₆-tagged wild type EcoDam and all variants generated here were expressed in *E.coli* and purified using Ni-NTA agarose to >95% purity as judged by SDS-polyacrylamide gel electrophoresis.

Catalytic activity of EcoDam and its variants on hemimethylated DNA substrate

After protein purification, we determined the DNA methylation activity of WT EcoDam and its variants using the hemimethylated DNA substrate (GATC-substrate, see materials section) in vitro. The hemimethylated DNA sites are the physiological substrate for EcoDam. The single turnover methylation rate constant of wild type EcoDam was 0.8 min⁻¹, which is similar to the values determined earlier (Horton *et al.*, 2006). The relative turnover rates of EcoDam and its variants are given in Fig. 1. The V133S and V133L showed strong reduction in activity. The activities of L122A, L122S, L122I, L122V, L122T, P134A, L122A/P134A and V133I were only slightly reduced or similar to WT. L122G, V133A and L122A/V133L even showed some increase in activity.

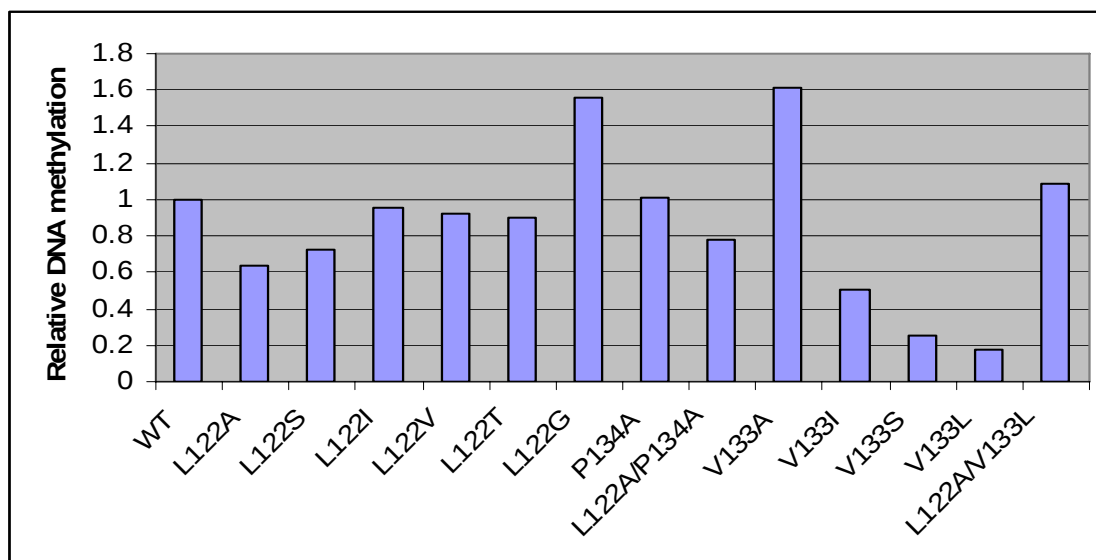


Fig. 1. Catalytic activity of WT EcoDam and its variants on the hemimethylated DNA substrate. Rates are given relative to the wild type enzyme activity

Kinetic characterization of the L122 substitution variants

We studied methylation of oligonucleotide substrates containing an unmethylated (Un-GATC substrate, see materials section) and hemimethylated GATC site to determine the activity of the variants and study their preference for the hemimethylated substrate. All experiments were carried out in triplicate and the initial slopes were averaged. The deviations of the relative mutant activities were below $\pm 20\%$ in all cases. The results of some EcoDam variants are shown in (Fig. 2 (a)). The L122 residue was substituted to Val, Ile, Ser, Thr, and Gly. The L122V, L122I, L122T, and L122S EcoDam variants had some preference for hemimethylated substrate (< 20 fold which was achieved by L122A variant, Horton *et al.*, 2006). However, the L122G variant had no preference for hemimethylated substrate and it was about 1.5 fold more active than the wild type enzyme on the hemimethylated DNA substrate. The relative DNA methylation activity of L122A and L122S on the hemimethylated DNA was dropped by about 35-40% when compared to that of wild type enzyme (Fig. 1). L122V, L122T and L122I showed roughly similar rates relative to wild type enzyme.

Kinetic characterization of the P134A EcoDam variant

Horton *et al.*, 2005 reported that the P134 residue contacts the third base pair of the GATC recognition sequence and that it is central to the recognition of the TA base pair (Horton *et al.*, 2005). Here, we substituted the P134 residue by Ala and determined the rate of DNA methylation activity of the P134A variant on both unmethylated and hemimethylated substrates. The variant showed some preference for the hemimethylated substrate and it displayed roughly similar activity on hemimethylated DNA as the wild type enzyme. At this stage we wanted to know what the preference for hemimethylated substrate will be if we combine P134A and L122A. The P134A/L122A variant had no preference for hemimethylated substrate and its activity dropped by about 20% relative to the P134A variant. So far, the molecular mechanism of why the L122A EcoDam variant methylates only the hemimethylated substrate is unclear. In this respect, the behavior of the P134A/L122A double mutant could not be explained.

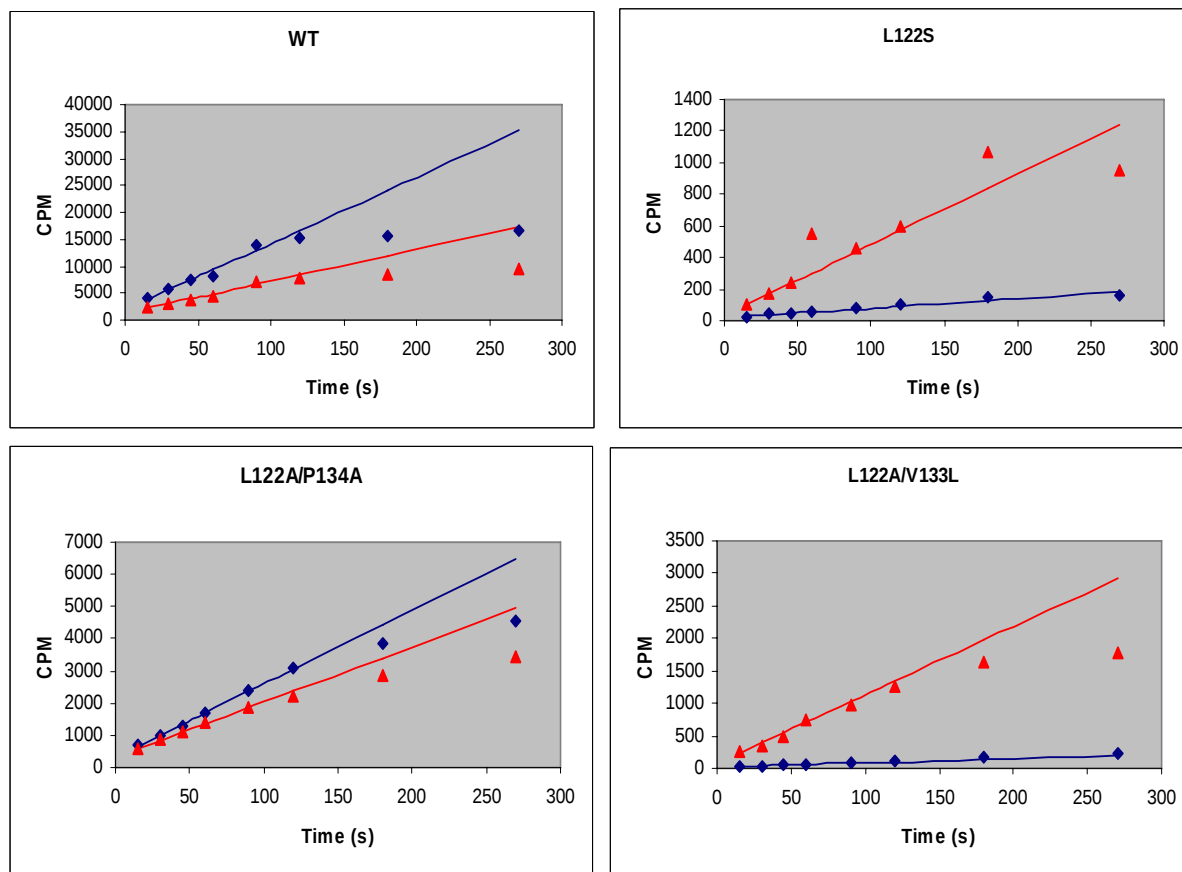
Kinetic characterization of the V133 substitutions variants

The crystal structure of EcoDam showed that the side chain of the L122 residue interacts with that of V133. Here we wanted to study if changing V133 to other amino acids will affect the preference for the hemimethylated substrate. We mutated V133 to Ala, Ser, Ile and Leu. The V133A, V133S, and V133I EcoDam variants had some preference for hemimethylated substrate and V133A was about 1.6 fold more active on hemimethylated DNA when compared with the wild type enzyme. Interestingly, V133L was almost inactive on the unmethylated substrate and it preferred the hemimethylated substrate 19 fold. However, the V133L variant showed an 80% reduced activity at hemimethylated sites (Fig. 1). At this stage, we wanted to see the effect of combining L122A and V133L. The L122A/V133L EcoDam variant showed some preference for the hemimethylated substrate but there was no big difference between it and the L122A variant (Fig. 2 (b)). The activity of L122A/V133L toward the hemimethylated substrate was roughly similar to that of the wild type enzyme.

In conclusion, by site –directed mutagenesis of L122, P134, and V133 residues, followed by quantitative determination of their catalytic activity in vitro, it was demonstrated that

L122A and to some extent L122A/V133L variants are able to sense the methylation status of the GATC target recognition site and methylate only the hemimethylated DNA. However, it was not possible to exceed the level of discrimination observed with L122A.

A.



B.

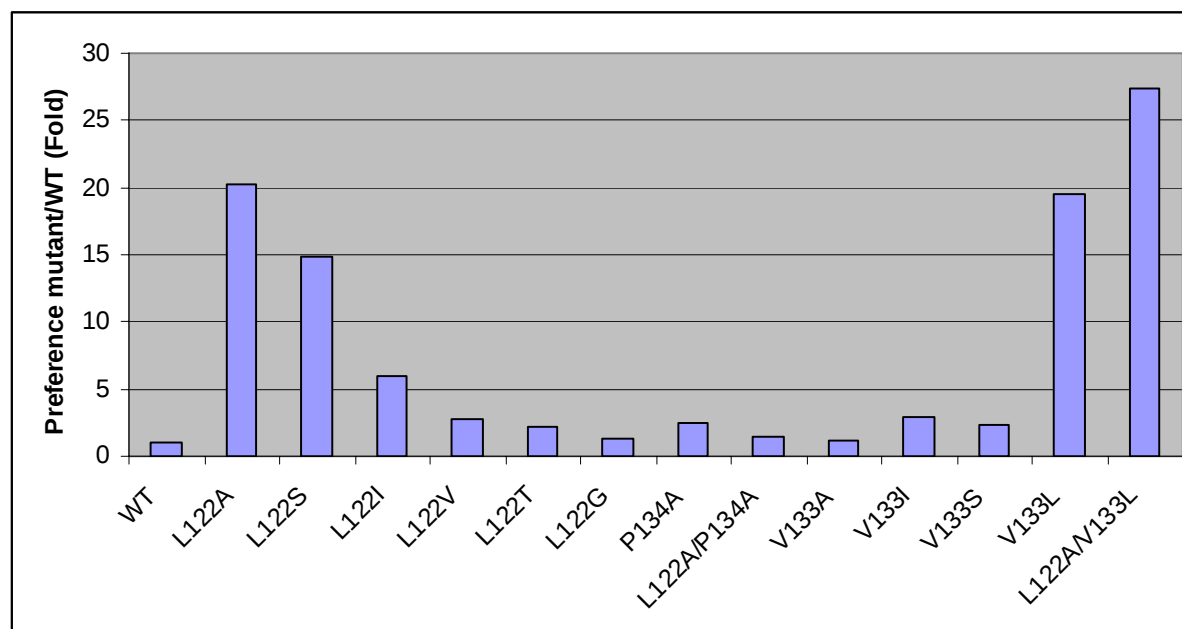


Fig. 2. **Discrimination between unmethylated and hemimethylated DNA.** A) Methylation of unmethylated (blue diamonds) and hemimethylated (red triangles) oligonucleotide substrates by the WT EcoDam, L122S, L122A/P134A and L122/V133 variants. B) Relative preference of variants for hemimethylated DNA when compared to wild type (Fold).

IV. Conclusions

Two complementary strategies of protein engineering are currently available: rational design and directed evolution. Although both approaches have been applied with great success, each has its limitations (Chen 2001).

Several features make DNA MTases an ideal model system for design projects aiming to change the specificity of enzymes (Roth & Jeltsch 2001). (1) All DNA MTases whose structures are known share one common fold that consists of two domains; catalytic and DNA binding domains. (2) Structure is available for some of them. (3) They share one mechanistic feature, namely that methylation of the target base is always preceded by flipping the base out of the DNA helix. The flipped base is bound into a hydrophobic binding pocket within the catalytic domain of the DNA MTase where catalysis takes place.

There are some previous efforts to change the specificity of DNA methyltransferases:

A) An attempt to alter the specificity of a bacteriophage methyltransferase by the fusion of two TRDs (Target recognition domain) with different DNA specificities led to an enzyme with a degenerate (relaxed) specificity that was not predicted from the specificities of the two parent TRDs (Lange *et al.*, 1996)

B) Roth and Jeltsch (Roth & Jeltsch 2001) used rational de novo protein design to change the substrate specificity of the EcoRV-(adenine N6)-methyltransferase to a DNA-(cytosine N4)-methyltransferase. The M.EcoRV variant, which had mutations in the conserved catalytic domain, displayed a 22-fold preference for cytosine as the target base, but only when located in a CT mismatch and with no increase in catalytic efficiency compared with methylation of the same substrate by the wild-type enzyme. In fact, the designed mutations abolish methylation of adenine without improving methylation of cytosine.

C) A library of mutants of M.SinI, target site GGWCC (W= A or T), was selected for the ability to methylate the degenerate site, GGNCC (N= A, C, G or T) (Kiss *et al.*, 2001). The best variant isolated by this method showed a more relaxed specificity at the central base, although the original target site was still methylated twice as efficiently as the new site and the methylation activity on both sites was 5 to 10-fold lower than the activity of the wild-type enzyme on the canonical site.

D) A directed evolution approach was successfully used to alter the sequence specificity of HaeIII methyltransferase using in vitro compartmentalization (IVC) (Cohen *et al.*, 2004). M.HaeIII methylates the internal cytosine of the canonical sequence GGCC, but there is promiscuous methylation of a variety of non-canonical sites, notably AGCC, at a reduced rate. Using IVC to select M.HaeIII libraries for the ability to methylate AGCC, they isolated a variant with a 670-fold improvement in catalytic efficiency ($k_{\text{cat}}/k_{\text{m}}^{\text{DNA}}$) on AGCC site.

E) Recently, Gerasimaite *et al.*, 2009 (Gerasimaite *et al.*, 2009) have used a structure-guided rational protein design combined with random mutagenesis and selection to change the specificity of the HhaI C5-MTase from GCGC to GCG. The specificity change was brought about by a five-residue deletion and introduction of two arginine residues within and nearby one of the target recognizing loops. DNA protection assays, bisulfite sequencing and enzyme kinetics showed that the best selected variant is comparable to wild-type M.HhaI in terms of sequence fidelity and methylation efficiency, and supersedes the parent enzyme in transalkylation of DNA using synthetic cofactor analogs. The designed C5-MTase can be used to produce hemimethylated CpG sites in DNA, which are valuable substrates for studies of mammalian maintenance MTases.

The basis of all screening and selection methodologies is a linkage between the gene and the enzyme it encodes. With DNA MTases, the nature of activity and specificity of an enzyme variant can be detected on the DNA coding for that particular protein. In this way, the linkage of phenotype (enzymatic properties) and genotype (DNA sequence), which is essential for the evolution of protein function, is readily available.

Directed evolution does not require information about how enzyme structure relates to function. This technique employs a random process in which error prone PCR or other mutagenesis methods are used to create a library of mutagenized genes. The sorted genes might be subjected to further cycles of mutation and screening to enhance the original beneficial mutation. Selection is different from screening. Selection involves transformation of the library into the host cell and then plating the transformants onto some sort of selective medium which only allows survival of the organism if the protein is present. It is much less time-intensive than screening, allowing for the analysis of 10^9 -

10^{11} colonies per round of selection. Additionally, selection depending on the method can be very powerful – less than 1 % activity over background may be sufficient for organism survival. However, due to the adaptive nature of the host organism, selection can sometimes result in false positives (i.e. the bacterium changes in ways not related to the evolved protein to prevent its death) and it merely provides a yes/no answer- catalytic rates must still be measured in a separate experiment. Finally, not all activities of interest can be tied to host cell survival.

Screening entails analyzing each colony for its level of the desired reaction. Screening involves three steps: the separation of individual colonies into their own assay compartment, the generation of signal (methylation), and the detection of the signal (restriction protection). The process is much more resource and time intensive, so only about 10^5 - 10^7 colonies can be screened in each round of evolution. The capacity limitation means that only a fraction of the generated library is screened in any round, making it likely that some highly active mutants may be missed.

In this study, selection of the desired mutants with altered specificity was performed from the mutant pool. In the cell, each enzyme variant modifies its own expression plasmid. The selection was performed by serial selection of those mutants that survive the restriction digestion by BspEI restriction enzyme. BspEI is an adenine methylation sensitive enzyme and its activity is blocked by methylation of the adenine within its sequence (TCCGGA). The mutant protein population was expressed in expression host cells and the propagated plasmids (carrying methylation at respective recognition sequences) were assessed for protection against BspEI. Only one round of selection was performed because we always observed the generation of plasmids that lost the restriction site. Afterwards we screened for the desired variants. To be sure that the restriction site was not lost, we performed PCR on the protected plasmids to amplify the restriction site followed by digestion and used only plasmids that their restriction sites were digested for the next round of design.

In this study, we successfully used a directed evolution approach to change the DNA recognition specificity of *E.coli* DNA adenine methyltransferase from GATC to GATT. One critical step for directed evolution is to have a starting point, where the desired

activity is already observed, at least in small amounts. In our case to change specificity from GATC to GATT we needed a starting enzyme with some activity towards the new target site. First we applied semi-rational design by using site saturation mutagenesis at R124 and P134 residues, which interact with the fourth and third base pair of the recognition sequence, respectively and obtained the R124S/P134S mutant which was the starting point of the directed evolution experiments. Four rounds of directed evolution (error-prone PCR, DNA shuffling and again error-prone PCR) were then used to create enzyme variants with new specificity for GATT.

One of the main limitations of directed evolution is the inability to exhaustively search the vast sequence space of protein variants. For a protein of typical size (300 amino acids), the number of variants containing three simultaneous mutations exceeds 10^{10} , which is too large to be screened experimentally. Conversely, it is also an overwhelming challenge for rational or computational design to identify multiple mutations needed for creation of a novel protein function. Thus, a combined directed evolution and rational design approach has been increasingly used to create novel protein functions because the engineering of some beneficial mutations can dramatically reduce the size of the sequence space to be explored in the design experiment.

In rational design, precise changes in amino acid sequence are made using a detailed knowledge of protein structure, function and mechanism. The substitutions are introduced using site-directed mutagenesis. We applied this strategy on EcoDam and T4Dam where the recognition of the first base pair is an interesting deviation between them. In EcoDam, Gua1 is contacted by K9 (which corresponds to A6 in T4Dam) whereas this contact is mediated by R130 in T4Dam (which corresponds to Y138 in EcoDam). In this study, we successfully swapped Gua1 recognition mode from T4Dam to EcoDam and showed that EcoDam K9A/Y138R double mutant is highly active and specific. The evolutionary intermediates of this transition were prepared and show that K9A has low activity and loss of the recognition of the first base pair whereas Y138R is fully active and specific. It has been reported that evolution go through intermediates of low or zero activity which is not true here, as observed by the behavior of Y138R variant. Our detailed understanding

of EcoDam and T4Dam structure and catalytic mechanism are reasons for this successful rational design.

However, we have an example of a failed attempt to design an enzyme as observed in our attempts to enhance the preference of L122A for hemimethylated DNA. What probably accounts for that failure is that the mechanism of L122A behavior is not known. One of the requirements for successful directed evolution experiment is the existence of a rapid screen or selection that reflects the desired function. Here, we are not able to apply directed evolution, because of lacking of screening system for preferential methylation of hemimethylated DNA in *E.coli*.

In summary, enzyme properties can be improved by rational design or directed evolution. Choosing the most effective approach for a particular enzyme-engineering task depends on the level that the mechanistic base of the desired property is understood, and if an effective selection scheme is available.

V. Materials and Methods

V. A. Materials

V. A. 1. Bacterial strains

HMS 174 (Novagen GmbH, Madison, USA)

Genotype

F⁻ *recA* 1 *hsdR*(*r*_{K12}⁻ *m*_{K12}⁺) *Rif*^r

XL-1 blue MRF⁺ (Stratagene)

Genotype

Δ (*mcrA*)183 Δ (*mcrCB*-*hsdsMR*-*mrr*)173 *endA*1 *supE44* *thi*-1 *recA* 1 *gyrA*96 *relA* 1 *lac*
[F⁺*proAB* *lacI*Z Δ M15 Tn10 (*Tet*^r)]

JM 110 (Stratagene)

Genotype

rpsL (*Str*^r) *thr* *leu* *thi*-1 *laY* *galK* *gal* *Tara* *tonA* *tsx* *dam* *dcm* *supE44* Δ (*lac*-*proAB*)
[F⁺*traD36* *proAB* *lac* Z Δ M15]

V. A. 2. Buffers and solutions

10x TPE Buffer

1M Tris, 225 mM phosphoric acid, 20 mM EDTA, pH 8.2

Loading dye solution for DNA electrophoresis analysis

250 mM EDTA, 25% sucrose, 1.2% SDS, 0.1% bromophenol blue, 0.1% xylancyanol FF,
pH 8.0

dNTPs stock solution

Mixture of 10 mM each dNTP (N=A, C, G, or T) in H₂O, stored at -20°C

10 x STE buffer

100 mM Tris-HCL (pH 8.0), 1 mM EDTA, 1 M NaCl

10 x PBST buffer

140 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.4 mM K₂HPO₄, 0.05% Tween 50, pH
9.6

10 x Serratia buffer

50 mM Tris-HCl pH 8.0, 0.5 mM MgCl₂

EcoDam protein purification buffers:

1. Washing buffer

20 mM Hepes, 0.1 mM DTT, 500 mM NaCl, 10% glycerol, 20 mM imidazole, pH 8.0

2. Elution buffer

20 mM Hepes, 0.1 mM DTT, 500 mM NaCl, 200 mM imidazole, 10% glycerol, pH 8.0

3. Dialysis buffer I

20 mM Hepes, 0.1 mM DTT, 1 mM EDTA, 300 mM NaCl, 10% glycerol, pH 7.5

4. Dialysis buffer II

20 mM Hepes, 0.1 mM DTT, 1 mM EDTA, 300 mM NaCl, 77% glycerol, pH 7.5

SDS- running buffer

250 mM Tris-base, 1.9 M glycine, 1% SDS, pH 8.3

SDS- loading buffer

2% SDS, 5% mercaptoethanol, 40% glycerol, 0.1 % bromophenol blue, 160 mM Tris-HCl pH 6.8

Ethidium bromide stock solution

10mg/ml ethidium bromide in H₂O

APS stock solution

40% Ammonium persulfate, stored at -20°C

Coomassie stain

0.1% Coomassie Brilliant Blue G-250, 2% phosphoric acid, 5% aluminum sulfate, 10% ethanol

Dam methylation buffer

100 mM Hepes pH 8.0, 1 mM EDTA, 50 mM NaCl, 0.5 mM DTT

50x TAE buffer

2 M Tris-base, 2 M acetic acid, 50 mM EDTA, pH 8

10 x TBE buffer

1 M Tris-base, 1 M boric acid, 25 mM EDTA, pH 8.3

V. A. 3. Chemicals and kits

V. A. 3. 1. Chemicals

Chemical	Company
Isopropyl- β -D-thiogalactopyranoside (IPTG)	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
Acrylamide 4k 40% (29:1)	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
S-Adenosyl-L-methionine-p-toluenesulfonate salt (SAM)	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
Ammonium persulfate (APS)	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
1,4-Dithiothreitol (DTT)	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany

[³ H] S-adenosyl-L- methionine	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
Sodium chloride	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
Tetracycline hydrochloride	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
TritonX-100(Octylphenol-polyethelenglycol)	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
Sinefugin	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
Avidin	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
Bovine serum albumin	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
Manganese chloride tetrahydrate	Sigma-Aldrich chemie GmbH P.O Steinheim, Germany
LB medium	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Agar bacteriological grade	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Kanamycin sulfate	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Agarose electrophoresis grade	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Boric acid	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Bromophenol blue sodium salt	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Chloramphenicol	Carl Roth GmbH & CO KG, Karlsruhe, Germany
EDTA	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Ethidium bromide	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Ethanol crude 99%	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Glycine	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Isopropanol	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Tris ultra pure	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Xylencyanol FF	Carl Roth GmbH & CO KG, Karlsruhe, Germany
Ampicilin sodium salt	Fluka, Sigma-Aldrich GmbH P.O. Steinheim, Germany
Coomassie Brilliant Blue G-250	Applichem GmbH, Darmstadt, Germany
β- Mercaptoethanol	Applichem GmbH, Darmstadt, Germany
Methanol	Applichem GmbH, Darmstadt, Germany
Sodium dodecyl sulfate (SDS)	Applichem GmbH, Darmstadt, Germany
Sulfuric acid 95-97%	Amersham
Nickel NTA Agarose	Qiagen, Germany
Glycerol, 87%	Macherey-Nagel GmbH& Co, Dueren, Germany.
Ortho phosphoric acid 85%	
2'-Dideoxyribonucleoside-5-triphosphate (dATP, dCTP, dGTP, dTTP)	NEB, New England Biolabs GmbH, Frankfurt- Main- Germany
N,N,N',N'-Tetramethylethylenediamine (TEMED)	Gibco, Invitrogen GmbH, Karlsruhe, Germany
Tween20-(Polyoxyethylen-sorbitan-Monolaureate)	PAA
[γ- ³² P] ATP	Hartmann Analytic GmbH, Braunschweig, Germany
Hepes	Promega GmbH, Mannheim, Germany

V. A. 3. 2. Kits

Kits	Company
Plasmid DNA purification kit	Macherey-Nagel GmbH & Co, Dueren, Germany
NucleoSpin® Gel extraction kit	Macherey-Nagel GmbH & Co, Dueren, Germany
QIAquick Gel Extraction kit	Qiagen GmbH, Germany
QIAquick PCR purification kit	Qiagen GmbH, Germany

V. A. 4. Enzymes

Enzymes	Company
DNase I	NEB New England Biolabs, Frankfurt, Main, Germany
BspEI	NEB New England Biolabs, Frankfurt, Main, Germany
DpnI	NEB New England Biolabs, Frankfurt, Main, Germany
Taq DNA polymerase	NEB New England Biolabs, Frankfurt, Main, Germany
Lambda exonuclease	NEB New England Biolabs, Frankfurt, Main, Germany
BglII	NEB New England Biolabs, Frankfurt, Main, Germany
<i>Pfu</i> DNA polymerase	MBI Fermentas

V. A. 5. Media and antibiotics**V. A. 5. 1. Media**

LB (Luria-Bertani) medium

20 g LB medium in 1L H₂O and autoclave to sterilize.

SOC medium

2% Bactotryptone, 0.5% Yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose and autoclave to sterilize (Sambrook *et al.*, 1989).

LB agar

1.5 g agar in 100 ml LB medium and autoclave to sterilize.

V. A. 5. 2. Antibiotics

Kanamycin stock solution

25 mg/ml, 2.5 g kanamycin was dissolved in 100 ml autoclaved ddH₂O and filtered using 0.2 µm filter. 1 ml aliquots were prepared and stored at -20°C

Tetracyclin stock solution

12 mg/ml, 0.6 g tetracycline was dissolved in 50 ml 50% ethanol (needn't further sterile). 1 ml aliquots were prepared and stored at -20°C

Streptomycin stock solution

10 mg/ml, 1 g streptomycin was dissolved in 100 ml autoclaved ddH₂O and filtered using 0.2 µm filter. 1 ml aliquots were prepared and stored at -20°C

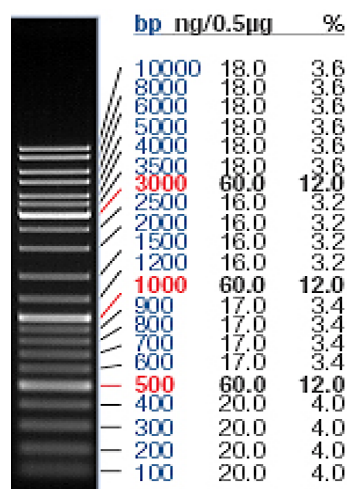
For the preparation of selective media, sterilized antibiotic stock solution is added to the autoclaved medium. The stock solution's concentration and final concentrations are:

Antibiotic	Stock Conc.	Final Conc.
Kanamycin	25 mg/ml	25 µg/ml
Tetracycline	12 mg/ml	10 µg/ml
Streptomycin	10 mg/ml	50 µg/ml

V. A. 6. Molecular weight markers

V. A. 6. 1. DNA molecular weight marker

The DNA size marker was mixed in 6x Loading Dye solution (MBI, Fermentas) to the final concentration of 0.083 µg/µl and stored at 4°C.



DNA molecular weight marker

(<http://www.fermentas.com/catalog/electrophoresis/generulers.htm#Mix>)

V. A. 6. 2. Protein molecular weight marker

It is a mixture of 7 purified proteins that resolve into sharp bands in the range of 14.4 kDa to 116.0 kDa when analyzed by SDS-PAGE and stained with Coomassie Brilliant Blue.

Protein	Source	MW, kDa
β- galactosidase	<i>E.coli</i>	116.0
Bovine serum albumin	bovine plasma	66.2
Ovalbumin	chicken egg white	45.0
Lactate dehydrogenase	Porcine muscle	35.0
Restriction endonuclease Bsp98I	<i>E.coli</i>	25.0
β- lactoglobulin	bovine milk	18.4
Lysozyme	chicken egg white	14.4

V. A. 7. Primers for EcoDam gene amplification

Dam-nested-fr Primer

5'-TTA ACT TTA AGA AGG AGA TAT ACC-3'

Dam-nested-rev Primer

5'-TTA GAG GCC CCA AGG GGT TAT G-3'

Primers for gene shuffling

Shuffling (pet)-fr Primer

5'-GCG CGG CAG CCA TAT GAA GA-3'

Shuffling (pet)-rev Primer

5'-GTG CCG CCG TTG CTG CTT A-3'

Primers for BspEI site amplification from pET-28a vector

SubSeq-fr Primer

5'-CTT TGT ACA AAC CAG GAG TCG TTT CAC C-3'

SubSeq-rev Primer

5'-CGA GAT AGG GTT GAG TGT TGT TCC AG-3'

V. A. 8. Dam substrates

Oligonucleotides were purchased from Thermo Hybaid (Ulm, Germany) in HPLC-purified grade. The concentrations of the oligonucleotides were determined from the absorbance at 260 nm using the extinction coefficients provided by the supplier. The following double-stranded oligodeoxynucleotide substrates obtained by annealing of complementary single-stranded oligonucleotides were used (**bold**, underlined = recognition site; M = N⁶-methyladenosine and Bt = Biotin).

To anneal double-stranded substrates, equimolar (20 μM) amounts of each strand were mixed, heated to 95°C for 5 min, and slowly cooled to room temperature.

GATC substrate	5`-GCGACAGT <b><u>GATC</u></b> GGCCTGTC-3` 3`-CGCTGTCA <b><u>CTMG</u></b> CCGGACAG-Bt-5`
Un-GATC substrate	5`-GCGACAGT <b><u>GATC</u></b> GGCCTGTC-3` 3`-CGCTGTCA <b><u>CTAG</u></b> CCGGACAG-Bt-5`
GATT substrate	5`-GCGACAGT <b><u>GATT</u></b> GGCCTGTC-3` 3`-CGCTGTCA <b><u>CTMA</u></b> CCGGACAG-Bt-5`
AATC substrate	5`-GCGACAGT <b><u>AATC</u></b> GGCCTGTC-3` 3`-CGCTGTCA <b><u>TTMG</u></b> CCGGACAG-Bt-5`
TATC substrate	5`-GCGACAGT <b><u>TATC</u></b> GGCCTGTC-3` 3`-CGCTGTCA <b><u>ATMG</u></b> CCGGACAG-Bt-5`
CATC substrate	5`-GCGACAGT <b><u>CATC</u></b> GGCCTGTC-3` 3`-CGCTGTCA <b><u>GTMG</u></b> CCGGACAG-Bt-5`

V. A. 9. Vectors

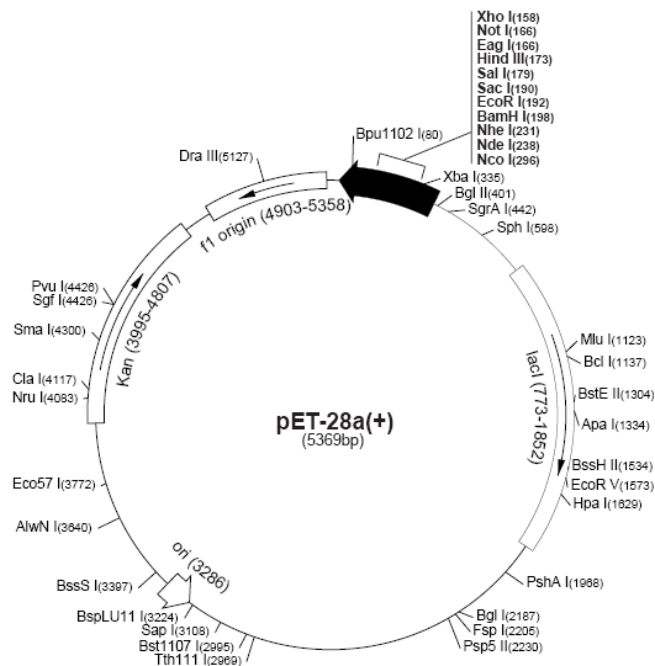
pET 28a

The pET-28a vector carries an N-terminal His-Tag/thrombin/T7Tag configuration plus an optional C-terminal His-Tag sequence. Unique sites are shown on the circle map. The sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circular map. The fl origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, single-stranded sequencing should be performed using the T7 terminator primer.

pET-28a(+) sequence landmarks

T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites (<i>Bam</i> H I - <i>Xho</i> I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
<i>lacI</i> coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
f1 origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-28c(+) is a 5367bp plasmid; subtract 2bp from each site beyond *Bam*H I at 198.



(Novagen, manual support)

(<http://www.merckbiosciences.co.uk/html/NVG/alltables.html>)

V. B. Methods:

V. B. 1. Microbiological methods

All media and heat-stable solutions were sterilized by autoclaving for 20 min at 121°C. The heat sensitive solutions were sterilized by filtering through a 0.2 µm membrane filter (Sartorius). Glassware was heated for 20 min at 160 °C.

V. B. 1. 1. Culture and storage of *Escherichia coli* strains

E.coli strains were incubated overnight in LB medium at 37°C with shaking at 150-200 rpm. The *E.coli* cells were cultured on LB-agar plates overnight at 37°C. These plates could be kept as working stock plates at 4°C for up to 3-4 weeks. For the selection for plasmids with specific resistance, appropriate antibiotics were added into the medium.

V. B. 1. 2. Growth curve

The growth of *E.coli* cells includes several distinct phases: lag phase (0-2 h after dilution), the logarithmic (log) phase (3-5 hours after dilution), stationary phase (~ 16 hours after dilution) and decline phase. Eventually the culture enters the phase of decline as cells start to lyse, the number of viable bacteria falls, and DNA becomes partly degraded.

The growth curve of a bacterial culture can be monitored photometrically by reading the optical density at 600 nm. High OD₆₀₀ readings are calculated by diluting the sample in culture medium to enable photometric measurement in the linear range between 0.1-0.5 OD₆₀₀.

V. B. 1. 3. Preparation of log-phase *E.coli* cells and protein induction

E.coli cells from glycerol stock or working stock plate were inoculated into 3 ml LB medium in a sterile tube, then cultured at 37°C on culture rotator overnight.

The overnight culture was diluted into LB medium in Erlenmeyer flasks at the ratio of 1:100, incubated at 37°C with shaking at 150-200 rpm until the optical density at a wavelength of 600 nm (OD₆₀₀) reached the value of 0.4-0.6 (about 2-3 h.). This value indicated the log-phase *E.coli* culture.

For protein induction, IPTG was added to a final concentration of 1 mM and the cells were further cultured at 37°C for 3 h. The culture was centrifuged at 4200 rpm (Heraeus Multifuge 3) for 15 min at 4°C to harvest the cells.

V. B. 1. 4. Transformation

Transformation is the genetic alteration of a cell by the introduction, uptake and expression of foreign DNA. Competent cells are cells which are capable of uptaking extra-chromosomal DNA such as plasmids.

There are two main methods for preparation of competent bacterial cells, the calcium chloride and the electroporation method (Dagert & Ehrlich 1979; Okamoto *et al.*, 1997; Topcu 2000). Electroporation is a highly efficient transformation method.

V. B. 1. 4. 1. Preparation of CaCl₂ treated competent cells for the heat-shock transformation

For single plasmid transformation, heat-shock of calcium-chloride treated competent cells with appropriate DNA is a convenient method, as it needs no special device, and multitransformations can be carried out in parallel.

1. Pick a single bacterial colony (2-3 mm in diameter) from a plate that has been streaked with appropriate strain and incubated for 16-20 h. at 37°C, transfer into 10 ml LB medium and incubate on culture rotator overnight at 37°C.
2. Dilute 2000 µl of the overnight culture into 200 ml of fresh LB medium (1:100) in a 1 L flask. Incubate for 2-3 hours at 37°C with shaking at 150-200 rpm until the culture reached an OD₆₀₀ of 0.45-0.55.
3. Transfer the culture into 50 ml Falcon-tube and centrifuge it at 4200 rpm for 15 min at 4°C. Decant the supernatant and resuspend the cell pellet in 40 ml of sterile, ice-cold 20 mM CaCl₂-80 mM MgCl₂ solution. Incubate on ice for 60 min.
4. For each 50 ml of original culture, centrifuge and resuspend the pellet by gentle swirling in 1-2 ml of sterile, ice-cold 100 mM CaCl₂ containing 10 % glycerol.
5. Transfer 200 µl competent cells to chilled eppendorf tube (e-tube) for the plasmid transformation and store at – 70°C.

V. B. 1. 4. 1. 1. Heat-shock transformation of the chemically competent cells

For the heat-shock transformation, add 10-500 ng DNA (ca. 10 ng of supercoiled plasmid, or 500 ng ligation-reaction) to 200 µl aliquots of the CaCl₂-treated competent cell suspension.

Incubate the mixture of DNA and competent cells on ice for 20 min, subject it to 2 min heat shock at 42°C, and then put back on ice for 2 min to reduce damage to the *E.coli*.

Immediately add 1 ml of fresh LB or SOC medium to each sample and recover the cells by incubating at 37°C with shaking at 150-200 rpm for 60 min.

Plate out 50-250 µl transformants on the LB agar plates containing appropriate antibiotic to select the transformed bacteria.

V. B. 1. 4. 2. Preparation of the electro-competent cells

1. Streak out 10 µl from frozen glycerol stock of appropriate bacterial strain onto LB agar plate, incubate it for 12-16 hours at 37°C. Single colony was transferred into 50 ml LB medium and incubated on rotator overnight at 37°C.
2. Dilute 5 ml of the overnight culture into 500 ml LB medium (1:100) in 2L flask and incubate at 37°C with shaking at 150-200 rpm until the OD₆₀₀ reaches 0.45-0.55 (about 2.5-3.5 hours).
3. Chill the culture at 0°C for 15 minutes and centrifuge (Beckman Avanti High-Speed Centrifuge with JLA-8.1000 rotor, USA) at 2000 rpm at 4°C for 20 min to pellet the cells.
4. Wash the pellet twice with 500 ml ice-cold sterile H₂O.
Each wash step includes: re-suspend the cells; incubate cells on ice for 15 minutes; centrifuge, and discard the supernatant.
5. Gently resuspend the pellet, from each 500 ml culture, in 40 ml ice-cold sterile 10% glycerol and centrifuge at 4200 rpm, at 4°C for 20 minutes.
6. Finally resuspend the cells in 1-2 ml ice-cold sterile 10% glycerol. Freeze them by placing in crushed dry ice/ethanol (96%) bath and store at – 70°C until needed.

V. B. 1. 4. 2. 1. Transformation of the electro-competent *E.coli* cells (Electroporation)

1. Mix 100 µl of the prepared electro-competent *E.coli* cells with 10-500 ng DNA in pre-chilled e-tube and transfer the mixture into a pre-chilled electroporation cuvette (gap width 1 mm). Keep these materials on ice.
2. Set the *E.coli* pulser (Bio-Rad, Muenchen, Germany) at voltage 1.8 kV for 4 mS (for 2 mm gap cuvette, set the voltage to 2.5 kV).
3. Place the electroporation cuvette, with the cell and DNA mixture, into the holder of the cell-pulser and pulse the cells till the gene pulser beeps.

4. Immediately add 1 ml LB or SOC medium into the transformed cells and incubate for 60 min at 37°C with shaking at 150-200 rpm.
5. Plate the cell suspensions onto plates containing appropriate antibiotic to select the transformed bacteria.

V. B. 2. Molecular biology methods

V. B. 2. 1. DNA concentration determination

The DNA concentration can be determined by spectrophotometric measurement of the absorption in UV- range. The readings were taken at 260 nm and 280 nm. The OD at 260 nm allows calculation of the concentration of nucleic acid in the sample. An OD of 1 corresponds to 50 µg/ml for double stranded DNA.

The purity of the sample can be estimated from the ratio of OD₂₆₀ to OD₂₈₀. As proteins have maximum absorption at 280 nm because of the existence of aromatic rings of Trp. and Tyr. A protein-free DNA preparation usually has a relative coefficient of OD₂₆₀ to OD₂₈₀ above 1.8.

In this study, DNA concentration was estimated by using the NanoDrop ND-1000 spectrophotometer.

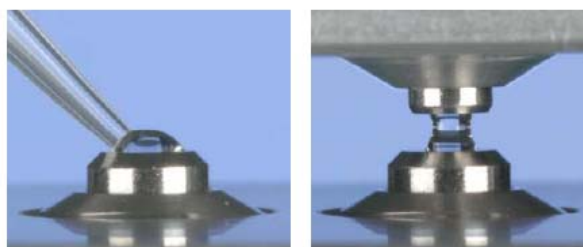


Figure 1a: Loading 1ul sample.

Figure 1b: Retention system.

(Application note, NanoDrop Technologies, Wilmington, Delaware USA)

<http://www.biocompare.com/review/284/NanoDrop-Technologies-Nanodrop-ND-1000-UVVisible-Spectrophotometer.html#links>

By using fiber optic technology and the inherent surface tension properties of liquid samples, NanoDrop Technologies microvolume instrumentation can accurately quantitate a wide range of biomolecules in volumes as small as 1 µl. The patented sample retention system enables absorbance measurements to be performed without traditional containment devices such as cuvettes or capillaries.

The system uses liquid surface tension to hold a droplet of sample in place between two optical surfaces during the measurement cycle. In order to make a measurement, 1 μ l of DNA sample was pipetted directly onto the lower optical (measurement) surface (Figure 1a). An upper optical surface automatically engaged the sample to form a liquid column of mechanically-controlled path length (Figure 1b). Once the measurement was complete, both optical surfaces were simply cleaned with a standard laboratory wipe to prepare for the next sample.

V. B. 2. 2. DNA purification using silica membrane spin columns

Plasmid DNA was isolated using a NucleoSpin plasmid kit, Macherey-Nagel (Dueren, Germany) according to the instructions of the supplier.

In principle, DNA is bound to the silica membrane spin column in the presence of a high concentration of chaotropic salt, contaminants are washed away, and the DNA is then eluted from the silica membrane in water or a low-salt buffer.

The major advantage of silica membrane spin columns is the fact that the silica is bound to a solid support, which eliminates the problem of glass-bead contamination of the DNA sample. This method is quick and convenient, and can produce a high yield of pure DNA.

V. B. 2. 3. PCR-product purification

DNA purification was carried out according to the manual of a NucleoSpin Extract II kit. With the NucleoSpin Extract II method, DNA binds to a silica membrane in the presence of chaotropic salt. The binding mixture is loaded directly onto NucleoSpin Extract II columns. Impurities are removed by a simple washing step with ethanolic buffer. Pure DNA is finally eluted under low ionic strength conditions with slightly alkaline buffer (5 mM Tris/HCl, pH 8.5).

V. B. 2. 4. Extraction of DNA from agarose gel

In order to isolate a desired DNA fragment after PCR or restriction enzyme digestion, agarose gel electrophoresis was carried out. The band corresponding to the desired DNA fragment was cut using a scalpel on a UV-lamp table. To avoid damage of DNA by UV this step should be done very quickly. The DNA fragment was extracted as described in the manual of QIAquick Gel Extraction kit.

V. B. 2. 5. DNA digestion with restriction enzymes

Restriction digestion is the process of cutting DNA molecules with special enzymes called Restriction Endonucleases (or Restriction Enzymes, REs). Restriction Enzymes recognize specific sequences in the DNA molecule (for example GATATC) wherever that sequence occurs in the DNA. The RE's activity is measured in Units where 1 Unit is the amount of enzyme needed to digest 1 µg of DNA in 1 hour. REs are often shipped at 10 Units/µl.

Digestion of DNA was performed in its respective buffer and temperature recommended by the supplier. In general, about 0.5-5 µg double stranded DNA (~5000bp) was digested with 2-20 units of a restriction endonuclease. When reaction conditions for two restriction enzymes are similar, digestion with both enzymes can be carried out simultaneously (double digestion). The amount of DNA and RE used depends on the task. A general example of components in RE digestion reaction is given in table V. B. 1.

Table V. B. 1. General composition of restriction digestion

DNA	0.5-5 µg (ca. 1-10 µl)
10x RE Buffer	2-20 µl
Restriction endonuclease/s	2-20 Units
H ₂ O	x µl
Total volume	20-200 µl

Mix the ingredients gently, incubate the mixture at the recommended temperature. The reaction time depends on the activity of REs, amount of RE and DNA used, normally, 1.5- 2 h. Completeness of the reaction can be judged by size or pattern of DNA bands on agarose-gel. To cut the restriction sites of PCR products a slight excess of enzyme or a longer incubation time can be applied to ensure complete digestion. The digested DNA products then are separated and purified by agarose gel electrophoresis for further experiments.

V. B. 2. 6. Agarose gel electrophoresis of DNA

Agarose gels were prepared by mixing 1x TPE buffer with agarose powder to the desired concentration. The solution was heated in a microwave oven for 3-5 min. until it

completely melted. The dissolved agarose was poured into a casting tray containing a sample comb and allowed to solidify at room temperature.

After the gel has solidified, the comb was removed, using care not to rip the bottom of the wells. The gel, still in its plastic tray, was inserted horizontally, into the electrophoresis chamber and just covered with 1x TPE buffer. Samples containing DNA mixed with loading dye were then pipeted into the sample wells, the lid and power leads were placed on the apparatus, and a current was applied.

When adequate migration has occurred, DNA fragments were visualized by staining with ethidium bromide. To visualize DNA, the gel is placed on a ultraviolet transilluminator and images were captured using Biometra camera.

V. B. 2. 7. Polymerase chain reaction (PCR)

Polymerase Chain Reaction (PCR) is a repetitive bidirectional and exponential DNA synthesis via primer extension of a region of nucleic acid. Amplification of a DNA template requires two oligonucleotide primers, four deoxynucleotide triphosphates (dNTPs) and DNA polymerase to perform the synthesis.

V. B. 2. 7. 1. Composition of the PCR mixture

A. Template DNA

Usually the template DNA amount is in the range of 50 pg-1 ng for plasmid and 0.1-1 µg for genomic DNA, for a total reaction mixture of 50 µl.

B. Primers

PCR primers are usually 20-30 nucleotides in length. Longer primers provide sufficient specificity. The GC content should be 40-60 %. More than three G or C nucleotides at the 3'-end of the primer should be avoided, as nonspecific priming may occur. The primer should not be self-complementary or complementary to any other primer in the reaction mixture, in order to avoid primer-dimer and hairpin formation. The melting temperature of flanking primers should not differ by more than 5 °C, so the GC content and length must be chosen accordingly.

If the primer is shorter than 25 nucleotides, the approximate melting temperature (T_m) is calculated using the following formula:

$$T_m = 4(G + C) + 2(A + T)$$

G, C, A, T –number of respective nucleotides in the primer

If the primer is longer than 25 nucleotides, the interactions of adjacent bases, the influence of salt concentration etc. should be evaluated. The melting temperature may be calculated using the following formula:

$$T_m [^{\circ}\text{C}] = 69.3 + 0.41 (\% \text{ GC} - 650 / \text{length of nucleotide})$$

Optimal annealing temperature is generally 5°C lower than the melting temperature of the primer-template DNA duplex.

C. dNTPs

The concentration of each dNTP in the reaction mixture is 200 µM. It is very important to have equal concentration of each dNTP (dATP, dCTP, dGTP and dTTP), as inaccuracy in the concentration of even a single dNTP dramatically increases the misincorporation level. Concentration of dNTPs (100-250 µM each) result in optimal balance between product yield (greater at higher dNTP concentration) and specificity.

V. B. 2. 7. 2. Temperature Cycling

A. Initial Denaturation Step

The complete denaturation of the DNA template at the start of the PCR reaction is of key importance. Incomplete denaturation of DNA results in inefficient utilization of template in the first amplification cycle and a poor yield of PCR product. The initial denaturation should be performed over an interval of 1-3 min at 95°C if the GC content is 50% or less, and this interval should be extended for up to 10 min for GC-rich templates or denaturation temperature may be increased up to 97°C.

B. Denaturation Step

Usually 0.5-2 min denaturation at 94-95°C is sufficient, since the PCR product synthesized in the first amplification cycle is significantly shorter than the template DNA and is completely denatured under these conditions.

The temperature of the denaturation step also depends on the DNA polymerase used: *Taq* polymerase being less thermostable, *Pfu* polymerase –more thermostable (Dieffenbach and Dveksler 1995).

C. Primer Annealing Step

Usually the optimal annealing temperature is 5°C lower than the temperature of primer-template DNA duplex. Incubation for 0.5-2 min is usually sufficient. If non-specific PCR products are obtained in addition to the expected product, the annealing temperature should be optimized by increasing it stepwise by 1-2 °C.

D. Extension Step

Usually the extension step is performed at 70-75°C. *Pfu* DNA polymerase exhibits lower extension rate than *Taq* DNA Polymerase (0.5kb/min versus 1kb/min) so 2 min extension time is recommended for every 1 kb to be amplified.

E. Number of Cycles

The number of PCR cycles depends on the amount of template DNA in the reaction mixture and on the expected yield of the PCR product. For most amplification reactions, 25-35 cycles are usually sufficient. In general, using fewer cycles can ensure less amount of non-specific background product.

F. Final Extending Step

After the last cycle, the samples are usually incubated at 72°C for 5-10 min to fill-in the protruding ends of newly synthesized PCR products (Skerra 1992).

There are a lot of variations of PCR method, differing in the type of DNA Polymerase used, primers, templates as a starting material and desired products to be obtained, in the following section are methods used in this work.

Amplification PCR

A general composition of PCR reaction used in this work was shown in Table (V. B. 2) ; the amount of reaction was set up to 50 µl. Table (V. B. 3) is a general program for PCR cycling.

Table V. B. 2. General composition of PCR reaction

Reagent	Final concentration
Template DNA	50 ng
dNTPs	200 µM
Primers	0.4 µM each
Reaction buffer	1x
DNA polymerase	0.5-2.5 U
PCR-grade H ₂ O	x µl

Table V. B. 3. Cycling profile for a typical amplification PCR

Step	Segment	Time	Temperature (°C)
1	Initial denaturation	1-3 min	94-96
2	Denaturation	30-45 sec.	92-95
3	Annealing	30sec.-1 min	45-65
4	Extension	1-2 min	72
5	Cycle step 2-4	25-30 times	
6	Final extension	2-5 min	
7	Final hold	∞	8

V. B. 2. 8. Site directed mutagenesis (Jeltsch & Lanio 2002)

Site-directed mutagenesis is a technique in which a mutation is created at a defined position in DNA molecule with known sequence. It is carried out by using two PCR reactions. In the first reaction, the mutation, together with a marker restriction site, is introduced by a PCR primer into a first PCR product. This first product is employed as a megaprimer in a second PCR reaction using the circular WT plasmid as template to amplify a linear DNA fragment with large, complementary, single-stranded ends, whose size corresponds to the length of the megaprimer. These ends can hybridize to form stable circular DNA molecules, which contain the mutation and a nick in each DNA strand. To avoid unexpected mutations in the plasmid, *Pfu* polymerase is used in PCR because of its higher accuracy compared to *Taq* polymerase. Because the newly synthesized DNA strands are based on the mutagenic primers, a mutated plasmid is produced. In order to remove the wild type plasmid, the second PCR product is treated with DpnI, which digests DNA specifically methylated at GATC sites. DNA isolated from *E.coli* is methylated, in contrast to the synthesized DNA. Finally, the newly synthesized, mutated DNA plasmid is transformed into *E.coli* XL1-Blue cells.

The components of the reaction mixture were pipetted together according to the following scheme:

Component	Volume (µl)	Final concentration
10 x polymerase buffer	5	1 x
dNTPs (10 mM)	1	0.2 mM
Forward primer (4 µM)	5	0.4 µM
Reverse primer (4 µM)	5	0.4 µM
DNA template (5 ng/µl)	1	
<i>Pfu</i> polymerase	1	2.5 U
Add H ₂ O to final volume of 50 µL		

The first PCR reaction was carried out in a thermal cycler using the following program:

Step	Segment	Time	Temperature (°C)
1	Initial denaturation	3 min	95
2	Denaturation	30 sec.	95
3	Annealing	30 sec.	62
4	Extension	1 min	72
5	Cycle step 2-4	30 times	
6	Final extension	5 min	72
7	Final hold	∞	8

V. B. 2. 9. Error prone PCR (Cadwell & Joyce 1992)

Error-prone PCR is the most commonly used random mutagenesis method, which introduces random mutations during PCR by reducing the fidelity of DNA polymerase. The fidelity of DNA polymerase can be reduced by adding manganese ions or by biasing the dNTP concentration. Use of the compromised DNA polymerase causes mis-incorporation of incorrect nucleotides during the PCR reaction, yielding randomly mutated products.

Error-prone PCR was carried out in 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 7 mM MgCl₂, 0.5 mM MnCl₂, 0.2 mM dATP, 0.2 mM dGTP, 1 mM dCTP, 1 mM dCTP, 0.4 μM each primer, 1 pg Template DNA and 2.5 U Taq DNA polymerase in a total volume of 50 μl. PCR was carried out using a program of 95°C for 3 min, 30 cycles of 94°C for 45 s; 60°C for 35 s, and 72°C for 3 min, and a final incubation at 72°C for 10 min.

To convert the product to a suitable form for transformation of a host strain, at least three steps are required: digestion of the product with restriction enzymes, separation of the fragments by agarose gel electrophoresis and ligation into a vector.

Although these steps do not constitute special techniques, they require almost an entire day of handling time. Further, the ligation step can sometimes be troublesome because low ligation efficiency can cause loss of the library. For these reasons, Rolling circle amplification (RCA) method is used to simplify these steps.

V. B. 2. 10. Megaprimer PCR of Whole Plasmid (MEGAWHOP) (Miyazaki & Takenouchi 2002)

MEGAWHOP is a powerful method for creating random mutagenesis megalibraries. It is illustrated in Figure V. B. 2.

A mutated gene fragment was prepared by random mutagenesis, using either a PCR-based (e.g., error-prone PCR or DNA shuffling) procedure or a non-PCR based procedure (e.g., chemical mutagenesis) (step a). Fragments were then annealed to a template plasmid that was propagated in *E.coli* dam⁺ strain (step b).

Next, whole plasmid PCR amplification was carried out to synthesize a nicked circular plasmid (step c). Incorporation of new mutations during the long-range PCR was minimized by using a high-fidelity DNA polymerase, such as *Pfu*. The product was then treated with DpnI (step d). Because DpnI digests only dam-methylated DNA (target sequence: 5'-G^{m6}ATC-3') (McClelland & Nelson 1992), the template plasmid propagated in dam⁺ strain was specifically digested.

On the other hand, the newly synthesized plasmid is tolerant, thereby allowing us to eliminate background template plasmid.

The DpnI-treated mixture was then introduced into *E.coli*. After transformation, nicks in the newly synthesized plasmid were repaired in vivo (step e).

RCA was carried out in 20 mM Tris-HCl, pH 8.8, 10 mM (NH₄)₂SO₄, 10 mM KCl, 0.1 % Triton-x-100, 0.1 mg/ml BSA, 2 mM MgSO₄, 0.2 mM dNTPs, 600-650 ng/μl mega primer, 50 ng of DNA Template and 2.5 U *Pfu* DNA Polymerase in a total volume of 50 μl.

PCR was carried out using a program of 95°C for 3 min, 25 cycles of 95°C for 1 min; 60°C for 50 s and 72°C for 13 min, and a final incubation at 72°C for 15 min.

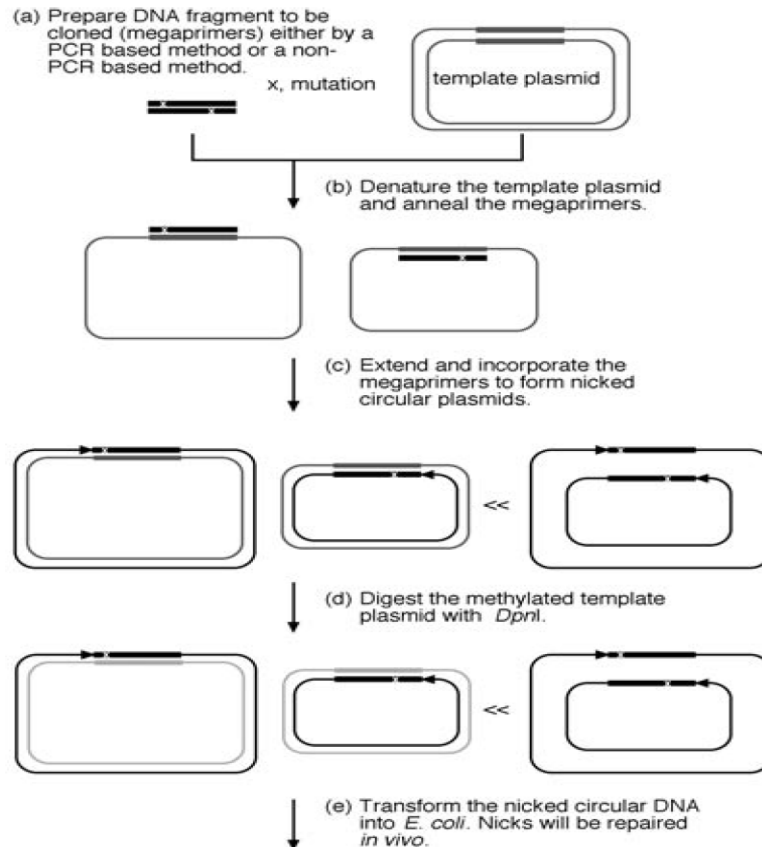


Figure V. B. 2. Outline of the MEGAWHOP method (Miyazaki & Takenouchi 2002).

V. B. 2. 11. DNA shuffling (Stemmer 1994)

The method of DNA shuffling, or sexual PCR, is used to recombine homologous DNA sequences during *in vitro* molecular evolution. While randomly recombining the DNA sequences, the technique also introduces new point mutations at a relatively high rate (0.7%). Though these point mutations may provide useful diversity for some *in vitro* evolution applications, they are problematic for others, especially when the mutation rate is this high.

DNA shuffling consists of four steps: (i) preparation of genes to be shuffled, (ii) fragmentation with DNaseI, (iii) reassembly by thermocycling in the presence of a DNA polymerase, and (iv) amplification of the reassembled products by a conventional PCR.

The method was carried out using the following protocol:

A. Fragmentation

1. Digest 2-5 µg of DNA substrate(s) with 2 units of DNase I per µl in 1 ml of 50 mM Tris-HCl, pH 7.4/ 1 mM MgCl₂ for 30 min at room temperature.
2. Electrophorese the digested products on a 2% agarose gel in 1x TAE buffer.
3. Place the gel stained with ethidium bromide on a preparative UV illuminator to visualize the DNA.
4. Excise the region of the gel containing DNA of the desired molecular weight range (50 bp).
5. Purify the DNA fragments from the agarose gel with the QIAEX II gel extraction kit, according to the manufacturer's instructions.
6. After gel extraction, the supernatant (DNA) is purified using QIAquick PCR purification kit, according to the manufacturer's instructions.

B. Assembly

The reaction mixture contained 10 mM KCl, 10 mM (NH₄)₂SO₄, 20 mM Tris-HCl pH 8.8, 2 mM MgSO₄, 0.1% Triton-x-100, 0.2 mM dNTPs, 10 µl of the purified fragments and 2.5 U Taq DNA polymerase in a total volume of 50 µl. No primers are added.

PCR was carried out using a program of 94°C for 2 min, 45 cycles of 94°C for 30 s; 55°C for 30 s, and 72°C for 30 s, and a final incubation at 72°C for 5 min.

C. Amplification

The reaction mixture contained 20 mM Tris-HCl, pH 8.8, 10 mM (NH₄)₂SO₄, 10 mM KCl, 0.1 % Triton-x-100, 0.1 mg/ml BSA, 2 mM MgSO₄, 0.2 mM dNTPs, 0.4 µM each primer (shuffled primers), 1 µl of Assembly reaction and 2.5 U *Pfu* DNA Polymerase in a total volume of 50 µl.

PCR was carried out using a program of 95°C for 3 min, 30 cycles of 95°C for 30 s; 58°C for 30 s and 72°C for 2 min, and a final incubation at 72°C for 6 min.

The product was subjected to 1 % agarose gel electrophoresis, a single product of the correct size is typically obtained and was ready for cloning into an expression vector.

V. B. 2. 12. DNA Sequencing

DNA sequencing was done in SeqLab in Goettingen- Germany.

V. B. 2. 13. Electrophoretic Mobility Shift Assay (EMSA) (Hellman & Fried 2007)

Electrophoretic mobility shift assay, also called gel shift or gel retardation assay, is a rapid and sensitive technique used to study protein-nucleic acid interactions. It is based on the observation that the electrophoretic mobility of a protein-nucleic acid complex is typically less than that of the free nucleic acid. This technique can be used to determine the affinity of the protein to the DNA, as well as to compare the binding of the protein of interest to different DNA or RNA substrates. It can also be used to analyze the specificity of the DNA-binding by proteins using competition experiments with DNA fragments containing a binding site for the protein of interest or other unrelated DNA. The specificity of DNA binding is indicated by the band pattern observed in EMSA experiments. EMSA experiment involves a number of steps:

A) End-labeling of DNA substrate.

B) Protein-DNA binding reaction.

A) End-labeling of DNA substrate

The DNA used for gel shift experiments can be labeled either radioactively or fluorescently. In this thesis, radioactive labeling was used. For radioactive-end labeling reaction, 100-200 ng DNA, 5 µl of 10 x T4 polynucleotide kinase buffer [70 mM Tris-Cl (pH 7.6), 10 mM MgCl₂, 10 mM DTT], 3 µl of T4 polynucleotide kinase (10 U/µl), and 3 µl of [γ -³²P]ATP (500 µCi) were mixed together in a final volume of 50 µl. The labeling was performed at 37°C for 1 hour.

B) Protein-DNA binding reaction.

In 20 µl, the reaction mixtures contained 10 mM Hepes, 0.1 mM EDTA, 5 mM NaCl, 50 µM DTT, 1 µg of acetylated bovine serum albumin, 0.5 mM sinefungin (Sigma), 20 nM DNA substrate and (0.5-5 µM) protein. After incubation for 30 min at room temperature, reactions were mixed with 4 µl of 87% glycerol and loaded onto a native polyacrylamide gel in 0.5 X TBE buffer, followed by overnight electrophoresis in the cold room (150V, 65 mA). The gel was dried and the separated complexes were visualized by autoradiography using a phosphorimager system (Fuji).

V. B. 3. Protein biochemical methods

V. B. 3. 1. Determination of protein concentration by UV absorption.

Amino acids tryptophan, tyrosine, phenylalanine and cysteine (disulfide bridges) absorb light within a range from 250 to 300 nm. On this basis, a formula for calculation of the molar extinction coefficient of proteins at 280 nm was set up (Pace *et al.*, 1995).

$$\epsilon_{280}(\text{M}^{-1} \text{cm}^{-1}) = \sum \text{Try. } 5500 + \sum \text{Tyr. } 1490 + \sum \text{Cys. } 125$$

Dividing the measured absorbance of a protein at 280 nm by the calculated or known molar extinction coefficient yields the molar concentration of the protein solution (Lambert-Beer law).

$$A = \epsilon \cdot c \cdot d \quad c = A / \epsilon \cdot d$$

A= Absorption

C= Concentration (M)

d= cell length (cm)

The concentration of the EcoDam was determined using an extinction coefficient of $\epsilon_{280\text{nm}} = 39.935 \text{ M}^{-1} \text{cm}^{-1}$.

V. B. 3. 2. Protein expression and purification

V. B. 3. 2. A. IPTG induction for protein expression

Expression of the Dam protein was carried out by using of pET-28a Vector which contains an inducible T7 promoter, IPTG (Isopropyl-beta-D-thiogalactopyranoside) served as an inducer.

The desired clones were cultured overnight in 20 ml LB medium supplemented with 25 mg/ml kanamycin at 37°C. In the ratio of 1:100 the cells were diluted with medium and shaken at 37°C. When OD₆₀₀ reached 0.6-0.8, IPTG was added to a final concentration of 1 mM and the cells were further cultured at 37°C for 3h. The culture was centrifuged at 2000 rpm (Beckman Avanti High-Speed Centrifuge with JLA-8.1000 rotor, USA) for 20 min at 4°C, the pellet was washed with 80 ml 1x STE buffer for each 1L culture. The washed pellet was subjected to the next processing step and stored at -20°C.

V. B. 3. 2. B. Soluble protein fraction from host cell

- a. On ice, the washed pellet was dissolved in 60 ml washing buffer.

- b. Sonicate on ice on Branson Sonifier R-250 (Branson Sonic Power Company, USA) for 12 cycles, where each cycle consists of 15 s pulses followed by 45 s incubation time on ice.
- c. Centrifuge to remove cell membrane and debris at 20000 rpm (Beckman Avanti High-Speed Centrifuge with JA-25-50 rotor, USA) for 60 min at 4°C.

The obtained soluble protein fraction was subjected to purification, e.g. IMAC purification.

V. B. 3. 2. C. Immobilized metal-ion affinity chromatography (IMAC) (Porath & Olin 1983).

Short peptide tags genetically fused to recombinant proteins have been widely used to facilitate detection or purification without the need to develop specific procedures. The His-tag, either N- or C-terminal, consisting of four to six consecutive histidine residues binding selectively to immobilized metal-ion (Ni^{2+}) (Hochuli *et al.*, 1988).

A nitrilotriacetic acid (NTA) adsorbant is developed for metal- chelate affinity chromatography.

Histidine is the amino acid that exhibits the strongest interaction with immobilized metal ion matrices (Ni-NTA), as electron donor groups on the histidine imidazole ring readily form coordination bonds with the immobilized transition metal. The His-tagged proteins are efficiently retained on IMAC. Following washing of the matrix material, the His-tagged proteins can be easily eluted by either adjusting the pH of the column buffer or by adding free imidazole. Elution is effective within a range of 20-250 mM imidazole (Hefti *et al.*, 2001; Janknecht *et al.*, 1991). Imidazole is subsequently removed by dialysis because the presence of imidazole can result in aggregation of the protein (Hefti *et al.*, 2001).

IMAC was carried out using the following protocol:

All the following steps have to be performed in a cold room (4°C).

1. Pack 800 µl of Ni-NTA agarose into column and equilibrate with 40 ml washing buffer.
2. Load the protein solution on Ni-NTA column, non-specific bound molecules are removed by washing two times with 60 ml washing buffer.

3. Elute the bound 6x His-tagged protein with 5 ml elution buffer.
4. To test which fractions should be pooled, load 10 μ l of each elution fraction onto whatman paper, stain with coomassie blue and wash to visualize your protein.
5. Dialyze the pooled protein fractions for 4 h against 1L dialysis buffer I, using dialysis tubing cellulose membrane (MW 10,000) which was prepared following the methods provided by the manufacturer.
6. Dialyze overnight against 1L dialysis buffer II

The resulting protein solution is stored at -20°C for further investigations.

V. B. 3. 3. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were analyzed over SDS- polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli 1970). By adding of the detergent sodium dodecyl sulfate (SDS), the proteins are denatured and the own charges of the proteins are masked, so that the proteins are separated in electrophoresis according to their molecular weight.

The gel casting includes two steps. First, casting the resolving gel, isopropanol was added onto the top to give a level surface. After polymerization of the resolving gel, the isopropanol was decanted and the gel was washed carefully. Then, the stacking gel was casted on the top of the resolving gel. The composition of the gels is shown in the following table:

	Resolving gel (15%)	Stacking gel (5%)
40% acrylamide	18.75 ml	6.025 ml
ddH ₂ O	10.75 ml	36.75 ml
1.5M Tris HCl pH 8.8	14 ml	-
1 M Tris HCl pH 6.8	-	6.5 ml
10 % SDS	0.5 ml	0.5 ml
To 10 ml and 3ml of resolving and stacking gels, respectively, the following components should be added just before pouring the gel.		
40% APS	15 μ l	10 μ l
TEMED	10 μ l	10 μ l

The proteins were loaded with an equal volume protein sample buffer (loading buffer) and denatured 5 min at 100°C. The proteins were isolated with a constant current of 35 mA in SDS-PAGE running buffer.

After electrophoresis, the protein bands were visualized by dipping the gel into fresh” Coomassie Brilliant Blue “staining solution for 2-3 h at 37°C with agitation or overnight at room temperature also with agitation. Destain with water.

V. B. 3. 4. Enzymatic activity assay

Biotin-avidin microplate assay is a method used to measure methylation of DNA (biotinylated oligonucleotide substrates) by DNA MTases (Roth & Jeltsch 2000).

The methylation reaction was carried out in solution using [methyl-³H]-AdoMet. Afterwards, the oligonucleotides were immobilized on an avidin coated microplate, where the incorporation of [³H] labeled methyl group into DNA was stopped by addition of unlabeled AdoMet to the binding buffer. Separation of radioactively labeled DNA from unreacted AdoMet and enzyme was performed by washing steps. Subsequently, the radioactivity incorporated into the DNA was released by a nucleolytic digestion of the DNA. By liquid scintillation counting the radioactivity could be determined.

Time course of methylation of different HPLC-purified Dam substrates was performed.

To coat microplates with avidin, 1 µg avidin dissolved in 100 µl of 100 mM NaHCO₃ (pH 9.6) was dispensed into each well and incubated overnight at 4°C. The wells were washed five times with 200 µl PBST buffer. Coated plates can be stored several days at 4°C. Methylation reaction were carried out using 0.5 µM oligonucleotide and 0.6 µM purified His₆-tagged EcoDam in 100 mM Hepes, pH 8.0, 1mM EDTA, 50 mM NaCl, 0.5 mM DTT, 0.2 mg/ml BSA in the presence of 5.5 µM labeled [methyl-³H]-AdoMet (250 µCi, 3.03 TBq/mmol) at ambient temperature. To measure the time course of the DNA methylation, aliquots comprising 2 µl were removed from the reaction mixture and pipetted into the wells of a microplate that contains 5 µl of 10 mM unlabeled AdoMet in 10 mM H₂SO₄ to quench the incorporation of [³H] into the DNA. PBST was added to a total volume of 40 µl and the mixture incubated for 30 min to allow binding of the oligonucleotides to the microplate.

The wells were washed five times with 200 µl PBST to remove the unreacted AdoMet and the enzyme. Subsequently, the DNA was degraded using 0.7 µg (700 U) *Serratia marcescens* nuclease in 100 µl of 50 mM Tris/HCl pH 8, 0.5 mM MgCl₂ for 30 min at

ambient temperature. The released radioactivity was analyzed by liquid scintillation counting of the reaction mixture after adding 160 μ l Liquid Scintillator cocktail.

VI. References

Aharoni, A., Griffiths, A. D. and Tawfik, D. S. (2005) High-throughput screens and selection of enzyme-encoding genes. *Curr. Opin. Chem. Biol.* 9, 210-216.

Becker, S., Schmoldt, H. U., Adams, T. M., Wilhelm, S. and Kolmar, H. (2004) Ultra high-throughput screening based on cell-surface display and fluorescence-activated cell sorting for the identification of novel biocatalysis. *Curr. Opin. Biotechnol.* 15 (4), 323-329.

Bickle, T. A. and Kruger, D. H. (1993) Biology of DNA restriction. *Microbiol. Rev.* 57, 434-450.

Bornscheuer, U. T. (2002) Methods to increase enantioselectivity of lipases and esterases. *Curr. Opin. Biotechnol.* 13, 543-547.

Bornscheuer, U. T. and Pohl, M. (2001) Improved biocatalysts by directed evolution and rational protein design. *Curr. Opin. Chem. Biol.* 5, 137-143.

Brooks, J. E., Blumenthal, R. M. and Gingeras, T. R. (1983) The isolation and characterization of the Escherichia coli DNA adenine methylase (dam) gene. *Nucleic Acids Res.* 11, 837-851.

Bujnicki, J. M. (2002) Sequence permutations in the molecular evolution of DNA methyltransferases. *BMC Evol. Biol.* 2 : 3

Bujnicki, J. M. and Radlinska, M. (1999) Molecular evolution of DNA-(cytosine-N4) methyltransferases: evidence for their polyphyletic origin. *Nucleic Acids Res.* 27, 4501-4509.

Cadwell, R. C. and Joyce, G. F. (1992) Randomization of genes by PCR mutagenesis. *PCR Meth. Appl.* 2, 28-33.

Chen, L., Paulsen, D. B., Scruggs, D. W., Banes, M. M., Reeks, B. Y. and Lawrence, M. L. (2003) Alteration of DNA adenine methylase (Dam) activity in *Pasteurella multocida* causes increased spontaneous mutation frequency and attenuation in mice. *Microbiology.* 149, 2283-2290.

Chen, R. (1999) A general strategy for enzyme engineering. *Trends Biotechnol.* 17, 344-345.

Chen, R. (2001) Enzyme engineering: rational redesign versus directed evolution. *Trends Biotechnol.* 19, 13-14.

Chen, R. (2001) Enzyme Technologies for Pharmaceutical and Biotechnological Application. (Kirst, H. A., Yeh, W. K. & Zmijewski, M. J. eds) Marcel Dekker Inc.

- Chen, R. et al.** (1996) Redesigning secondary structure to invert coenzyme specificity in isopropylmalate dehydrogenase. *Proc. Natl. Acad. Sci. USA.* 93, 12171-12176.
- Cohen, H. M., Tawfik, D. S. and Griffiths, A. D.** (2004) Altering the sequence specificity of HaeIII methyltransferase by directed evolution using in vitro compartmentalization. *Protein Eng. Des. Sel.* 17 (1), 3-11.
- Colot, V. and Rossignol, J. L.** (1999) Eukaryotic DNA methylation as an evolutionary device. *Bioessays.* 21, 402-411.
- Dagert, M. and Ehrlich, S. D.** (1979) Prolonged incubation in calcium chloride improves the competence of *E. coli* cells. *Gene.* 6 (1), 23-28.
- Damborsky, J.** (2007) Protein design and evolution for biocatalysis. *Biotechnol. J.* 2, 176-179.
- Dieffenbach, C. W. and Dveksler, G. S.** (1995) PCR primer. A laboratory Manual. Cold Spring Harbor Laboratory Press, New York, p. 714.
- Dryden, D. T. F.** (1999) In S-adenosylmethionine-dependent Methyltransferases: Structures and Functions (Eds.: Cheng, X. and Blumenthal, R. M.), World Scientific Publishing, Singapore, p. 283.
- Eiden-Plach, A., Zagorc, T., Heintel, T., Carius, Y., Breinig, F. and Schmitt, M. J.** (2004) Viral prepro toxin signal sequence allows efficient secretion of green fluorescent protein by *Candida glabrata*, *Pichia pastaris*, *Saccharomyces cerevisiae*, and *Schizosaccharomyces pombe*. *Appl. Environ. Microbiol.* 70 (2), 961-966.
- Farinas, E., Schwaneberg, U., Anton, G. and Arnold, F. H.** (2001) Directed evolution of a cytochrome P450 monooxygenase for alkane oxidation. *Adv. Synth. Catal.* 343, 601-606.
- Faumann, E. B., Blumenthal, R. M. and Cheng, X.** (1999) in S-adenosylmethionine-dependent methyltransferases: Structures and functions (Eds.: Cheng, X. and Blumenthal, R. M.), World Scientific Publishing, Singapore, p.1.
- Gerasimaite, R., Vilkaitis, G., and Klimasauskas, S.** (2009) A directed evolution design of a GCG-specific DNA hemimethylase. *Nucleic Acids Res.* 1-10.
- Getzoff, G. D. et al.,** (1992) Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. *Nature.* 358, 347-351.
- Glick, B. R. and Pasternak, J. J.** (2003) Molecular biotechnology: principles and applications of recombinant DNA. 3rd ed. Washington (DC) : ASM press.

- Goedecke, K., Pignot, M., Goody, R. S., Scheidig, A. J. and Weinhold, E.** (2001) Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. *Nat. Struct. Biol.* 8, 121-125.
- Hattman, S., Brooks, J. E. and Masurekar, M.** (1997) Sequence specificity of the P1 modification methylase (M.Eco P1) and the DNA methylase (M.EcoDam) controlled by the *Escherichia coli* dam gene. *J. Mol. Biol.* 126, 367-380.
- Hattman, S., Kenny, C., Berger, L. and Pratt, K.** (1978) Comparative study of DNA methylation in three unicellular eukaryotes. *J. Bacteriol.* 135, 1156-1157.
- Hattman, S., Wilkinson, J., Swinton, D., Schlagman, S., Macdonald, P. M. and Mosig, G.** (1985) Common evolutionary origin of the phage T4Dam and host *Escherichia coli* dam DNA-adenine methyltransferase genes. *J. Bacteriol.* 164, 932-937.
- Hefti, M. H., Caroline, J. G., der Toorn, V. V., Dixon, R. and Vervoot, J.** (2001) A novel purification method for histidine-tagged proteins containing a thrombin cleavage site. *Anal. Biochem.* 295, 180-185.
- Hellman, L. and Fried, M.** (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature protocols.* 2 (8), 1849-1861.
- Hernday, A., Krabbe, M., Braaten, B. and Low, D.** (2002) Self-perpetuating epigenetic pili switches in bacteria. *Proc. Natl. Acad. Sci. USA.* 99 (Suppl 4), 16470-16476.
- Heusipp, G., Falker, S. and Alexander-Schmidt, M.** (2007) DNA adenine methylation and bacterial pathogenesis. *Int. J. Med. Microbiol.* 297, 1-7.
- Hochuli, E., Bannwarth, W., Doebeli, H., Gentz, R. and Stueber, D.** (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbant. *Bio. Technol.* 6, 1321-1325.
- Hoekstra, M. F. and Malone, R. E.** (1985) Expression of the *Escherichia coli* dam methylase in *Saccharomyces cerevisiae*: Effect of in vivo adenine methylation on genetic recombination and mutation. *Mol. Cell. Biol.* 5, 610-618.
- Honma, Y., Fernandez, R. E. and Maurelli, A. T.** (2004) A DNA adenine methylase mutant of *Shigella flexneri* shows no significant attenuation of virulence. *Microbiology.* 150, 1073-1078.
- Horton, J. R., Liebert, K., Bekes, M., Jeltsch, A. and Cheng, X.** (2006) Structure and substrate recognition of the *Escherichia coli* DNA Adenine methyltransferase. *J. Mol. Biol.* 358, 559-570.

- Horton, J. R., Liebert, K., Hattman, S., Jeltsch, A. and Cheng, X.** (2005) Transition from non-specific to specific DNA interaction along the DNA recognition pathway of Dam methyltransferase. *Cell*. 121, 349-61.
- Hurley, J. H. et al.**, (1996) Determinants of cofactor specificity in isocitrate dehydrogenase: structure of an engineered NADP⁺ → NAD⁺ specificity-reversal mutant. *Biochemistry*. 35, 5670-5678.
- Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R., Nordheim, A. and Stunnenberg, H. G.** (1991) Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Proc. Natl. Acad. Sci. USA*. 88, 8972-8976.
- Jeltsch, A.** (1999) Circular permutations in the molecular evolution of DNA methyltransferases. *J. Mol. Evol.* 49, 161-164.
- Jeltsch, A.** (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *ChemBioChem*. 3, 274-293.
- Jeltsch, A.** (2003) Maintenance of species identity and controlling speciation of bacteria: a new function for restriction/modification systems?. *Gene*. 317, 13-16.
- Jeltsch, A.** (2008) Reading and writing of DNA methylation. *Nat. Struct. Mol. Biol.* 15, 1003-1004.
- Jeltsch, A. and Gumpert, R. I.** (2004) DNA modification: Bacterial methyltransferases. In: Encyclopedia of Biological Chemistry (Lennarz, W.J. & Lane, D.M., eds.) Academic Press, San Diego, CA, ISBN 0124437109.
- Jeltsch, A. and Jurkowska, R. Z.** (2009) DNA methylation. In: The chemical biology of nucleic acids (Guenter Mayer, ed.), Wiley-VCH, in press.
- Jeltsch, A. and Lanio, T.** (2002) Site directed mutagenesis by polymerase chain reaction. *Methods Mol. Biol.* 182, 85-94.
- Jeltsch, A., Christ, F., Fatemi, M. and Roth, M.** (1999) On the substrate specificity of DNA methyltransferases: adenine-N⁶ DNA methyltransferases also modify cytosine residues at position N⁴. *J. Biol. Chem.* 274, 19538-19544.
- Jeltsch, A., Jurkowska, R. Z., Jurkowski, T. P., Liebert K., Rathert, P. and Schlickenrieder, M.** (2007) Application of DNA methyltransferases in targeted DNA methylation. *Appl. Microbiol. Biotechnol.* 75, 1233-40.
- Jeltsch, A., Sobotta, T. and Pingoud, A.** (1996) Structure prediction of the *EcoRV* DNA methyltransferase based on mutant profiling, secondary structure analysis, comparison to

known structures of methyltransferases and isolation of catalytically inactive single mutants. *Protein Engineering*. 9, 413-423.

Joseph, N., Duppatla, V. and Rao, D. N. (2006) Prokaryotic DNA mismatch repair. *Prog. Nucleic. Acid. Res. Mol. Biol.* 81, 1-49.

Joyce, G. F. (1994) *In vitro* evolution of nucleic acids. *Curr. Opin. Struct. Biol.* 4, 331-336.

Joyce, G. F. (2004) Directed evolution of nucleic acid enzymes. *Annu. Rev. Biochem.* 73, 791-836.

Julio, S. M., Heithoff, D. M., Sinsheimer, R. L., Low, D. A. and Mahan, M. J. (2002) DNA adenine methylase overproduction in *Yersinia pseudotuberculosis* alters YopE expression and secretion and host immune responses to infection. *Infect. Immun.* 70, 1006-1009.

Jurkowski, T. P., Anspach, N., Kulishova, L., Nellen, W. and Jeltsch, A. (2007) The M.EcoRV DNA-(Adenine N⁶)-methyltransferase uses DNA bending for recognition of an expanded EcoDam recognition site. *J. Biol. Chem.* 282 (51), 36942-36952.

Kiss, A., Posfai, G., Zsurka, G., Rasko, T and Venetianer, P. (2001) Role of DNA minor groove interactions in substrate recognition by the M.SinI and M.EcoRII DNA (cytosine-5) methtransferases. *Nucleic Acids Res.* 29 (15), 3188-3194.

Klimasauskas, S., Kumar, S., Roberts, R. J. and Cheng, X. (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell.* 76, 357-369.

Kumar, S., Johnson, W. S. and Tomasz, M. (1993) Orientation isomers of the mitomycin-c interstrand cross-link in non-self complementary DNA – differential effect of the 2 isomers on restriction endonuclease cleavage at a nearby site. *Biochemistry.* 32, 1364-1472.

Lacks, S. and Greenberg, B. (1977) Complementary specificity of restriction endonucleases of *Diplococcus Pneumoniae* with respect to DNA methylation. *J. Mol. Biol.* 114, 153-168.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 227 (52950), 680-685.

Lange, C., Wild, C. and Trautner, T. A. (1996) Identification of subdomain within DNA-(cytosine-C5)-methyltransferases responsible for the recognition of the 5' part of their DNA target. *EMBO J.* 15, 1443-1450.

- Leung, D. W., Cheng, E. and Goedel, D. W.** (1989) A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Techniques*. 1, 11-15.
- Lin, H. and Cornish, V. W.** (2002) Screening and selection methods for large-scale analysis of protein function. *Angew. Chem. Int. Ed. Engl.* 41 (23), 4402-4425.
- Lobner-Olesen, A., Skovgaard, O. and Marinus, M. G.** (2005) Dam methylation: coordinating cellular processes. *Curr. Opin. Microbiol.* 8, 154-160.
- Mahan, M. J. and Low, D. A.** (2001) DNA methylation regulates bacterial gene expression and virulence. *ASM News*. 67, 356-361.
- Mahan, M. J., Heithoff, D. M., Sinsheimer, R. L. and Low, D. A.** (2000) Assessment of bacterial pathogenesis by analysis of gene expression in the host. *Ann. Rev. Genet.* 34, 139-164.
- Matsuura, T. and Yomo, T.** (2006) In vitro evolution of proteins. *J. Biosci. Bioeng.* 101 (6) 449-456.
- Matsuura, T., Ernst, A., Zechel, D. L. and Plueckthun, A.** (2004) Combinatorial approaches to novel proteins. *ChemBioChem*. 5, 177-182.
- McClelland, M. and Nelson, M.** (1992) Effect of site-specific methylation on DNA modification methyltransferases and restriction endonucleases. *Nucleic Acids Res.* 20, 2145-2157.
- Mills, D. R., Peterson, R. L. and Spiegelman, S.** (1967) An extracellular darwinian experiment with a self-duplicating nucleic acid molecule. *Proc. Natl. Acad. Sci. USA*. 58, 217-224.
- Miyazaki, K. and Takenouchi, M.** (2002) Creating random mutagenesis libraries using megaprimer PCR of whole plasmid. *Biotechniques*. 33, 1033-1038.
- Noyer-Weidner, M. and Trautner, T. A.** (1993) in DNA methylation: Molecular Biology and Biological Significance (Eds.: J. P. Jost, H. P. Saluz), Birkhauser, Basel, p. 40.
- Ogino, H. and Ishikawa, H.** (2001) Enzymes which are stable in the presence of organic solvents. *J. Biosci. Bioeng.* 91, 109-116.
- Okamoto, A., Kosugi, A., Koizumi, Y., Yanagida, F. and Udaka, S.** (1997) High efficiency transformation of *Bacillus brevis* by electroporation. *Biosci. Biotechnol. Biochem.* 61 (1), 202-203.

- Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. and Gray, T.** (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 4 (11), 2411-2423.
- Pingoud, A. and Jeltsch, A.** (2001) Structure and function of type II restriction endonucleases. *Nucleic Acids Res.* 29, 3705-3727.
- Pingoud, A., Fuxreiter, M. and Pingoud, V.** (2005) Type II restriction endonucleases: structure and mechanism. *Cell. Mol. Life. Sci.* 62, 685-707.
- Porath, J. and Olin, B.** (1983) Immobilized metal affinity absorption and metal ion affinity chromatography. *Biochemistry.* 22, 1621-1630.
- Ratel, D., Ravanat, J. L., Berger, F. and Wion, D.** (2006) N⁶-methyladenine: the other methylated base of DNA. *Bioessays.* 28, 309-315.
- Reinisch, K. M., Chen, L., Verdine, G. L. and Lipscomb, W. N.** (1995) The crystal structure of HaeIII methyltransferase covalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. *Cell.* 82, 143-153.
- Roodveldt, C., Aharoni, A. and Tawfik, D. S.** (2005) Directed evolution of proteins for heterologous expression and stability. *Curr. Opin. Struct. Biol.* 15, 50-56.
- Roth, M. & Jeltsch, A.** (2000) Biotin-Avidin microtiter plate assay for quantitative analysis of DNA methylation by DNA methyltransferases. *Biol. Chem.* 381, 269-272.
- Roth, M. and Jeltsch, A.** (2001) Changing the target base specificity of the EcoRV DNA methyltransferase by rational de novo protein-design. *Nucleic Acids Res.* 29 (15) 3137-3144.
- Sambrook, J., Fritsch, F. and Maniatis, T.** (1989) Molecular Cloning: A laboratory manual. 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Schmidt-Dannert, C.** (2001) Directed evolution of single proteins, metabolic pathways and viruses. *Biochemistry.* 40, 13125-13136.
- Schmidt-Dannert, C. and Arnold, D. S.** (1999) Directed evolution of industrial enzymes. *Trends Biotechnol.* 17, 135-136.
- Sen, S., Venkata, D. V. and Mandal, B.** (2007) Developments in directed evolution for improving enzyme functions. *Appl. Biochem. Biotechnol.* 143, 212-223.
- Sidhu, S. S.** (2000) Phage display in pharmaceutical biotechnology. *Curr. Opin. Biotechnol.* 11, 610-616.

- Skerra, A.** (1992) Phosphorothioate primers improve the amplification of DNA sequences by DNA polymerases with proofreading activity. *Nucleic Acids Res.* 20, 3551-3554.
- Stemmer, W. P. C.** (1994) DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA.* 91,10747-10751.
- Tao, H. and Cornish, V. W.** (2002) Milestones in directed enzyme evolution. *Curr. Opin. Chem. Biol.* 6, 858-864.
- Taylor, V. L., Titball, R. W. and Oyston, P. C. F.** (2005) Oral immunization with a dam mutant of *Yersinia pseudotuberculosis* protects against plague. *Microbiology.* 151, 1919-1926.
- Topcu, Z.** (2000) An optimized recipe for cloning of the polymerase chain reaction-amplified DNA inserts into plasmid vectors. *Acta Biochem. Polon.* 47 (3), 841-846.
- Vogit, C. A., Kauffman, S., and Wang, Z. G.** (2000) Rational evolutionary design: the theory of in vitro protein evolution. *Adv. Protein Chem.* 55, 79-160.
- Waldo, G. S.** (2003) Genetic screens and directed evolution for protein solubility. *Curr. Opin. Chem. Biol.* 7, 33-38.
- Waston, M. E., Jarisch, J. and Smith, A. L.** (2004) Inactivation of deoxyadenosine methyltransferase (dam) attenuates *Haemophilus influenzae* virulence. *Mol. Microbiol.* 55, 651-654.
- Williams, G. J., Nelson, A. S. and Berry, A.** (2004) Directed evolution of enzymes for biocatalysis and the life sciences. *Cellul. Mol. Sci.* 61, 3034-3046.
- Wilson, D. S. and Szostak, J. W.** (1999) *In vitro* selection of functional nucleic acids. *Annu. Rev. Biochem.* 68, 611-647.
- Wilson, G. G.** (1992) Amino acid sequence arrangements of DNA-methyltransferases. *Methods Enzymol.* 216, 259-279.
- Wion, D. and Casadesus, J.** (2006) N6-methyladenine: an epigenetic signal for DNA-protein interactions. *Nat. Rev. Microbiol.* 4, 183-192.
- Wong, T. S., Tee, K. L., Hauer, B. and Schwaneberg, U.** (2004) Sequence Saturation Mutagenesis (SeSaM): A novel method for directed evolution. *Nucleic Acids Res.* 32, 3e26.

Yang, Z., Horton, J. R., Zhou, L., Zhang, X. J., Dong, A. Zhang, X. *et al.* (2003) Structure of the bacteriophage T4 DNA adenine methyltransferase. *Nature Struct. Biol.* 10, 849-855.

Zhao, H., Giver, L., Shao, Z., Affholter, J. A. and Arnold, F. H. (1998) Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat. Biotechnol.* 16 (3), 258-261.

Zweiger, G., Marczynski, G. and Shapiro, L. (1994) A *Caulobacter* DNA methyltransferase that functions only in the predivisinal cell. *J. Mol. Biol.* 235, 472-485.

Supplement (1)

Changing the DNA recognition specificity of the EcoDam DNA-(adenine
N6)-methyltransferase by directed evolution



Available online at www.sciencedirect.com



Changing the DNA Recognition Specificity of the EcoDam DNA-(Adenine-N6)-Methyltransferase by Directed Evolution

Sanjay Chahar, Hany Elsaywy, Sergey Ragozin and Albert Jeltsch*

School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, 28725 Bremen, Germany

Received 22 April 2009;
received in revised form
16 July 2009;
accepted 14 September 2009

EcoDam is an adenine-N6 DNA methyltransferase that methylates the GATC sites in the *Escherichia coli* genome. We have changed the target specificity of EcoDam from GATC to GATT by directed evolution, combining different random mutagenesis methods with restriction protection at GATT sites for selection and screening. By co-evolution of an enzyme library and a substrate library, we identified GATT as the best non-GATC site and discover a double mutation, R124S/P134S, as the first step to increase enzyme activity at GATT sites. After four generations of mutagenesis and selection, we obtained enzyme variants with new specificity for GATT. While the wild-type EcoDam shows no detectable activity at GATT sites in *E. coli* cells, some variants prefer methylation at GATT over GATC sites by about 10-fold in cells. *In vitro* DNA methylation kinetics carried out under single-turnover conditions using a hemimethylated GATC and a GATT oligonucleotide substrate confirmed that the evolved proteins prefer methylation of GATT sites to a similar degree. They show up to 1600-fold change in specificity *in vitro* and methylate the new GATT target site with 20% of the rate of GATC methylation by the wild-type enzyme, indicating good activity. We conclude that the new methyltransferases are fully functional *in vivo* and *in vitro* but show a new target-site specificity.

© 2009 Published by Elsevier Ltd.

Edited by F. Schmid

Keywords: enzyme design; directed evolution; DNA methylation; DNA recognition

Introduction

Synthetic biology aims to design new organisms to produce desired compounds including fuel, food, or bioactive molecules.^{1–3} This aim requires the redesign of genetic elements, metabolic pathways, and individual enzymes with respect to their substrate specificity. When combined with rational and computational design, directed evolution is a very promising approach for enzyme design.^{4,5} To develop this method, we applied directed evolution to modify the DNA recognition specificity of the EcoDam DNA methyltransferase, which specifically recognizes GATC sequences and methylates the adenine residue within.^{6,7} DNA methylation has a number of important epigenetic

roles in bacteria including the control of gene expression, coordination of DNA replication and the cell cycle, and post-replicative mismatch repair.^{8–11} In addition, it is involved in restriction-modification systems to protect bacteria from foreign DNA, and DNA methyltransferases are pathogenicity factors in several pathogens.^{9,12,13} DNA MTases are particularly prone to directed evolution because they modify DNA at specific sites. Thus, the blueprint of the activity and specificity of an enzyme variant can be detected on the DNA coding for that particular protein, giving a unique coupling of genotype (DNA sequence) and phenotype (enzymatic properties) on individual DNA molecules.¹⁴ Given this connection, DNA MTases compete with ribozymes^{15,16} for being the enzyme model system best suited for *in vitro* evolution. The special properties of DNA MTases so far have been exploited in projects aiming to redirect the specificity of DNA MTases^{17–19} and to broaden our knowledge on natural evolution pathways.^{20,21}

*Corresponding author. E-mail address: a.jeltsch@jacobs-university.de.

Abbreviation used: epPCR, error-prone PCR.

Results and Discussion

The principle of the selection of EcoDam variants for new specificities is illustrated in Fig. 1a. After random mutagenesis, the pool of MTase variants is transformed into the HMS174(DE3) *Escherichia coli* cells for the expression of the enzyme. In the cell, each enzyme variant modifies its own expression plasmid. The modified plasmids are isolated and digested with an appropriate restriction enzyme that is inhibited by adenine-N6 methylation. Uncleaved (i.e., methylated) plasmids are enriched by retransformation (Fig. 1b). Practical experiences showed us that it was necessary to treat the cleaved mixture with λ -exonuclease prior to the retransformation. This enzyme degrades the restriction enzyme cleavage products and, thereby, prevents their re-ligation in the cell after transformation.

Since we always observed false positives in the form of plasmids that lost the BspEI restriction site or gained a GATC site overlapping with the BspEI site, we performed only one round of selection after mutagenesis and screened several hundred colonies for most highly protected candidates, which were then used as starting material in the next round of mutagenesis. For screening, the EcoDam variants were expressed in *E. coli* cells, the plasmids were isolated, and the methylation was analyzed by cleavage with BspEI (see below). For changing the specificity of EcoDam, we employed a combination of site-saturation mutagenesis using randomized oligonucleotides for site-directed mutagenesis,²² error-prone PCR (epPCR),²³ and DNA shuffling.²⁴ In addition, we developed a new system for mutual co-evolution of enzyme variants and new target sites (Fig. 1c and d). We used a BspEI site (TCCGGA) followed by two randomized bases cloned into the

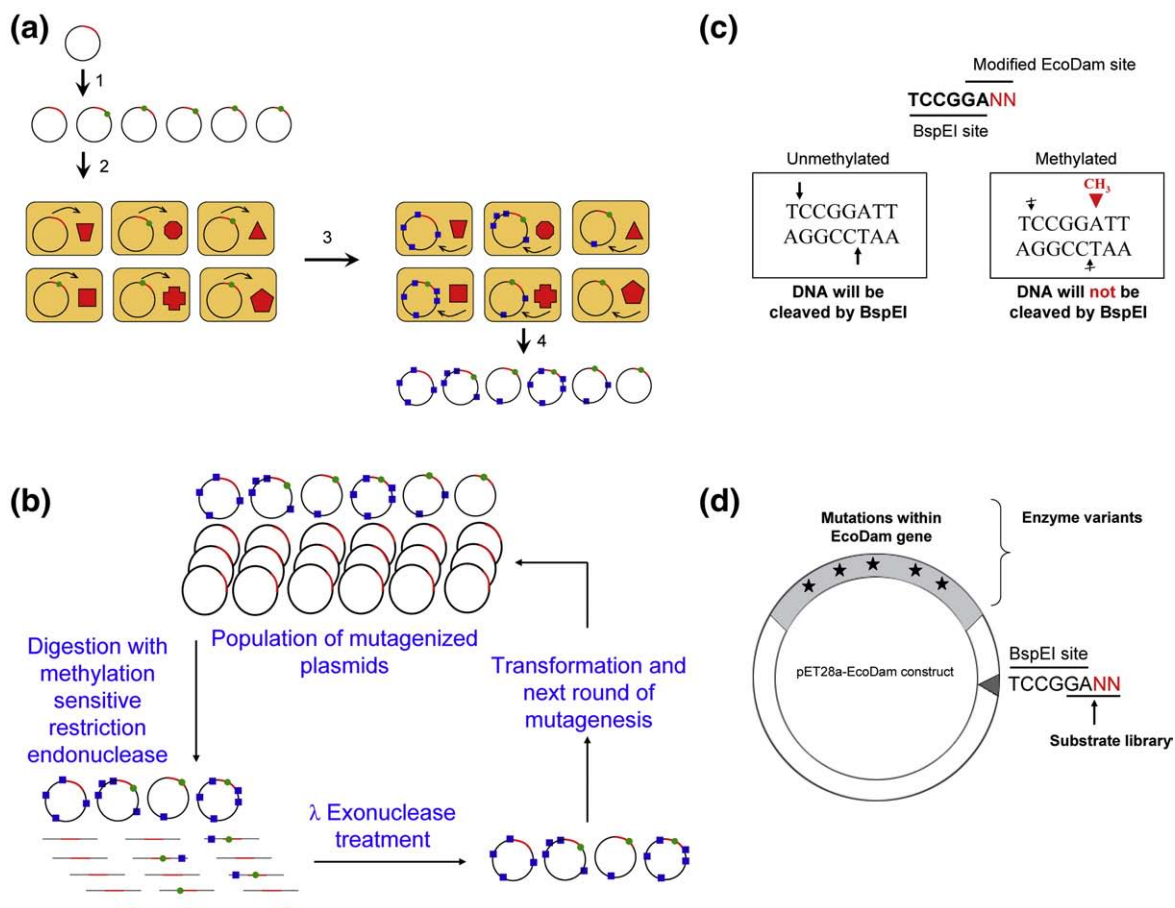


Fig. 1. Principles of *in vitro* evolution experiments with DNA MTases. (a) Coupling of genotype and phenotype with DNA methyltransferases. Step 1: a plasmid containing the gene for a DNA MTase (red color) is subjected to random mutagenesis (mutations indicated by green dots). Step 2: the library of MTase expression plasmids is transformed into bacteria, and the MTase variants get expressed (red symbols). Step 3: the MTases methylate the DNA in the cell (blue squares). Step 4: plasmids can be prepared, the methylation pattern of which reflects the activity and specificity of the MTase encoded by the corresponding plasmid. (b) The plasmid library can be screened using restriction digestion and retransformation. Finally, the mutant genes encoding variants with interesting properties can be retrieved. (c) Example of screening for methylation at modified GATC sites by using overlapping BspEI and GATT sites. Methylation of the terminal A residue in the upper strand of the BspEI site will inhibit DNA cleavage by this enzyme. (d) Scheme of the mutual co-evolution of enzyme activity and target site by cloning a library of enzyme variants and modified target site in the same plasmid and selection for DNA methylation at the modified target site.

EcoDam expression vector. Thereby, the BspEI site (the cleavage of which is inhibited by the methylation of the adenine residue) overlaps with modified EcoDam sites (GANN instead of GATC). All substrate libraries were prepared such that the original GATC site was excluded. By combination of different enzyme libraries with the substrate library, we were able to identify pairs of modified enzymes and matching new target sites in one selection step.

First-generation combined enzyme–substrate library prepared by site-saturation mutagenesis

Structural and biochemical data showed that most of the specific interactions of EcoDam with DNA are formed by residues in a short β -hairpin structure with R124 interacting with the fourth base pair, L122, and P134 interacting with the third one.^{25,26} The first guanine is recognized by K9. Based on the crystallographic data, we started the design of EcoDam variants towards new specificity by site-saturation mutagenesis at the two residues (R124 and P134) that are directly involved in DNA recognition of the TC part of the GATC site (RP library) combined with a GANN substrate library (Fig. 2). Successful library generation was confirmed by sequencing of some clones. Already after one selection cycle, several clones that provided partial protection against BspEI cleavage were obtained (Fig. 3). We screened 200 clones for BspEI protection and found 18 of them protected to at least 40%. Sequencing revealed that, in 9 of them, a GATC site overlapping with the BspEI site appeared and 2 of them had mutations within the BspEI site. These alterations were not present in the randomized oligonucleotide mixture that was used to prepare

the library; thus, they occurred spontaneously in the *E. coli* cells. Our ability to select spontaneous mutations that are known to occur at very low frequency underscores the very strong selection pressure established in the selection procedure. Seven of the remaining clones contained a GATT site (Table 1). Among the clones methylating GATT sites, we observed a strong enrichment for the R124S and P134S exchanges. The R124S exchange disrupts the contact to the cytosine, explaining the change in specificity at the fourth base pair. The P134S mutation had been identified 20 years back in a natural isolate of the phage T4 Dam enzyme and shown to increase its activity,²⁷ which might explain why it was enriched here.

Two clones (R124R/P134A and R124S/P134A) were found to recognize other altered sites (GAAC and GATG, respectively). The R124R/P134A clone keeps the R124 contact to the C4 but shows an altered target site at position 3. The R124S/P134A clone was found with both GATT and GATG targets, suggesting that it might be promiscuous with respect to the recognition of the fourth base of the target sequence, which would make it an interesting candidate for further evolutionary optimization.²⁰

First-generation enzyme–GATT library prepared by site-saturation mutagenesis

We speculated that a smaller library combining the site saturation at R124 and P134 (RP library) with a fixed GATT target site might allow for the identification of clones that showed a higher level of protection and prepared such library (Fig. 2). Again, after one round of selection, protected clones were collected and sequenced, but we did not observe a better

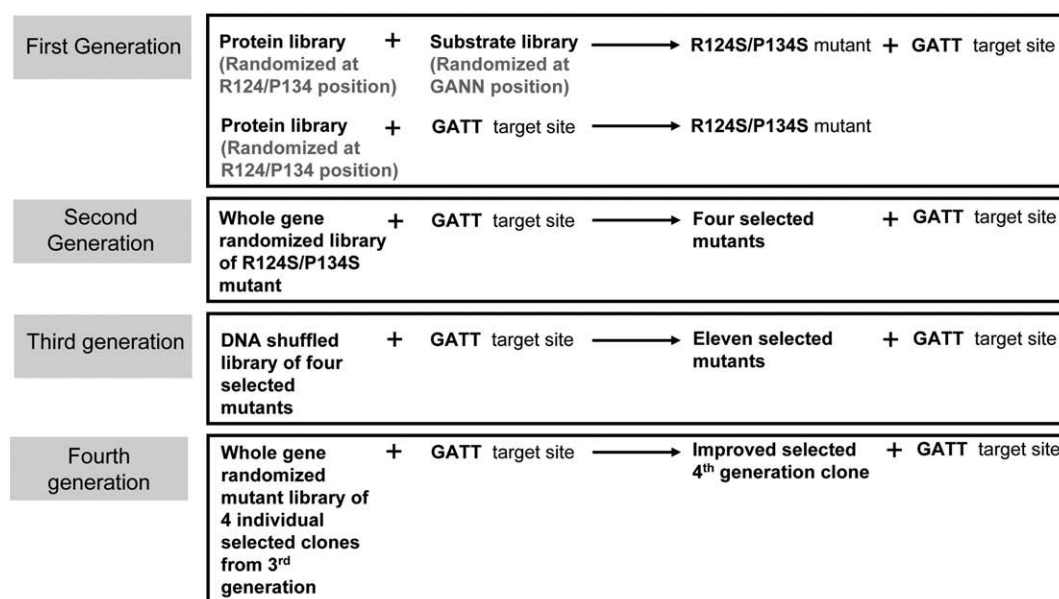


Fig. 2. Experimental setup of this study. We conducted four successive rounds of mutagenesis and selection. In the first round, site-saturation mutagenesis was performed at R124 and P134 and the enzyme variant library was combined with a substrate library, which led to the identification of GATT as the best non-GATC target. In the second round, whole-gene randomization was carried out, followed by DNA shuffling in the third and another step of whole-gene randomization in the fourth round.

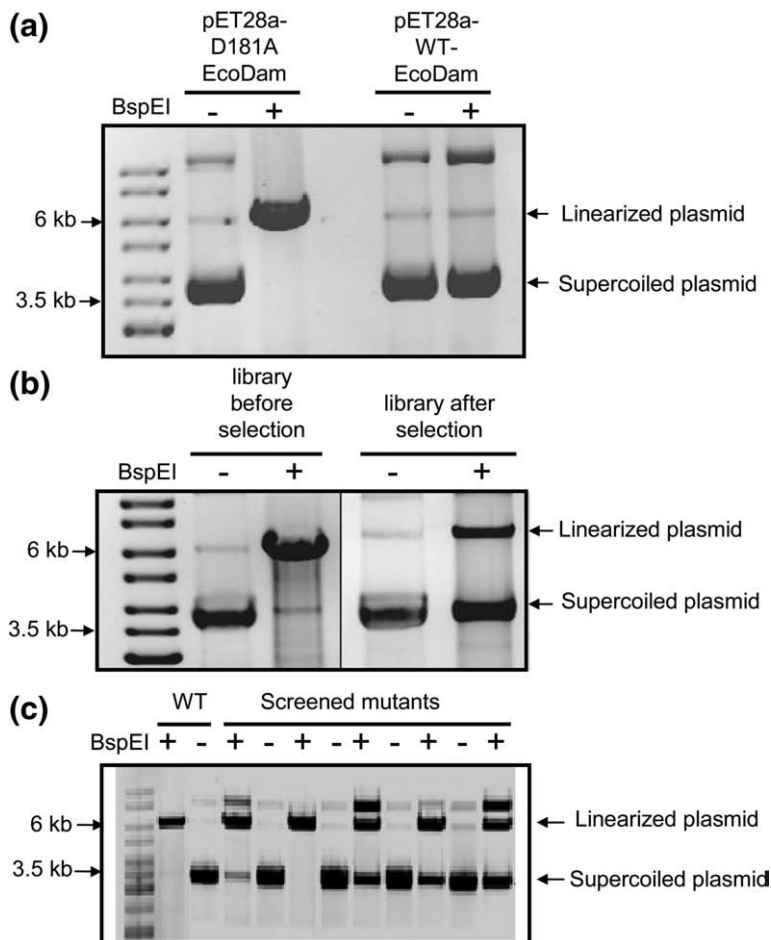


Fig. 3. Screening of the first-generation libraries. (a) Example of methylation sensitivity of BspEI restriction cleavage used for selection. Wild-type *EcoDam* and an inactive D181A variant were expressed in *dam*-negative *E. coli* cells; the plasmids containing a single overlapping BspEI/GATC site were isolated and cleaved by BspEI (for details, cf. Fig. 1c). (b) An example of the improvement of restriction protection at TCCGGANN sites after one round of selection of the RP-GANN library. (c) Example of the screening of clones obtained after the first round of selection. In this example, clone 1 shows a low level and clones 3–5 show a high level of protection against BspEI cleavage.

protection. Sequencing of several clones confirmed the enrichment of R124S and P134S exchanges as observed before. After a second and third selection cycle of the RP-GATT library, the level of protection

did not increase (data not shown). Given the small size of the initial library, we concluded that selection is saturated already after the first cycle and better enzymes could be only obtained after randomization of additional parts of the enzyme.

Table 1. Summary of the results obtained from sequencing of independent *EcoDam* variants that show partial methylation at GATT sites

	Exchange at position R124	Exchange at position P134	Source
Clone 1	A	A	RP-GANN library
Clone 2	Y	S	RP-GANN library
Clone 3	Y	S	RP-GANN library
Clone 4	S	G	RP-GANN library
Clone 5	S	A	RP-GANN library
Clone 6	S	S	RP-GANN library
Clone 7	S	A	RP-GANN library
Clone 2-1	A	G	RP-GATT library
Clone 2-2	S	S	RP-GATT library
Clone 2-3	S	A	RP-GATT library
Clone 2-4	S	S	RP-GATT library
Clone 2-5	G	S	RP-GATT library
Clone 2-6	C	S	RP-GATT library
Clone 2-7	S	S	RP-GATT library
Clone 2-8	A	S	RP-GATT library
Clone 2-9	S	S	RP-GATT library
Clone 2-10	G	S	RP-GATT library
Clone 2-11	S	S	RP-GATT library

Clones 1–7 were obtained from the R124/P134-GANN substrate library. Clones 2-1 to 2-11 were obtained from the R124/P134-GATT library.

Second-generation library prepared by epPCR

We prepared a second-generation library by whole-gene randomization of the R124S/P134S variant (Fig. 2). After selection, around 400 clones were screened and 6 clones showed an improved level of methylation at GATT sites as judged from the comparison of the amount of superhelical and open circle *versus* linear DNA after BspEI cleavage (Fig. 4a). Sequencing showed that clones 3 and 4 were identified twice. Overall, the 4 clones contained 18 different mutations, in addition to the original R124S/P134S mutations (Table 2).

Third-generation library prepared by DNA shuffling and *in vivo* enzyme specificity analysis

We wanted to investigate the combinatorial effect of the mutations identified so far and prepared a third-generation library by DNA shuffling.²⁴ After DNA shuffling and one round of selection, 200 clones were screened for improved DNA methylation at GATT sites by the BspEI restriction protection

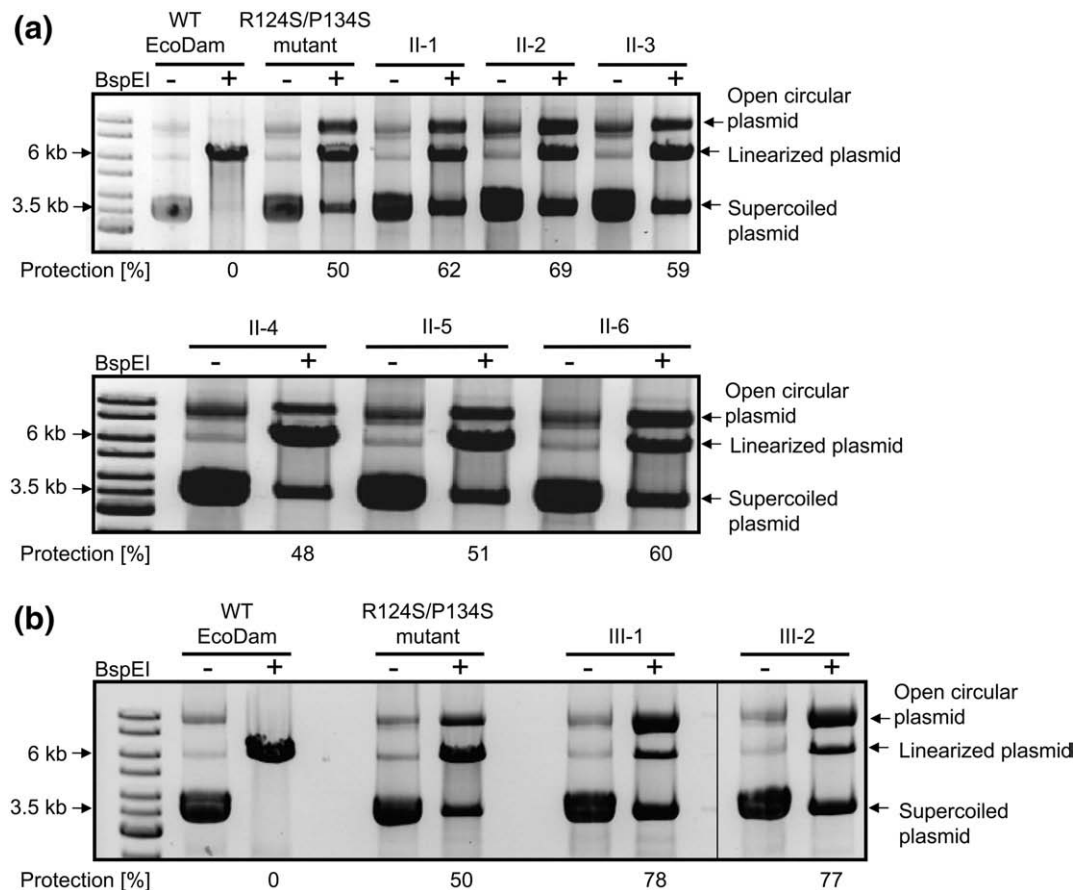


Fig. 4. Examples of the screening of clones from second-generation (a) and third-generation (b) library with BspEI after the first round of selection. The level of protection was estimated on the basis of the scanned intensities of the supercoiled and open circular plasmid fractions, when compared with the linearized band.

assay. Two clones were found to show increased protection against BspEI cleavage (clones III-1 and III-2 in Fig. 4b and Table 3); several others showed similar protection as observed in generation II (Table 3). Sequencing revealed the presence of 14 mutations, 10 of them shuffled from the previous rounds and 4 new ones (Table 3).

To investigate the methylation preferences of the various EcoDam variants in *E. coli* cells, we expressed the mutant enzymes in dam-negative

cells (JM110). Under these conditions, a second BspEI site on the plasmid that overlaps with a GATC site became available for cleavage. In dam-positive cells, this site is always protected by the endogenous EcoDam. In the dam-negative host, the EcoDam variants methylate the original GATC site or the new GATT site depending on their intrinsic preferences. The specificity of methylation was analyzed by a double digestion, indicating that the clones of the third generation showed an about 10-fold preference for DNA methylation at GATT sites *in vivo*, which is a remarkable result considering that wild-type EcoDam does not show detectable protection at GATT sites (Fig. 5), indicating an at least 50-fold preference for GATC methylation *in vivo*.

Fourth-generation library prepared by epPCR

So far, we evolved enzyme variants towards methylation at GATT sites. Our next step was to evolve the protein against GATC site methylation, while keeping the activity at GATT sites. To this end, a mixture of the genes of the four isolated third-generation mutants was used as template to generate a fourth-generation mutant library by whole-gene randomization. Proteins were expressed in dam-negative host and the plasmid pool was

Table 2. Sequence analysis of six clones from the second-generation library (all clones contained the R124S and P134S exchanges)

II-1	II-2	II-3	II-4
M1T	N4S	Q75R	Y138H
K67R	S40P	E84V	
K139E	P104L	K241E	
F159L	L127Q		
G256S	L203H		
	M231L		
	K246N		
	S250N		
	K260Q		

Clones 3 and 4 were identified twice.

Table 3. Sequence analysis of 11 individual clones after DNA shuffling (all clones contained the R124S and P134S exchanges)

III-1	III-2	III-3	III-4	III-5	III-6	III-7	III-8	III-9	III-10	III-11
	Q75R	Q75R E84V	E84V			E84V				Q75R
K139E F159L	F159L	K139E S200G	P104L K139E F159L	K139E F159L	K139E F159L	K139E F159L	K139E	K139E F159L	C87R P104L K139E	
K241E	K241E G256S		K241E	K241E		M231L	K241E	M231L K241E		M231L
							S274H A276R			G256S

isolated. After double digestion with *Bgl*II and *Bsp*EI, the 4.2-kb DNA fragment (originating from plasmids methylated at the GATT site but not methylated at the GATC site) was excised from the gel, the *EcoDam* gene was amplified by PCR, and the DNA was re-cloned. Two hundred clones were

screened for improved specificity of methylation at GATT sites. One variant showed an increased specificity as evidenced by the loss of the 5- and 1.2-kb bands, indicative of methylation at the GATC sites, and the further improvement of the ratio of the 4.2- and 5-kb bands (Fig. 5c).

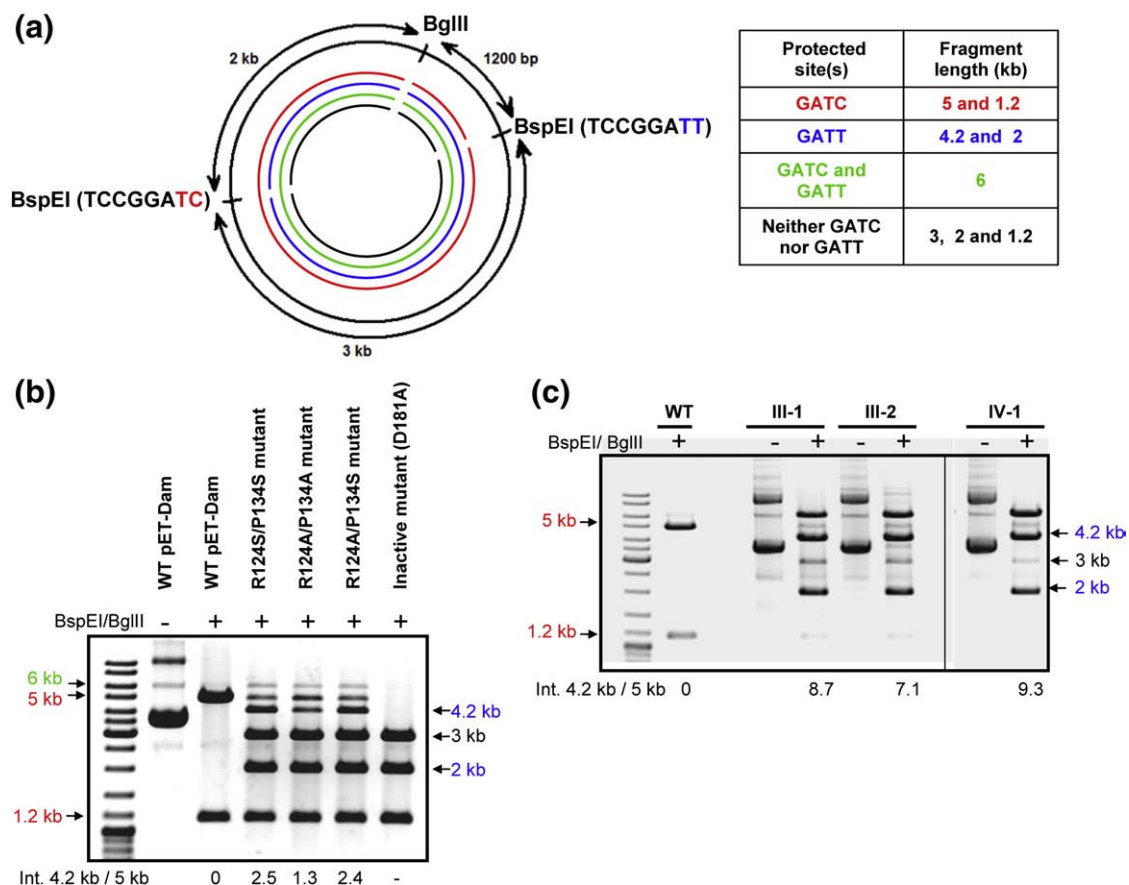


Fig. 5. *In vivo* specificity of *EcoDam* and its variants. The mutant *dam* expression plasmid containing two *Bsp*EI sites (one overlapping with GATC and the other with a GATT sequence) was grown in *dam*-negative *E. coli* cells. After expression of the *EcoDam* variants, plasmids were isolated and the methylation state at GATC and GATT sites was determined by digestion with *Bsp*EI. (a) Schematic drawing of the plasmid and the expected fragments after double digestion with *Bgl*II and *Bsp*EI. (b) Examples from the first-generation variants, which clearly show methylation at the GATT site. As an estimate of the specificity of the enzyme variants, the ratio of the intensities of the 4.2- and 5-kb bands (which are characteristic of GATT and GATC methylation, respectively) was determined. (c) Examples of the III-1 and III-2 clones from the shuffled library (third generation) and the IV-1 clone from the fourth generation, which clearly showed preferential methylation at the GATT site. Relative intensities of the 4.2- and 5-kb bands were indicated as described in (b).

In vitro investigation of enzyme specificity

Finally, we purified the initial R124S/P134S variant and the best variants obtained in the third and fourth generation and determined their DNA methylation specificity *in vitro* by DNA methylation kinetics using two oligonucleotide substrates containing one GATC or GATT site in an otherwise identical sequence context (Fig. 6). The GATC substrate was used in hemimethylated form to allow enzymatic methylation in only one strand and make the results comparable with the non-palindromic GATT substrate that has only one target adenine (in the upper DNA strand). Because of slow rates of GATT methylation by wild-type EcoDam and low activities of some of the selected variants, kinetics were performed under single-turnover conditions. The results indicate a complete switch in target-site specificity when changing from the wild-type enzyme (which shows a more than 60-fold preference for methylation at GATC sites) to the variants, such as III-1, which shows a more than 20-fold preference for methylation at GATT, overall corresponding to a 1600-fold change in specificity. The activity of the variants is not largely affected, and the enzyme variants methylate the new target-site GATT with rates comparable to the rate of wild-type EcoDam at GATC sites. For example, the activity of mutant IV-1 at GATT sites is 20% of wild-type activity at GATC sites, indicating that mutant IV-1 is a good methyltransferase. At the same time, the mutants show much lower activity at the original GATC site than wild type has at GATT sites; variant III-1 is virtually inactive at GATC sites.

Molecular role of the selected mutations

All mutations identified in this work after three cycles of random mutagenesis and selection for methylation of GATT sites (Table 3) are located on the surface of the enzyme, which is not unexpected, since exchange of surface residues has a lower risk to disrupt the protein structure. Two mutated residues (K139 and G256) are located close to the DNA (Fig. 7a), suggesting that they may affect DNA interaction. Interestingly, these residues approach the DNA sequence in the 5' direction of the GATC site and not close to the T4 of the GATT site, the recognition of which was altered here, suggesting that they are not directly involved in recognition of the T4. Other exchanges may affect protein structure and stability. For example, it has been reported that amino acid exchanges which introduce the consensus residue observed this position in related proteins often stabilize proteins.²⁸ We identify three such examples in our data set (Q75R, C87R, and S200G).

Conclusions

By combining directed evolution and rational design, it was possible to redesign the target sequence of an enzyme that specifically interacts with DNA and generate a new enzyme with high activity and

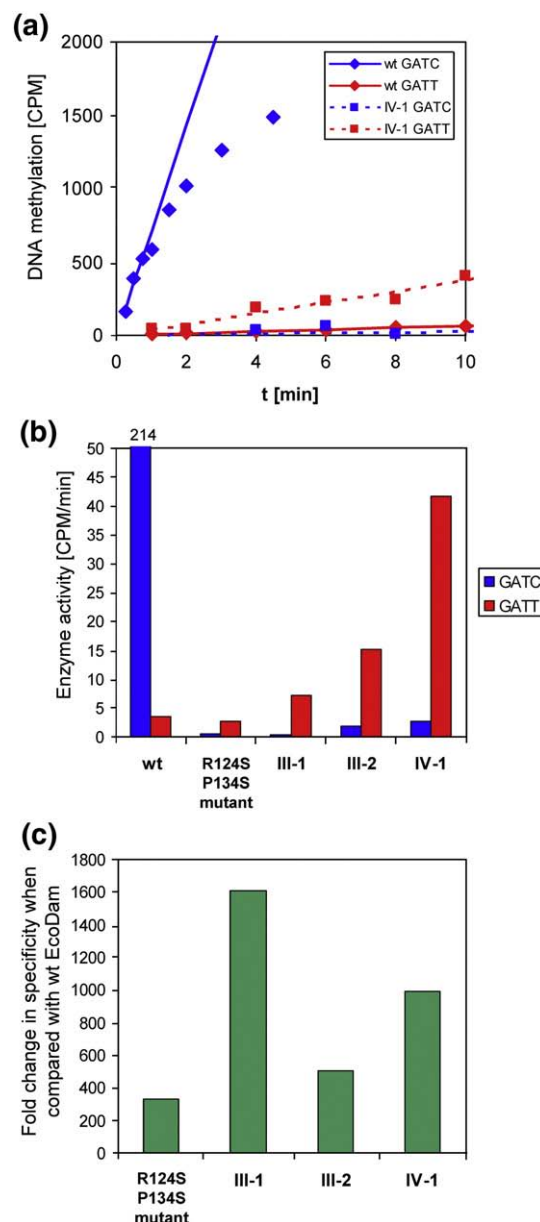


Fig. 6. Kinetic properties of wild-type EcoDam and some of its mutants. (a) Examples of DNA methylation kinetics with the purified wild-type EcoDam and mutant IV-1 tested using synthetic oligonucleotides containing a GATC or GATT site. All kinetics were carried out using 3 μ M enzyme, except GATC methylation by wild-type EcoDam, which was done using 0.6 μ M enzyme. (b) Comparison of the averaged initial slopes of reaction progress curves obtained from at least three independent methylation kinetics experiments. Deviations in slopes between the experiments were below $\pm 20\%$. For this comparison, the slope of GATC methylation by wild-type EcoDam was multiplied by a factor of 5, to account for the lower enzyme concentration used in the assay. (c) The fold change in specificity of the mutants in respect to wild type was calculated using the normalized methylation rates shown in (b).

specificity not present in nature. Epigenetic systems control important processes in bacteria including regulation of gene expression, phase variation, and

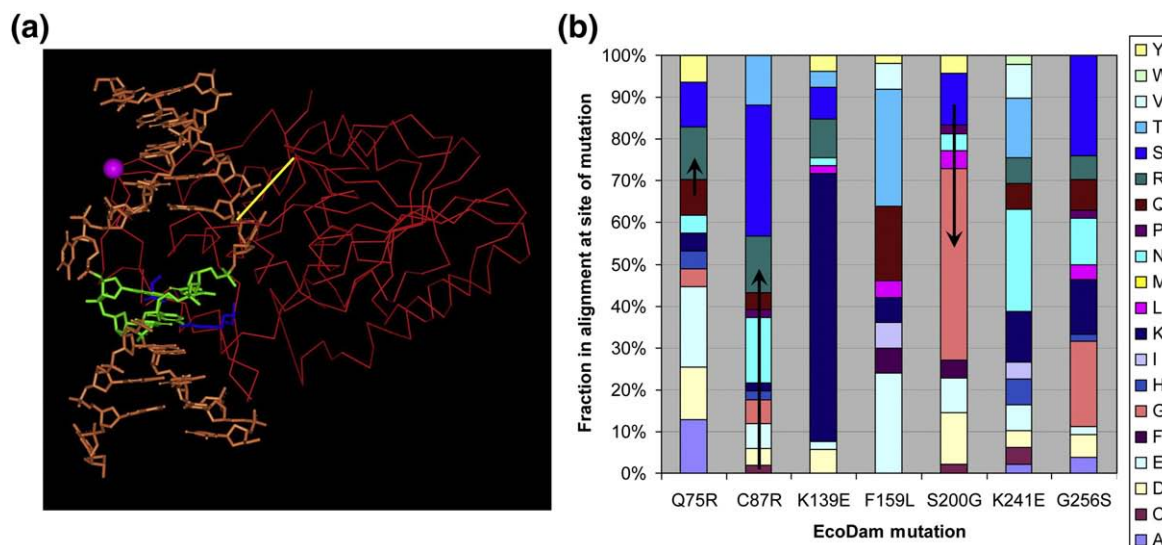


Fig. 7. Molecular interpretation of the identified mutations. (a) Structure of EcoDam in complex with specific DNA. The DNA is shown in orange with the third and fourth base pair of the GATC sequence highlighted in green. The enzyme is shown as C α trace in red. R124 and P134 are shown in blue; the C α position of K139 is highlighted by a pink sphere, and the position of the disordered loop containing G256 is shown in yellow. (b) Conservation of the mutated residues in an alignment of 54 EcoDam homologues. In three cases, we observed mutations that change the residue originally present in EcoDam to one more often observed in the multiple sequence alignment (Q75R, C87R, and S200G, indicated by arrows). The alignment quality was not good enough to draw conclusions on the conservation of EcoDam residues E84, P104, M231, S274, and A276.

pathogenicity.^{9,11} The generation of 6-methyladenine changes the structural properties of the DNA and influences DNA-protein interactions,²⁹ providing a general biological readout of the modification in the context of a new target sequence. DNA MTases with novel specificities might allow to construct new regulatory epigenetic systems in bacteria or constitute an artificial restriction-modification system.¹⁹

Materials and Methods

Site-directed and site-saturation mutagenesis

The RP protein library and the GANN substrate library were generated using randomized oligonucleotides by site-directed mutagenesis and PCR-megaprimer methods as described previously.²² Primers were purchased in purified form from MWG (Ebersbach, München). Mutagenesis was confirmed by restriction marker site analysis and DNA sequencing. Primers for the protein library contained NNB sequences at the sites of randomization (with N = A, G, C, or T and B = G, C, or T), which encode for all amino acids but exclude two of the three possible stop codons: CGGTTA-CAACGGCCTGTGTNNBTACAATCTCCGCGGTGAGT-TTAACGTGNNBTTCGGCCGCTA.

The sequences of primers used to generate the GANN substrate library was as follows:

primer 1, AACTATATCCGGANDGGCGAATGGGAC-GCGCCCTG (with D = A, G, or T)

primer 2, AACTATATCCGGAVCGGCGAATGGGA-CGCGCCCTG (with V = A, G, or C).

Primers 1 and 2 were used in 4:1 molar ratio to obtain an equal mixture of all GANN sites but exclude the GATC sequence.

Whole-genome randomization by epPCR

Whole-genome randomization was performed using epPCR by employing biased nucleotide composition, high Mg²⁺, and addition of Mn²⁺.²³ In order to get a high mutational load, we used 1 pg of template DNA and carried out 35 cycles of epPCR reaction. The epPCR reaction mixture (50 μ l) contained 10 mM Tris/HCl, pH 8.3, 50 mM KCl, 7 mM MgCl₂, 0.5 mM MnCl₂, 0.2 mM each of deoxyadenosine triphosphate and deoxyguanosine triphosphate, 1 mM each of deoxycytidine triphosphate and deoxythymidine triphosphate, 25 pmol each of the oligonucleotide primers (forward, CAT CAC AGC AGC GGC CTG GTG CCGC; reverse, TTT GTT AGC AGC CGG ATC CAG TG), 1 pg of template DNA, 1% dimethyl sulfoxide, and 2.5 U of Taq DNA polymerase (NEB). PCR conditions were 1 $\times$ 94 $^{\circ}$ C for 4 min; 35 $\times$ 94 $^{\circ}$ C for 45 s, 60 $^{\circ}$ C for 35 s, and 72 $^{\circ}$ C for 3 min; and, finally, 1 $\times$ 72 $^{\circ}$ C for 10 min. The epPCR product was excised from agarose gel and re-amplified by normal semi-nested PCR using 100 pg epPCR product (100 pg in 50 μ l) as template, forward primer (TTA ACT TTA AGA AGG AGA TAT ACC) and reverse primer (TTA GAG GCC CCA AGG GGT TAT G) (25 pmol each), deoxyribonucleotide triphosphate (0.2 mM each), 10 $\times$ ThermoPol buffer (NEB) (5 μ l), and Taq DNA polymerase (NEB) (2.5 U). PCR conditions were 1 $\times$ 94 $^{\circ}$ C for 4 min; 30 $\times$ 94 $^{\circ}$ C for 45 s, 60 $^{\circ}$ C for 35 s, and 72 $^{\circ}$ C for 1 min; and, finally, 1 $\times$ 72 $^{\circ}$ C for 10 min.

DNA shuffling

The third-generation mutant library was generated by *in vitro* DNA shuffling.²⁴ The EcoDam gene (1 kb) was PCR amplified from selected second-generation mutants (II-1, II-2, II-3, and II-4) using specific primers. DNA

(4 µg) containing equimolar amounts of all PCR products was digested with 0.15 U of DNaseI (NEB) in 100 µl of 50 mM Tris/HCl, pH 7.4, and 1 mM MgCl₂ for 10–20 min at room temperature. Fragments of 10–50 bp were purified from a 2% low melting agarose gel using the QIAEX II Gel Extraction Kit (QIAGEN). The purified fragments were resuspended in PCR mixture (0.2 mM each dNTP, 2.2 mM MgCl₂, 50 mM KCl, 10 mM Tris/HCl, pH 9.0, and 0.1% Triton X-100) at a concentration of 10–30 ng/µl. No primers were used at this point. *Taq* DNA polymerase (2.5 U) was added per 100 µl reaction mixture. A PCR program of 94 °C for 1 min, 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s (45 times) and 72 °C for 5 min was used to assemble full gene products. Finally, a conventional PCR was set up to amplify the EcoDam gene using a semi-nested primer set.

Protein expression and purification

EcoDam gene was cloned as His₆-fusion protein in pET28a (Clontech) expression vector with T7 inducible promoter. Wild-type EcoDam and all its variants were expressed in *E. coli* HMS174(DE3) (*F-recA 1 hsdR(r_{K12}m_{K12}) Rif^r*) (Novagen) and purified using Ni-NTA agarose (QiaGen) as described previously.³⁰

Restriction protection analyses

For restriction protection analysis, each transformed mutant was grown overnight in LB media in the presence of kanamycin (25 µg/ml) at 37 °C with shaking at 200 rpm. Overnight culture was diluted 1:100 into fresh LB media containing kanamycin and grown until an OD₆₀₀ (optical density at 600 nm) of 0.7. The culture was induced by addition of 1 mM IPTG and grown for 2 h, and the plasmid was isolated using Macherey–Nagel plasmid purification kit. Plasmids (1 µg) were digested with 10 U BspEI (New England BioLabs) in buffer recommended by the supplier. The reaction products were separated on agarose gels, DNA was stained with ethidium bromide, and images were taken using a gel documentation system. Images were analyzed with AIDA Image Analysis software (Raytest, Straubenhard, Germany) for semiquantitative assessment of the methylation levels.

DNA methylation kinetics

DNA methylation was analyzed under single-turnover conditions in 50 mM Hepes (pH 7.5), 50 mM NaCl, 1 mM ethylenediaminetetraacetic acid, 0.5 mM DTT, and 0.2 g/l bovine serum albumin containing 0.76 µM [methyl-³H]AdoMet (NEN) at 37 °C as described previously³¹ using 0.5 µM oligonucleotide substrate and 3 µM enzyme. Due to fast reaction rates, methylation of GATC sites by wild-type EcoDam was assayed using 0.6 µM enzyme. Initial turnover rates were derived by linear regression of the initial part of the reaction progress curves. The sequences of the 20-mer oligonucleotide substrates used for kinetics are as follows:

GATC-substrate, 5′-tatatccggatcgcgcaatg-3′/5′-Bt-cattcgccgMtcgggatata-3′
GATT-substrate, 5′-tatatccggattggcgcaatg-3′/5′-Bt-cattcgccaatccgggatata-3′
(with M=N6-methyladenosin and Bt=biotin).

Acknowledgements

This work has been supported by the Deutsche Forschungsgemeinschaft Priority program SPP 1170 (JE 252/5) and a stipend of the Ministry of Higher Education of the Arab Republic of Egypt to H.E.

References

- Benner, S. A. & Sismour, A. M. (2005). Synthetic biology. *Nat. Rev., Genet.* **6**, 533–543.
- Pleiss, J. (2006). The promise of synthetic biology. *Appl. Microbiol. Biotechnol.* **73**, 735–739.
- Channon, K., Bromley, E. H. & Woolfson, D. N. (2008). Synthetic biology through biomolecular design and engineering. *Curr. Opin. Struct. Biol.* **18**, 491–498.
- Meyer, A., Pellaux, R. & Panke, S. (2007). Bioengineering novel in vitro metabolic pathways using synthetic biology. *Curr. Opin. Microbiol.* **10**, 246–253.
- Jackel, C., Kast, P. & Hilvert, D. (2008). Protein design by directed evolution. *Annu. Rev. Biophys.* **37**, 153–173.
- Lacks, S. & Greenberg, B. (1977). Complementary specificity of restriction endonucleases of *Diplococcus pneumoniae* with respect to DNA methylation. *J. Mol. Biol.* **114**, 153–168.
- Hattman, S., Brooks, J. E. & Masurekar, M. (1978). Sequence specificity of the P1 modification methylase (M.Eco P1) and the DNA methylase (M.Eco dam) controlled by the *Escherichia coli* dam gene. *J. Mol. Biol.* **126**, 367–380.
- Jeltsch, A. (2002). Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *ChemBioChem*, **3**, 274–293.
- Wion, D. & Casadesus, J. (2006). N6-methyl-adenine: an epigenetic signal for DNA–protein interactions. *Nat. Rev., Microbiol.* **4**, 183–192.
- Jeltsch, A., Jurkowska, R. Z., Jurkowski, T. P., Liebert, K., Rathert, P. & Schlickerrieder, M. (2007). Application of DNA methyltransferases in targeted DNA methylation. *Appl. Microbiol. Biotechnol.* **75**, 1233–1240.
- Casadesus, J. & Low, D. (2006). Epigenetic gene regulation in the bacterial world. *Microbiol. Mol. Biol. Rev.* **70**, 830–856.
- Low, D. A., Weyand, N. J. & Mahan, M. J. (2001). Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infect. Immun.* **69**, 7197–7204.
- Heusipp, G., Falker, S. & Schmidt, M. A. (2007). DNA adenine methylation and bacterial pathogenesis. *Int. J. Med. Microbiol.* **297**, 1–7.
- Jeltsch, A., Sobotta, T. & Pingoud, A. (1996). Structure prediction of the EcoRV DNA methyltransferase based on mutant profiling, secondary structure analysis, comparison with known structures of methyltransferases and isolation of catalytically inactive single mutants. *Protein Eng.* **9**, 413–423.
- Breaker, R. R. & Joyce, G. F. (1994). Inventing and improving ribozyme function: rational design versus iterative selection methods. *Trends Biotechnol.* **12**, 268–275.
- Famulok, M., Hartig, J. S. & Mayer, G. (2007). Functional aptamers and aptazymes in biotechnology, diagnostics, and therapy. *Chem. Rev.* **107**, 3715–3743.
- Cohen, H. M., Tawfik, D. S. & Griffiths, A. D. (2004). Altering the sequence specificity of HaeIII methyl-

- transferase by directed evolution using in vitro compartmentalization. *Protein Eng. Des. Sel.* **17**, 3–11.
18. Timar, E., Groma, G., Kiss, A. & Venetianer, P. (2004). Changing the recognition specificity of a DNA-methyltransferase by in vitro evolution. *Nucleic Acids Res.* **32**, 3898–3903.
 19. Timar, E., Venetianer, P. & Kiss, A. (2008). In vivo DNA protection by relaxed-specificity SinI DNA methyltransferase variants. *J. Bacteriol.* **190**, 8003–8008.
 20. Aharoni, A., Gaidukov, L., Khersonsky, O., McQ, G. S., Roodveldt, C. & Tawfik, D. S. (2005). The 'evolvability' of promiscuous protein functions. *Nat. Genet.* **37**, 73–76.
 21. Peisajovich, S. G., Rockah, L. & Tawfik, D. S. (2006). Evolution of new protein topologies through multi-step gene rearrangements. *Nat. Genet.* **38**, 168–174.
 22. Jeltsch, A. & Lanio, T. (2002). Site-directed mutagenesis by polymerase chain reaction. *Methods Mol. Biol.* **182**, 85–94.
 23. Cadwell, R. C. & Joyce, G. F. (1992). Randomization of genes by PCR mutagenesis. *PCR Methods Appl.* **2**, 28–33.
 24. Crameri, A., Raillard, S. A., Bermudez, E. & Stemmer, W. P. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature*, **391**, 288–291.
 25. Horton, J. R., Liebert, K., Hattman, S., Jeltsch, A. & Cheng, X. (2005). Transition from nonspecific to specific DNA interactions along the substrate-recognition pathway of dam methyltransferase. *Cell*, **121**, 349–361.
 26. Horton, J. R., Liebert, K., Bekes, M., Jeltsch, A. & Cheng, X. (2006). Structure and substrate recognition of the *Escherichia coli* DNA adenine methyltransferase. *J. Mol. Biol.* **358**, 559–570.
 27. Miner, Z., Schlagman, S. L. & Hattman, S. (1989). Single amino acid changes that alter the DNA sequence specificity of the DNA-[N6-adenine] methyltransferase (Dam) of bacteriophage T4. *Nucleic Acids Res.* **17**, 8149–8157.
 28. Bershtein, S., Goldin, K. & Tawfik, D. S. (2008). Intense neutral drifts yield robust and evolvable consensus proteins. *J. Mol. Biol.* **379**, 1029–1044.
 29. Marinus, M. G. & Casadesus, J. (2009). Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiol. Rev.* **33**, 488–503.
 30. Urig, S., Gowher, H., Hermann, A., Beck, C., Fatemi, M., Humeny, A. & Jeltsch, A. (2002). The *Escherichia coli* dam DNA methyltransferase modifies DNA in a highly processive reaction. *J. Mol. Biol.* **319**, 1085–1096.
 31. Roth, M. & Jeltsch, A. (2000). Biotin-avidin microplate assay for the quantitative analysis of enzymatic methylation of DNA by DNA methyltransferases. *Biol. Chem.* **381**, 269–272.

Supplement (2)

Transition from EcoDam to T4Dam DNA recognition mechanism without
loss of activity and specificity

Transition from EcoDam to T4Dam DNA Recognition Mechanism without Loss of Activity and Specificity

Hany Elsaywy,^[a] Sveatoslav Podobinschi,^[b] Sanjay Chahar,^[a] and Albert Jeltsch^{*[a]}

The EcoDam and T4Dam DNA-(adenine N6)-methyltransferases both methylate the adenine residue in GATC sites. These enzymes are highly related in amino acid sequence, but they deviate in their contact to the first base pair of the target sequence. EcoDam contacts Gua1 with K9 (which corresponds to T4Dam A6), while T4Dam contacts Gua1 with R130 (which corresponds to EcoDam Y138). We have “transplanted” the T4Dam DNA recognition into EcoDam and show that the EcoDam K9A/Y138R double mutant is highly active and specific. We also studied the intermediates of this transition: The EcoDam K9A variant showed low activity and loss of recognition of Gua1 [Horton, et al., *J. Mol. Biol.* **2006**, 358, 559–570]. In

contrast, the EcoDam Y138R variant, which carries both Gua1 recognition elements (K9 from EcoDam and R138 corresponding to R130 from T4Dam), is fully active and specific. This result indicates that a smooth evolutionary pathway exists for changing the EcoDam DNA recognition mode to T4Dam without loss of activity and without generation of evolutionary intermediates with reduced activity. We consistently observed increased activity of EcoDam variants containing Y138R; this suggests that the transition from EcoDam (Gua1 recognition through K9) to T4Dam (Gua1 recognition through R130) was driven by selective pressure towards increased catalytic activity.

Introduction

Post-replicative methylation of DNA is catalyzed by DNA methyltransferases (MTases) that transfer methyl groups from AdoMet (S-adenosyl-L-methionine) to nucleobases in DNA to produce S-adenosyl-L-homocysteine (AdoHcy) and 5-methylcytosine, N⁴-methylcytosine, or N⁶-methyladenine.^[1–3] Most prokaryotic DNA MTases are components of restriction modification (RM) systems that serve to protect bacteria against bacteriophage infection and control uptake and incorporation of exogenous DNA. In RM systems, DNA entering a cell is cleaved by a restriction endonuclease at specific, often palindromic sites, while the endogenous DNA is protected by methylation of a nucleobase within the restriction site.^[4] However, DNA from bacteriophages grown in a bacterial strain that harbors one RM system is methylated; this means that no protection is provided. For this reason, RM systems are enormously divergent in nature and have adapted to cleave DNA at different sites; this is similar to the high diversification of genes in the mammalian immune system. The great target sequence variability of RM systems is illustrated by the finding that by 2007 about 8000 such systems had been identified or bioinformatically predicted and more than 250 different specificities were known.^[5] The DNA methyltransferases of all these systems are clearly related and are the outcome of divergent evolution. (The restriction enzymes also are related but amino acid sequence similarity often is much lower.) Hence, the question appears of how these enzymes managed to change their DNA recognition specificity through molecular evolution. Experimental studies suggest that evolution went through intermediates of reduced specificity,^[6] but it remains unclear if such intermediates could have been functional.

In addition to RM-based DNA MTases, many bacteria (and their phages) contain orphan MTases, like the Dam-related MTases found in γ -proteobacteria, which methylate the N6 atom of the exocyclic amino group of adenine residues in the palindromic sequence GATC.^[7,8] Dam-dependent methylation considerably affects bacterial gene expression, replication initiation, postreplicative mismatch repair and pathogenicity.^[2,3,9–11] Interestingly, bacteriophage T4 (and T2) also encode Dam MTases,^[12] most likely to ensure efficient methylation of the bacteriophage DNA during the massive viral DNA replication. EcoDam has considerable similarity to T4Dam in amino acid sequence and structure.^[13,14] However, the mechanism of recognition of the first base pair is an interesting deviation between them (Figure 1). In T4Dam, the Gua1 base is contacted by R130, which is located at the basis of the β -hairpin module that mediates all other sequence specific contacts to the third and fourth base pair of the recognition sequence.^[15] EcoDam carries a Tyr at the corresponding position (Y138) that does not play an important role in the Gua1 recognition.^[14] Instead, EcoDam contains a Lys in the N-terminal loop of the protein (K9) that contacts the Gua1. T4Dam contains an Ala at the cor-

[a] H. Elsaywy, S. Chahar, Prof. Dr. A. Jeltsch
Biochemistry Laboratory, School of Engineering and Science
Jacobs University Bremen
Campus Ring 1, 28725 Bremen (Germany)
Fax: (+ 49) 421-200-3249
E-mail: a.jeltsch@jacobs-university.de

[b] S. Podobinschi
Biochemistry and Cell Biology Program
School of Engineering and Science, Jacobs University Bremen
Campus Ring 1, 28725 Bremen (Germany)

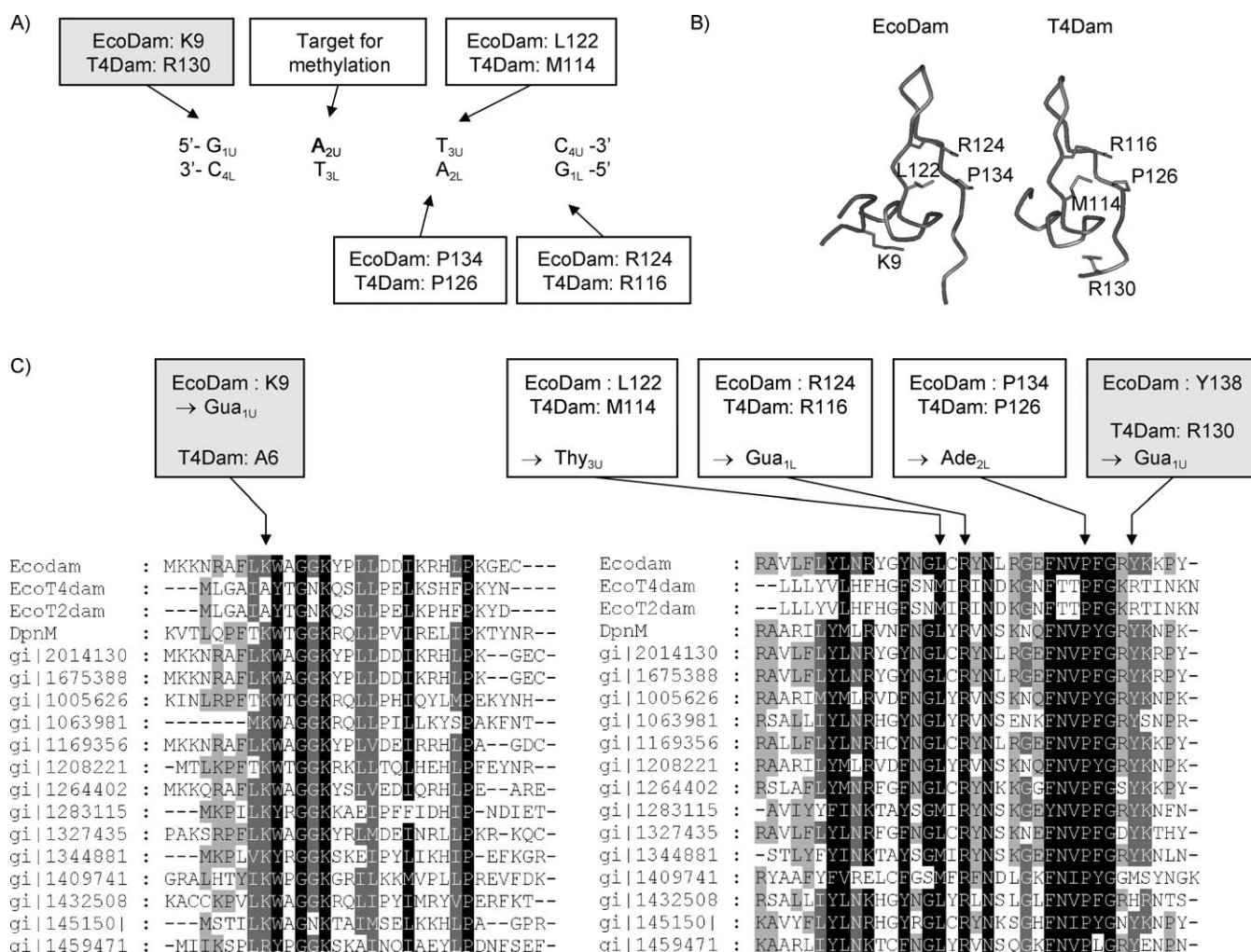


Figure 1. DNA recognition by EcoDam and T4Dam. A) Schematic drawing of the DNA recognition elements of EcoDam and T4Dam. The upper DNA strand is labeled with "U", the lower DNA strand with "L". B) Structure of T4Dam and EcoDam showing the β -hairpin of both enzymes and the N-terminal loop of EcoDam. The protein backbone is shown in α -tube representation. The side chains of the residues involved in the recognition of the first, third and fourth base pair of the GATC sequence are displayed. (A colored version of this figure also displaying the DNA is shown in the table of contents.) C) Multiple sequence alignment of EcoDam, T4Dam, T2Dam and DpnM with Dam methyltransferases from various bacteria (specified by their GI number).

responding position. The EcoDam/T4Dam pair represents a very interesting model case to study molecular evolution. Multiple sequence alignments of members of the EcoDam family indicate that the EcoDam mechanism is evolutionary ancient and the T4Dam setup has evolved from it because most homologues of EcoDam in other bacteria have a Lys at the K9 position, while the T4Dam configuration is unique for the T2Dam and T4Dam enzyme pair (Figure 1C). In addition, we have shown that M.EcoRV (a distant homologue of EcoDam) also recognizes DNA with a Lys residue at the position equivalent to K9.^[16] Here, in an attempt to "transplant" the T4Dam DNA recognition modules into EcoDam, the residues making contact with the first base pair in EcoDam were mutated to the corresponding residues in T4Dam. In addition, we prepared the evolutionary intermediates of this transition and studied their catalytic activity and DNA recognition specificity.

Results

Experimental design

To study the evolutionary transition of DNA recognition from EcoDam to T4Dam, we prepared the EcoDam Y138R variant, which contains an Arg at position 138 where it contacts Gua1 in T4Dam (Figure 1). (Throughout this paper amino acid positions are given for EcoDam, which is the experimental model system used. The corresponding T4Dam positions are specified in Figure 1). This variant corresponds to an evolutionary intermediate of the EcoDam to T4Dam transition and contains both residues responsible for Gua1 recognition in the two systems (Lys9 and Arg138). The EcoDam K9A variant was prepared previously^[14] and investigated here again. This variant corresponds to an evolutionary intermediate that had lost both residues for Gua1 recognition carrying an Ala at position 9 and Tyr at position 138. Finally, we created the K9A/Y138R double mutant in an attempt to completely transplant the T4Dam Gua1 recognition elements into EcoDam. In addition,

because we wanted to study if the amino acid residues flanking residue 138 would influence DNA recognition, we prepared an EcoDam variant in which Tyr138 together with its two flanking residues (Arg137 and Lys139) were changed to their counterparts in T4Dam creating an EcoDam R137 Y138 K139→KRT variant (KRT variant). Finally, the KRT variant was combined with K9A, which again transplanted the T4Dam recognition elements into EcoDam.

Site-directed mutagenesis and protein purification

All mutants were generated by site-directed mutagenesis. Successful mutagenesis and the absence of additional mutations was confirmed by sequencing of the complete EcoDam genes. The His₆-tagged wild-type EcoDam, K9A variant^[14] and all variants generated here were expressed in *E. coli* and purified using Ni-NTA agarose to >95% purity. All protein preparations were of comparable purity and protein concentration (Figure 2A).

Kinetic characterization of the K9A and Y138R EcoDam variants

Next we studied the methylation of oligonucleotide substrates containing a single GATC, AATC, TATC and CATC site to determine the activity of the variants and study their recognition of the first base pair of the EcoDam recognition sequence. We used oligonucleotides that were methylated in the lower DNA strand to prevent enzymatic methylation of the lower strand. This strategy allows the association of the experimental methylation rates with the defined target sequence, which carries alterations in Gua1. This clear-cut interpretation of results would not have been possible with unmethylated substrates because, for example, AATC methylated in the lower strand corresponds to the enzyme interacting with a GATT target site. Examples of the DNA methylation reactions are shown in Figure 2B. All experiments were carried out in triplicate and the initial slopes were averaged. The deviations of the relative

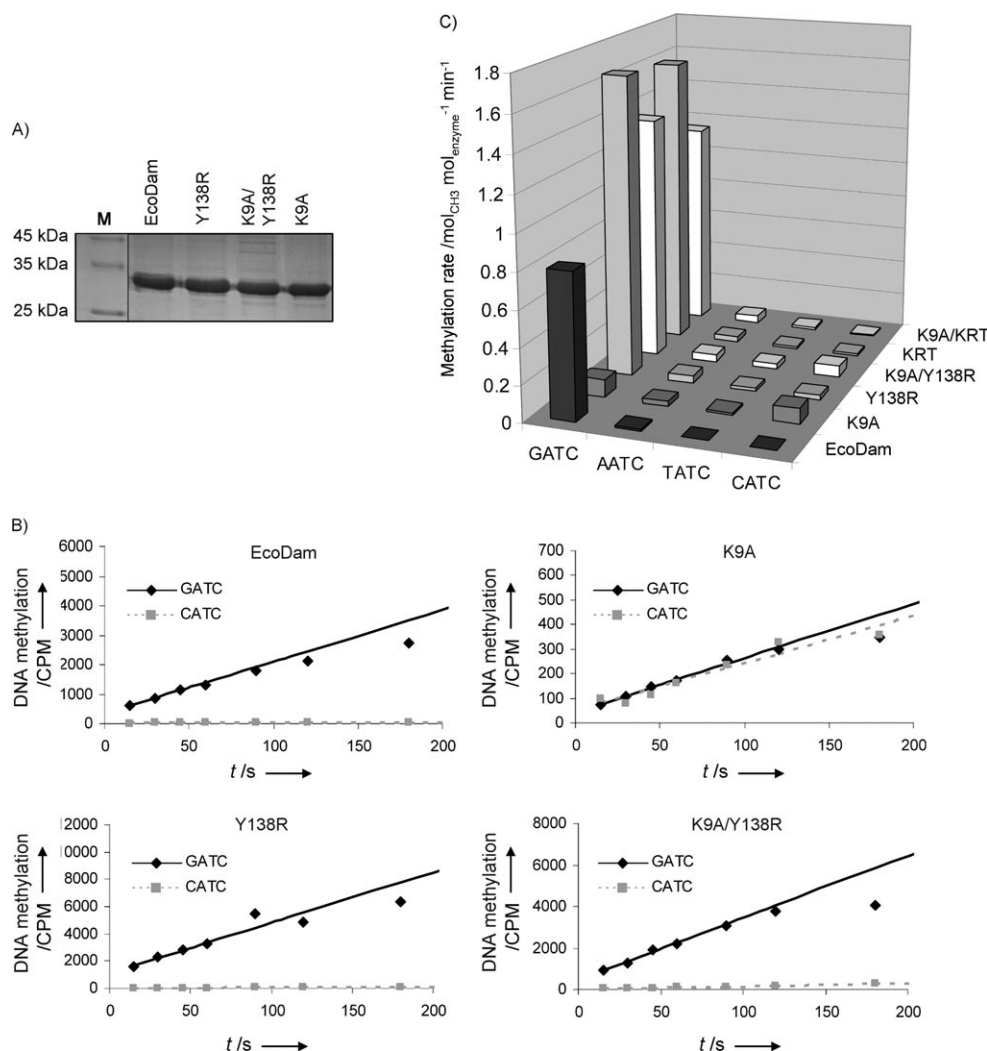


Figure 2. Kinetic properties of EcoDam variants. A) Coomassie stained SDS polyacrylamide gel showing examples of the purified EcoDam variants. B) Examples of the time courses of DNA methylation (measured in CPM) by wild-type EcoDam and the EcoDam variants. In total, the activity of wild-type EcoDam and all five variants was investigated with four substrates. All kinetics were carried out in triplicate and normalized to wild-type reactions performed in parallel on the same plate and averaged. C) Compilation of the catalytic activities of all mutants with all substrates. The deviations of relative mutant activities were below $\pm 5\%$ in all cases.

mutant activities were below $\pm 5\%$ in all cases. All results are summarized in Figure 2C.

The results obtained here with wild-type EcoDam closely match previous data with respect to absolute rates and specificity.^[14,17] The rate of methylation of GATC substrates by the wild-type enzyme corresponded to 0.8 turnovers min^{-1} . The enzyme was highly specific and substrates modified at Gua1 were methylated with at least 80-fold reduced rates. Similarly to results obtained previously, K9A had reduced activity and showed strong reduction of recognition of the first base pair. The K9A/Y138R was slightly more active than wild-type EcoDam and the variant was fully specific. This result indicates that it was possible to move the T4Dam DNA recognition element R130 into EcoDam, where it can completely replace the EcoDam K9 module. Interestingly, the Y138R variant was fully active and specific, although it harbored both Gua1 recognition elements (K9 of EcoDam and R130 from T4Dam). This result shows that there exists a continuous evolutionary pathway from EcoDam to T4Dam that does not include intermediates of reduced activity or specificity.

Kinetic characterization of the KRT and K9A/KRT EcoDam variants

In the KRT and K9A/KRT variants, the Arg138 was moved from T4Dam into EcoDam together with its two neighboring amino acid residues to investigate if the local context could help R138 to function in *E. coli*. As described in the last paragraph, we found that even in the single Y138R context, R138 was

fully functional in *E. coli*. This result was confirmed by the findings with the KRT and K9A/KRT variants (Figure 2C), which behave very similar to the Y138R and K9A/Y138R variants. Interestingly, the Y138R and KRT variants were about twofold more active than the wild-type enzyme. In cases in which these mutations are combined with K9A, activities dropped by about 20–25% but still stayed above wild-type activity. The only small difference in the behavior between Y138R and KRT was observed in the methylation of CATC. While the K9A/Y138R double mutant showed some residual activity at CATC sites, the K9A/KRT variant did not; this indicates that the specificity of the DNA interaction is improved by the positioning of R138 in a T4Dam-like context.

DNA binding of wild-type EcoDam and the K9A/Y138R EcoDam variant

We have investigated DNA binding of EcoDam by gel retardation experiments (Figure 3). The results showed the appearance of a ladder of bands, which is indicative of relatively non-specific DNA interaction. It is likely that this additional nontarget binding occurs at near-cognate sites, which differ from GATC in only one base pair similarly as we have shown for M.EcoRV recently.^[16] Addition of the AdoMet analogue sinefungin increased DNA binding of wild-type EcoDam about 2.5-fold, as indicated by the finding that the amount of bound DNA with 5 μM EcoDam roughly equals the result observed at 2 μM EcoDam in the presence of sinefungin. We have shown previously that AdoMet is required for base flipping by

EcoDam [μM]	-	0.5	1	2	5		0.5	1	2	5
Sinefungin	-	-	-	-	-		+	+	+	+

K9A/Y138R [μM]	-	0.5	1	2	5		0.5	1	2	5
Sinefungin	-	-	-	-	-		+	+	+	+

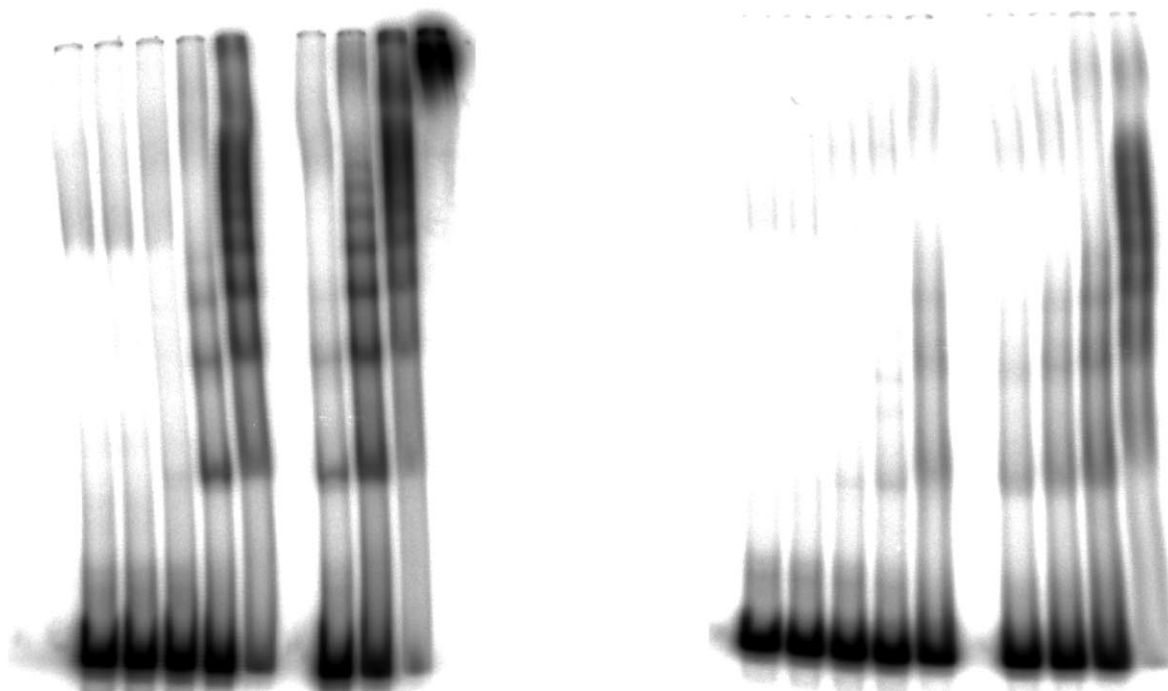


Figure 3. DNA binding by wild-type EcoDam (left panel) and the K9A/Y138R variant (right panel) studied by gel-retardation experiments. The free DNA is shown in the first lane in both panels, the samples without and with sinefungin were separated by an empty lane.

EcoDam,^[18] and this suggests that the formation of a tighter EcoDam–DNA complex is accompanied by base flipping. DNA binding of the K9A/Y138R variant was weaker than wild type, but still nonspecific. DNA binding of the K9A/Y138R variant was stimulated by addition of sinefungin to a similar degree as observed with wild-type EcoDam. This observation suggests that the transition of weakly bound DNA into tightly bound DNA with a flipped target base is similar in the K9A/Y138R variant as in wild-type EcoDam.

Discussion

Understanding the pathways of natural evolution is a major scientific challenge. In addition, it will help our efforts to redesign protein function by rational and evolutionary approaches. The EcoDam and T4Dam DNA-(adenine N6)-methyltransferases both methylate the adenine residue in GATC sites and they are highly related in amino acid sequence.^[14,15] We have shown previously that they interact with the first base pair of their target sequence with different structural modules.^[14] EcoDam contacts Gua1 with K9 where T4Dam carries an Ala, while T4Dam contacts Gua1 with Arg138, where EcoDam carries a Tyr. Thus, the EcoDam/T4Dam pair represent a very interesting model case to study molecular evolution. Multiple sequence alignments of members of the EcoDam family indicate that the EcoDam like mechanism of DNA recognition is evolutionary ancient and the T4Dam setup has evolved from it, because most homologues have a Lys at the K9 position. In addition, we have shown that M.EcoRV (a distant homologue of EcoDam) also recognizes DNA with a Lys residue at the position equivalent to K9.^[16]

In an attempt to “transplant” the T4Dam DNA recognition into EcoDam, we have prepared an EcoDam K9A/Y138R double mutant that established the T4Dam DNA interaction modules in an EcoDam framework and showed that the variant is active and specific. Previously, we showed that the EcoDam K9A variant (which lacks both Gua1 recognition elements from EcoDam and T4Dam) shows low activity and loss of recognition of the Gua1.^[14] Here we show that the EcoDam Y138R variant, which carries both Gua1 recognition elements at the same time (K9 from EcoDam and R138 corresponding to R130 in T4Dam), is fully active and specific and shows very similar DNA binding behavior as wild-type EcoDam. This is an interesting example of an evolutionary change of enzyme functional elements that is not accompanied by a loss in activity and specificity. It shows that there is a smooth evolutionary pathway from the EcoDam to the T4Dam DNA recognition mode without loss of activity and without generation of intermediates with reduced activity.

It is striking that all variants containing the Y138R exchange displayed an increase in activity as compared to the wild-type enzyme. We have previously shown that in EcoDam, formation of the Gua1 interaction is a late step in the conformational rearrangement that occurs during target sequence interaction and requires the ordering of the N-terminal loop of EcoDam.^[17] In contrast to K9, R138 is located at the base of the β -hairpin and it can interact with Gua1 without the requirement for con-

formational change; this might explain the increase in activity in all variants containing R138. We conclude that the evolutionary transition from EcoDam to T4Dam might be driven by a selection pressure for increased catalytic activity.

Experimental Section

Oligonucleotides: Oligonucleotides were purchased from Thermo Hybaid (Ulm, Germany) in HPLC-purified grade. The following double-stranded oligonucleotide substrates obtained by annealing of complementary single-stranded oligonucleotides were used (bold and underlined corresponds to recognition site; M = N⁶-methyladenosine and Bt = Biotin):

GATC substrate:

5'-GCGACAGT**GATC**GGCCTGTC-3'
3'-CGCTGTCA**CTMG**CCGGACAG-Bt

AATC substrate:

5'-GCGACAGT**AATC**GGCCTGTC-3'
3'-CGCTGTCA**ATTMG**CCGGACAG-Bt

TATC substrate:

5'-GCGACAGT**TATC**GGCCTGTC-3'
3'-CGCTGTCA**ATMG**CCGGACAG-Bt

CATC substrate:

5'-GCGACAGT**CATC**GGCCTGTC-3'
3'-CGCTGTCA**CTMG**CCGGACAG-Bt

To anneal double-stranded substrates, equimolar (20 μ M) amounts of each strand were mixed, heated to 95 °C for 5 min, and slowly cooled to RT.

Site-directed mutagenesis, protein expression, and protein purification: The EcoDam variants were generated by the megaprimer site-directed mutagenesis method as described.^[19] Wild-type EcoDam and its variants were purified as described.^[18]

Kinetic analysis: Quantitative determination of the catalytic activity of the wild-type EcoDam and its variants was carried out in vitro with a biotin/avidin methylation assay as described.^[20] Methylation experiments were performed in reaction mixtures (40 μ L) containing oligonucleotide (0.5 μ M), enzyme (0.2 μ M), [³H]AdoMet (5.5 μ M), bovine serum albumin (0.2 μ g/ μ L), Hepes pH 8.0 (100 mM), EDTA (1 mM), NaCl (50 mM), and DTT (0.5 mM), at ambient temperature. All kinetics were done at least in triplicate and the initial slopes averaged. The results were normalized to reactions performed with wild-type EcoDam on the same plate to define the relative mutant activity.

Electrophoretic mobility shift assay (EMSA): To investigate the DNA binding by EcoDam, a 150 bp DNA fragment containing a single EcoDam target site (GATC) was used. The DNA substrate was amplified by PCR amplification from pET28a(+) by using the following oligonucleotides: 5'-ATC CCA CTA CCG AGA TAT CCG CAC C-3' and 5'-TGT CCG GTT TTC AAC AAA CCA TGC-3' and radioactively labeled with [γ -³²P]ATP with T4 polynucleotide kinase (New England Biolabs, Ipswich, MA, USA). Gel retardation experiments were basically carried out as described.^[21] The reaction mixtures (20 μ L) contained Hepes pH 8 (10 mM), EDTA (0.1 mM), NaCl (5 mM), DTT (50 μ M), acetylated bovine serum albumin (1 μ g), sinefungin (Sigma) (0.5 mM), DNA substrate (20 nM) and different amounts of enzyme (0.5–5 μ M). After incubation for 30 min at RT, reactions were mixed with glycerol (4 μ L of 87%) and loaded onto a native polyacrylamide gel (8%) in TBE buffer (0.5 $\times$), followed by

overnight electrophoresis in a cold room (150 V, 65 mA). The gel was dried and the separated complexes were visualized by using autoradiography phosphorimager system (Fuji Film Life Science, Woodbridge, CT, USA).

Acknowledgements

This work was supported by DFG (JE 252/2 and JE 252/5) and by a grant of the Egyptian government to H.E.

Keywords: DNA methyltransferase • DNA • enzyme specificity • enzymes • molecular evolution

- [1] A. Jeltsch, *ChemBioChem* **2002**, 3, 274–293.
- [2] J. Casadesus, D. Low, *Microbiol. Mol. Biol. Rev.* **2006**, 70, 830–856.
- [3] D. Wion, J. Casadesus, *Nat. Rev. Microbiol.* **2006**, 4, 183–192.
- [4] A. Pingoud, A. Jeltsch, *Nucleic Acids Res.* **2001**, 29, 3705–3727.
- [5] R. J. Roberts, T. Vincze, J. Posfai, D. Macelis, *Nucleic Acids Res.* **2007**, 35, D269–270.
- [6] H. M. Cohen, D. S. Tawfik, A. D. Griffiths, *Nucleic Acids Res.* **2002**, 30, 3880–3885.
- [7] S. Lacks, B. Greenberg, *J. Mol. Biol.* **1977**, 114, 153–168.
- [8] S. Hattman, J. E. Brooks, M. Masurekar, *J. Mol. Biol.* **1978**, 126, 367–380.
- [9] A. Lobner-Olesen, O. Skovgaard, M. G. Marinus, *Curr. Opin. Microbiol.* **2005**, 8, 154–160.
- [10] G. Heusipp, S. Falker, M. A. Schmidt, *Int. J. Med. Microbiol.* **2007**, 297, 1–7.
- [11] A. Jeltsch, R. Z. Jurkowska, T. P. Jurkowski, K. Liebert, P. Rathert, M. Schlickerrieder, *Appl. Microbiol. Biotechnol.* **2007**, 75, 1233–1240.
- [12] S. L. Schlagman, S. Hattman, *Gene* **1983**, 22, 139–156.
- [13] S. Hattman, J. Wilkinson, D. Swinton, S. Schlagman, P. M. Macdonald, G. Mosig, *J. Bacteriol.* **1985**, 164, 932–937.
- [14] J. R. Horton, K. Liebert, M. Bekes, A. Jeltsch, X. Cheng, *J. Mol. Biol.* **2006**, 358, 559–570.
- [15] J. R. Horton, K. Liebert, S. Hattman, A. Jeltsch, X. Cheng, *Cell* **2005**, 121, 349–361.
- [16] T. P. Jurkowski, N. Anspach, L. Kulishova, W. Nellen, A. Jeltsch, *J. Biol. Chem.* **2007**, 282, 36942–36952.
- [17] K. Liebert, J. R. Horton, S. Chahar, M. Orwick, X. Cheng, A. Jeltsch, *J. Biol. Chem.* **2007**, 282, 22848–22855.
- [18] K. Liebert, A. Hermann, M. Schlickerrieder, A. Jeltsch, *J. Mol. Biol.* **2004**, 341, 443–454.
- [19] A. Jeltsch, T. Lanio, *Methods Mol. Biol.* **2002**, 182, 85–94.
- [20] M. Roth, A. Jeltsch, *Biol. Chem.* **2000**, 381, 269–272.
- [21] L. M. Hellman, M. G. Fried, *Nat. Protocols* **2007**, 2, 1849–1861.

Received: July 16, 2009

Published online on September 1, 2009

Curriculum Vitae

Personal Data



Name: Hany Amin Elsaywy Mostafa
Address: Biochemistry Division,
Chemistry department,
Tanta university,
Tanta, Egypt.
E-mail: hanielsawy@yahoo.com h.elsawy@jacobs-university.de
Date of Birth: November 30th 1973
Nationality: Egyptian
Marital status: Married, 3 children
Military service: Exempted
Languages: English, German (ZD), Arabic (mother tongue)

Education:

- PhD in Biochemistry, Jacobs University Bremen, School of Science and Engineering, thesis supervisor Prof. Dr. Albert Jeltsch, 2009.
- Master of Science (Biochemistry), Alexandria University, Egypt, 2001.
- Bachelor of Science (Biochemistry, special degree), Alexandria University, Egypt, 1995.

Qualifications:

- Elective training in Rakta Co. for papers, 1993.
- Elective training in Medical Research Institute, Alex. University.
- Preparatory studies in medical science (for preparation of medical research) organized by medical research institute, Alex. University, during the academic year 1996/1997 with grade (Excellent).

- Attended a course in (Principles of genetic engineering) organized by institute of graduate studies and research, Alex. University, during the academic year 2000/2001.

Jobs and previous positions:

1996 Domenstrator at Medical Research Institute, Biochemistry department, Alexandria, Egypt.

1997-2001 Domenstrator at Tanta University, Faculty of Science, Biochemistry department, Tanta, Egypt.

Since 2001 Assistant lecturer at Tanta University, Faculty of Science, Biochemistry department, Tanta, Egypt.

Publications:

Chahar, S., **Elsawy, H.**, Ragozin, S. and Jeltsch, A. (2009) *Journal of Molecular Biology*, in press.

Elsawy, H., Podobinschi, S., Chahar, S. and Jeltsch, A. (2009) *ChemBioChem*. 10, 2488-2493.

Balbaa, M., Mansour, H., **Elsawy, H.** and El-Ashry, E. S. H. (2002) *Nucleosides, Nucleotides and Nucleic acids*, 21(10), 695-708.