

Investigations into the mode of action of the
DNA uridine endonuclease Mth212 of
Methanothermobacter thermautotrophicus ΔH

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August Universität zu Göttingen

Vorlegt von
Elena Ciirdaeva (geb. Jivotovscaia)
aus Chişinău, Moldova

Göttingen 2009

D7

Referent:

Prof. Dr. Hans-Joachim Fritz

Korreferent:

PD Dr. Wilfried Kramer

Tag der mündlichen Prüfung:

22. Januar 2010

Table of contents

1 Introduction	
1.1 DNA structure, stability, and dynamics	1
1.1.1 Hydrolytic DNA damage	2
1.1.2 Hydrolytic cytosine deamination	4
1.1.3 Phylogenetic distribution of uracil-DNA glycosylases	5
1.2 Uracil-DNA glycosylases in Archaea	7
1.2.1 <i>Methanothermobacter thermautotrophicus</i> as a model organism	8
1.2.2 Discovery of Mth212 as a DNA uridin endonuclease in <i>M.thermautotrophicus</i>	8
1.3 Aim of project	10
2 Materials and Methods	11
2.1 Materials	11
2.1.1 Bacterial strains	11
2.1.2 Plasmids	11
2.1.3 Primers for polymerase chain reaction	12
2.1.4 2'-Desoxyriboseoligonucleotides for enzymatic activity assay	13
2.1.5 Molecular Ladders and Markers	14
2.1.6 Enzymes und Proteins	14
2.1.7 Chemicals and Reagents	15
2.1.8 Molecular Biology Kits	16
2.1.9 Buffers and Solutions	16
2.1.10 Bacterial Growth Media	20
2.1.11 Molecular Biology Equipment and computer hardware	20
2.1.12 Other materials	21
2.1.13 Software	21
2.1.14 Databanks	22
2.2 Methods	22
2.2.1 Microbiological methods	22
2.2.2 Molecular biology methods	23
2.2.3 Protein biochemical methods	31
3 Results and discussion	36
3.1 Construction and properties of Mth212/D151N mutant	36
3.1.1 Rationale for mutation of conserved Asp-151	36
3.1.2 Mth212/D151 mutants: construction and expression	37
3.1.3 Endonuclease assays with Mth212/D151 variants	41
3.1.4 Gel-retardation assays with Mth212/ D151 variants	43

3.1.4.1 EMSA with Mth212/D151N and Substrate I.....	43
3.1.4.2 EMSA with Mth212/D151N, Substrate I and pET-vector as a competitor DNA ...	45
3.1.4.3 EMSA with Mth212/D151A and Mth212/D151S, Substrate I and pET- vector as a competitor DNA	47
3.1.4.4 EMSA with Mth212/D151N, Substrate II and 20-mer oligonucleotides as a competitor DNA	48
3.1.4.5 EMSA with Mth212/D151N, Substrate III and 20-mer oligonucleotides as a competitor DNA	50
3.1.4.6 EMSA with Mth212/D151N, Substrate III (with U/C, U/T and U/A) and 20-mer oligonucleotides as a competitor DNA	52
3.2 Semi-quantitative investigation of wt Mth212 substrate preferences	55
3.3 Mth212- structure analysis	58
4 Summary	62
5 Abbreviations	64
6 References	67
7 Appendix.....	73
7.1 Sequences (attached CD).....	73
7.1.1 pET_B_001 sequence and restriction map.....	73
7.1.2 Mth212 nucleotide and amino acid sequences, restriction map.....	73
7.1.3 Mth212/D151X (X: N, A, S) mutants verification.....	73
7.1.3.1 pET_B001_Mth212/D151N sequencing data	73
7.1.3.2 pET_B001_Mth212/D151A sequencing data.....	73
7.1.3.3 pET_B001_Mth212/D151S sequencing data.....	73
Acknowledgments.....	74
Curriculum vitae.....	75

1 Introduction

1.1 DNA structure, stability, and dynamics

DNA is the material bearer of hereditary information in all living cells and DNA viruses. This information is written in a sequence of deoxyribonucleotide bases - the structural monomers of DNA molecule. DNA is a polymer composed of covalently bound mononucleotides of two types: pyrimidines (thymine (T), cytosine (C)) and purines (adenine (A) guanine (G)). In solution DNA molecule exists in form of a double helix. Two polynucleotide chains are held together by hydrogen bounding between complementary bases (A:T and G:C) and by hydrophobic interaction between adjacent stacked base pairs.

The realisation that DNA directs all cell activities suggests that this molecule is relatively stable. However, studies of DNA metabolism processes such as replication, transcription, repair, and recombination revealed dynamical state of DNA molecule in the cell. Furthermore, different extrinsic and intrinsic factors, also known as mutagens, induce structural changes in DNA molecule, evidencing its limited chemical stability. The major extrinsic mutagens that impair DNA integrity are harmful chemicals (e.g. oxidative and alkylating agents), ionizing radiation, and UV-light. The intrinsic mutagenic factors include reactive cellular metabolites (e.g. reactive oxygen species (ORS), **S**-adenosyl-L-methionine (SAM) and water (Nickoloff, 1998).

Structural changes in DNA, if not repaired rapidly and appropriately, represent a serious threat to genomic stability and cell survival. Fortunately, during evolution all living organisms have evolved efficient mechanisms to repair various DNA damages. There are three major DNA repair pathways: 1) Direct reversion of DNA damage by enzymatic dealkylation and photoreactivation; 2) Excision repair, where DNA lesion is restored through enzymatic excision of the damage region and re-synthesis prior to DNA replication and cell division. Excision repair include **b**ase **e**xcision repair (BER), **n**ucleotide **e**xcision repair (NER), and **m**ismatch repair MMR); 3) Recombinational repair, where damaged DNA may be exchanged, reciprocal or nonreciprocal, with homologous region of DNA (**h**omologous **r**ecombination (HR) and **n**on-**h**omologous **e**nd **j**oining (NHEJ)) (Grogan, 2004). Some of the principal agents that can impair DNA integrity as well as examples of DNA damages and possible mechanisms of their restoration are summarised in Table 1.

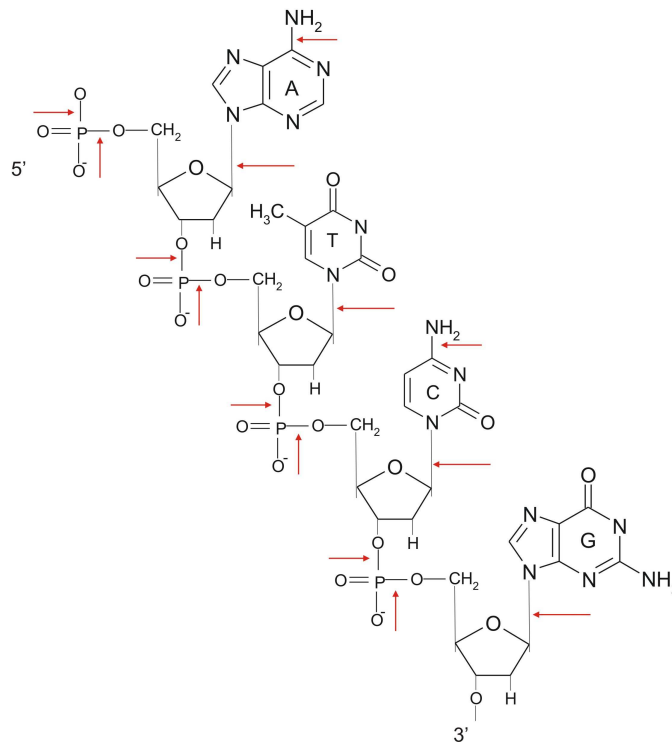
Table 1: Mutagens, types of DNA damage and possible repair mechanisms (D. Hartl and W.Jones, 2002)

Mutagen	Example of DNA alterations	Repair mechanisms
Water	Depurination or deamination of nucleobases; single-strand breaks (SSB)	BER, MMR, NER
Reactive intracellular metabolites	Oxidation, alkylation of nucleobases; single-strand breaks;	BER, MMR, NER
Chemical agents	Deamination, alkylation or oxidation of nucleobases; intramolecular crosslink	BER, MMR, NER, Recombinational repair
UV-light	Intermolecular crosslink (T-T dimers); strand breaks	Photoreactivation, BER, MMR, NER
Ionizing radiation	Single- and double-strand breaks (DSB); alteration of the nucleobases	HR, BER, MMR, NER

1.1.1 Hydrolytic DNA damage

Apart from harmful chemicals, ionizing radiation and UV-light known to promote deleterious effect on DNA structure, water - the universal milieu for all intracellular chemistry - is considered as a major mutagen causing spontaneous hydrolysis of DNA.

It has been found that in vitro nucleic acids undergo rapid spontaneous decomposition in solutions. Hence in metabolically active cell, whereby DNA exists in fully hydrated form, it would be particularly prone to spontaneous hydrolysis (Friedberg, 2003; Lindahl, 1993). The target sites within desoxyribonucleotide sequence susceptible to intrinsic hydrolytic attack are shown in Figure 1.

**Figure 1: Target sites of spontaneous DNA hydrolysis (T. Lindahl, 1993)**

A fragment of primary DNA structure with four normal nucleobases: A-adenine, T-thymine, C-cytosine and G-guanine. Red arrows indicate sites undergoing spontaneous hydrolysis.

Diesters, like sugar-phosphate linkage in DNA backbone, are normally quite labile and subject to the hydrolytic attack. Direct cleavage of the phosphate backbone has been estimated to be the most frequent type of hydrolytic damage resulting in a single-strand nick (Table 2).

The sugar-base glycosidic bond is also relatively labile and is subject to spontaneous hydrolysis. Hydrolytic cleavage of glycosidic bond, termed depurination and depyrimidination, results in formation of baseless or abasic sites (AP-site). The replication system usually inserts an adenine nucleotide in the daughter strand opposite the AP-site. After another round of replication, the DNA duplex would undergo a transition or transversion mutation. AP-sites can undergo further hydrolytic cleavage, resulting in single-strand breaks. The later occur at rates similar to or slightly slower than depurination. DNA strand breaks are potentially mutagenic and can have deleterious consequences such as cell cycle arrest and cell death (Bernstein, 1991; Lindahl, 1977; Nickoloff, 1998).

Aside from the hydrolytic cleavage of phosphodiester and glycosidic bond, DNA bases (cytosine, adenine, and guanine) containing exocyclic amino groups can undergo hydrolytic deamination. The hydrolytic cleavage of their exocyclic amino groups results in conversion of affected bases into uracil, hypoxanthine and xanthine, respectively (Lindahl, 1977). The rates of spontaneous DNA damage estimated in mammalian cells at neutral pH and 37°C are represented in Table 2.

Table 2: The rates of spontaneous hydrolytic DNA damage in mammalian cells at 37°C (Bernstein C, 1991)

Damage	Events per cell per day
Single-strand breaks	55.200
Depurination	12.000
Depyrimidination	600
Cytosine deamination	192

The data presented in Table 2 evidence significant instability of DNA molecule even under physiological solvent conditions.

Changes in DNA structure can have deleterious effect on the integrity of cellular genetic material. Insertion of inappropriate or modified bases in particular may lead to the replication block and cell death or to genetic mutations, unless they are actively restored. Hence, cellular DNA-repair system is the principle factor defining the normal function of the DNA and genetic stability of the entire organism.

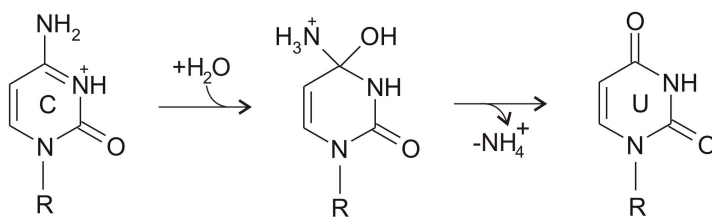
Base excision repair (BER) is the main pathway to repair spontaneous DNA damages including modified bases, AP-sites, strand breaks, and short gaps in DNA. BER is initiated by a specific DNA glycosylase that recognises and removes the damaged base leaving an

abasic site (AP-site) behind. Then an AP-site specific endonuclease incises DNA strand at the AP-site resulting in one-nucleotide gap. The original nucleotide sequence is then restored by consequent function of DNA-polymerase and DNA-ligase.

1.1.2 Hydrolytic cytosine deamination

Alongside with strand breaks and base loss, hydrolytic cytosine deamination to uracil is one of particularly common and genotoxic process in that it gives rise to pre-mutagenic U:G mismatch, which if left unrepaired after second round of replication produces GC→AT transition mutation in 50% of the progeny (Figure 2).

A



B

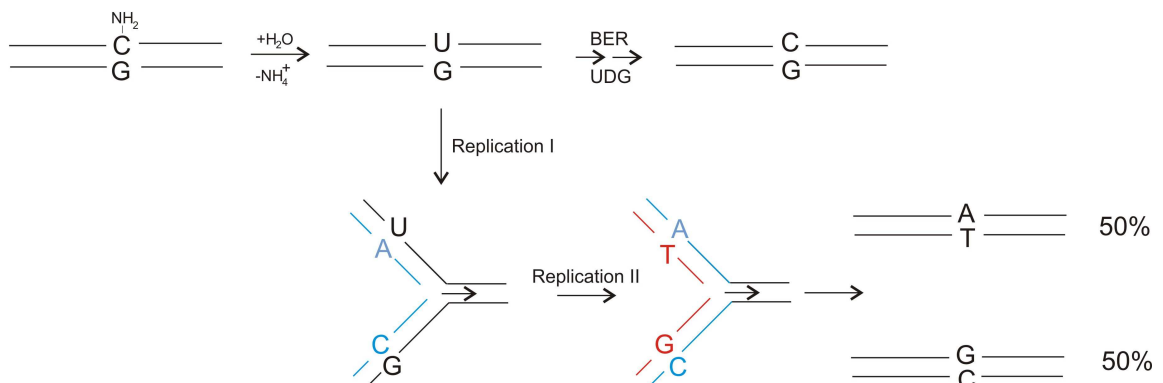


Figure 2: Schematic representation of hydrolytic cytosine deamination and outcomes of U:G mismatch

A: hydrolytic deamination of cytosine (C) to uracil (U) by nucleophilic attack of a primary amine on C4 of the pyrimidine ring. **B:** hydrolytic deamination of cytosine in dsDNA results in U:G mispair incorporation. U:G mismatch can be effectively repaired by base excision repair (BER) with the restoration of native C:G base pair, alternatively a CG→AT transition mutation appears in 50% of the progeny, as a result of the replication past U:G lesion. UDG: uracil-DNA glycosylase—an uracil specific enzyme, that recognises and excises the mispaired uracil from DNA initiating BER.

To counteract this mutagenic effect, most organisms eliminate uracil residues from DNA by means of the uracil-excision DNA repair pathway, which is initiated by a specific uracil-DNA glycosylase (UDG). UDGs exquisitely recognise and excise uracil from DNA as a free base by hydrolysing C1'-N-glycosidic bond leaving an AP-site behind. Uracil-DNA glycosylases are highly conserved and widespread DNA repair enzymes found in almost all living organisms (Krokan et al., 1997).

1.1.3 Phylogenetic distribution of uracil-DNA glycosylases

Numerous representatives of uracil specific glycosylases were classified into five families and later combined into UDG-superfamily based on substrate specificity and/or structural similarities, (Aravind, 2000). The phylogenetic overview of UDG-superfamily is represented in Table 3.

Table 3: The phylogenetic overview of UDG superfamily

Family number and Protein example Organism	F*1 UNG	F2 Mug/TDG	F3 sMUG	F4 tUDGa	F5 tUDGb
Eubacteria					
<i>E.coli</i>	+	+	-	-	-
<i>T.thermophilus</i>	+	+	-	+	+
Eukaryotes					
Mammals	+	+	+	-	-
<i>S.cerevisiae</i>	+	-	-	-	-
Archaea					
<i>A. fulgidus</i>	-	-	-	+	-
<i>M.thermautotrophicus</i>	-	+	-	-	-

Symbols and abbreviations are: *F - a family number, + - encoded in genome, - not detected, **Ung** - Uracil N-Glycosylase, **Mug/TDG** - Mismatch specific Uracil-DNA-Glycosylase/Mismatch-specific Thymine-DNA-Glycosylase, **sMUG** - Single-strand-specific Monofunctional Uracil-DNA-Glycosylase, **tUDGa** and **tUDGb** - Thermostable Uracil-DNA Glycosylases.

Uracil-N-Glycosylase (UNG) from *E.coli* is the first found (Lindahl, 1974) and most extensively studied enzyme within UDG super-family (Varshney and van de Sande, 1991). *E.coli* UNG is a prototype for Family 1 uracil DNA glycosylases (UNGs). Structural basis of UNG specific substrate recognition and catalysis is well studied (Renos Savva, 1995). *E.coli* UNG is able to remove uracil base from both single and double-stranded DNA irrespectively of nucleotide opposing the target uracil (Lindahl, 1977). Uracil base is flipped out from the base stack into the enzyme catalytic site, followed by hydrolysis of the N-glycosydic bond between base and 2'-desoxyribose (Kunkel and Wilson, 1996). Representatives of UNG Family can be effectively inhibited by Uracil-DNA glycosylase inhibitor (Ugi) - a competitive inhibitor encoded by *Bacillus subtilis* bacteriophages PBS1 and PBS2 (Mol et al., 1995a; Wang, 1989).

The second UDG family is comprised of TDG/MUG homologs. Mismatch-specific Thymine DNA Glycosylase (TDG) purified from HeLa cells is known to cleave both T:G and U:G mispairs (Neddermann and Jiricny, 1993). Mismatch Uracil-DNA Glycosylase (MUG) from *E.coli* was found as a functional homologue of human TDG (hTDG). *E.coli* MUG is primarily G:U mismatch specific, although it does display activity on T:G mismatch at high

enzyme concentrations (Barrett et al., 1998; Gallinari and Jiricny, 1996). Unlike UNG, both TDG and MUG enzymes are found to repair mismatched bases exquisitely on the double-stranded DNA substrates (Neddermann and Jiricny, 1994).

First representative of the Family 3 UDGs was identified and characterised in *Xenopus* and denoted as a Single-strand-specific Monofunctional Uracil-DNA Glycosylase (SMUG) (Haushalter et al., 1999). Later on, SMUG homologous were identified in humans, insects and some bacteria. SMUG family representatives were first characterised as single-stranded DNA specific enzymes prone to remove uracil and 5-hydroxymethyluracil. However, recent studies of human SMUG homologue (hSMUG1) revealed that this enzyme was able to excise uracil from both double- and single-stranded DNA and had broader substrate specificity, including 5-hydroxymethyluracil (5hmUra) and 3,N⁴-ethenocytosine (Kavli et al., 2002; Pettersen et al., 2007).

The thermostable uracil-DNA glycosylases were combined into Family 4 tUDGs. First tUDG was discovered in thermophilic bacteria *Thermotoga maritima* and denoted as TmUDG. This enzyme is heat stable up to 75°C and has a weak sequence similarity to *E.coli* MUG. TmUDG can remove uracil from both single- and double-stranded DNA containing U:A or U:G base opposition (Sandigursky and Franklin, 1999). Homologs of TmUDG were found in both thermophilic eubacteria and archaea (Sandigursky and Franklin, 2000; Sartori et al., 2001; V. Starcuvienė and H.-J. Fritz, 2002). Representatives of tUDG family were later divided into two distinct sub-families, namely Family 4 tUDGA and Family 5 tUDGB, according to their different substrate specificities. Both tUDGA and tUDGB were discovered in *Thermus thermophilus* - a thermophilic eubacterium (V. Starcuvienė and H.-J. Fritz, 2002). tUDGA displays single- and double-stranded DNA specificity. It is highly selective for uracil residue regardless of the opposing base, but unable to remove thymine from G:T mismatch. tUDGB, unlike tUDGA, is able to remove uracil exquisitely from double-stranded DNA substrate (Hoseki et al., 2003; V. Starcuvienė and H.-J. Fritz, 2002). Some representatives of tUDGA family were found to possess the iron-sulphur (4Fe-4S) cluster. (Hinks et al., 2002; Hoseki et al., 2003). Representatives of fifth UDG family contain non polar residue in the active-site motif I (GLAPAG-X₁₀-F), which is proposed to activate a water molecule in other UDGs (Sartori et al., 2002; V. Starcuvienė and H.J. Fritz, 2002).

To summarize, enzymes of Family 1 UNGs are widely distributed in eubacteria, eukaryotes and even DNA viruses with the notable exception of *Drosophila melanogaster*. The Family 2/TDG homologous genes were found in all life domains: Archaea, Bacteria and Eukaryotes. SMUG family genes were found so far only in Eukaryotes and some Bacteria. Neither SMUG nor UNG family genes were found in Archaea.

Until recently UNG was considered to be the major enzyme responsible for general uracil repair, mostly based on its strong conservation from bacteria to humans, exquisite

specificity for the excision of 2'dU residue regardless of DNA sequence context and its high expression level in growing cells. Furthermore, the presence of at least one representative of UDG family in almost all organisms evidences an important role of uracil repair in the restoration of genome integrity (Pettersen et al., 2007). The most striking here is perhaps the lack of this otherwise highly conserved UNG family genes in Archaea - the third domain of life. The latter fact suggests that Archaea may have an alternative DNA-uracil repair pathway.

1.2 Uracil-DNA glycosylases in Archaea

Most Archaea studied so far are extremophiles that exist in extreme conditions (high temperatures, acidity or in the saturated salt solution) that affect significantly their genome integrity. Elevated temperatures in particular directly destabilise the primary and secondary structure of DNA and accelerate spontaneous DNA decomposition reactions (such as deamination and depurination) (Lindahl, 1993). For these reasons Archaea and particularly hyperthermophilic Archaea are expected to have more effective and optimal systems for DNA repair and in particular for hydrolytic DNA damage repair.

As was mentioned earlier, no genes for otherwise ubiquitous Family 1 uracil-DNA glycosylases have been identified in Archaea. However, Family4 tUDG glycosylases were isolated from some thermostable Archaea. A representative of thermostable uracil-DNA glycosylases was found in extreme thermophilic archaeon - *Archaeoglobus fulgidus* (optimal growth temperature 80°C). This enzyme, denoted as AfUDG, revealed a high degree of primary amino acids sequence similarity to Family 4 TmUDG from *T. maritima*. A homologous ORF with significant sequence similarity to TmUDG and AfUDG was identified in a hyperthermophilic archaeon - *Pyrobaculum aerophilum* (optimal growth temperature 100°C). All these enzymes can remove uracil both from single- and double-stranded DNA containing either U:G or U:A oppositions.

Interestingly, some Archaea with completely sequenced genome seem to be devoid of UDG family genes. Among them is a thermophilic archaeon - *Methanothermobacter thermautotrophicus* (optimal growth temperature 65°C). However, two mismatch-specific DNA glycosylases, namely Mig.MthI and Mig.MthII, members of Mig/MBD4 family DNA glycosylase, were isolated and characterised in *M. thermautotrophicus* (Horst, 1996; Starcuvienė, 2001). Mig/MBD4 homologs comprise the second, after UDG, superfamily of DNA glycosylases, namely the helix-hairpin-helix-GPD (HhH-GPD) structure superfamily. The name of this family derived from specific motif: HhH and Gly/Pro rich loop (GP) followed by a conserved aspartate (D) (Krokan et al., 2002). MBD4 (methyl CpG binding domain protein 4) was isolated and characterised in human and mouse. Mig/MBD4 substrate specificity is similar to the Family 2 TDGs recognising G:T and G:U mismatches within

definite context. MBD4 is a sequence specific enzyme that binds preferentially to CpG context, in particular to the 5-meCpG-TpG mismatches, presumably counteracting the mutagenic effect of deamination of 5me-C to thymine (Hendrich et al., 1999). The relating members of HhH-GPD superfamily including three from *E.coli*: 8-oxoG:A mispair-specific glycosidase MutY (Fromme et al., 2004), thymine glycolglycosylase EndoIII (Thayer et al., 1995), 3-methyladenine DNA glycosylase II (AlkA) (Hollis et al., 2000), and human 8-oxoguanin DNA glycosylase (OGG1) (Bruner et al., 2000) remove diverse types of damaged DNA bases, including uracil, oxidised and alkylated bases.

Unlike UNG family enzymes, members of Mig/MBD4 family are context-dependent enzymes with broader substrate specificity and with residual uracil-excision activity, if any at all. Hence, neither Mig.MthI nor Mig.MthII can be considered as a general uracil-DNA glycosylase initiating and coordinating efficient repair of a deaminated cytosine (or U:G mismatch) in *M. thermautotrophicus*. From all these, the nature of uracil repair in *Methanothermobacter thermautotrophicus* and possibly in other thermophilic and hyperthermophilic Archaea with similar distribution of DNA repair genes was, until recently, unclear.

1.2.1 *Methanothermobacter thermautotrophicus* as a model organism

M. thermautotrophicus str. ΔH belongs to the domain Archaea, *Euryarchaeota*, *Methanobacteria*. The archaeon *M. thermautotrophicus* (Mth) is a strict anaerobe that converts hydrogen and carbon dioxide to methane and grows optimally at 65°C (Zeikus and Wolfe, 1972). Mth genome was completely sequenced (Smith et al., 1997) and represents a single circular DNA molecule of 1.7 million base pair (bp) length. The G+C content comprises 49.5% of the genome (in comparison to *E.coli* (51%) and *S. cerevisiae* (38%)). 42% of ORFs were found to be similar to bacterial sequences and 13% of ORFs were similar to eukaryal sequences. Genomic analysis has confirmed that archaeal DNA-processing enzymes are more similar to those found in eukaryotes than in bacteria. In fact archaea share many features with eukaryotes and bacteria and therefore serve as alternative model systems (adapted to harsh environmental conditions) for the studying molecular diversity of DNA stability and repair as well as for the better understanding of these processes both in bacteria and eukaryotes (Grogan, 2004; Kelman and White, 2005; Majernik et al., 2004).

1.2.2 Discovery of Mth212 as a DNA uridin endonuclease in *M. thermautotrophicus*

A new type of uracil specific DNA repair enzyme, namely Mth212, an ExoIII homologue from *Methanothermobacter thermautotrophicus* (Mth), was discovered and biochemically characterised in our laboratory (Georg et al., 2006). Mth212, in addition to the numerous enzymatic activities inherent to ExoIII homologues, recognises DNA-U and cleaves the

phosphodiester backbone direct to the 5' side of the 2'dU residue. More recently, essential role of Mth212 in DNA-U repair initiation was established during in vitro reconstruction of U/G mismatch repair pathway in *Mth* cell extract (Schomacher et al., 2009). Amino acid sequence analysis revealed that Mth212 shares 30% sequence identity with *E.coli* exonuclease III (ExoIII) and 40% identity with human AP-endonuclease Ape1 (Figure 3).

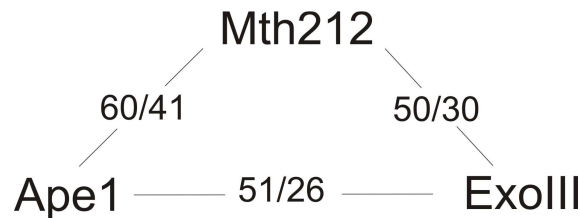


Figure 3: Amino acid sequence similarity and identity within homologous proteins

Mth212 from *M. thermautotrophicus*, ExoIII from *E.coli* and Ape1 from *H.sapiens*. Numbers on the left indicate sequence similarity and on the right sequence identity in percentage.

A number of conserved catalytically important amino acids known from human Ape1 and ExoIII (Barzilay et al., 1995; Rothwell et al., 2000) are conserved in Mth212 as well (Figure 4).

MTH212	1	-----MT
APE1_HUMAN	1	MPKRGKKGVAEDGDELRTPEAKKSKTAAKNDKEAAGEGPALYE DPPDQKTSPSGKPA
exoIII_E.coli	1	-----
MTH212	3	VLKIIISWNVNGLRAVHRKGLKWFMEF-KPDILCLQETIKAAPEQLPRKLRHVEG--YRSE
APE1_HUMAN	61	TLKICSWNVNGLRAWIKKGLDWVKEE-APDILCLQETKCSFNKLPALQELPCLSHQYN
exoIII_E.coli	1	-MKFVSFNINGLRA--RPHCLEAIVEKHQPDVIGLQETKVHDEMFPLEEVAKLGYNVFF-
MTH212	60	FTFAERKGYSGVAMYTKVPFSSSLREGFCVERFDTEGRIQIADFDFL----LYNIYFPNG
APE1_HUMAN	120	SAESDKEGYSGVGLLSRQCPLKVSYGICDEEHDQEGRVIVAEFDSFV----LVITAYVPNA
exoIII_E.coli	57	--HGQKGHYGVALLTKETPIAVRRGEPGDEEFAQRRIIMAEIPSPILGNVTVINGYFPOG
MTH212	116	KMSERLKY--KLEFYDAFLEDVNRERDSGRNVIICGDFNTAHREIDLARPKENSN----
APE1_HUMAN	176	GRGLVRLEY--RQRWDEAFRKEL-KGLASRKPLVICGDLNVAHEIDIRNPKGNKK----
exoIII_E.coli	114	ESRDHPITKFPKAAQFYQNLQNYLETETKRENPEVLTMGDMNISPGDLDIGIGENRKRWLR
MTH212	170	--VSGFLPVERAWIDKFTEEN-CYVDTFR-MFNSDPGQYTWSYRTRARERNVWRLDYFF
APE1_HUMAN	229	--NAGFTIPQERQGFGEILQAVFLADSFRLYPNTFYAYTFWTYMMNARSKNVWRLDYFL
exoIII_E.coli	174	TGKCSFLPEEREWMERLY-SWGLVDTFRHANPQTADRFSWFDYRSKGFDDNRGTRIDLIL
MTH212	226	VNEEFKGVKRS----WILSDVMGSDHCPIGLETFL-----
APE1_HUMAN	287	LSHSILPALQDS----KIRSKALGSDHCPITLYLAL-----
exoIII_E.coli	233	ASQPIAECCVETGIDYEIRSMKEKPSDHPVWATFRR-----

Figure 4: Multiple amino acid sequence alignment of Mth212 (*Methanothermobacter thermautotrophicus*) with ExoIII (*Escherichia coli*) and Ape1 (*Homo sapiens*)

Amino acid residues involved in the catalysis by Ape1 according to (Mol et al., 2000) and (Rothwell et al., 2000) are marked with an asterisk. Alignment was performed using WISCONSIN PACKAGE version 10.3 (Womble, 2000) and arranged with BOXSHADE version 3.21.

The three-dimensional structure (3D) of Mth212 apo-enzyme was determined (K. Lakomek, 2009). As the result of the 3D structure analysis, Mth212 was assigned to the DNase I structure family, who's most extensively characterized representatives are bovine DNase I (Weston et al., 1992), *E.coli* ExoIII (Mol et al., 1995b) and human Ape1 (Mol et al., 2000) (Figure 5).

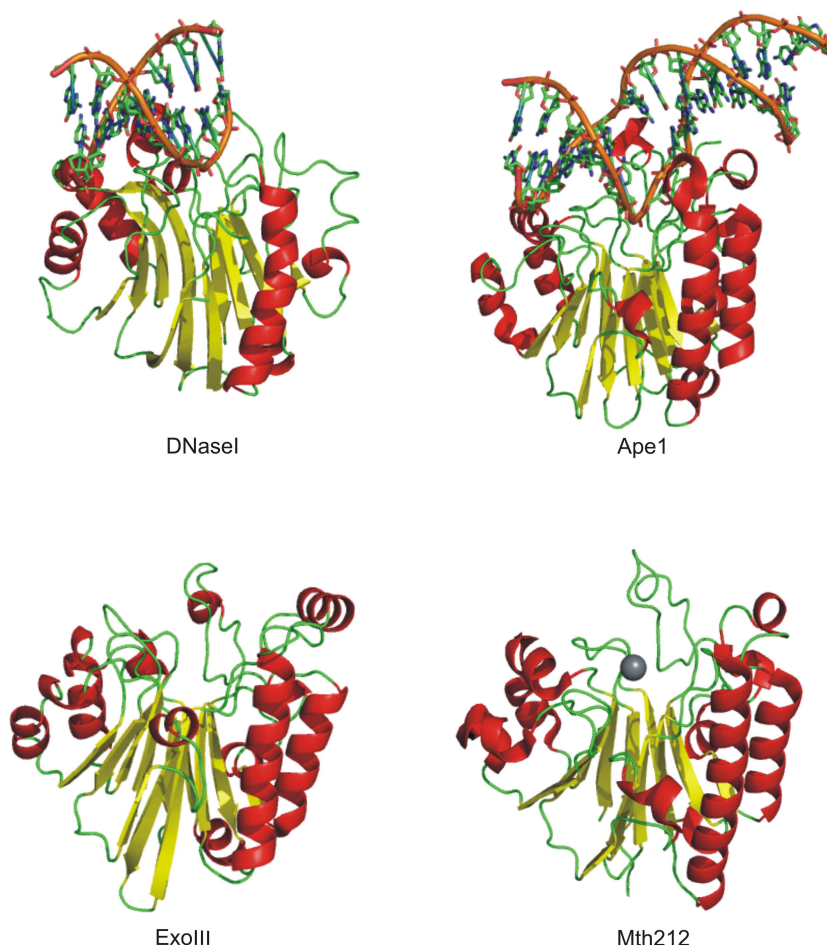


Figure 5: 3D structure comparison within DNase I structure family

The structural data for DNase I-DNA (1DNK; Weston et al., 1992) and Ape1-DNA complexes (1DEW; Mol et al., 2000) as well as apo-ExoIII (1AKO; Mol et al., 1995b) and apo-Mth212 (3G91, K. Lakomek, 2009) were processed using PyMOL (DeLano Scientific LLC) software. α -helices are shown in red coils, β -sheets are shown in yellow arrows, coils are shown in green tubes. DNA helix (orange) with the nucleotide bases represented as atom coloured sticks. A metal ion (here Mg^{2+}) bound in Mth212 active site is shown as a grey sphere.

1.3 Aim of project

Despite high sequence and structure similarity found within ExoIII family members characterised so far, Mth212 is the only representative able to recognise and repair DNA uridin. The questions is: does this newly discovered unique activity of Mth212 reside in the same active site as AP-endo and 3'→5' exonuclease activities?

The main purpose of this project was to determine whether Mth212 has unique structural features responsible for DNA-U specificity in addition to the other catalytic activities inherent to ExoIII homologues.

2 Materials and Methods

2.1 Materials

2.1.1 Bacterial strains (*Escherichia coli*)

2.1.1.1 **DH5 α** (Invitrogen, Carlsbad, CA)

F⁻, Φ 80*dlacZ* M15, *endA*1, *recA*1, *hsdR*1 (*rK-mK*+), *supE*44, *thi*-1, *gyrA*96 (NalR), *relA*1, (*lacZYA-argF*) U169

2.1.1.2 **BL21-CodonPlus (DE3)-RIL** (Stratagene, La Jolla, CA)

E. coli B, *F*⁻, *ompT*, *hsdS* (*rB-mB*-), *dcm*+, TetR, *gal*1(DE3), *endA*, Hte [*argU*, *ileY*, *leuW*, CmR]

2.1.1.3 **BL21_UX** (Georg *et al.*, 2006, this paper) (based on BL21-CodonPlus(DE3)-RIL (Stratagene, La Jolla, CA)

E. coli B, *F*⁻, *ompT*, *hsdS* (*rB-mB*-), *dcm*+, TetR, *gal*1(DE3), *endA*, Hte [*argU*, *ileY*, *leuW*, CmR], Δ *ung::kan*

2.1.1.4 **BL21_UXX**

E. coli B, *F*⁻, *ompT*, *hsdS* (*rB-mB*-), *dcm*+, TetR, *gal*1(DE3), *endA*, Hte [*argU*, *ileY*, *leuW*, CmR], Δ *ung*

2.1.1.5 **XL1-Bleu**

XL1-Bleu MRF' (Stratagene USA, (Bullock *et al.*, 1987))

*RecA*1 *endA*1 *gyrA*96 *thi*-1 *hsd*17 *supE*44 *relA*1 *lac*; *F'**lacI*^f *lacZ* Δ M15 *proA*⁺*B*⁺ *Tn*10 (*tet*^r) *Amy cam*^r

2.1.2 Plasmids

2.1.2.1 **pET_B_001** (Georg *et al.*, 2006)

pET_B_001 is a derivative of vector pET-21d (Novagen, San Diego, CA), in which the multiple cloning site between NcoI and XhoI is replaced by the following 14nt section: dTCTGCGGCCGCACA. The nucleotide sequence and restriction map of the vector can be seen in Appendix 7.1.1.

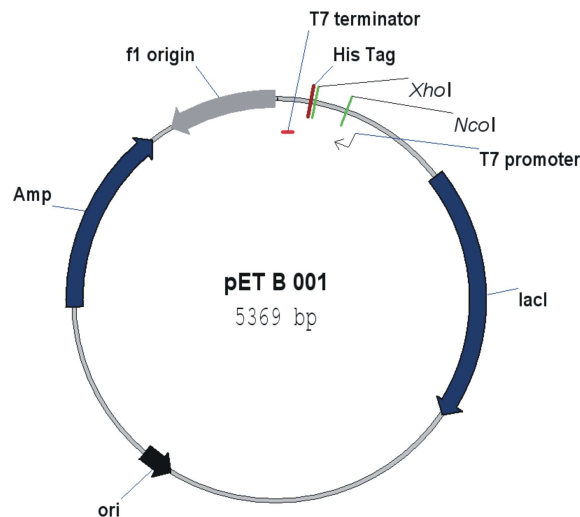


Figure 6: Schematic representation of pET_B_001 vector

Genes for ampicillin resistance mediation (Amp, β -lactamase), and lac repressor (lacI) are indicated in blue; Origin of replication of phage f1 is indicated in grey. Arrows indicate the direction of transcription and replication. Positions of selected restriction sites are marked with green bars and labelled by unique occurrence in the vector with grey letters. His-Tag (ruby stick): nucleotide sequence that encodes the 6-His tail. Positions of T7 promoter and T7 terminator are shown in blue and red correspondingly. Black arrow: indicates the origin of replication of the plasmid. The scheme was created using Vector NTI Software (Invitrogen) (see section 2.1.13).

2.1.3 Primers for polymerase chain reaction

2.1.3.1 Primers for site-directed mutagenesis:

Primers for Mth212/D151N mutant generation:

212_D151N_FOR (DNA, 27mer)	5'-GATTATATGTGGG AAC TTCAACACAGC-3'
212_D151N_REV (DNA, 27mer)	5'-GCTGTGTTGAA GTT CCCACATATAATC-3'

Primers for Mth212/D151A mutant generation:

212_D151A_FOR (DNA, 27mer)	5'-GATTATATGTGGG GCA TTCAACACAGC-3'
212_D151A_REV (DNA, 27mer)	5'-GCTGTGTTGAA TGC CCCACATATAATC-3'

Primers for Mth212/D151S mutant generation:

212_D151S_FOR (DNA, 27mer)	5'-GATTATATGTGGG TCA TTCAACACAGC-3'
212_D151S_REV (DNA, 27mer)	5'-GCTGTGTTGAA TGA CCCACATATAATC-3'

2.1.3.2 Primers for nucleotide sequence analysis

pET-vector sequencing primers (pET_B_001 und pET-28a):

PETS1_UP (18mer)	5' CAGCAGCCAACTCAGCTT 3'
PETS1_LO (18mer)	5' ATAGGGGAATTGTGAGCG 3'

pCR-Blunt II-Topo-vector sequencing primers:

M13 Forward (16mer) 5' GTAAAACGACGGCCAG 3'
 M13 Reverse (17mer) 5' CAGGAAACAGCTATGAC 3'

2.1.4 2'-Desoxyriboseoligonucleotides for enzymatic activity assays

The oligonucleotides were synthesised by either PURIMEX (Greibenstein) or Sigma (Munich).

Oligonucleotides for Endonuclease assay:

Prince-U (40mer) 5' (F)GGGTACTTGGCTTACCTGCCCTG**U**GCAGCTGTGGGCGCAG 3'
 40_PRI_AP(40mer) 5'(F)GGGTACTTGGCTTACCTGCCCTG(**AP**)GCAGCTGTGGGCGCAG3'
 PRINCE_G (35mer) 5' CTGCGCCACAGCTG**C**CAGGGCAGGTAAGCCAAG 3'
 23-M (23mer) 5'(F)GGGTACTTGGCTTACCTGCCCTG 3'

Oligonucleotides for EMSA:

40_PRI_AP_Dunkel 5'dGGGTACTTGGCTTACCTGCCCTG(**AP**)GCAGCTGTGGGCGCAG 3'
 PRINCE_C 5'dGGGTACTTGGCTTACCTGCCCTG**C**GCAGCTGTGGGCGCAG 3'
 PRINCE_U_Dunkel 5'dGGGTACTTGGCTTACCTGCCCTG**U**GCAGCTGTGGGCGCAG 3'
 Prince_A_Blunt 5' dCTGCGCCACAGCTG**C**ACAGGGCAGGTAAGCCAAGTACCC 3'
 Prince_T_Blunt 5' dCTGCGCCACAGCTG**C**TACAGGGCAGGTAAGCCAAGTACCC 3'
 Prince_C_Blunt 5' dCTGCGCCACAGCTG**C**CCAGGGCAGGTAAGCCAAGTACCC 3'
 Prince_G_Blunt 5' dCTGCGCCACAGCTG**C**GCAGGGCAGGTAAGCCAAGTACCC 3'
 40_Prince-G 5' dCCCACAGCTG**C**GCAGGGCAGGTAAGCCAAGTACCCTACGT 3'
 20_UP_Prince 5' dCTGCCCTGCGCAGCTGTGGG 3'
 20_LO_Prince 5' dCCCACAGCTGCGCAGGGCAG 3'
 40_Prince_C 5' dCCCACAGCTG**C**CCAGGGCAGGTAAGCCAAGTACCCTAGCT3'
 40_Prince_A 5' dCCCACAGCTG**C**ACAGGGCAGGTAAGCCAAGTACCCTAGCT3'
 40_Prince_T 5' dCCCACAGCTG**C**ICAGGGCAGGTAAGCCAAGTACCCTAGCT3'

Nucleotides shown in bold build a mismatch pair in substrate oligonucleotides used in enzymatic tests. F: fluorescein (6-isomer), P: phosphate, U: 2'-d-Uridin residue, [AP]: model of a stable AP site (see section 3.1.3, Figure 14, A and B)

2.1.5 Molecular Ladders and Markers

2.1.5.1 DNA size marker

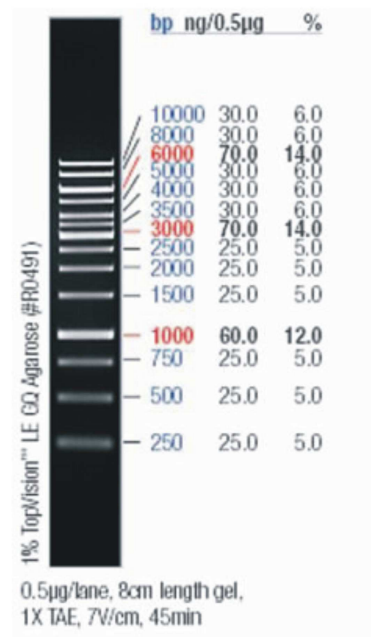


Figure 7: GeneRuler™ 1kb (Fermentas, Burlington, Ontario)

The DNA molecular length standards were adjusted with TE buffer (2.1.9) and 6x loading dye solution (Fermentas) to the DNA concentration of 0.1 µg/ µl and stored at 4°C. Lengths are indicated in base pairs (bp).

2.1.5.2 Protein size marker

The unstained protein molecular weight marker (Fermentas, Burlington, Ontario) was used in all SDS-PAGE analysis.

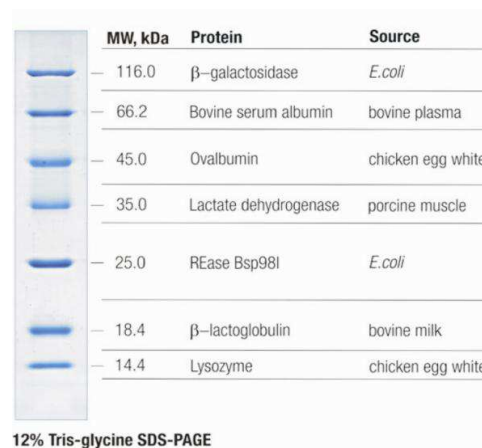


Figure 8: Unstained protein molecular weight marker (Fermentas, Burlington, Ontario)

In addition to the relative molecular weights $\times 10^{-3}$ of the various proteins, their sources are given as well.

2.1.6 Enzymes und Proteins

BioGenes (Berlin)	Polyclonal <i>Anti</i> -Mth212-Anticorper
Boehringer (Mannheim)	Ribonuclease A (RNaseA)
Fermentas (Burlington, Ontario)	Restriction endonucleases
	<i>Pfu</i> -DNA-Polymerase (recombinant)
	T4-DNA-Ligase
	Calf Intestine Alkaline Phosphatase
New England Biolabs Inc. (Ipswich, MA)	Restriction endonucleases
	<i>Taq</i> -DNA-Polymerase
Sigma (St. Louis, MO)	Benzonase (298 U/μl)

2.1.7 Chemicals and reagents

AGS GmbH, Heidelberg:

Qualex Gold Agarose

AppliChem, Darmstadt:

Acrylamid 4k – 30% solution; Ampicillin (Na-salt); Bisacrylamid
4K–2% solution; Chloramphenicol; Coomassie Brilliant Blue G250 und R250; Dithiothreitol (DTT); Methanol; 2-(*N*-morpholino) ethanesulfonic acid (MES); Nickel (II)-Chloride; Phenylmethansulfonylfluorid (PMSF);

Fermentas:

Isopropyl-β-D-Thio-Galactopyranosid (IPTG)

Fluka, Neu-Ulm:

Bromphenol blue; Dimethylpimelinidat (DMP); Glutathion (reduced); Glycerol, 87%; Formamide; Polyethylene glykol (PEG) 6000; Silver nitrate

GE Healthcare, Uppsala, Sweden:

Chelating Sepharose™ Fast Flow

Invitex, Berlin:

dNTP (2'-Desoxyribosenucleosidtriphosphate)

Invitrogen GmbH, Carlsbad, CA:

SYBR® Gold Nucleic Acid Gen Stain;
Dynabeads® Protein G- Immunoprecipitation kit

Merck, Darmstadt:

2-Mercaptoethanol; Calcium chloride dehydrate; Formaldehyde; Magnesium chloride hexahydrat

MJ Research, Waltham, MA:

Chill-out™ liquid wax

Oxoid, Cambridge, UK:

Yeast extract; Tryptone; Bacteriological agar

Roth GmbH, Kalsruhe:

Ammonium acetate, Ammonium chloride, Ammonium sulphate, Ampicillin sodium salt, Chloramphenicol, Citric acid monohydrate, Disodium phosphate, Ethanol, Ethidium bromide, Acetic acid, Urea, (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 2-Propanol, Kanamycin, Potassium chloride, Potassium hydroxide, Methanol, Sodium acetate, Sodium chloride, Monosodium phosphate, Sodium hydroxide, Sodium thiosulfate, Phenol, Phosphoric acid, Rotiphorese® Gel30 (Acrylamide:Bisacrylamide 37.5:1), Sucrose, Hydrochloric acid (smoking), Tris (hydroxymethyl) aminomethane

Scharlau Chemie, Barcelona, Spain:

Chloroform, Boric acid, Imidazole

Serva, Heidelberg:

Ammonium persulfate, Ethylenediaminetetraacetic acid (EDTA), Glycine, Sodium laurylsulphate (SDS), N,N,N',N'- Tetramethylethylenediamine(TEMED)

Sigma, Steinheim:

Tetrasodium pyrophosphate, Triethanolamine, Triton X-100 (Octylphenol-Polyethylenglycol), Polyethylenglycol-Sorbitan-Monolaurat (TWEEN® 20), Xylencyanol FF, L-Arabinose

2.1.8 Molecular biology Kits

Genomed, Bad Oyenhausen

Macherey-Nagel, Düren

Invitrogen, Carlsbad, CA

JETSTAR Plasmid Purification Kit

NucleoTrap® and NucleoSpin® Gel extraction Kit

Zero Blunt® TOPO® PCR Cloning Kit

2.1.9 Buffers and solutions

ALF-marker

95% Formamide, 20mM EDTA pH 8.0, 5 mg/ml

Dextran-Blue, stored at 4°C

APS stock solution, 10%	Ammonium persulfate in H ₂ O, stored in aliquots at -20 °C
dNTP stock solution	10 mM of each dNTP in H ₂ O, stored at -20°C
EDTA stock solution	500 mM EDTA dissolved in H ₂ O with addition of solid NaOH
DTT stock solution	1M DTT in H ₂ O, sterilised (0.45 µm pore size filter) and stored at -20°C
Ethidium bromide stock solution	10 mg/ml Ethidium bromide
Isopropyl-β-D-thiogalactopyranosid (IPTG) stock solution	1M IPTG in H ₂ O, sterilised (0.45 µm pore size filter) and stored at -20°C
PBS-Buffer (1x)	10 mM Na ₂ HPO ₄ , 1.8 mM KH ₂ PO ₄ pH 7.3, 140 mM NaCl, 2.7 mM KCl
Phenol/Chloroform	25 vol. of phenol/TE, 24 vol. of Chloroform, 1 vol. of Isoamyl alcohol
Phenol/TE	Phenol, saturated with TE-Buffer, 0.1% (w/v) 8-Hydroxyquinolin
SDS-PAGE Loading buffer	98% Formamide, 10 mM EDTA pH 8.0, 0.025% Xylencyanol FF, 0.025% Bromphenol blue
RNase A stock solution	10 mg RNaseA in 1 ml of 10 mM Tris/HCl pH 7.5, 15 mM NaCl. The mix was preheated for 15 min at 100°C and cooled down at RT
SSC-Buffer	50 mM NaCl, 15 mM Sodium citrate
Sucrose marker (DNA loading buffer)	60% (w/v) Sucrose, 0.05% (w/v) Bromphenol blue, 0.05% (w/v) Xylencyanol FF in TAE-Buffer
50x TAE buffer	2M Tris/Acetate, 50 mM EDTA
10x TBE buffer	0.89M Tris/Borate pH 7.9, 25 mM EDTA
T4-DNA-Ligase buffer	400 mM Tris/HCl pH 7.8, 100 mM MgCl ₂ , 100 mM DTT, 5 mM ATP
TE-buffer	10 mM Tris/HCl pH 8.0, 0.5 mM EDTA
1XTB-buffer (EMSA running buffer)	0.89M Tris/Borate pH7.9
10x Taq-DNA-polymerase buffer	100 mM Tris/HCl pH 9.0, 500 mM KCl, 15 mM MgCl ₂ , 1% (v/v) Triton X-100

Antibiotika-Stammlösungen:

Ampicillin stock solution	100 mg/ml Ampicilin (Na-Salt)
Chloramphenicol stock solution	25 mg/ml in Ethanol, sterilised (0.45 µm pore size filter) and stored at -20°C
Kanamycin stock solution	50 mg/ml in H ₂ O, sterilised (0.45µm pore size filter) and stored at -20°C
Tetracycline stock solution	25 mg/ml in H ₂ O, sterilised (0.45 µm pore size filter) and stored at -20°C

Activity assay buffers:

Endonuclease buffer (1x)	20 mM H ₂ KPO ₄ /HK ₂ PO ₄ buffer pH 6.2, 50 mM KCl, 1 mM MgCl ₂ , 0.1 mg/ml BSA
--------------------------	--

Bradford assay buffer:

Bradford dye solution	100 mg Coomassie brilliant blue G-250, 50 ml 95% Ethanol, 100 ml 85% Phosphoric acid, add H ₂ O up to 1 L
BSA stock solution	0, 25, 50, 75, 100, and 150 µg of BSA in 100µl 1XPBS buffer

Buffers for chromatography:

IMAC wash buffer	25 mM HEPES /KOH pH 7.6, 0.5M NaCl
IMAC-elution buffer	30 mM, 60 mM, 70 mM, 80 mM, 90 mM, 100 mM, 300 mM, and 500 mM Imidazole in IMAC wash buffer
Heparin column wash buffer	20 mM HEPES/KOH pH 7.6, 5 mM 2-Mercaptoethanol (sterile filtrated)
Wash buffer for making salt gradient	3M NaCl, sterile filtrated

JETSTAR buffers (Genomed, Bad Oeynhausen):

E1	50 mM Tris/HCl pH 8.0, 10 mM EDTA
E2	200 mM NaOH, 1% (w/v) SDS
E3	3.1M Potassium acetate pH 5.5
E4	100 mM Sodium acetate pH 5.0, 600 mM NaCl, 0.15% Triton X-100
E5	100 mM Sodium acetate pH 5.0, 800 mM NaCl,
E6	100 mM Tris/HCl pH 8.5, 1.25 M NaCl

Buffers for restriction endonucleases (1x, Fermentas, Burlington, Ontario):

B+ buffer	10 mM Tris/HCl pH 7.5, 10 mM MgCl ₂ , 0.1 mg/ml BSA
G+ buffer	10 mM Tris/HCl pH 7.5, 10 mM MgCl ₂ , 50 mM NaCl, 0.1 mg/ml BSA
O+ buffer	50 mM Tris-HCl pH 7.5, 10 mM MgCl ₂ , 100 mM NaCl, 1 mg/ml BSA
R+ buffer	10 mM Tris/HCl pH 8.5, 10 mM MgCl ₂ , 100 mM KCl, 0.1 mg/ml BSA
Buffer Y+/Tango™	33 mM Tris/Acetate pH 7.9, 10 mM magnesium acetate, 66 mM potassium acetate, 0.1 mg/ml BSA
<i>Bam</i> HI-buffer	10 mM Tris/HCl pH 8.0, 5 mM MgCl ₂ , 100 mM KCl, 0.02% Triton X-100, 0.1 mg/ml BSA
<i>Eco</i> RI-buffer	50 mM Tris/HCl pH 7.5, 10 mM MgCl ₂ , 100 mM NaCl, 0.02% Triton X-100, 0.1 mg/ml BSA

Buffers for restriction endonucleases (1x, New England Biolabs, Ipswich, MA):

NEB 1	10 mM Bis-Tris Propane/HCl pH 7.0, 10 mM MgCl ₂ , 1mM DTT
NEB 2	10 mM Tris/HCl pH 7.9, 10 mM MgCl ₂ , 50 mM NaCl, 1 mM DTT
NEB 3	50 mM Tris/HCl pH 7.9, 10 mM MgCl ₂ , 100 mM NaCl, 1 mM DTT
NEB 4	20 mM Tris/Acetate pH 7.9, 10mM Magnesium acetate, 50 mM Potassium acetate, 1 mM DTT

SDS-PAGE buffers:

Coomassie R250 dye solution	0.7% (w/v) Coomassie R-250 in Methanol, 20% Acetic acid, mixed at 1:1 before use
Loading dye (or sample buffer)	200 mM Tris/HCl, 8M urea, 200 mM DTT, 2% (w/v) SDS, 0.05% Bromphenol blue
Laemmli buffer (1x)	25 mM Tris/HCl pH 8.4, 200 mM Glycin, 0,1% (w/v) SDS
SDS	10% (w/v) SDS in H ₂ O
Stacking gel buffer	1.25M Tris/HCl pH 6.8
Separating gel	1.875M Tris/HCl pH 8.8

2.1.10 Bacterial Growth Media

dYT- (double Yeast Tryptone)	1.6% (w/v) Trypton, 1% (w/v) yeast extract, 0.5% (w/v) NaCl in H ₂ O, autoclaved
dYT-Agar	1.6% (w/v) Trypton, 1.5% (w/v) Agar, 1% (w/v) yeast extract, 0.5% (w/v) NaCl in H ₂ O, autoclaved

Antibiotics were added after the media were autoclaved and cooled to below 60°C.

2.1.11 Molecular Biology Equipment and computer hardware

Automated Laser Fluorescence

DNA Sequencer (A.L.F.-sequencer)	GE Healthcare, Uppsala, Sweden
Incubator	W. C. Heraeus GmbH, Hanau
Constant Cell Disruption System	Constant Systems Ltd, Northants, England
Electrophoresis unit	model 2050
Midget	GE Healthcare, Uppsala, Sweden
Gel Jet Imager	Intas, Göttingen
Gene Pulsar® and Pulse Controller	BioRad, Munich
Metal block thermostat	Institute of Microbiology and Genetics, University of Göttingen
Milli-Q® Water Purification System	MILLIPORE, Eschborn
Pipetman® Model P1000, P200, P20	Gilson, Bad Camberg
pH-Meter-526	Schütt Labortechnik, Göttingen
Rotary shaker	Infors AG, Bottmingen, Switzerland
Thermocycler Primus 96plus	MWG Biotech, Ebersberg
Branson Sonifier W-250	Heinemann, Schwäb. Gmünd
UV-VIS Spectrophotometer	UV-1601 SHIMADZU Corporation, Kyoto, Japan
UVT2035 UV transilluminators	Herolab, Wiesloch
Precision Balances L 420 P	Sartorius, Goettingen
Precision Balances U 4800 P	Sartorius, Goettingen
Vision Workstation, BioCad® Family	Applied Biosystems, Foster City, CA
Vortex Genie 2™	Bender & Hobein AG, Zurich, Switzerland

Electrophoresis power supplies:

ECPS 3000/150	GE Healthcare, Uppsala, Sweden
LNGs 350-06	Heinzinger, Rosenheim

Centrifuges:

Centrikon T-1055	Kendro, Langenselbold
------------------	-----------------------

Mikroliter	Hettich, Tuttlingen
Mikro Rapid/K	Hettich, Tuttlingen
Rotanta/RPC	Hettich, Tuttlingen
Roto Silenta/RP	Hettich, Tuttlingen
Sorvall® RC5C (Rotor SS34)	Kendro, Langenselbold

2.1.12 Other materials

Baktolin 5.5 (Disinfection solution)	Bode Chemie, Hamburg
Dialysis tubes VISKING®	SERVA, Heidelberg
Disposable Syringes	Perumo®Syringe, Leuven, Belgium
Glass flasks and test tubes	Schott, Mainz
Glass pipettes	Brand, Wertheim
Glass plates for SDS gels	GE Healthcare, Uppsala, Sweden
Heparin column, Poros® HE 20	Applied Biosystems, Foster City, CA
Meliseptol (disinfection solution)	Braun Melsungen AG, Melsungen
Parafilm	American National Can., Chicago, USA
pH-Indicator stick	Merck, Darmstadt
PCR cups	Biozym, Hess.-Oldendorf
Petri dishes	Greiner, Nürtingen
Pipette tips	Sarstedt, Nümbrecht
Precision Cells-Quartz glass cuvettes	Hellma, Mühlheim/Baden
Reaction Vessels (1,5 ml, 2 ml, 50 ml)	Sarstedt, Nümbrecht
Scalpel blades	Bayha GmbH, Tuttlingen
Ultrafiltration tubes	Vivaspin Vivascience®, Hannover

2.1.13 Software

ALF-Manager	Version 3.02 (1995), GE Healthcare, Uppsala, Sweden
BOXSHADE	Version 3.21, K. Hofmann and M. Baron (www.ch.embnet.org/software/BOX_form.html)
Chromas®	Version 1.45 (32 bit), Version 2.01 and Version 2.31 Technelysium Pty. Ltd.
CLUSTAL W	Service of European Bioinformatics Institute (EBI) http://www.ebi.ac.uk/Tools/clustalw/
CorelDRAW®X3	Version 13.0, Corel GmbH, Unterschleissheim
Endnote	Version 7, Thomson ISI Research Soft, Carlsbad, CA
ISIS™/Draw	Version 2.1.3d, MDL Information Systems Inc., San Ramon, CA

Fragment Manager	Version 1.2 (1995), GE Healthcare, Uppsala, Sweden
OligoCalc	Service of Northwestern University Medical School http://www.basic.northwestern.edu/biotools/oligocalc.html
Protein Molecular Weight	Service of Bioinformatics Organisation Inc. http://www.bioinformatics.org/sms/prot_mw.html
PyMOL™	Version 0.98 (2005), DeLano Scientific LLC, San Carlos, CA
RNAfold	Vienna RNA Package, Vienna, Austria
SigmaPlot	Version 6.0, SPSS Inc., Chicago, IL
TreeView	Version 1.6.6, Roderic D.M. Page, 2001

Vision Workstation Perfusion

Chromatography	Version 2.0, Applied Biosystems, Foster City, CA
Wisconsin Package	Version 10.0, Genetic Computer Group (GCG), Madison, WI
Microsoft® Office Word	Version 2007, Microsoft GmbH, Unterschleißheim
Vector NTI	Version 10.3, Invitrogen, Carlsbad, CA

2.1.14 Databanks

NCBI-Databank for protein, nucleotide, and genomic sequences:

<http://www.ncbi.nlm.nih.gov/>

Repertoire of T/G-, dU-Repair Enzymes in Different Organisms:

<http://www.gobics.de/repairenzymes/maintable>

(User name: gast, Password: Sac2haromyces)

2.2 Methods

2.2.1 Microbiological methods

2.2.1.1 Bacterial media and stocks preparation

Bacterial growth media, buffers and thermostable solutions were sterilized by autoclaving for 20 min at 212°C. Antibiotics were sterile filtered using a syringe and 0.2 µm sterile filters. For selective media preparation, the autoclaved media solutions were cooled down to approximately 50°C and sterile antibiotic solution was added.

2.2.1.2 *E.coli* culture and storage

For the initiation of a new cell culture, 50 µl of *E.coli* glycerine culture were inoculated into 50 ml of dYT medium and incubated at 37°C over night. *E.coli* cells were grown in dYT medium (2.1.10) at 37°C on a shaker. Cell growth was monitored by absorption at 600 nm (OD₆₀₀). In order to obtain single cell colonies, cultures were diluted and plated on dYT/agar or LB/agar plates. *E.coli* glycerine culture stocks were made by re-suspending of 1 ml over night culture in glycerine at 1:1 ratio (v/v) and stored at –20°C. Alternatively, 9% DMSO culture stock was made and stored at –80°C.

2.2.2 Molecular biology methods

2.2.2.1 Ethanol precipitation of DNA

This method allows concentration of DNA samples, as well as removal of salts and small organic molecules. 1/10 vol. of 7M ammonium acetate and 3 vol. of 96% ethanol were added to DNA samples followed by incubation at –20°C for 30min. DNA was pelleted by centrifugation for 30 min at 15000 rpm and 4°C (Mikro Rapid/K Hettich). DNA pellets were washed with 70% ethanol, centrifuged for 15 min at 4°C, and dried at 37°C. Finally, dried DNA pellets were dissolved in double distilled water (ddH₂O) and stored at –20 °C.

2.2.2.2 Phenol/Chlorophorm DNA extraction

Phenol/Chlorophorm extraction was mostly used to remove protein contaminants from DNA samples subjected to a restriction endonuclease digestion. DNA samples (min 50 µl) were consecutively mixed with corresponding volumes of phenol (pH8.0), phenol/chlorophorm/isoamylalcohol (25:24:1, v/v), and chlorophorm. Before each solution change, samples were centrifuged for 5 min at 15000 rpm and RT and DNA containing top aqueous phase was transferred into a new tube. DNA was finally precipitated with ethanol (see 2.2.2.1).

2.2.2.3 Plasmid DNA mini-preparation (alkaline lysis)

2 ml of bacterial culture grown over night at 37°C were used to prepare plasmid DNA by modified alkaline/SDS lysis method using JetStar E1, E2, E3-solutions (Birnboim and Doly, 1979). *E.coli* cells were pelleted by centrifugation for 1 min at 13000 rpm and RT. Pellet was resuspended in 150 µl of E1 solution by vortexing and 150 µl of E2 lysis solution was added followed by incubation for 5 min at RT. After neutralization with 150 µl of E3 solution, samples were incubated for 10 min on ice, and centrifuged for 15 min at 15000 rpm and 4°C. RNase A solution (10 µg/ml final concentration) was added to the supernatant followed by incubation

for 30 min at 37°C. Finally, plasmid DNA was phenol/chlorophorm extracted and precipitated with ethanol (2.2.2.1).

2.2.2.4 Plasmid DNA midi-preparation

JETstar[®]Kit was used for plasmid DNA preparation from 50 ml of over night *E.coli* culture. Cells were pelleted by centrifugation (4000 rpm, 10 min, 4°C (Rotanta RPK, Hettich)) and resuspended in 4 ml of E1 solution until homogeneous. Then E2 solution was added with gentle mixing and samples were incubated for 5 min at RT. After addition of E3 solution, samples were mixed by multiple inversions and cell debris were removed by centrifugation for 10 min at 4000 rpm and RT. Cleared supernatant was applied to a JETstar column equilibrated with E4 solution followed by washing twice with 10 ml of E5 solution. DNA was eluted with 5 ml of E6 solution, precipitated with 0.7 volume of isopropanol and pelleted by centrifugation for 15 min at 15000 rpm and 4°C. Finally, DNA pellet was washed with 70% ethanol, dried at 37°C, and redissolved in 10 µl of double distilled water (ddH₂O).

2.2.2.5 Polymerase Chain Reaction (PCR)

PCR process usually consists of a series of 25-50 cycles carried out in an automated thermal cycler and comprises three major steps: DNA denaturation, primer annealing, and primer extension. During denaturation step at 95°C, the double-stranded DNA opens into single-stranded DNA. Then two specific primers complementary to sequence, which flanks the target DNA loci, hybridize to a single-stranded template at primer annealing temperature (T_A). Finally, a thermostable DNA polymerase builds up a complementary DNA strand beginning from 3' end of the primer. During PCR amplification, the amount of target DNA fragment grows exponentially. The following general PCR mixture was used for amplification on plasmid and genomic DNA templates:

Reaction mix	
Template DNA	50-100 ng
Primer N1 (10 pmol/µl)	1-2 µl
Primer N2 (10 pmol/µl)	1-2 µl
Polymerase	1-2 U
10x Polymerase buffer	5 µl
dNTPs (10 mM)	1 µl
dd H ₂ O	up to 50 µl

The following typical cycling protocol was used with modifications of annealing temperature, elongation time, and/or elongation temperature depending on the primer pair used and on the length of PCR fragment:

Step	Temperature	Time
1.	95°C	3 min
2. Polymerase addition	85°C	2 min
3.	94°C	1 min
4.	T _A *	1 min
5.	72°C	1 min
6.	72°C	10 min
Step 2-5: 30x repeat cycles		

* T_A: Optimal primer annealing temperature calculated by following equation:

$T_A = (T_{m1} + T_{m2}) / 2 - 3^\circ\text{C}$, where T_{m1} and T_{m2} are melting temperatures of Primer 1 and Primer 2, accordingly;

T_m of PCR primers were calculated using the following equation: $T_m [^\circ\text{C}] = 69.3 + 0.41(\%G+C) - 650/N$ (Chester and Marshak, 1993)

30 µl of Chill-out™ Wax (2.1.7) was added on top of each PCR mix in order to avoid the sample evaporation during PCR in thermal cycler (2.1.11).

2.2.2.6 Spectrophotometric determination of DNA concentration

Absorption of ultra-violet (UV) light by the ring structure of purines and pyrimidines is the basis of the spectrophotometric DNA analysis. DNA samples were diluted in ddH₂O (1:10 or 1:100) depending on estimated concentrations. DNA concentration were determined by measuring the absorbance at $\lambda=260$ nm (A₂₆₀) using a spectrophotometer and a quartz cuvette. The concentration was calculated based on the assumption that A₂₆₀ = 1.0 is equal to 50 µg/ml of double-stranded (ds) DNA, to 40 µg/ml of single-stranded (ss) DNA and RNA, and to 20 µg/ml of oligonucleotides. The purity of DNA was assessed by the A₂₆₀/A₂₈₀ ratio. Ratios between 1.8 and 2.0 are indicative of pure, protein-free DNA.

2.2.2.7 Agarose gel electrophoretic analysis of DNA

DNA molecules can be separated according to their size by electrophoretic migration. For preparative and analytical DNA analysis 1%-2% agarose gel electrophoresis was used. The agarose was dissolved in 1xTBE buffer (2.1.9) in a microwave. Then 0.5 µg/ml of ethidium bromide (2.1.9) was added and the gel was poured into a horizontal gel-forming chamber. Samples were mixed with 0.5 vol. of loading buffer (2.1.9). Gels were run in 1xTBE buffer at the constant electric field power of 5-10 V/cm. DNA in the gel was visualized under UV light at 305 nm using UV transilluminator (2.1.11).

2.2.2.8 Denaturing polyacrylamide gel electrophoresis (A.L.F-PAGE)

Automated Laser Fluorescence DNA Sequencer (A.L.F.-DNA sequencer 2.1.11) allows direct detection of fluorescently labelled DNA. During electrophoresis DNA fragments migrate

downwards through the gel. The laser beam excites fluorescently labelled DNA and the emitted light is detected by fluorescent detection system. Collected photo-detector signals are then digitized and sent to a computer for further processing. Results are represented in the form of intensity peaks plotted against the running time scale. Reaction products of all activity assays performed with DNA modifying enzymes were analyzed by A.L.F.-sequencer (2.1.11). DNA samples were mixed with A.L.F.-marker (2.1.9) and applied onto 11% denaturing polyacrylamide/urea gel (30 x 28 x 0.5 cm).

11% A.L.F.-PAGE protocol:

Reagent	Amount
Urea	29.43 g (7M)
Acrylamide (30%)	24.84 ml (10.65%)
Bisacrylamide (2%)	12.42 ml (0.35%)
10XTBE	8.40 ml
APS (10%)	700 µl
TEMED	70 µl
ddH ₂ O	up to 70 ml

Sterile filtered A.L.F.-PAGE was poured in between two cleaned (water, ethanol, isopropanol) and thermostable glass plates and a plastic comb was inserted to form wells. Gel was polymerized for approximately 30 min at RT. Then comb was removed, wells were rinsed with water, gel was placed into electrophoresis chamber, and buffer reservoirs were filled with 1XTBE buffer (2.1.9). Before applying samples, the gel was preheated for 20 min. DNA samples were mixed with A.L.F. marker (½ sample volume), preheated for 5 min at 95°C, and applied onto the gel loading up to 15 µl into each well. The gel was run for 250 min at a constant power of 52W, 52°C, a laser power of 4mW, and a sampling power of 2s. Electrophoresis data were processed using Fragment Manager (GE Healthcare, see 2.1.13).

2.2.2.9 DNA extraction from agarose gel

To extract DNA from agarose gel, the desired band was excised with a sterile scalpel under UV light (305 nm). DNA was purified using NucleoTrap® (Macherey & Nagel, section 2.1.8) purification kit, where agarose gel slices were dissolved in the presence of chaotropic salts (buffer NT1 and NT2) and DNA was bound to a silica matrix. After several washing steps of silica matrix, pure DNA was finally eluted under low ionic strength conditions (NE buffer). NT1, manufacturer provided NT2 and NE buffers.

2.2.2.10 Site-directed mutagenesis using PCR (modified Stratagene protocol)

PCR based mutagenesis allows to introduce site-specific mutations in double-stranded plasmid DNA. This procedure utilizes a supercoiled dsDNA vector with an insert of interest

and a pair of complementary oligonucleotide primers containing a desired mutation (Figure9). During amplification in an automated thermocycler, *Pfu* DNA polymerase extends the primers and linear mutated plasmid is generated. Then parental DNA is eliminated by enzymatic digestion using *DpnI* restriction nuclease specific for methylated and hemimethylated DNA. Finally, newly synthesized mutated plasmid is transformed into competent *E.coli* cells where it undergoes repair by endogenous bacterial machinery (Figure9).

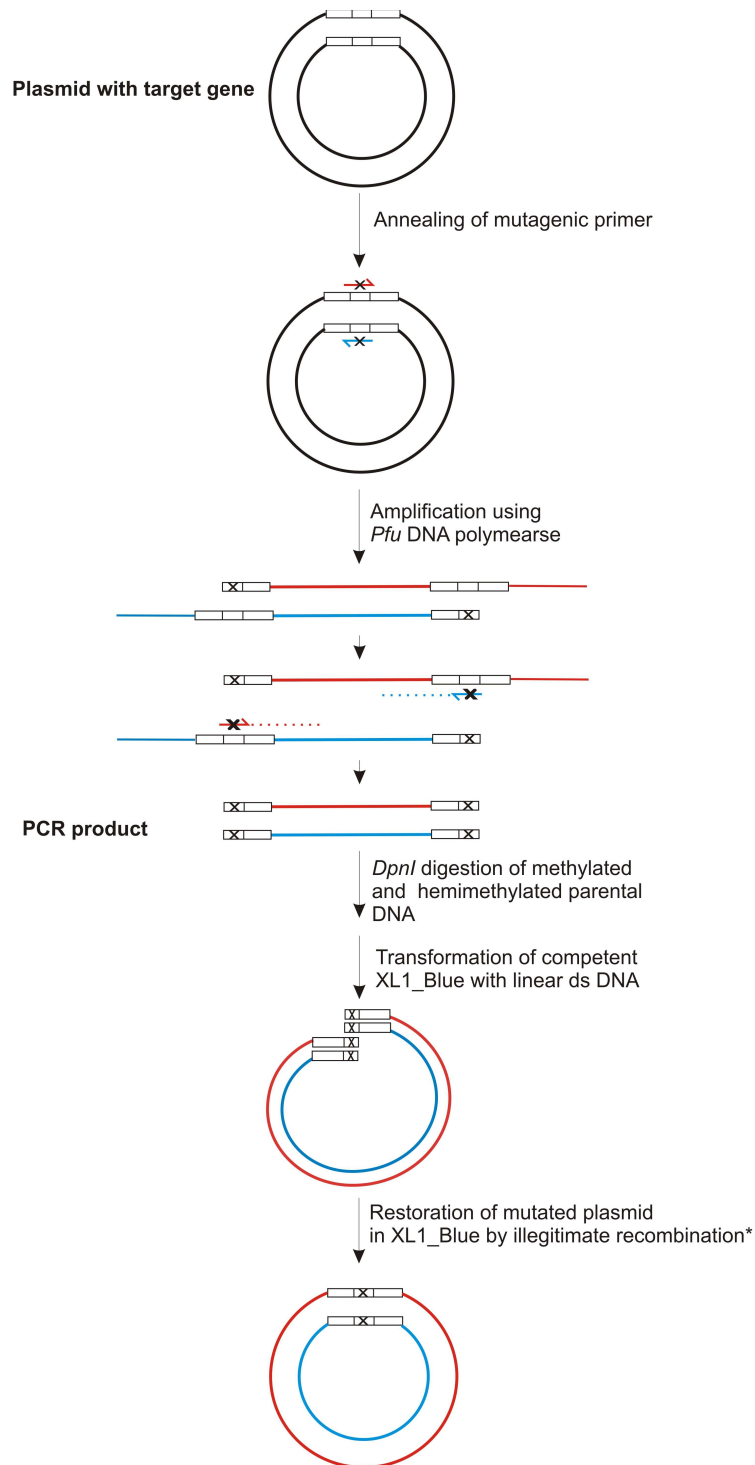


Figure 9: Schematic representation of the site-directed mutagenesis

*(Conley et al., 1986; Jones, 1994; Yamashita et al., 1999)

The following general PCR mixture and cycling protocol were used for linear amplification on plasmid DNA templates:

PCR mix components	Amount
Template DNA	50 ng
Primer N1 (100 nmol/μl)	125 ng
Primer N2 (100 nmol/μl)	125 ng
<i>Pfu</i> polymerase	2.5 U
10x <i>Pfu</i> buffer (+MgSO ₄)	5 μl
dNTPs (10 mM)	1 μl
ddH ₂ O	up to 50 μl

PCR protocol:

Step	Temperature	Time
1.	95°C	1 min
2. Polymerase addition	85°C	2 min
3.	95°C	50 sec
4.	65°C T _A *	50 sec
5.	72°C	10 min
6.	72°C	15 min
Step 3-5: 18x repeat cycles		

T_A* indicates an annealing temperature for mutagenic primer estimated using following formula:

$$T_A = T_m - 10$$

$$T_m = 81.5 + 0.41(\%G+C) - 675/N - \% \text{ mismatch}, \text{ where}$$

N is the primer length in bases and values for %G+C and % mismatch are the whole numbers.

Usually 5X 50 μl PCR mixtures were linear amplified in an automated thermocycler (2.1.11) using *Pfu* DNA polymerase (2.1.6, Fermentas), pooled down and ethanol precipitated (2.2.2.1, Figure 1, A). DNA was resuspended in 8 μl of ddH₂O and subjected to *DpnI* (2.1.6, Fermentas) digestion (1 μl of *DpnI* for 1h at 37°C) in 10 μl total volume, followed by preparative DNA agarose gel electrophoresis and DNA gel extraction (2.2.2.7 and 2.2.2.9, Figure 11, B and C). 50 μl of electro competent XL1-Blue *E. coli* cells (2.1.1.6) were transformed with 2 μl of gel purified mutated plasmid DNA and incubated overnight on dYT/Amp¹⁰⁰/Tet²⁵ (2.1.10) agar plates. 10-20 single colonies were used for plasmid preparation (2.2.2.3 and 2.2.2.4), followed by a preliminary restriction analysis with *XhoI*/*NcoI* restriction endonucleases (Fermentas 2.1.6, Figure 11, D). Finally, DNA from a minimum of 10 transformants was verified for desired point mutation by sequencing. One of the positive clones was subsequently used for mutant protein production (2.2.3.3.).

2.2.2.11 Transformation of competent *E. coli* by electroporation (Dower et al., 1988).

200 µl of over night *E.coli* culture were inoculated into 50 ml of dYT medium (2.1.10) and grown to mid-log phase ($OD_{600} \sim 0,6$) at 37°C on a shaker. Cells were harvested by centrifugation in 50 ml falcon tubes for 10 min at 4.000 r.p.m. and 4°C (Rotanta/RPK, Hettich), carefully resuspended in 50 ml of ice-cold H₂O and incubated on ice for 20 min. Then cell suspension was centrifuged and cell pellet was washed 4 times by resuspension in ice-cold H₂O (40 ml, 30 ml, 20 ml and 10 ml), followed by centrifugation after each wash. Finally, pellet was dissolved in 1-2 ml of H₂O and left on ice. For the transformation, 1-2 µl of DNA solution (2µg) were mixed with 75 µl of electrocompetent cells and incubated on ice for 5 min. DNA-cell mixture was transferred into a chilled electroporation cuvette and subjected to the pulse at 2.5kV, 200Ω and 2.5µF (2.1.11). Then transformants were immediately resuspended with 1-2 ml of dYT medium (2.1.10), incubated for 1h at 37°C on a shaker and 100 µl aliquots (with or without dilution) were plated out onto dYT agar plates containing appropriate antibiotics. Untransformed competent cells used as a negative control were also plated on antibiotics containing agar plates.

2.2.2.12 DNA cleavage by restriction endonucleases

For effective DNA cleavage optimal temperatures and buffers recommended by manufacturer were used (2.1.6). 50 U of enzyme was used for preparative DNA cleavage (minimum 5 µg) in 100 µl assay volume for 1h or maximal 16h. For analytical DNA cleavage (300-500 ng), 5 U of restriction enzyme in 20 µl assay volume were used. In case of double digestion, if no single buffer could be found where both enzymes would show 50-100% activity, two subsequent digestion reactions were performed. After first digestion, DNA was ethanol precipitated (2.2.2.1).

2.2.2.13 Preparation of oligonucleotide substrates for activity assays

For endonuclease assays, 0.01 pmol/µl substrate stocks were prepared by hybridization of 5 pmol of a fluorescein-labelled oligonucleotide with 25 pmol of an opposite strand oligonucleotide in 100 µl of 1XSSC buffer (2.1.9) using an automated thermocycler (Program: 90°C, 15 sec; 80°C, 3 min; 50°C, 15 min; 20°C, 15 min), followed by a 1:5 dilution with water. In case of a single-stranded (ss) substrate preparation, 5 pmol of fluorescein-labelled oligonucleotides were mixed with 100 µl of 1XSSC buffer and 400 µl of water subsequently. For EMSA substrates preparation, 100 pmol of 40_PRI_AP_Dunkel, PRINCE_U_Dunkel, PRINCE_C_Dunkel, PRINCE_T_Dunkel (not fluorescein-labelled) were hybridized with 100 pmol of PRINCE_G_35mer or PRINCE_G_40mer opposite strand in 100 µl of 1XSSC buffer (2.1.9).

2.2.2.14 Electrophoretic Mobility Shift Assay

The Electrophoretic Mobility Shift Assay (EMSA) also referred to as the gel shift assay is based on the fact that DNA-protein complexes migrate slower relative to free DNA molecules when subjected to a non-denaturing polyacrylamid gel electrophoresis. We used EMSA to investigate protein-DNA binding activity.

Gel-shift assays were performed by incubating of purified protein with 1 pmol (~26ng) of a respective unlabelled oligonucleotide substrate (EMSA substrate 2.2.2.13) for 5 min at 65°C in 10 µl of EMSA buffer (2.1.9) containing 10% glycerol. Reaction products were analysed on 10% TE-PAGE (see Table 4). Gels were run at room temperature for 50 min at 10 V/cm in 1xTE (2.1.9) running buffer and stained with SYBRO®-Gold (Invitrogen, 2.1.7) according to the manufacturer protocol. DNA was visualised by UV-Transilluminator (2.1.11) at 302 nm.

Competitive EMSAs were carried out as described previously, except that EMSA substrate oligonucleotides were mixed with increasing amount of unlabelled competitor DNA for 5 min at 65°C in EMSA buffer followed by adding appropriate amount of pure protein to a final volume of 15 µl and further incubation for 5-10 min at 65°C. The amount of competitor DNA added was calculated as molar equivalents of nucleotide relative to substrate in case of entire pET-vector and as molar equivalents of DNA ends in case of blunt 20-mer dsDNA oligonucleotides correspondently.

Table 4: 10% TE-PAGE (60 ml) recipe:

Reagent	Volume
Acrylamid/Bisacrylamid 30% (37.5:1)	20 ml
10XTBE buffer	6 ml
10% ammonium persulfate (APS)	400 µl *
TEMED	30 µl*
H ₂ O	33.9 ml

2.2.2.15 DNA sequence analysis

DNA sequence analysis was made by Goettingen Genomics Laboratory (G2L) or by SEQLAB (Goettingen) using chain termination method by Sanger (Sanger et al., 1977). DNA samples were prepared by mixing a template DNA, a sequencing primer and ddH₂O to a final volume and concentrations recommended by G2L or SEQLAB. Sequence data were analyzed using Chromas®.

2.2.3 Protein biochemical methods

2.2.3.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Gel electrophoresis provides a means of separating molecules that migrate through a porous matrix in response to an electric field. In denaturing sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970), sodium dodecyl sulphate (SDS) - an anionic detergent - confers a negative charge to polypeptides in proportion to their length and β -mercaptoethanol - a reducing agent - reduces disulfide bridges in proteins. Consequently, proteins are separated based on their molecular mass. Laemmli gels are composed of two different gels (stacker and running gel). SDS-PAGE (mini-gels: 7.5 cm x 8 cm and 0.75 cm thick) were prepared as shown in Table 5:

Table 5: 15% mini SDS-PAGE recipe:

Components	15% Resolving gel (60 ml)	5% Stacking gel (35 ml)
Acrylamid/Bisacrylamid 30% (37.5:1)	30 ml	5.9 ml
1.875 M Tris-HCl, pH 8.8	12 ml	-
1.25 M Tris-HCl, pH 6.8	-	3.5 ml
10% sodium dodecyl sulphate (SDS)	600 μ l (1% final)	350 μ l (1% final)
dH ₂ O	17.2 ml	-
10% ammonium persulfate (APS)*	200 μ l	120 μ l
TEMED*	30 μ l	35 μ l

* Added just before pouring the gel

15% resolving gels were poured between two glass plates, overlaid with isopropanol to compose the flat gel surface and polymerized for 30 min at room temperature (RT). Then 5% stacking gel mix was made, poured onto top of set resolving gel, a comb was inserted and gel was let to polymerize for 30 min at RT. Protein samples were mixed with 1/5 volume of loading buffer (2.1.9) and denatured by heating for 5 min at 95°C just before loading onto the gel. Gels were run in 1x Laemmli buffer (2.1.9) at 20 mA constant current, followed by staining with Coomassie brilliant blue R250 (2.1.9).

2.2.3.2 Silver staining of SDS-PAGE

To detect low amounts of proteins (as little as 1 ng), silver staining was used to visualize protein bands in SDS-PAGE (modified protocol by (Blum et al., 1987). After electrophoresis, gels were incubated with fixation solution (2.1.9) for 1h at RT with gentle agitation, washed three times with 50% ethanol (v/v) for 20 min, followed by incubation with wash solution (0.01% sodium thiosulphate, 2.1.9) for 1 min. Then gels were immediately washed with dH₂O (2 times x 30 s), incubated with silver stain solution (2.1.9) for 20 min, washed again with dH₂O (2 times x 30 s) and incubated in developing solution (2.1.9) for about 10 min.

Developing reaction was stopped by incubation the gels with 10 mM EDTA-containing stop solution for 10 min. Finally the gels were washed with water, placed between 2 plastic foils and sealed.

2.2.3.3 Heterologous protein expression

pET expression system (pET vector (2.1.2.1) and *E. coli* BL21 (DE3) strain (2.1.1.2)) was chosen for over-expression of recombinant proteins. In pET vector target gene is regulated by strong T7 promoter (bacteriophage origin), which is selectively recognised only by T7 RNA polymerase. BL21 (DE3) expression host possesses T7 RNA polymerase gene, *lac*-promoter and *lac*-operator incorporated in its genome. Expression of T7 RNA polymerase and consequently the target gene can be activated by addition of lactose or its analogue isopropyl- β -D-thiogalactopyranosid (IPTG) to a cell growth medium.

Usually 20 ml of appropriate *E.coli* expression strain (2.1.1.3 or 2.1.1.4) grown over night on a selective medium and freshly transformed with respective expression vector (2.1.2.1) was used to inoculate 1L of dYT selective medium (2.1.10) in 3L Erlenmeyer flasks. Culture was grown at 37°C with shaking to an OD₆₀₀ of 0.6. Gene expression was induced by addition of IPTG to a final concentration of 1 mM and cell culture was further incubated for 3h at 30°C. Cells were harvested by centrifugation for 30 min at 4000 r.p.m. and 4°C (Rota Silenta/RP, Hettich, Germany), resuspended in 25 ml of IMAC-washing buffer (2.1.9) and centrifuged for 30 min at 9000 r.p.m and 4°C (SS-34 Rotor, Sorvall® RC5C, Kendro). Cell pellet was frozen and stored at -80°C. After thawing and resuspension in 20 ml of IMAC-washing buffer (2.1.9), cells were homogenized by sonication on ice for 2X3 min using Branson Sonifier 250 (output level 5, duty cycle 50%) and disrupted by passage through a Constant Systems Ltd (Daventry, England) cell disruptor at 180 MPa and 14°C. Cell debris were precipitated by centrifugation for 1h at 15.000 r.p.m. and 4°C (SS-34 Rotor, Sorvall® RC5C, Kendro). To purify proteins originated from thermostable organisms, *E.coli* cell lysate was incubated for 40 min at 65°C and recentrifuged (15.000 r.p.m., 30 min, 4°C). Clear cell lysate was subjected to Immobilized Metal Ion Affinity Chromatography (IMAC) (2.2.3.4). Cell pellets were also analysed by SDS-PAGE (2.2.3.1) to check for possible aggregation of target protein. Test expression of recombinant protein was performed in 50 ml *E.coli* culture, where induced and non-induced *E.coli* cells were harvested by centrifugation, resuspended in 1x Laemmli buffer (2.1.9) and analyzed by 15% SDS-PAGE.

2.2.3.4 Immobilized Metal Ion Affinity Chromatography

Immobilized Metal Ion Affinity Chromatography (IMAC) is based on highly specific coordinate binding of amino acids to immobilized metal ions. Recombinant proteins cloned into an expression vector such as pET_B001 (2.1.2.1) are provided with polyhistidine-tag (here

6xHis) at their C- or N-terminus. Imidazole ring of histidine residue possess high affinity to Ni^{2+} ions that in case of IMAC are immobilized on a sepharose matrix. During IMAC purification, cell lysate is incubated with affinity matrix, washed with buffer and the target protein is then eluted with imidazole gradient.

IMAC column was prepared by pouring 5 ml of 50 % slurry of Chelating Sepharose™ Fast Flow (2.1.7), previously equilibrated in IMAC buffer, into 5 ml plastic syringe and let it settle down by gravity force. Then the column was equilibrated with 1 column volume (5ml) of 100 mM NiCl_2 , washed with 5 column volumes of water, and equilibrated with 3 column volumes of IMAC washing buffer (2.1.9). Clear bacterial cell lysate, prepared as described in 2.2.3.3, was applied onto the IMAC column followed by 2x washing each time with two column volumes of IMAC-washing buffer (2.1.9). Target protein was eluted from the column by passing two column volumes of IMAC elution buffer (2.1.9), one at a time (imidazole concentration: 30, 40, 50, 60, 70, 80, 90, 100, 300 and 500 mM). Collected IMAC fractions were analyzed by 15% SDS-PAGE (2.2.3.1). Fractions containing protein of interest were combined and concentrated to a final volume of 5 ml.

2.2.3.5 Heparin affinity chromatography

Heparin affinity chromatography was performed on a pre-packed POROS® HE20 (Perfusion chromatography, PerSeptive Biosystems) column using Vision Workstation (BioCad®Family, Applied Biosystems, 2.1.11) designed for automated control of essential chromatographic parameters such as: flow-rate, pressure, pH, elution volume, fraction volume *etc.*, and sophisticated computerized data analysis. Heparin coupled with a high number of anionic sulphate groups is a high-capacity cation exchanger that allows specific purification of positively charged DNA binding enzymes from protein mix. Specific proteins can be then selectively dissociated from heparin with a salt gradient. Heparin column was equilibrated with heparin washing buffer (2.1.9) at a flow rate of 4 ml/min. Then 5 ml of concentrated IMAC protein solution described in 2.2.3.4 was diluted 1:10 with heparin washing buffer (up to 50 ml final volume) and applied onto heparin column at same flow rate. The column was washed with 30 ml of heparin washing buffer. Proteins retained in the column were eluted with 15 column volumes of a continuous 0-1.5M NaCl salt gradient and collected in 1 ml fractions. Protein elution was monitored at 260 and 280nm by a computer-controlled spectrophotometer and results were displayed as a dual line chromatogram (Figure 13). Pick heparin fractions were analyzed by SDS-PAGE (2.2.3.1). Fractions containing protein of interest were combined and concentrated (2.2.3.8).

2.2.3.6 Spectrophotometric determination of protein concentration

Concentration of purified proteins was determined by measuring the absorbance at 250-300 nm (UV-Region). Aromatic amino acids (tyrosine, phenylalanine and tryptophan) exhibit a strong UV-light absorbance. Consequently, proteins absorb UV-light in proportion to their aromatic amino acids content and total concentration. The molar concentration of protein solutions were estimated using Lambert-Beer law and following equation:

$$A = \varepsilon \cdot c \cdot d \quad \Leftrightarrow \quad c = \frac{A}{\varepsilon \cdot d} \quad ; \text{Where} \quad \begin{array}{l} A = \text{Absorbance} \\ c = \text{molar concentration} \\ d = \text{light path (cm)} \\ \varepsilon_{280} = \text{molar extinction coefficient at 280 nm}^* \end{array}$$

*The molar extinction coefficient (ε_{280}) was calculated for each particular protein from its amino acid sequence using following equation (Pace et al., 1995):

$$\varepsilon_{280} (M^{-1}cm^{-1}) = \sum Trp \cdot 5500 + \sum Tyr \cdot 1490 + \sum Cystine \cdot 125$$

2.2.3.7 Protein concentration determination by Bradford assay

Bradford protein determination is based on a shift in absorbance maximum of free Coomassie Brilliant Blue G-250 dye (465 nm) relative to a dye-protein complex (595 nm). Furthermore, G-250 selectively binds to arginine, lysine, histidine and aromatic residues (Bradford, 1976). An increase in A_{595} of protein-bound dye is proportional to the amount of protein present in the sample. First, a standard curve is made using standard protein concentrations. Then the amount of protein in experimental samples is determined from the standard curve with respect to the samples volumes and dilution factors, if any.

Bovine serum albumin (BSA) was used as a protein standard. 100 μ l of standard BSA solutions (0, 25, 50, 75, 100, 150 μ g/ml) were prepared by dilution of 1 mg/ml or 10 mg/ml stocks in 1xPBS buffer. Then 5 ml of Bradford dye solution (2.1.9) were added to each 100 μ l BSA standard solution and the whole mix was left for 10 min in the dark at RT, followed by measurement of absorption at 595 nm in a spectrophotometer (2.1.11). Standard curve was made by plotting A_{595} values of standard samples against total amount of protein. Protein samples of unknown concentration were diluted in PBS buffer up to a final volume of 100 μ l and incubated with Bradford reagent as described above. Then absorption at 595 nm was measured and protein concentration was determined using the standard curve.

2.2.3.8 Protein samples concentration and storage

Purified proteins were concentrated up to 0.5 - 1 ml final volume using 20 ml centrifugal concentrators (Vivaspine, Vivascience®) with appropriate Molecular Weight Cut Off (MWCO) at 3000 rpm and 4°C (Rotanata/RPC, Hettich). The protein concentration was determined measuring absorbance at 280 nm using a spectrophotometer (2.2.3.6 and 2.1.11).

Concentrated protein solutions were mixed with glycerol at a 1:1 ratio and stored at -20°C or -80°C.

2.2.3.9 Endonuclease assay:

0.12 pmol of appropriate substrate (2.2.2.13) was pre-incubated in endonuclease buffer (2.1.9) for 10 min at the optimum temperature of the tested enzyme. Then an appropriate amount of enzyme was added to a final volume of 50 µl and reaction mix was incubated for 20 min unless otherwise specified. Then 25 µl of A.L.F.-marker (2.1.9) was added, samples were heated for 5 min at 95°C and applied onto the 11% A.L.F.-PAGE (2.2.2.8) in 7 µl total volume (~17 fmol of fluorescein-labelled material).

3 Results and discussion

3.1 Construction and properties of Mth212/D151N mutant

This thesis is concerned with analysis of the interaction between Mth212 and its U-containing DNA substrate. The biochemical methods such as substrate binding studies and catalytic assays were performed with both wild-type Mth212 enzyme and its constructed mutants. Furthermore, the materials developed and the insights gained during the project were used for structural studies employing X-ray crystallography in collaboration with Department of Molecular Structural Biology (Georg-August University, Goettingen).

3.1.1 Rationale for mutation of conserved Asp-151

To examine the molecular and structural bases of Mth212 substrate recognition and catalysis, a catalytically inactive Mth212 enzyme with retained substrate-binding capacity was constructed. The rationale was that such a mutant would provide the information as to whether the newly discovered U-endonuclease (U-endo) activity of Mth212 is served by the same active site as the other enzymatic activities common to all characterised ExoIII homologues. Given that Mth212 has a single catalytic site for both Ap-endo and U-endo activities, such catalytically inactive mutant should enable further studies of substrate recognition and catalysis using specific approaches such as: gel-retardation assay and co-crystallisation with DNA substrates containing an AP-site or a uracil residue. To identify Mth212 catalytically essential amino acid residues, a comparative amino acid sequence analysis was performed between Mth212 and other ExoIII homologues with known 3D structures.

By the time this project was started the crystal structure of Ape1, a human ExoIII homologue, and its active site essential amino acid residues had been reported (Barzilay et al., 1995; Mol et al., 2000). Furthermore, Rothwell and co-workers found that Asp-210 residue in human Ape1 is essential for the catalytic function of this enzyme but at the same time is dispensable for substrate recognition (Rothwell et al., 2000). Analysis of the three-dimensional structure (3D) of Ape1 in complex with AP-site-containing DNA showed that Asp-210 aligns with the scissile P-O3' bond resided in the active site of the enzyme (Gorman et al., 1997).

Amino acid sequence alignment of human Ape1 and Mth212 revealed Asp-151 in Mth212 as the equivalent residue to Asp-210 in human Ape1. Therefore, Asp-151 was chosen for the substitution experiment (Figure 10).

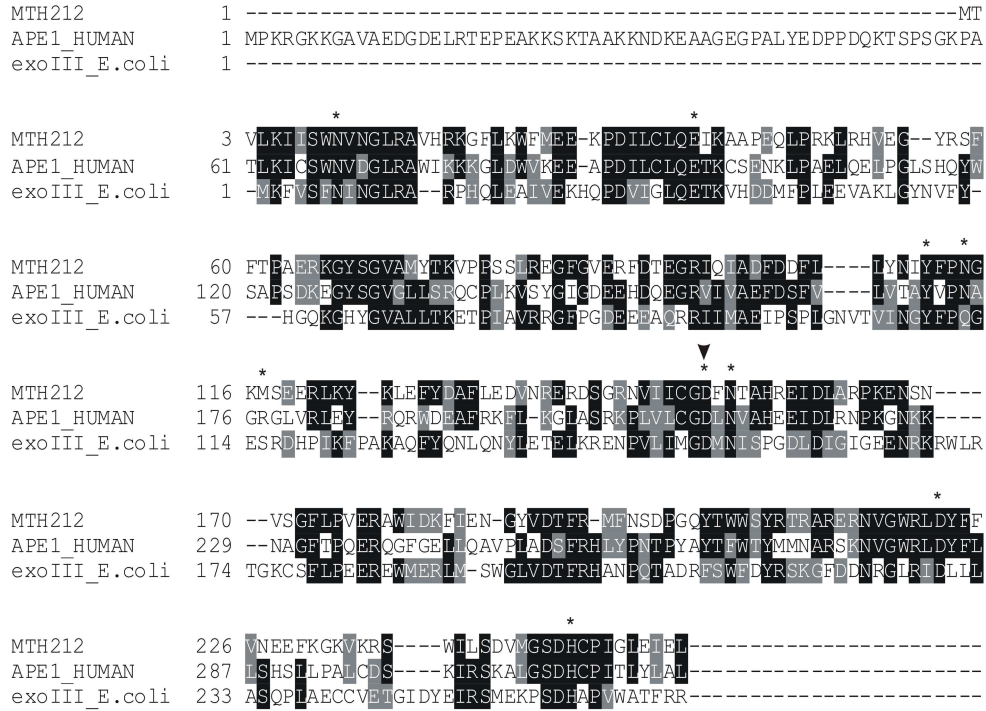


Figure 10: Multiple amino acid sequence alignment of Mth212 (*Methanothermobacter thermautotrophicus*) with ExoIII (*Escherichia coli*) and Ape1 (*Homo sapiens*)

Amino acid residues involved in the catalysis by Ape1 according to (Mol et al., 2000) and (Rothwell et al., 2000) are marked with an asterisk. Arrow indicates the conserved Asp-151 residue in Mth212 active site that was mutated to Asn, Ala or Ser. Alignment was performed using WISCONSIN PACKAGE version 10.3 (Womble, 2000) and arranged with BOXSHADE version 3.21.

3.1.2 Mth212/D151 mutants: construction and expression

mth212 gene, inserted in pET_B_001 expression vector (2.1.2.1) between *NcoI* and *XhoI* restriction sites by Jens Georg during his diploma project (Georg, 2005), was used for the site-directed mutagenesis. Mth212 gene was mutated to replace the codon for Asp-151 with one encoding an asparagine (Asn or N), alanine (Ala or A) or serine (Ser or S). There were two reasons for generation Ala and Ser variants. First, to exclude the possibility of spontaneously arising revertants in case of D151N variant, where only one of the three original nucleotides was changed (GAC→AAC, changed nucleotide shown in red). Substitution of Asp codon by Ala codon introduces two different nucleotides (GAC→GCA) and in case of Ser codon all three nucleotides are different from the original ones (GAC→TCA). The second reason was to eliminate remaining endonucleolytic activity resulting from Mth212/D151N during endonuclease assay with AP/G substrate (2.1.4, 2.2.2.13) and 100:1 enzyme/substrate ratio.

Three Asp-151 mutants were generated using PCR based site-directed mutagenesis as described in section 2.2.2.10. Generation of D151N variant is shown in Figure 11 as an example. The two other mutant variants, D151A and D151S were created using same procedure (data not shown). All site-specific mutations were verified by DNA sequence

analysis (see Appendix 7.1.3). The obtained mutants were designated as Mth212/D151N, Mth212/D151A and Mth212/D151S.

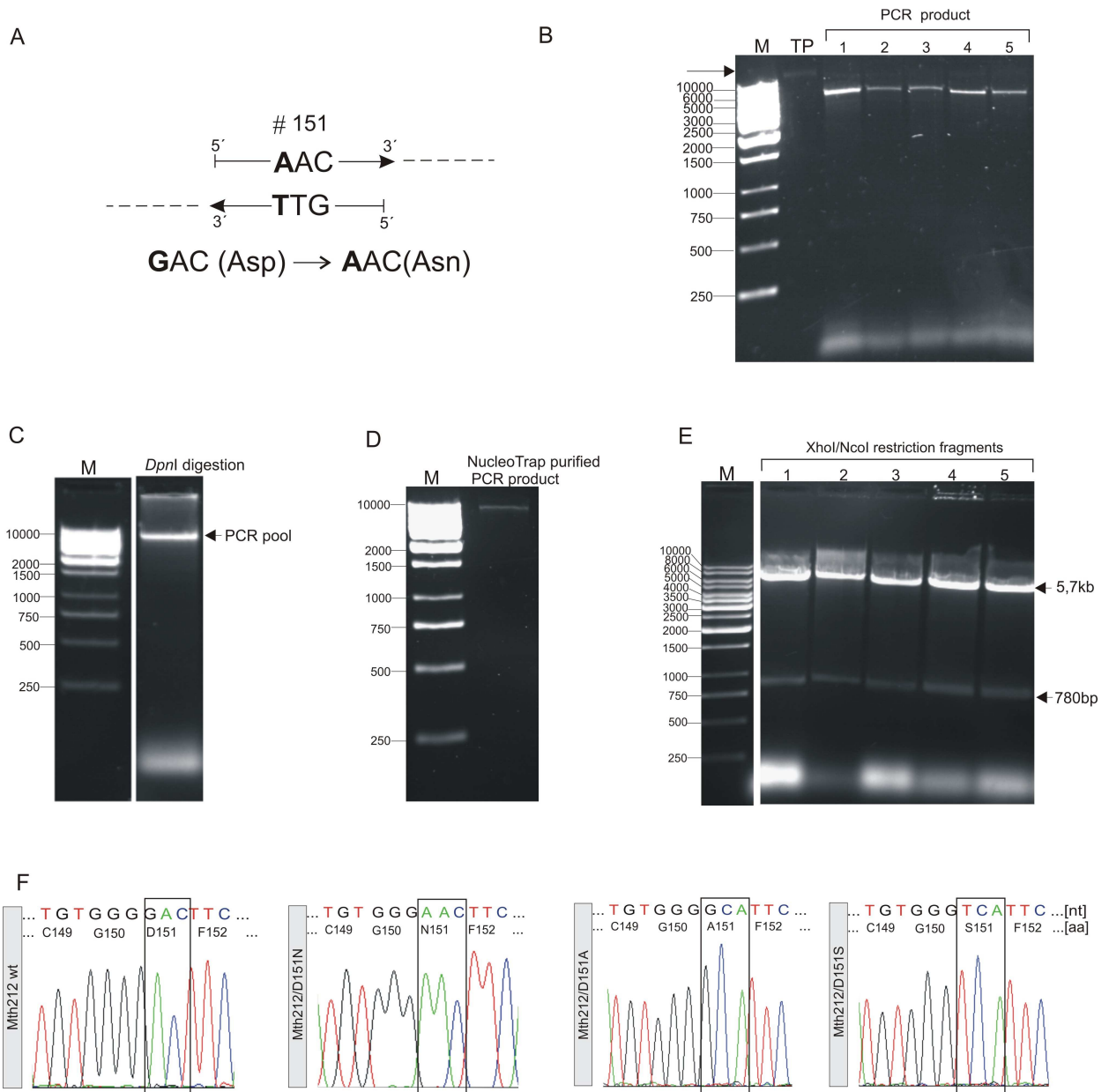


Figure 11: Quick-Change mutagenesis of Mth212

1% agarose gel analysis (2.2.2.7) during site-directed mutagenesis of *mth212* (2.2.2.10).

A: Schematic representation of mutagenic primers (black arrows): 212_D151N_For (top) and 212_D151N_Rev (bottom) used for substitution of Asp-151 to Asn (changed bases are outlined). **B:** PCR amplification products (Lane 1-5) were combined for *DpnI* digestion. **TP:** Template DNA (pET_B001_Mth212, 68 ng). Arrow indicates the band corresponding to the relaxed circular template DNA. **M:** Molecular length marker (GeneRuler™ 1kb DNA Ladder, Fermentas). **C:** PCR amplification products after *DpnI* digestion (2.2.2.10). **D:** NucleoTrap® purified PCR product used to transform *E.coli* XL1-Blue (2.1.1.5). **E:** *XhoI/NcoI* restriction endonuclease fragment analysis (2.2.2.12) of pET_B001_Mth212/D151N derivatives. Arrows indicate restriction fragments: ~5.7 kb corresponds to the linear pET_B001 vector and 780 bp to the length of Mth212 gene. **F:** Fragments of sequencing data of wild-type (wt) Mth212 and its three Asp-151 (D151) mutants represented in Chromas 2.33. Indications in grey bars are: Mth212 wt (wild-type Mth212) and Mth212/D151X, where X: N (Asn), A (Ala) and S (Ser). Black rectangle highlights wt and mutated codon in *mth212* nucleotide sequence [nt] and one letter symbol and number of changed amino acid in Mth212 protein sequence [aa], respectively.

Each of the three mutant variants was produced in *ung⁻ E.coli* BL21_UX (2.1.1.3) strain, isolated as described in 2.2.3.3 and purified through IMAC (2.2.3.4) (Figure 12), followed by Heparin affinity chromatography (2.2.3.5) (Figure 13A, B and C). The obtained proteins were checked for purity by SDS-PAGE (2.2.3.1) (Figure 13D, E and F) and concentrated to the final volume of 1 ml. Their final concentrations were calculated as described in section 2.2.3.6 and comprised 7.03 mg/ml, 8 mg/ml and 14 mg/ml for Mth212/D151N, Mth212/D151A and Mth212/D151S, correspondingly.

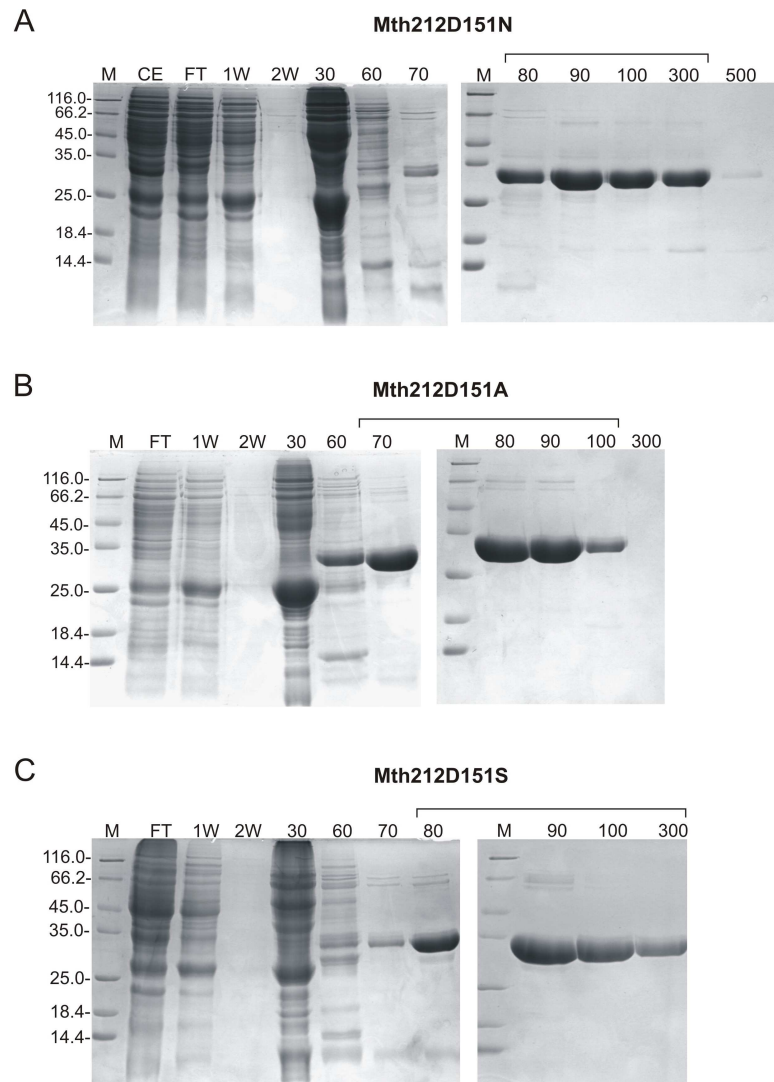


Figure 12: Immobilized-metal-ion affinity chromatography of His-tagged Mth212 mutants

IMAC were carried out using 5 ml columns packed with Chelating Sepharose™ Fast Flow (2.1.7) as described in section 2.2.3.4. **A-C:** 15% SDS-PAGE analysis (2.2.3.1) of IMAC fractions of Mth212/D151N (**A**), Mth212/D151A (**B**) and Mth212/D151S (**C**). **M:** Molecular weight marker (Fermentas, 2.1.5). **CE:** Clarified *E.coli* BL21_UX lysate containing Mth212/D151 variants prepared by sonication/homogenization as described in 2.2.3.3. **FT:** Flow-through fraction (cell lysate eluted from IMAC column). **1W** and **2W:** First and second column wash fractions. **30-500:** Protein elution with 80 ml step gradient of 30-500 mM imidazole in IMAC wash buffer (2.1.9), where 1 column volume comprised 5 ml. Brackets indicate fractions pooled for subsequent heparin affinity chromatography (2.2.3.5).

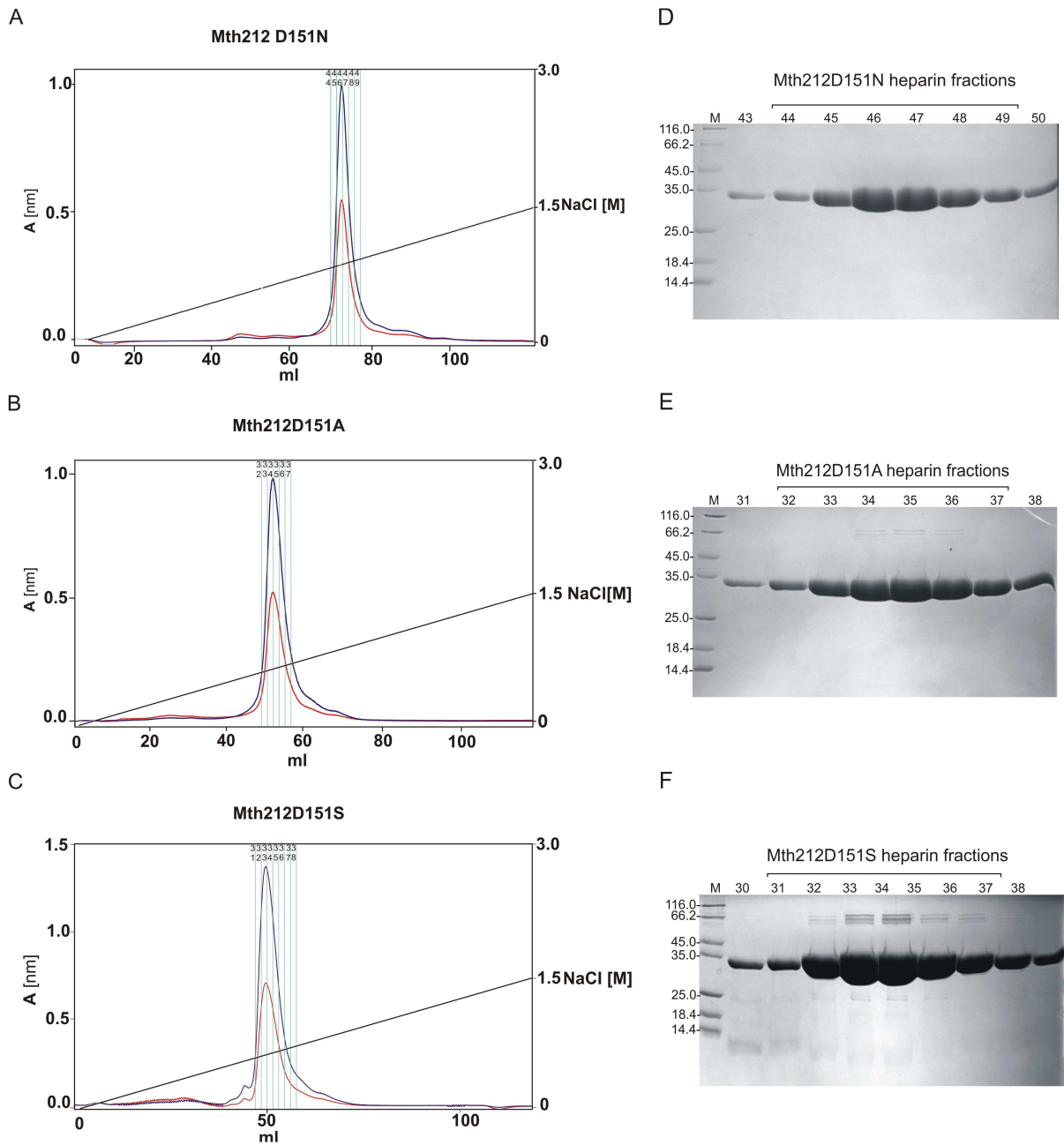


Figure 13: Heparin affinity chromatography of Mth212 mutants.

A-C: Heparin affinity chromatography elution pattern of Mth212/D151N (**A**), Mth212/D151A (**B**) and Mth212/D151S (**C**) (for details see section 2.2.3.5). Heparin chromatography was performed on BioCAD™ Workstation (Applied Biosystems) and POROS® HS 20 µm column (10 mm x 100 mm, 7.89 ml) at a 4 ml/min flow rate. Left ordinate: absorption at 260 nm (red) and 280 nm (blue); Right ordinate: concentration of NaCl in mol [M]; Abscissa: elution volume in ml. Green lines and numbers above the chromatogram indicate fractions that were analysed by 15% SDS-PAGE (2.2.3.1) shown in panels **D-F**. **M:** Protein molecular weight marker (2.1.5). Brackets indicate fractions pooled and concentrated (2.2.3.8) for subsequent enzymatic characterization.

All three heterologous produced mutant proteins were fully soluble in *E.coli* and had chromatographic properties and electrophoresis mobility parameters similar to those of wild-type (wt) Mth212 (for the details see (Georg et al., 2006)).

3.1.3 Endonuclease assays with Mth212/D151 variants

In order to study the effect of Asp-151 substitution on the enzymatic activity, Mth212/D151 mutant variants were analysed by endonuclease assay. The assay was made as described in 2.2.3.9 using 40-mer double-stranded DNA oligonucleotide substrates containing an AP/G (where AP is an abasic site analogue, Figure 14B) or a U/G mismatch at position 24. One of two DNA strands in substrate duplex was labelled with fluorescein group on its 5' end. A schematic representation of assay substrate is shown in Figure 14A. Reaction products obtained during endonuclease assays were analysed under denaturing conditions by 11% A.L.F.-polyacrylamid gel (A.L.F.-PAGE, see section 2.2.2.8) (Figure 14C, D, E and F).

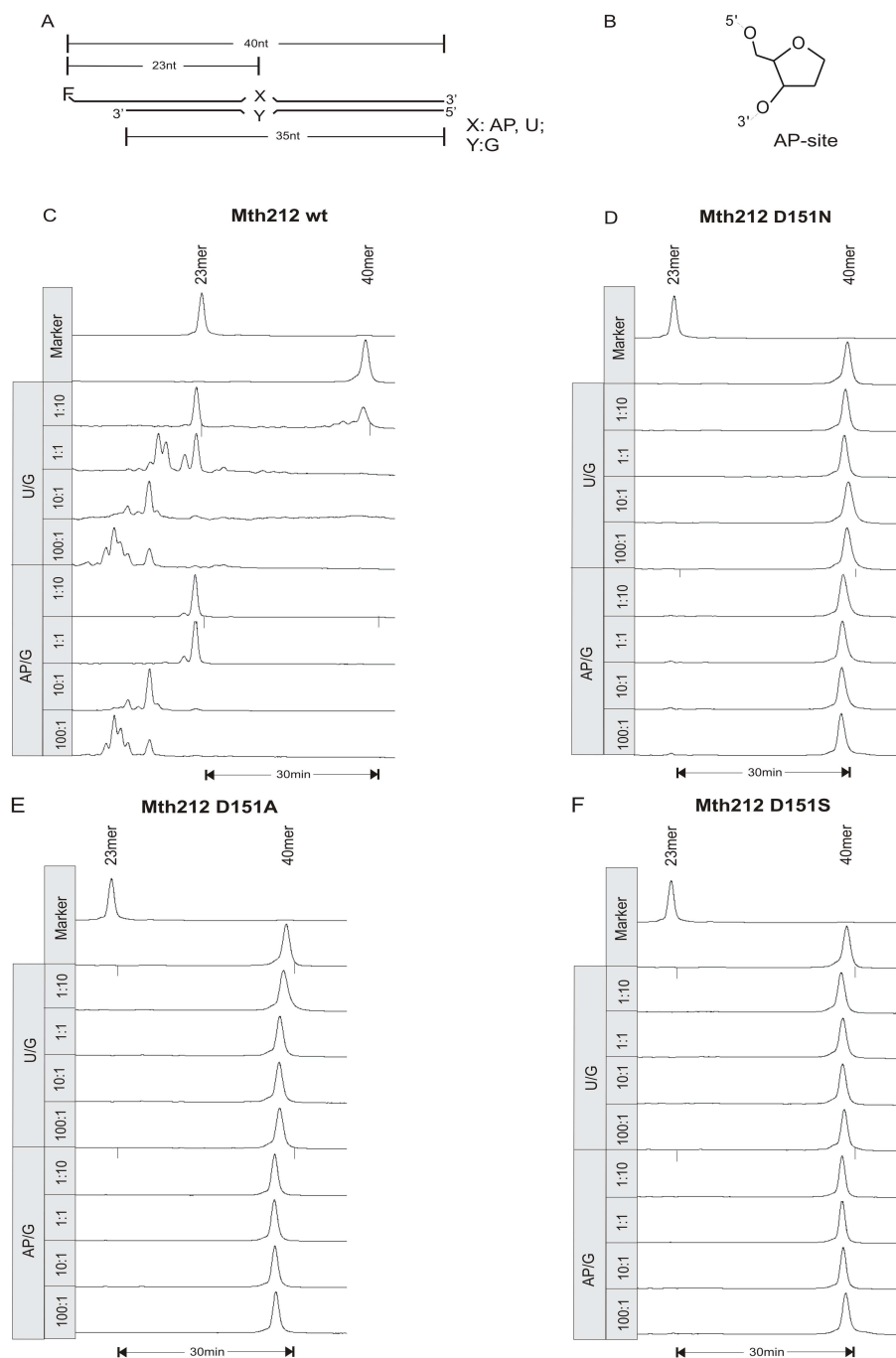


Figure 14: Endonuclease assay with wt Mth212 and Mth212/D151 mutants

Endonuclease assays were carried out with substrate containing an AP-site or a U/G mismatch at position 24 and conducted as described in 2.2.3.9. **A:** A schematic representation of substrate and expected length of reaction products (F: fluorescein, nt: nucleotides, AP: a stable AP-site model residue as shown in panel **B**). Prince-U, 40_PRI_AP and Prince-G oligonucleotides were used for substrate preparation (sections 2.1.4 and 2.2.2.13). **C-F:** 11% A.L.F-PAGE analysis (2.2.2.8) of endonucleolytic assays performed with AP/G- and U/G-containing substrates and wt Mth212 (**C**), Mth212D151N (**D**), Mth212D151A (**E**) and Mth212D151S (**F**). 0.12 pmol of each substrate was incubated with different amounts of purified enzymes for 15 min at 65°C in 50 µl of endo nuclease buffer (2.1.9). Molar enzyme/substrate ratios are indicated in gray bars. **30 min:** Gel running time difference between marker oligonucleotides.

Wild-type Mth212 showed strong activity with both AP/G and U/G specific substrates (Figure 14C). This confirms our previously published data (Georg et al., 2006). The product peak

with 23-mer marker electrophoretic mobility and a series of additional peaks shorter than 23-mer product were clearly detectable. Additional peaks were considered as a result of further degradation of main 23-mer products by Mth212 3'→5' exonucleolytic activity (Figure 14C). In contrast to wt Mth212, Mth212/D151 mutants did not yield 23-mer products neither with AP/G nor with U/G mismatch-containing oligonucleotide substrates (Figure 14D, E and F). It has to be noted, that small peak corresponding to 23-mer marker was observed in case of D151N mutant and AP-site-containing oligonucleotides at 100:1 enzyme/substrate ratio. This could result from Mth212/D151N remaining activity and/or from wt Mth212 that might have been produced by spontaneous revertants (see section 3.1.2).

Mth212 mutational analysis revealed that D-151 by analogy with Ape1/D-210 (Rothwell et al., 2000) is required for Mth212 enzymatic catalysis. The fact that Asp-151 substitution eliminated both AP-endonuclease and U-endonuclease activities indicates that both enzymatic activities must originate from the same active site. However, the presence of a single active site in Mth212 can only be proved by still ongoing structural studies.

It is known that an amino acid substitution may affect a specific catalytic function not only due to the lack of catalytically essential functional group, but also through impact on the global structural features, leading to a complete loss of enzymatic activity. To provide additional data to substantiate this point gel-retardation assays were performed

3.1.4 Gel-retardation assays with Mth212/D151 variants

The activity assay presented in section 3.1.3 revealed that all three Mth212/D151 mutants were completely deficient in endonuclease activity. To analyse the effects of Asp-151 substitution on the overall protein structure perturbations and substrate binding properties the gel-retardation assays were performed using purified D-151 mutants and double stranded DNA oligonucleotides. The Following issues were addressed:

- 1) Substrate binding efficiency of Mth212/ D151 variants
- 2) Binding of specific AP-site- and dU-containing substrates
- 3) Substrate binding strengths in the presence of unspecific competitor DNA
- 4) Binding to substrate's ends (3'recessed, 3'protruding or blunt)

3.1.4.1 EMSA with Mth212/D151N and Substrate I

To assess the potential impact of Asp-151 substitution on Mth212 substrate binding, gel-retardation assays (electrophoretic mobility shift assays (EMSAs) were performed. EMSAs were carried out as described in section 2.2.2.14, using constant amounts of unlabelled double-stranded (ds) DNA substrate, namely Substrate I (1xblunt and 1x3' recessed ends, Figure 15A) containing an AP-site analog or a U/G mismatch at position 24 (2.1.4) and

increasing amounts of purified Mth212/D151N mutant. C/G substrate containing a normal Watson-Crick base pair at position 24 was used as a control (Figure 15).

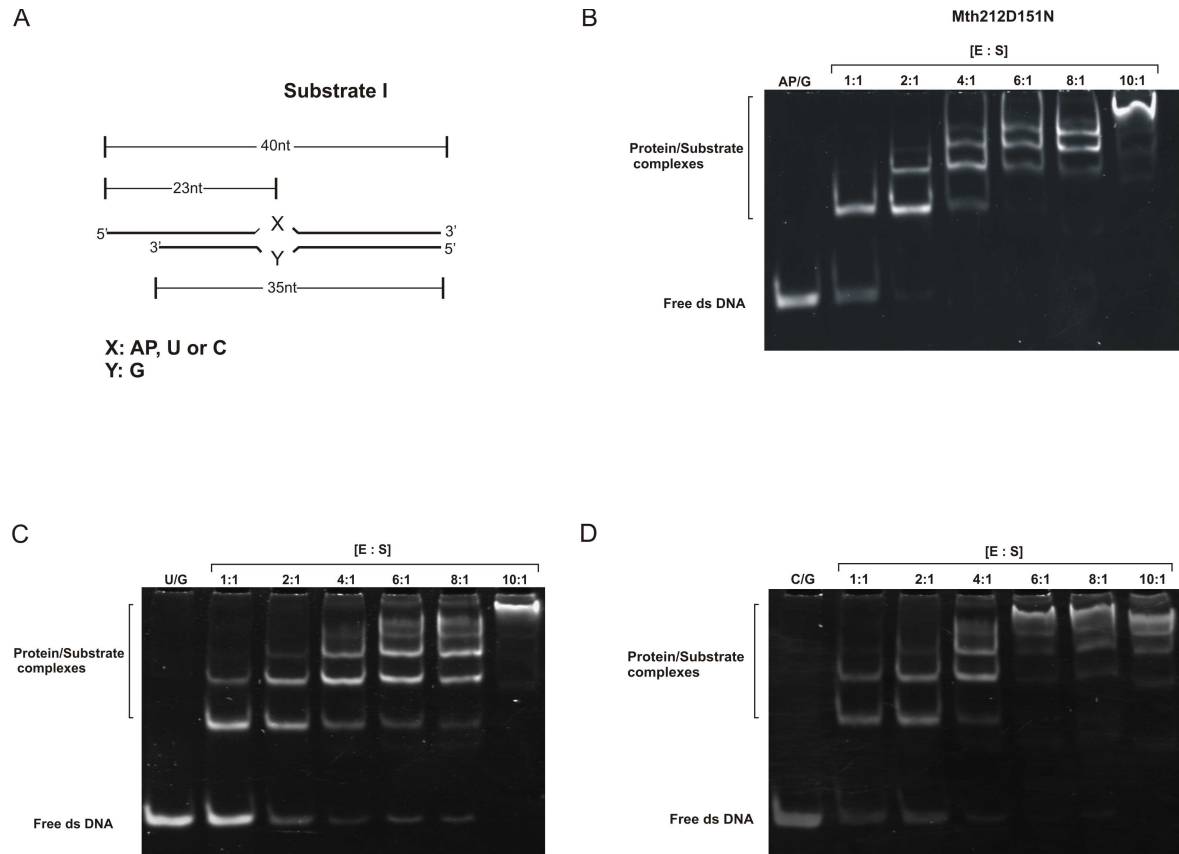


Figure 15: Mth212/D151N substrate binding test

A: A schematic representation of Substrate I used in band-shift assays shown in panels **B-D** (AP: a stable AP-site (for details see Figure 14), nt: nucleotides)). **B-D:** 10% TE-PAGE analysis of Electrophoretic Mobility Shift Assays (EMSA). EMSAs were performed as described in section 2.2.2.14 using Mth212/D151N and Substrate I containing an AP-site (**B**), a U/G mismatch (**C**) or a normal C/G base pair (**D**). Enzyme/substrate ratios are indicated above the gel tracks. **AP/G, U/G or C/G:** Substrate I containing an AP/G or a U/G misspairs or a normal C/G base pair incubated in EMSA buffer (2.1.9) without an enzyme and applied onto the gel as a control. Positions of bands corresponding to free ds DNA and protein/substrate complexes are indicated on the left.

When interpreting gel-retardation results, it is important to keep in mind that all assays were done with short DNA oligonucleotides, which elevate the probability of [Enzyme/DNA-end] complex formation. Also, all EMSA samples were incubated for 5 min with D151N mutant protein and immediately applied onto the TB-PAGE (for details see section 2.2.2.14). Moreover, use of catalytically inactive Mth212/D151N mutant in binding assays imposed its own restrictions on extrapolation of results on wt Mth212 behaviour.

Gel-retardation assays presented in Figure 15 revealed that D151N variant was proficient in binding of specific AP-site- and U/G mismatch-containing substrates. Mth212/D151N mutant protein was found to retain its substrate binding property, confirming that substitution of conserved Asp-151 in Mth212 analogically to Asp-210 in human Ape1

(Rothwell et al., 2000) does not influence substrate binding. Based on this result it was assumed that Mth212/D151N protein was properly folded and could be used for further functional and structural analysis.

The binding efficiency of Mth212/D151N to AP-site-containing substrate was obviously higher than to U/G substrate. At 4:1 enzyme/substrate ratio, almost all free AP-site-containing substrate was bound, while unbound free DNA bands were still present in case of U/G substrate. (Figure 15B, Lane 4:1). Comparing the gel shift patterns obtained with AP/G and U/G substrates, it was noticed that number of shift-bands was different even at 1:1 enzyme/substrate ratio (Figure 15B and C, Lanes 1:1). In case of AP-site-containing substrate, only one shift-band corresponding to protein/substrate complex was clear to detect, while U/G substrate gave a second shift-band with higher retention properties at the same enzyme/substrate ratio. The first shift-band could result from a 1:1 enzyme/substrate binding or mono binding, whereas the second and subsequent ones from a 1:X binding (where X= 2, 3...) or multiple binding. Similar shift-band patterns were identified in case of specific U/G mismatch-containing substrate and C/G duplex-containing substrates, indicating that D151N binds C/G duplex as strong as specific U/G substrate (Figure 15C and D). For better discrimination between specific U/G substrate and unspecific C/G duplex, the following gel-retardation assays were performed in presence of unspecific competitor DNA (see sections 3.1.4.2 and 3.1.4.4). The unspecific competitor DNA was expected to compete for the binding with substrate DNA oligonucleotides eliminating shifted bands.

3.1.4.2 EMSA with Mth212/D151N, Substrate I and pET-vector as a competitor DNA

Mth212/D151N binding efficiency to AP/G and U/G mismatch-containing DNA oligonucleotides was assessed by competitive gel mobility shift assays (EMSAs). The entire pET-vector (~5.4kb) was used as a competitor DNA. EMSAs were carried out as described in section 2.2.2.14, using Substrate I containing an AP-site or a U/G mismatch. Substrate oligonucleotides were pre-incubated with competitor DNA followed by addition of purified Mth212/D151N. The enzyme/substrate ratio used here was 4:1, at which almost all substrate DNA was bound (see Figure 15, Lane 4:1). The amount of competitor DNA added to the samples was calculated as molar equivalents of nucleotide relative to substrate correspondingly (Figure 16).

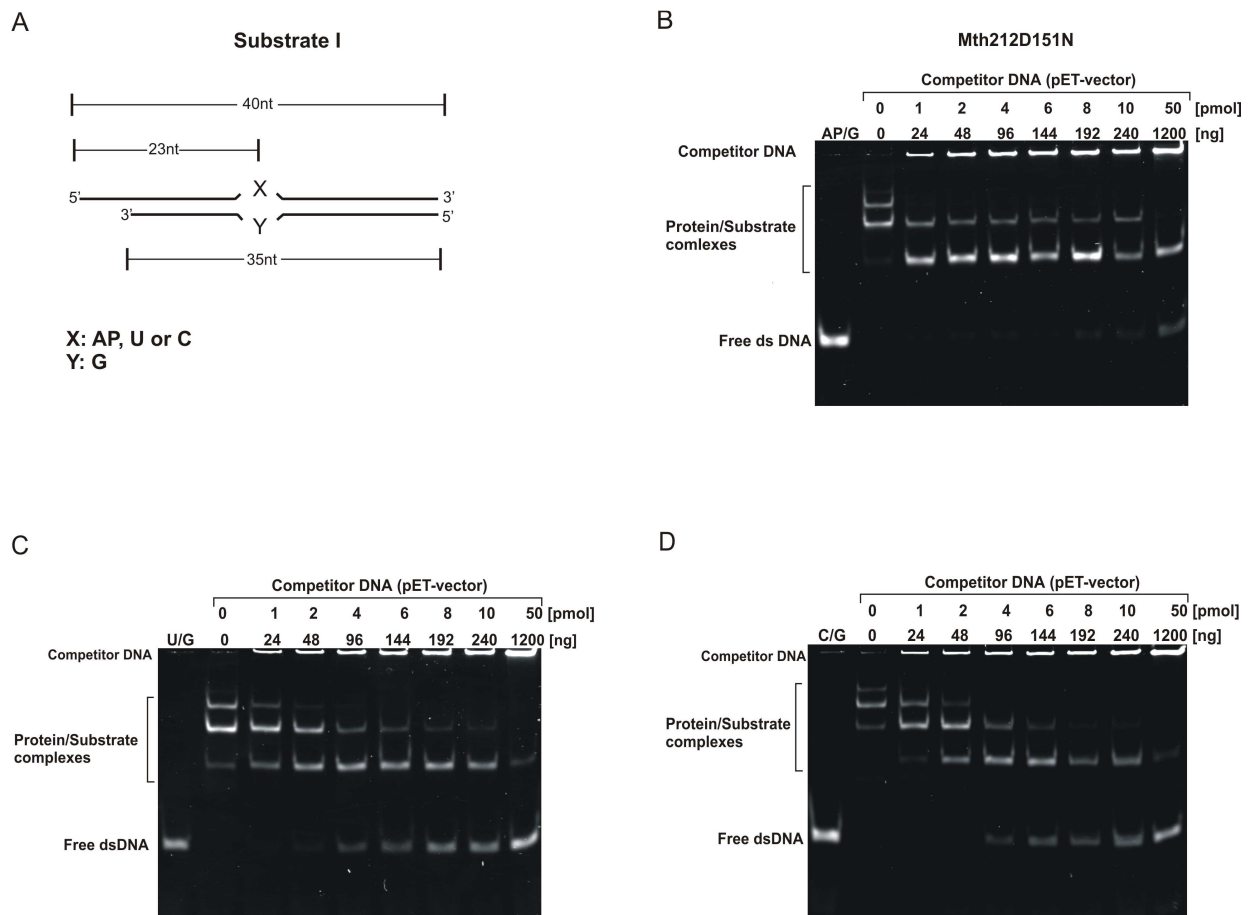


Figure 16: Mth212/D151N gel-retardation assay with Substrate I and pET vector as a competitor DNA

A: A schematic representation of Substrate I used in band-shift assays shown in panels **B-D** (AP: a stable AP-site, nt: nucleotides). **B-D:** 10% TE-PAGE analysis of Electrophoretic Mobility Shift Assays (EMSA). EMSAs were conducted as described in section 2.2.2.14 using 4 pmol of purified Mth212/D151N, 1 pmol (24 ng) of Substrate I containing an AP-site (**B**), a U/G mismatch (**C**) or a normal C/G base pair (**D**) and increasing amounts of competitor DNA (entire pET-B001 vector, 5.7kb). Amounts of competitor DNA were calculated as nucleotide equivalents relative to substrate oligonucleotides and indicated above the gel tracks in pmol. Positions of bands corresponding to competitor DNA, free dsDNA and protein/substrate complexes are indicated on the left.

The data presented in Figure 16B indicate strong specific binding of AP-site-containing substrate by Mth212/D151N mutant, where small amounts of unbound free substrate DNA appeared first at 8-fold excess (192 ng) of competitor DNA. Moreover, in presence of 50-fold excess (1200 ng) of competitor DNA the specific shift-band (first shift-band from the bottom) was still very intensive. U/G mismatch-containing substrate was bound by Mth212/D151N with relatively high efficiency, but weaker than AP-site-containing substrate. Here bands corresponding to free dsDNA appeared already at 2-fold excess (48 ng) of competitor DNA. Unexpectedly, competitive gel-retardation assays revealed that D151N mutant bound specific U/G mismatch-containing substrate insignificantly stronger than control C/G duplex. This result was reproducible, namely U/G- and C/G-containing oligonucleotides gave almost similar gel-retardation patterns (Figure 16C and D). The other two Mth212 mutants, namely

D151S and D151A (see section 3.1.2) were also tested for their binding properties with respect to U/G mismatch-containing substrate in comparison to D151N mutant.

3.1.4.3 EMSA with Mth212/D151A and Mth212/D151S, Substrate I and pET-vector as a competitor DNA

First round of gel-retardation analysis done with D151N mutant protein showed no difference in binding of specific U/G substrate and C/G-containing duplex. To determine which one from three Aps-151 mutants displays the highest affinity to U/G substrate for possibly better discrimination between U/G substrate and C/G duplex, a new series of band-shift assays were carried out. The assays were performed largely as described in section 3.1.4.2, using U/G mismatch-containing Substrate I and purified D151N, D151A and D151S mutant proteins (see section 3.1.2 and Figure 17).

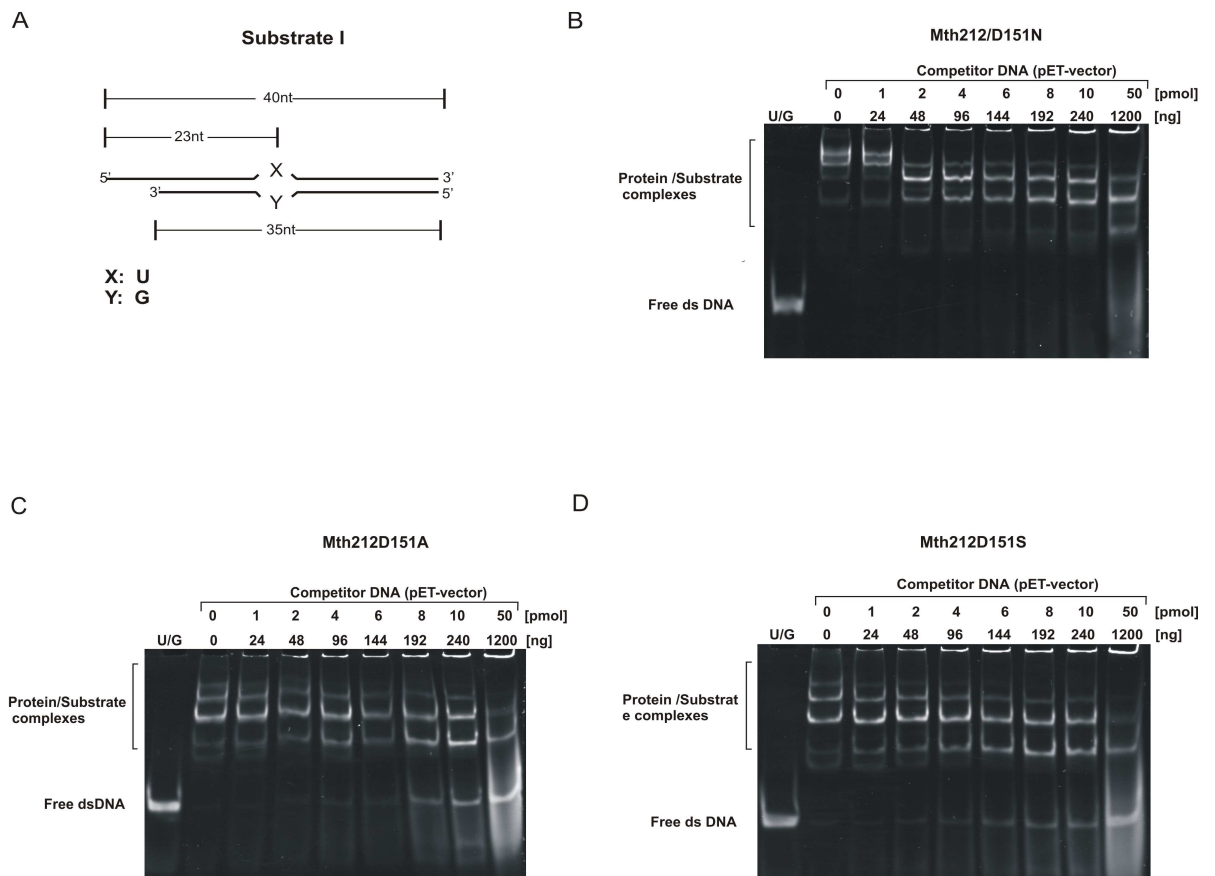


Figure 17: Mth212D151N, Mth212D151S and Mth212D151A gel-retardation assays with U/G mismatch-containing Substrate I and pET vector as a competitor DNA

A: A schematic representation of Substrate I used in band-shift assays shown in panels **B-D** (AP: a stable AP-site, nt: nucleotides). **B-D:** 10% TE-PAGE analysis of Electrophoretic Mobility Shift Assays (EMSA). EMSAs were performed as described in 2.2.2.14 and Figure 16, using Substrate I containing a U/G mismatch, increasing amounts of competitor DNA (entire pET_B001 vector, 5.7 kb) and purified Mth212D/151N (**B**), Mth212/D151A (**C**) and Mth212/D151S (**D**). Amounts of competitor DNA were calculated as nucleotide equivalents relative to substrate oligonucleotides and indicated above the gel tracks in ng and in pmol.

However, almost no difference in binding strength of D151A and D151S to the specific U/G substrate was found (Figure 17C and D). Mth212/D151N exhibited the highest binding efficiency to U/G substrate compared to other two Asp-151 mutants (Figure 17B, C, and D, Lane: 240 ng). Therefore, Mth212/D151N was used for further binding experiments and X-ray crystallographic analysis.

Some difficulties were encountered in interpretation of the multiple discrete shift-bands observed in samples with high protein concentration. Similar pattern of progressive gel-retardation could be interpreted as a multiple enzyme/substrate-binding (see section 3.1.4.1), where presumably more than one protein molecule is bound to the substrate oligonucleotide. Indeed, as an ExoIII homologue, Mth212 displays strong 3'→5' exonuclease activity on 3' recessed ends and also has high affinity to blunt DNA ends (Pfeifer and Greiner-Stoffele, 2005). Additionally the oligonucleotide substrate used in the first round of gel-retardation experiments (Figures 16A and 17A) had one 3' recessed and one blunt end. All these admit the assumption that catalytically inactive Mth212 with retained substrate binding properties is able to bind substrate ends as well. If indeed Mth212/D151N possessed high affinity to the substrate ends, the circular plasmid (pET-vector) with no ends would not be able to effectively compete for binding short substrate DNA oligonucleotides.

To test this hypothesis, two experimental parameters were changed during next round of binding assays: 1) a competitor DNA (short 20-mer DNA oligonucleotides with sequence context identical to that in substrate oligonucleotides) and 2) the structure of substrate ends (2x blunt or 2x3' recessed ends).

3.1.4.4 EMSA with Mth212/D151N, Substrate II and 20-mer oligonucleotides as a competitor DNA

Multiple shift-bands observed in binding assays presented in sections 3.1.4.3 suggest that Mth212/D151N can bind DNA oligonucleotide substrates not only at the mismatch position, but also at their ends, resulting in multiple enzyme/substrate complexes. Such strong DNA-ends affinity was expected to be better competed by short 20-mer ds DNA oligonucleotides. Obviously, at equal amounts of ds DNA oligonucleotides, 20-mer competitors should have 2-fold excess of ends compared to 40-mer substrate oligonucleotides. In case these 20-mer oligonucleotides better compete for binding to DNA ends, one should expect to see one or two reproducible shift-bands even at high amount of competitor DNA.

To substantiate this point, a second round of binding assays was carried out as described in section 3.1.4.1 and 2.2.2.14, using Substrate II (a 40-mer blunt dsDNA oligonucleotides) containing an AP-site or a U/G mismatch (Figure 18A), Mth212/D151N and 20-mer blunt oligonucleotides as a competitor DNA. The amount of competitor DNA added

was calculated as molar equivalents of DNA ends relative to the substrate ends accordingly. To exclude false interpretation of shift-bands observed, 20-mer oligonucleotides were incubated with mutant protein and applied onto the gel as a control (Figure 18).

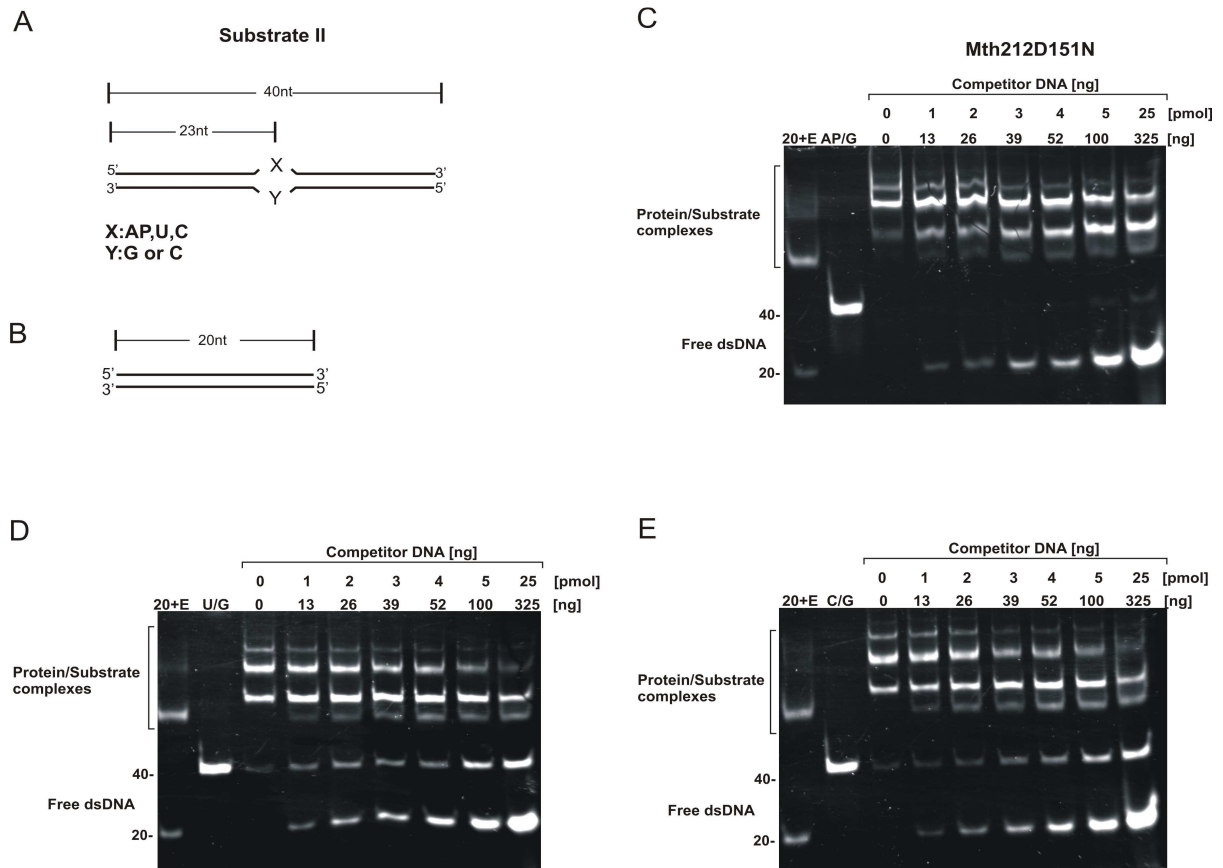


Figure 18: Mth212/D151N gel-retardation assay with Substrate II and 20-mer oligonucleotides as a competitor DNA

A-B: A schematic representation of Substrate II and 20-mer ds competitor DNA used in band-shift assays shown in panels **C-E** (AP: a stable AP-site, nt: nucleotides). **C-E:** 10% TE-PAGE analysis of band-shift assays. EMSAs were carried out as described in section 2.2.2.14 and Figure 16, using purified Mth212/D151N, Substrate II containing an AP-site (**C**), a U/G mismatch (**D**) or U/C base pairs (**E**) and increasing amounts of competitor DNA. **20+E:** 4 pmol of ds 20-mer competitors DNA were incubated with 4 pmol of Mth212/D151N for 5 min at 65°C in 15 µl of EMSA buffer (2.1.9) and applied onto the gel.

In binding assays with Mth212/D151N, Substrate II (2x blunt ends) and 20-mer competitor DNA, three shift-bands corresponding to enzyme/substrate complex were observed. It should be noted that shift-bands resulted from competitor-DNA/protein complex were ignored. Control samples made without competitor DNA (Figure 18, Lane 0) resulted in shift-bands of different thickness and intensity. In case of AP/G substrate, the second shift-band from the bottom was most intensive, while in case of U/G substrate both bands the first and the second one were equally intensive. Furthermore, comparison of Substrate I with Substrate II (compare Figure 16 and Figure 18, Lanes 0) suggested that Mth212/D151N bound Substrate II mostly in form of multi-enzyme/substrate complex.

Assays with Mth212/D151N, blunt U/G substrate and short competitor DNA revealed again no difference between C/G duplex and specific U/G substrate binding (Figure 18D and E). Consistent specific binding of Mth212/D151N to AP/G substrate was observed, where free dsDNA band appeared only at 25x molar excess of competitor DNA (325ng). In case of U/G substrate, free dsDNA bands appeared even at 1x molar excess of competitor DNA. It was also found that increasing amounts of competitor DNA lead to slow reduction of band intensity and eventually to elimination of the third and the second shift-bands, while the first shift-band became more intensive. These results confirm our hypothesis that short oligonucleotides compete more effectively for protein binding than coiled plasmid. At the same time presence of two intensive shift-bands even at high amount of competitor DNA provide evidence for strong specific binding of Mth212/D151N to the blunt ends.

It was experimentally found that wt Mth212 has no exonucleolytic activity on single stranded DNA or on DNA duplexes with 3' protruding ends (Svetlana Ber, this department, unpublished). If the structure of substrate ends influences Mth212/D151N binding strength, it can be assumed that 3' overhangs would prevent to some extent formation of protein/DNA-end complex (an *exo*-complex). To verify this assumption, a new substrate, namely Substrate III possessing two 3' protruding ends, was synthesised and used in the next series of binding assays.

3.1.4.5 EMSA with Mth212/D151N, Substrate III and 20-mer oligonucleotides as a competitor DNA

Previous results suggest that Mth212/D151N mutant possesses high affinity to the blunt substrate ends (section 3.1.4.4). This presumably caused the appearance of additional shift-bands making it more complicated to discriminate between specific U/G substrate and C/G duplex and to interpret the results. In order to decrease affinity of Mth212/D151N to the blunt ends, the third round of binding assays was made, using Substrate III (40-mer dsDNA with 5-nucleotide long 3' overhangs on both ends) containing an AP-site or a U/G mismatch (Figure 19A).

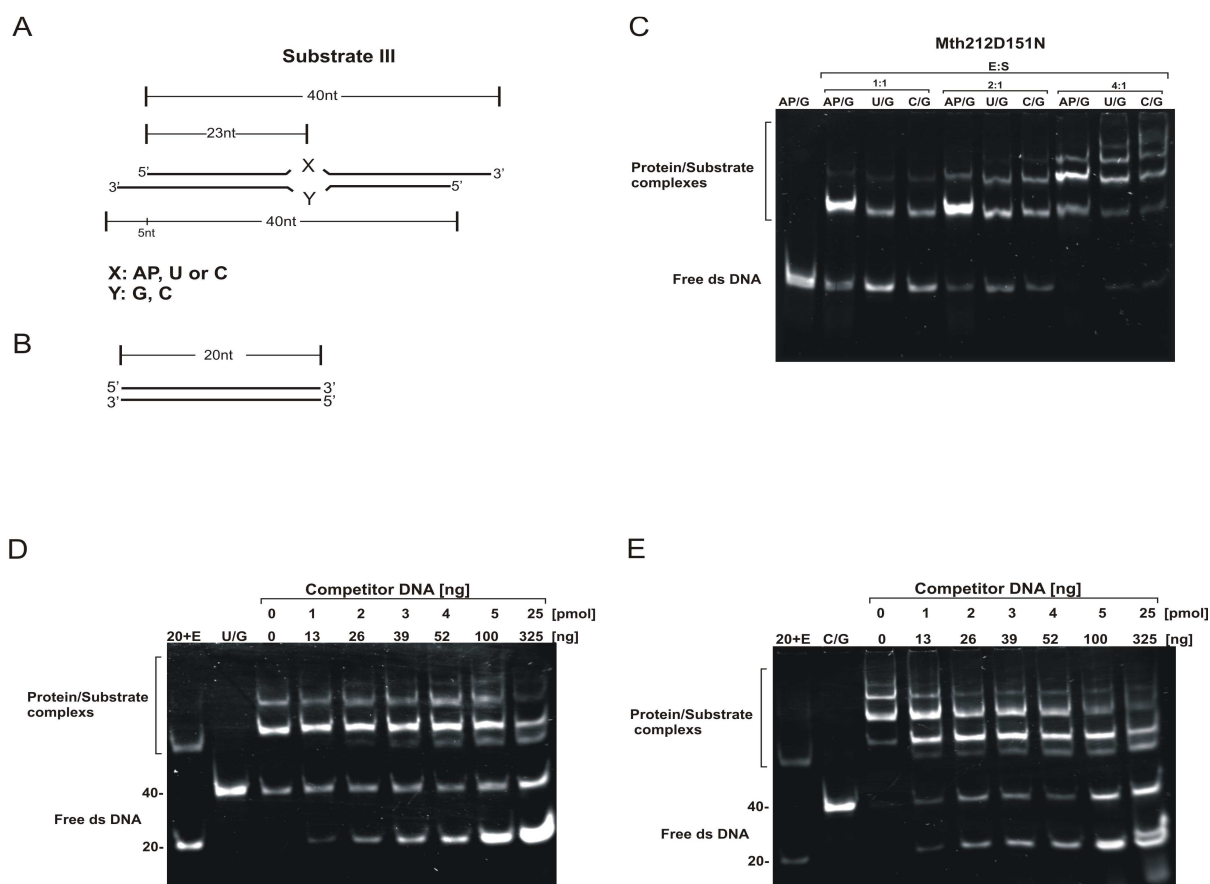


Figure 19: Mth212D151N gel-retardation assay with Substrate III and 20-mer oligonucleotides as a competitor DNA

A-B: A schematic representation of Substrate III and 20-mer ds competitor DNA used in band-shift assays shown in panels C-E (AP: a stable AP-site, nt: nucleotides). **C-E:** 10% TE-PAGE analysis of band-shift assays. **C:** Substrate binding assays were conducted as described in section 2.2.2.14 using 1 pmol (26 ng) of Substrate III containing an AP-site, a U/G mismatch or a C/G base pair and increasing amounts of Mth212/D151N. **D and E:** EMSAs were carried out as described in section 2.2.2.14 and Figure 16, using purified Mth212/D151N, Substrate III containing a U/G mismatch (**D**) or U/C base opposition (**E**) and increasing amounts of competitor DNA. The amount of competitor DNA added was calculated as molar equivalents of DNA ends relative to the substrate and indicated above the gel tracks in ng and in pmol. **20+E:** 4 pmol of double-stranded 20-mer competitors DNA were incubated with 4 pmol of Mth212/D151N for 5 min at 65°C in 15 µl of EMSA buffer (2.1.9) and applied onto the gel.

The third round of binding assays with Substrate III showed strong specific binding of AP/G substrate by D151N mutant (Figure 19C). Since this result was reproducible, only U/G substrate and C/G control duplex were compared in the following series of gel-retardation assays. Competitor DNA effectively displaced Substrate III even at 1x molar excess resulting in appearance of free ds DNA bands (Figures 19D and E, Lane 1). When comparing band-shift assays done with all three substrate types (I, II and III), Mth212/D151N was found to bind Substrate III (2x3' protruding ends) less efficient than Substrate I and Substrate II (Figures 16A and 18A). These results confirm the assumption that 5-nucleotide long 3'protruding end can be indeed the obstacle for an *exo*-Enzyme/Substrate complex formation. However, Mth212/D151N affinity to the specific U/G substrate possessing 2x3' protruding ends was again not significantly stronger than to the control C/G duplex.

Taken together, these results demonstrate that D151N binds U/G-mismatch-containing substrate weaker than AP/G substrate and DNA ends. At the same time, D151N binds control C/G duplex and specific U/G substrate with same efficiency at least under these experimental conditions. Binding assays with AP/G-, U/G- substrates and C/G duplex suggest that binding strengths between Mth212/D151N and short dsDNA substrates decrease in following order: AP/G > DNA ends [3' recessed or blunt ends] > U/G.

3.1.4.6 EMSA with Mth212/D151N and Substrate III (with U/C, U/T, U/A) and 20-mer oligonucleotides as a competitor DNA

To shed light on Mth212 specificity of DNA-U recognition and binding energetics, further EMSA experiments were done with DNA duplexes containing following non-canonical base oppositions: U/A, U/T and U/C. Gel-retardation analyses were done as described in section 3.1.4.5, using purified D151N and Substrate III (2x3' protruding ends) at 4:1 substrate to enzyme ratio (Figure 20).

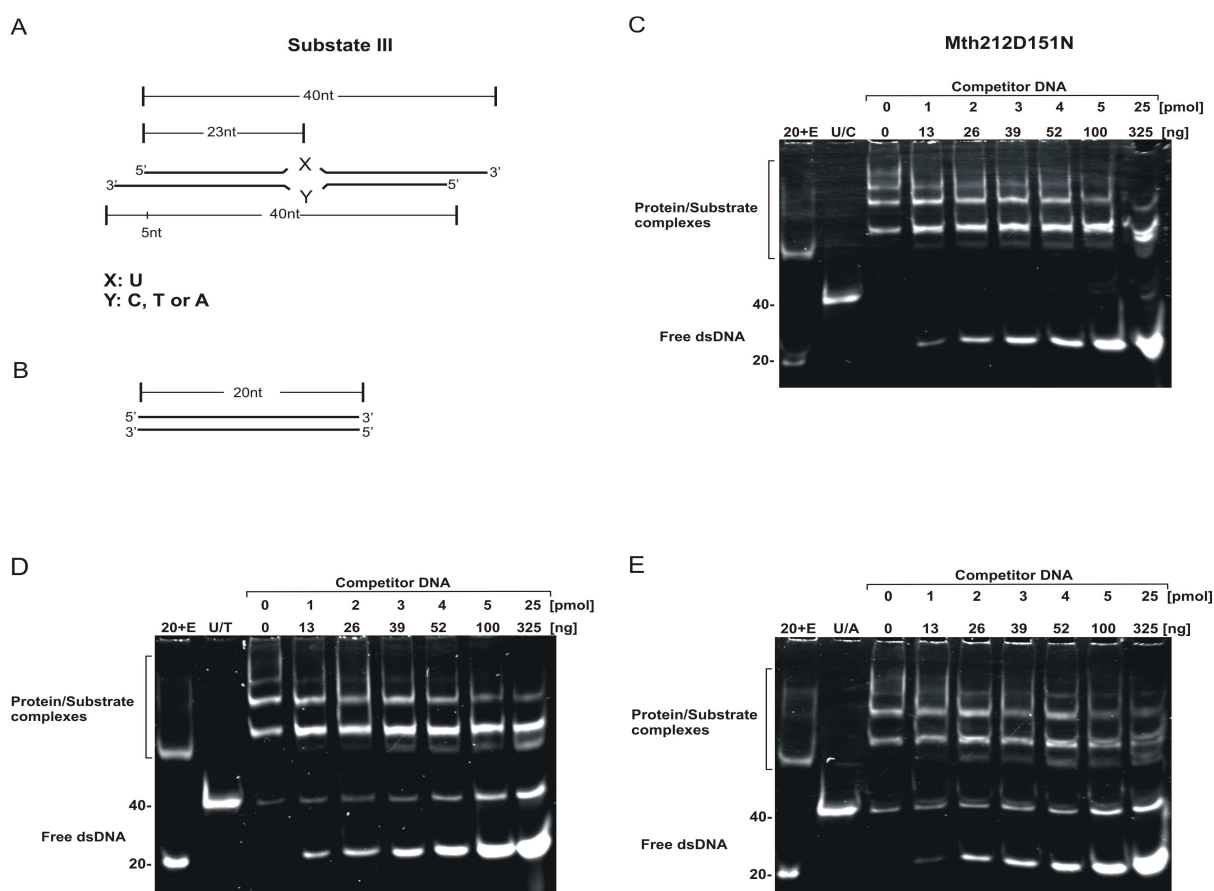


Figure 20: Mth212/D151N gel-retardation assay with Substrate III (U/C, U/T and U/A) and 20-mer oligonucleotides as a competitor DNA

A and B: A schematic representation of Substrate III and 20-mer ds competitor DNA used in band-shift assays shown in panels **C-E** (AP: a stable AP-site, nt: nucleotides). **C-E:** 10% TE-PAGE analysis of binding assays. EMSAs were carried out as described in section 2.2.2.14 using 4 pmol of purified Mth212D151N, 1 pmol of Substrate III containing U/C (**C**), U/T (**D**) or U/A (**E**) base oppositions and increasing amounts of competitor DNA. **20+E:** 4 pmol of ds 20-mer competitor DNA were incubated with 4 pmol of Mth212D151N for 5 min at 65°C in 15 µl of EMSA buffer (2.19) and applied onto the gel.

Binding assays with U/A, U/T and U/C base oppositions showed that D151N variant bound U/C-containing duplex obviously stronger than U/A duplex and surprisingly stronger than U/T. Comparing these results with previous gel-retardation data, it was observed that U/C duplex was bound by D151N almost as strong as AP-site-containing substrate (compare Figure 18C with Figure 20C). In both cases free substrate DNA band was detected only after addition of high amounts of competitor DNA (100-325 ng). Furthermore, U/T- and U/A-containing duplexes were bound by D151N with almost same affinity as U/G substrate (compare Figure 18D with and 20D, E). It should be noted that gel-retardation assays with U/T, U/A and U/C were repeated at least 3 times with each base opposition and revealed consistently strong binding of D151N to U/C duplex. Unlike U/C-containing duplex, U/A- and U/T-containing duplexes were bound by D151N obviously weaker resulting in similar gel-retardation pattern. Based on strong binding affinity of D151N to AP-site-containing substrate and U/C-containing duplex, it was decided that both AP/G and U/C oppositions would be the

proper candidates for co-crystallization with Mth212/D151N and subsequent structural analysis of Protein-DNA complexes (for details see section 3.3).

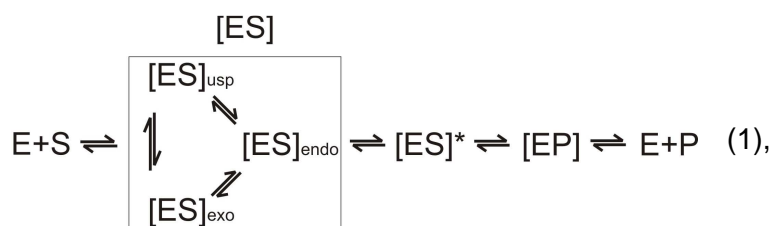
The observed difference in Mth212/D151N substrate binding strengths can be explained by the assumption that Mth212, like many other mismatch DNA-repair enzymes, is equipped with base-flipping mechanism for mismatch recognition and catalysis (Mol et al., 1999). This mechanism involves two stages. At the first stage the enzyme binds DNA non-specifically building a transient protein-DNA complex. Thereafter the enzyme begins to slide along the DNA strand exploring the local helix instabilities resulting from base damage or base mispair accompanied with base flipping until the target mismatch is detected. Upon encountering the target base, the transient non-specific complex is converted into a catalytically competent or enzymatically productive complex where the target base is extruded from the DNA base stack into the substrate-binding pocket. This second stage is characterised by strong specific binding of the enzyme to the mismatch prior to the catalytic cut (Parikh et al., 1998). The observed D151N substrate binding strength (with exception of U/T) seem to hint a flipped-out conformation of the 2'dU residue being crucial for making a stable Enzyme/Substrate complex.

It has already been found that various U/X base oppositions (where X is any of nucleobases) differ in their base pairing properties. Hence, in the context of a DNA double helix, different U/X base oppositions are expected to require different energy input for dU flipping to occur. The expected order of energy request for extruding the target uracil from DNA base stack is: U/A>U/G>U/C (Parikh et al., 1998; Pingfang Liu et al., 2008; V. Starcuvienne and H.J. Fritz, 2002). Indeed this order is reflected in Mth212/D151N binding assays except for U/C, which was bound by D151N stronger than U/T (Figure 20). This phenomenon cannot be explained by proposed base-flipping mechanism, since U/C and U/T duplexes were expected to have similar base pairing properties. This particular case led to the second assumption that Mth212 can interact with a nucleotide placed opposite to the target uracil. According to the binding assay results shown in Figure 20 interaction with cytosine opposite uracil was more stable than with thymine at least at these experimental conditions.

Based on the structure of ds DNA oligonucleotide substrates used in binding assays several *in vitro* substrate binding modes for Mth212 were proposed:

1. Binding to oligonucleotide ends, forming so-called *exo*-complex ($[ES]_{\text{exo}}$, see Equation 1);
2. Binding at a target mismatch, building an *endo*-complex ($[ES]_{\text{endo}}$), in which the target base is flipped-out towards the active site;
3. Binding in between a substrate end and a mismatch position, resulting in an unspecific complex ($[ES]_{\text{usp}}$) formation;

Given that all three [ES] complexes exist in reaction system in parallel and their correlation can be described as dynamic equilibrium, the following scenario for Mth212 *in vitro* catalysis can be considered:



Where:

E indicates an enzyme, **S** a substrate and **P** a reaction product;

$[\text{ES}]_{\text{usp}}$ indicates an unspecific enzyme/substrate complex;

$[\text{ES}]_{\text{endo}}$ indicates a specific enzyme/substrate complex in the *endo* binding mode (e.g. at dU residue);

$[\text{ES}]_{\text{exo}}$ indicates a specific enzyme/substrate complex in the *exo* binding mode (i.e. at physical DNA ends);

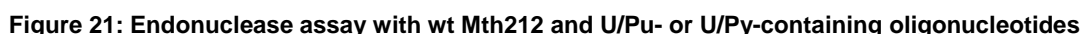
$[\text{ES}]^*$ indicates a transition state of enzyme/substrate complex (a high-energy intermediate or activated complex);

$[\text{EP}]$ indicates enzyme/product complex;

Given that transformation rate of $[\text{ES}]_{\text{endo}}$ into $[\text{EP}]$ complex is similar for all U/X base oppositions and according to Equation 1, the equilibrium of $[\text{ES}]_{\text{endo}}$ complex formation is expected to determine the overall reaction rate. If reaching the flipped-out conformation in productive $[\text{ES}]_{\text{endo}}$ complex is a rate-limiting factor, it would be expected that the order of substrate binding strengths in binding assays to be reflected in similar order of substrate cleavage rates in enzymatic activity assays. This hypothesis was tested by semi-quantitative determination of the substrate cleavage efficiency in enzymatic cleavage assays.

3.2 Semi-quantitative analysis of wt Mth212 substrate cleavage rate

In order to test if the gradual changes in substrate binding strength reflect the substrate cleavage rates, endonuclease assays with wt Mth212 and DNA duplexes containing U/A, U/T, U/C and U/G base oppositions were carried out. The assays were performed as described in 2.2.3.9 at 1:1 enzyme/substrate ratio. DNA duplexes were incubated with the enzyme for 5 and 10 min that correspond to the incubation times used in EMSAs (see section 2.2.2.14, Figure 21).



wt Mth212 was able to incise dsDNA duplexes containing uracil base oppositions (U/A, U/G, U/C, U/T) at position 24, which is consistent with our previously published data (Georg et al., 2006). An endonucleolytic cut next to DNA uracil residue was evidenced by appearance of product peak with 23-mer marker electrophoretic mobility (Figure 21B). However, a gradual change in substrate cleavage rates was observed. In case of U/C and U/T duplexes 23-mer product peaks were clearly detectable after 5 min incubation with wt Mth212, while U/G substrate was less efficiently degraded and there was no 23-mer substrate peak observed in case of U/A substrate. Even more pronounced difference in substrate cleavage rates was observed after 10 min incubation time. U/C substrate was almost completely cleaved to 23-mer intermediate products (compare height of U/C substrate peaks after 5 and 10 min, Figure 21B), which were further degraded by 3'→5' exonuclease activity of Mth212 (see series of peaks faster than 23-mer marker, Figure 21B).

U/T-containing duplex was slightly less degraded in 10 min than U/C, but obviously more efficiently than U/G substrate. U/A duplex remained almost undigested after 10 min incubation with the enzyme. Thus, the rates of recognition and processing of different U/X oppositions by Mth212 differed only at the early stage (i.e. mismatch recognition and base-flipping stages) of the enzymatic reaction, whereas after 20 min incubation time at 1:1 enzyme/substrate ratio Mth212 efficiently degraded all four uracil-containing oligonucleotides to their 23-mer intermediate products (Georg et al., 2006). These results support our hypothesis of a positive relationship between substrate binding strength and substrate cleavage rate (see section 3.1.4.6), in that the order of substrate cleavage rates by wt Mth212 reflects the order of substrate binding strengths by D151N, namely U/C>U/T>U/G>U/A.

Since in all DNA duplexes used in endonuclease assays U/X oppositions had identical sequence context, the observed difference in relative substrate cleavage rates suggests that the nature of nucleotide placed opposite the target uracil (i.e. a purine or a pyrimidine) affects Mth212 enzymatic catalysis as well. In confirmation of this, Mth212 processed uracil opposite a cytosine more efficiently than uracil opposite a guanine or adenine. Moreover, Mth212 processed even U/C and U/T oppositions slightly different (Figure 21).

Given that Mth212 recognises and cleaves specific substrate using nucleotide-flipping mechanism, DNA base-pairing properties in U/Pu or U/Py pairs were expected to affect the rate of enzymatic base flipping from DNA backbone and ultimately the overall enzymatic reaction rate. The obtained results confirmed this assumption in that U/Py base oppositions were more efficiently cleaved by wt Mth212 than U/Pu (Figure 21). Trends observed in wt Mth212 were similar to those found in other enzymes known or suspected to involve base-flipping mechanism for substrate recognition and catalysis. Representatives of UDG family enzymes, namely human UDG (Parikh et al., 1998), UNG and MUG from *Escherichia coli* (Pearl, 2000; Pingfang Liu et al., 2002; Pingfang Liu et al., 2008), TTUDGA and TTUDGB from *Thermus thermophilus* (V. Starcuvienne and H.-J. Fritz, 2002), and TDG from *M. thermautotrophicus* were tested for substrate selectivity. It was found that a uracil is less efficiently repaired from base pairs that approximate Watson-Crick configuration, e.g. a U/A, whereas it is more efficiently repaired from U/C or from the wobble U/G base pairs. Such mode of enzymatic activity was attributed primarily to the difference in base pairing properties of various U/X oppositions and to the energetic costs required for the extruding of the target uracil from DNA base stack. Hence, the observed decrease in wt Mth212 substrate cleavage rates (U/C>U/T>U/G>U/A), in correlation with the energy requirements, argues for the base flipping-out as a rate-limiting factor.

In conclusion, complete loss of enzymatic activity by Mth212 Asp-151 mutants

alongside with retention of substrate binding properties provides evidence for the unique catalytic site for both AP-endo and U-endonuclease activities. Mth212/D151N variant demonstrated the strongest binding affinity to AP-site-containing DNA that is in accordance with the data that wt Mth212 is the main AP-site-specific endonuclease of *M. thermautotrophicus* (Pfeifer and Greiner-Stoffele, 2005). Based on the gel-retardation analysis the following decrease in Mth212 substrate binding strength was proposed: AP-site > DNA ends [3'recessed or blunt ends] > U/Py > U/Pu. It seems that binding strength between Mth212 and its DNA substrate is determined by the mode of substrate binding, which ultimately determines the efficacy of enzymatic catalysis.

Differences in Mth212 activity (binding strength and cleavage rate) towards different U/X oppositions within identical sequence context implicate a putative base-flipping mechanism of Mth212 enzymatic catalysis. Given that Mth212 is equipped with base-flipping mechanism, potential rate-limiting factors of Mth212 enzymatic catalysis *in vitro* would be the following:

- Mode of substrate binding (see Equation 1);
- Stability of uracil in its stacked conformation in DNA duplex.

Concerning Mth212 mode of action, it is more likely that Mth212, similar to other damage-specific DNA repair enzymes (Mol et al., 1999), binds DNA distortion unspecifically building a non-productive transient complex and slides on DNA strand until the target mismatch is encountered.

To shed light on Mth212 active-site structure and mechanism of enzymatic catalysis, X-ray crystallographic analysis of wt Mth212 together with its catalytically inactive variant bound to DNA substrate was performed (see section 3.3).

3.3 Mth212 structure analysis

Mth212/D151N was over-expressed, purified as described in Materials and Methods (see section 2.2) and co-crystallised with a double-stranded DNA oligonucleotide substrate containing a U/G, U/T or U/C base oppositions. wt Mth212 used for the crystallographic analysis was produced in appropriate amounts by L. Schomacher and C. Preiß (this department). X-ray crystallographic analysis as well as production of some protein batches has been performed by Kristina Lakomek (Department of Molecular Structural Biology, Institute of Molecular Biology and Genetics, Georg-August University of Goettingen).

Structurally conserved amino acids residues constituting the active site pocket of Mth212 enzyme have been found by comparative analysis of 3D structures of Mth212, ExoIII and Ape1. Glu-38, Asp-151, Asp-222 and His-248 in Mth212 (K. Lakomek, 2009) were proposed to be the functional equivalents of catalytically essential Glu-34, Asp-151, Asp-229

and His-259 in ExoIII (Mol et al., 1995), and Glu-96, Asp-283, Asp-210 and His-309 in Ape1 (Gorman et al., 1997) (see Figure 22).

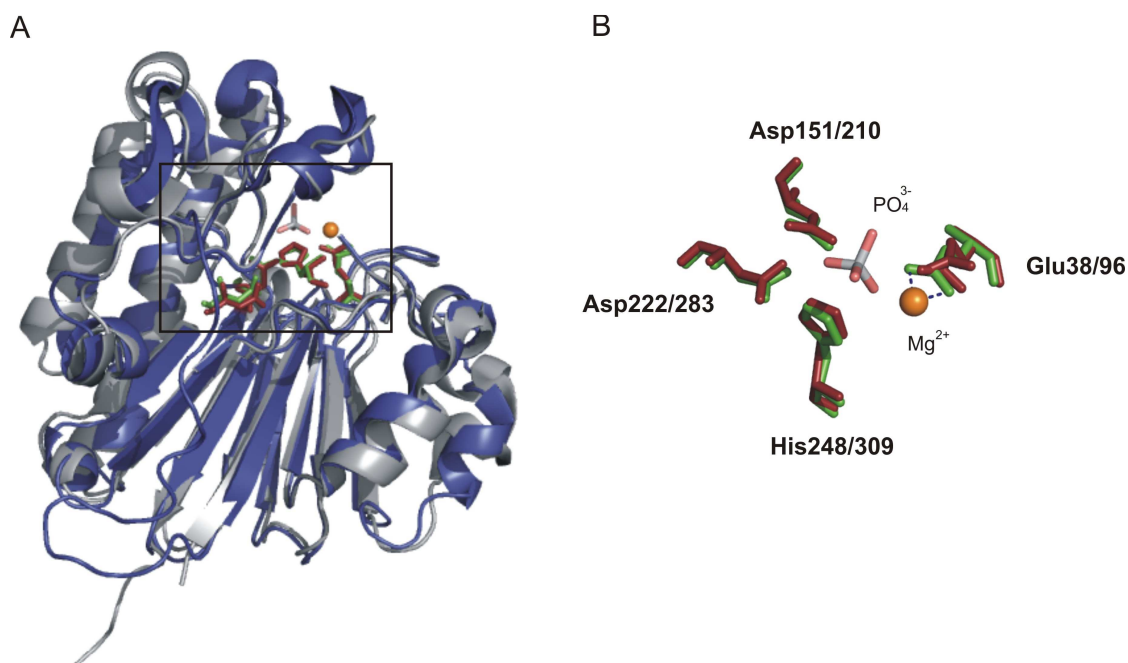


Figure 22: 3D structure of Mth212 superimposed with human Ape1

A: A cartoon representation of overall view of superimposed Mth212 apo-structure (3G91) shown in grey and Ape1 (1DEW) shown in dark blue. The catalytically essential amino acid residues of Ape1 are represented as ruby sticks and of Mth212 as green sticks. The active site comprising amino acid residues enclosed in a black square are shown in enlarged view in panel B. **B:** A close-up view of Mth212 active site residues: Asp-151, Glu-38, His-248 and Asp-222 (green sticks) superimposed with Ape1 active site residues: Asp-210, Glu-96, His-309 and Asp-283 (ruby sticks). A divalent metal ion (here Mg²⁺) and a phosphate ion (PO₄³⁻) bound in Mth212 active site are shown as an orange sphere and grey-pink sticks, respectively. The cartoon was slightly tilted and rotated with respect to the overall view shown in panel A.

In several Mth212 structures exhibiting wild-type Asp-151 residue a divalent metal ion Mg²⁺ was bound in the active site by the side chains of Asn-12 and Glu-38. In contrast, none of Mth212/D151N structures revealed a metal ion in the active site. These data, together with the complete loss of enzymatic activity by D151N mutant indicate that both Asp-151 and Mg²⁺ are prerequisite for the proper orientation of the target PO3' scissile bond in Mth212 active site prior to the enzymatic cleavage (Figure 22).

Structure analysis of wt Mth212 and Mth212/D151N in complex with specific DNA substrates identified specific mode of DNA binding. Both wt Mth212-DNA and Mth212/D151N-DNA complexes were found in form of an *exo*-complex, where 3' end of the DNA duplex was fixed in the active site through numerous ionic interactions involving Asn-10, Glu-38, Tyr-111, Asp-151/Asn-151, Asn-153 and His-248 (Figure 23).

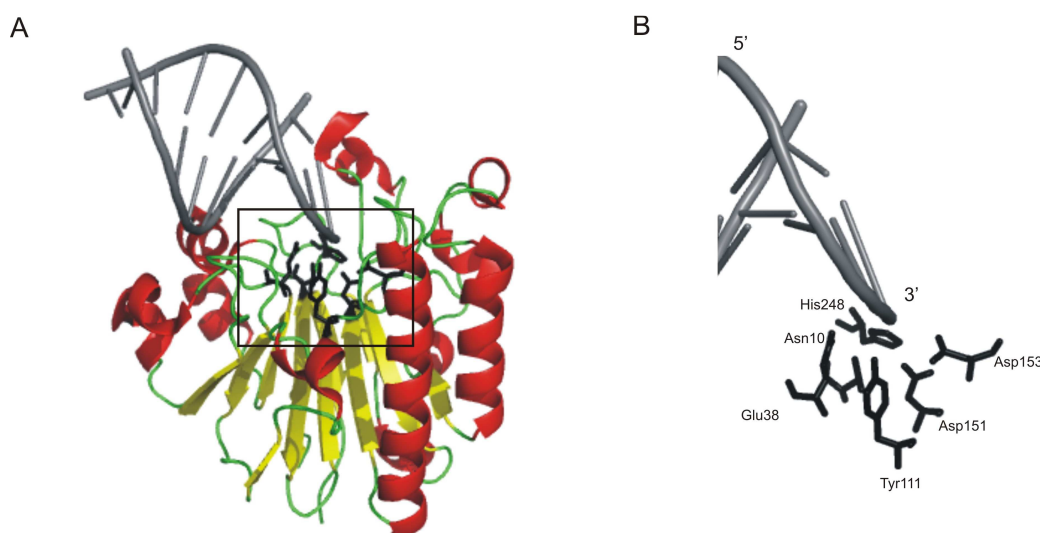


Figure 23: Mth212/D151N-DNA an exo-complex

A: An overview of Mth212/D151N (K. Lakomek, 2009) in complex with DNA showing one of the 3' ends of DNA duplex bound in the active site. Shown are: α -helices (red coils), β -sheets (yellow arrows), coils (green tubes) and DNA helices (grey tubes). **B:** A close-up view of the 3' end of DNA duplex (grey) fixed in Mth212/D151N active site by Asn-10, Glu-38, Tyr-111, Asn-151, Asp-153 and His-248 residues (black sticks).

Owing to the difficulty in obtaining Enzyme-DNA complex in *endo* binding mode, the exact mechanism of Mth212 U/G mismatch recognition and catalysis is unclear. Therefore, the obtained results were interpreted based on structural data known for other ExoIII homologue, human Ape1. From the crystal structure of human Ape1 bound to AP-site-containing DNA and discovered extra-helical AP-site binding mode Mol and co-workers proposed the nucleotide flipping mechanism as an intrinsic property of AP-site-specific endonucleases (Mol et al., 2000). More recently, Jingyang Chen *et al.* have published an experimental study on structure and dynamics of DNA nucleotides with A, T, G, and C opposite an AP-site. The authors pointed out that the base placed opposite of target AP-site intrinsically influenced DNA conformation during Ape1-AP site interaction (Jingyang Chen, 2007).

The three-dimensional structure analysis revealed that Mth212/D151N-DNA exo-complex (K. Lakomek, 2009) mostly resembles Ape1-DNA *endo*-complex (Mol et al., 2000) (Figure 24). These results suggest that Mth212 can be equipped with Ape1-like mechanism of substrate recognition and catalysis. This resemblance however gives no clue as to what are the structural bases of Mth212 uracil specificity, as Ape1 lacks such U-endonuclease activity. Meanwhile, presence of a single active site in Mth212 for major enzymological hallmarks suggests that recognition of an AP-site, 3' end and DNA-U residue occurs via same set of structural elements providing Mth212 with definite mechanism of molecular catalysis.

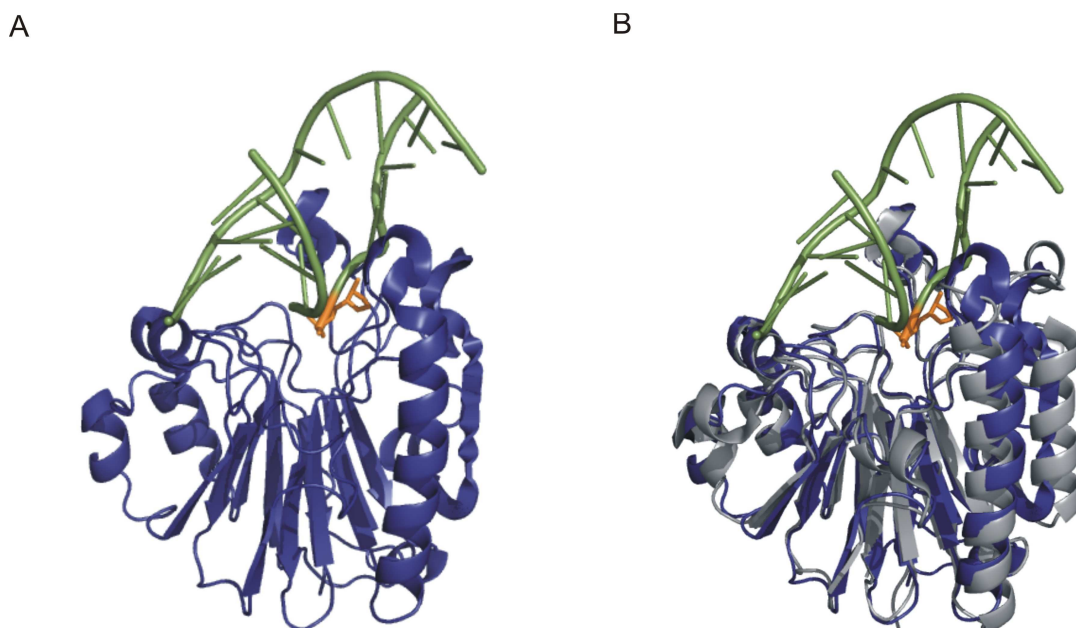


Figure 24: 3D structure of human Ape1-DNA complex superimposed with Mth212D151N-DNA complex

A: An overall view of Ape1 complex with AP-site containing DNA in a cartoon representation. DNA (in green) with the target AP-site (orange) bound to Ape1 (dark blue) (from the co-crystal structure of Ape1-DNA (1DEW), Mol et al., 2000). **B:** Superimposed human Ape1 (dark blue) in complex with AP-site (orange) containing DNA (green) and Mth212/D151N-DNA complex (grey) (3G38, here same structure as shown in Figure 23A, except for DNA molecule that was hidden). Structure data were processed using PyMOL (DeLano Scientific LLC) software.

In conclusion, based on the functional and structural analysis a putative mechanism of Mth212 specific substrate recognition and catalysis could be proposed. It is most likely that Mth212 similar to the other damage-specific DNA repair enzymes discriminates between damaged and undamaged DNA by flipping-out mechanism for the mismatched base recognition and cleaves damaged base extra-helically. However, the structural basis of the specific uracil recognition and discrimination against cytosine and thymine remains unclear. In the absence of the crystal structure of Mth212-DNA complex in *endo* binding mode, the assignment of Asp-151, as well as other conservative amino acid residues in Mth212 specific uracil residue recognition and catalysis is speculative and requires further studies.

4 Summary

Spontaneous hydrolytic deamination of cytosine to uracil is one of the most frequent DNA modifications during cellular everyday metabolism. To counteract this mutagenic effect, most organisms eliminate uracil residues from DNA by means of the base-excision repair pathway (BER), which is initiated by a specific uracil-DNA glycosylase enzyme (UDG). UDGs are highly conserved and widespread DNA repair enzymes that are found in almost all living organisms. The most striking is perhaps the lack of UDG family genes in *Methanothermobacter thermautotrophicus*, a thermophilic Archaeon, suggesting an alternative DNA uracil repair system.

Mth212 from *M. thermautotrophicus* is the only enzyme within ExoIII family members, which besides all enzymatic activities inherent in ExoIII homologues recognises uracil in ds DNA and cleaves the phosphodiester linkage located to the 5' side of 2'dU residue.

The purpose of this project was to gain an insight into the mode of Mth212 specific substrate recognition and structural features providing Mth212 with uracil specificity. To address this, a conserved Asp-151 residue in Mth212 was substituted for Asn (D151→N). Mth212/D151N mutant was purified and analysed *in vitro* for enzymatic activity and substrate binding properties. D151→N substitution eliminated both AP-endo and U-endonuclease activities but did not influence the substrate binding. These data indicate that Mth212 has a single catalytic active site for both AP-endo and U-endonuclease activities. Furthermore, retention of the substrate binding properties provides evidence for the proper folding of Mth212/D151N mutant protein.

Analysis of D151N binding to the ds DNA substrates containing either an AP-site or U/X oppositions (X: Py or Pu base type) revealed the following decrease in substrate binding strength: AP-site > DNA ends [3'recessed or blunt ends] > U/Py > U/Pu. These results suggest that *in vitro* Mth212 can bind DNA substrate in either *endo*- or *exo*-binding mode. It is most likely that the mode of substrate binding determines the strength of substrate binding and the efficacy of the overall enzymatic reaction.

The data on wt Mth212 substrate cleavage efficiency show that wt Mth212 processes different U/X oppositions with different preferences at the early stages of enzymatic reaction. U/Py base oppositions were cleaved faster than U/Pu and the order of substrate cleavage rates reflected the order of substrate binding strengths. Such mode of enzymatic activity can be attributed to the difference in base pairing properties of various U/X oppositions and to the energetic costs required for the extruding of the target uracil from DNA base stack. Moreover, the observed difference in substrate cleavage rates, in correlation with the energy requirements, argues for the base flipping-out as a rate-limiting factor in the overall enzymatic reaction. Based on binding assays and in substrate cleavage results a putative

nucleotide-flipping mechanism for the uracil-containing mismatch discrimination and processing was proposed.

The crystallographic analysis of Mth212/D151N in complex with specific DNA substrates revealed all Protein-DNA complexes in form of *exo*-binding mode, where one of the 3' ends of DNA duplex is fixed in the active site. Since free DNA ends *in vivo* is a rare event, it is more likely that Mth212, similar to other damage-specific DNA repair enzymes, binds DNA distortion unspecifically building a transient complex and slides along DNA strand until the target mismatch is encountered.

5 Abbreviations:

aa amino acids

Amp Ampicillin

AP- apyrimidinic/apurinic site

APS Ammoniumperoxodisulfat

bp base pair

BSA *bovine serum albumin*

i.e. *id est* (that is; in other words)

e.g. *exempli gratia* (for example)

etc. *et cetera* (and so forth)

ORF open reading frame

°C degree Celsius

Cm Chloramphenicol

C-terminal carboxy-terminal

DNA Deoxyribonucleic acid

DNA-U-Rep DNA-Uridin-Repair

dNTP deoxyribonucleoside triphosphate

ddNTP dideoxyribonucleoside triphosphate

ssDNA double stranded DNA

DTT Dithiothreitol

dYT *double yeast tryptone*

ϵ_{280} molar extinction coefficient at 280 nm

EDTA ethylenediaminetetraacetate

etc. *et cetera*

F Farad, electric capacitance

g gramm

h hours

HEPES 4-(2-Hydroxyethyl)-1-piperazin-ethan-sulfonic acid

HhH Helix-hairpin-helix,

GPD Glycin-Prolin-Aspartat

IMAC immobilized metal ion affinity chromatography

IPTG Isopropyl- β -D-thiogalaktopyranosid

kb Kilo-base pair

l Liter

LB Luria Bertani

M molar

Mr relative molecular mass

ml, μ l milli- (10^{-3}), mikro- (10^{-6}) liter

min minutes

μ m, nm mikro- (10^{-6}), nano- (10^{-9}) meter

N-terminal amino-terminal

ODx Optical density x nm

Ω Ohm, electrical resistance

ori origin of replication

p Pico (10^{-12})

PA / PAGE Polyacrylamid / Polyacrylamid-Gel electrophoresis

PBS phosphate buffered saline

PCR polymerase chain reaction

PEG Polyethylenglykol

Pfu-Polymerase, DNA-Polymerase B from *Pyrococcus furiosus*

Pol B, DNA-Polymerase B from *Methanothermobacter thermautotrophicus* Stamm ΔH

R universal gas constant, R= 8,314 J/K mol

RT room temperature

rpm rotation per minute

sec second

SDS Sodiumdodecylsulfat

ssDNA single stranded DNA

T temperature

t time

Taq-Polymerase, DNA-Polymerase from *Thermus aquaticus*

TEMED N,N,N',N'-Tetramethylethylenamin

Tris Tris(hydroxymethyl)aminomethan

u enzymatic activity unit

U-Endo DNA-Uridin-Endonuclease

UDG Uracil-DNA-Glycosylase

Ugi Uracil-DNA-Glycosylase Inhibitor

Ung Uracil-N-Glycosylase from *E.coli*

UV Ultraviolet

V Volt

Vol. volume

v/v volume/volume

w/v weight/volume

3D three-dimensional

DNA/RNA-Bases:

A Adenine

C Cytosine

G Guanine

T Thymine

U Uracil

Amino acids:

A / Ala Alanine

C / Cys Cysteine

D / Asp Aspartic Acid

E / Glu Glutamic Acid

F / Phe Phenylalanine

G / Gly Glycine

H / His Histidine

I / Ile Isoleucine

K / Lys Lysine

L / Leu Leucine

M / Met Methionine

N / Asn Asparagine

P / Pro Proline

Q / Gln Glutamine

R / Arg Arginine

S / Ser Serine

T / Thr Threonine

V / Val Valine

W /Trp Tryptophan

Y / Tyr Tyrosine

Organism: Mth *Methanothermobacter thermautotrophicus*

6 References

- Aravind, L.K., E. V. (2000) The alpha/beta fold uracil DNA glycosylases: a common origin with diverse fates. *Genome Biol*, **1**, RESEARCH0007.
- Barrett, T.E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J. and Pearl, L.H. (1998) Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. *Cell*, **92**, 117-129.
- Barzilay, G., Mol, C.D., Robson, C.N., Walker, L.J., Cunningham, R.P., Tainer, J.A. and Hickson, I.D. (1995a) Identification of critical active-site residues in the multifunctional human DNA repair enzyme HAP1. *Nat Struct Biol*, **2**, 561-568.
- Barzilay, G., Walker, L.J., Robson, C.N. and Hickson, I.D. (1995b) Site-directed mutagenesis of the human DNA repair enzyme HAP1: identification of residues important for AP endonuclease and RNase H activity. *Nucleic Acids Res*, **23**, 1544-1550.
- Bernstein, C.a.B., H. (1991) Aging, Sex and DNA Repair, Academic press,Inc.,San Diego. *Academic press,Inc.,San Diego*.
- Blum, H., Beier, H. and Gross, H. (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis*, 93-99.
- Bodil Kavli, O.S., Mansour Akbari, Marit Otterlei, Hilde Nilsen, Frank Skorppen, Per Arne Aas, Lars Hagen, Hans E. Krokan, and Geir Slippaug. (2002) hUNG2 Is the Major Repair Enzyme for Removal of Uracil from U:A Matches, U:G Mismatches, and U in Single-stranded DNA, with hSMUG1 as Broad Specificity Backup. *The J.of Biol. Chem.*, **277**, 39926-39936.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, **72**, 248-254.
- Bruner, S.D., Norman, D.P. and Verdine, G.L. (2000) Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature*, **403**, 859-866.
- Bullock, W.O., Fernandez, J.M. and Short,J.M. (1987) XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *Biotechniques*, **5**, 376-387.
- Chester, N. and Marshak, D.R. (1993) Dimethyl sulfoxide-mediated primer T_m reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction. *Anal Biochem*, **209**, 284-290.
- Conley, E.C., Saunders, V.A., Jackson, V. and Saunders, J.R. (1986) Mechanism of intramolecular recyclization and deletion formation following transformation of *Escherichia coli* with linearized plasmid DNA. *Nucleic Acids Res*, **14**, 8919-8932.
- Dower, W.J., Miller, J.F. and Ragsdale, C.W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res*, **16**, 6127-6145.

- Erzberger, J.P. and Wilson, D.M., 3rd. (1999) The role of Mg²⁺ and specific amino acid residues in the catalytic reaction of the major human abasic endonuclease: new insights from EDTA-resistant incision of acyclic abasic site analogs and site-directed mutagenesis. *J Mol Biol*, **290**, 447-457.
- Friedberg, E.C. (2003) DNA damage and repair. *Nature*, **421**, 436-440.
- Fromme, J.C., Banerjee, A., Huang, S.J. and Verdine, G.L. (2004) Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase. *Nature*, **427**, 652-656.
- Gallinari, P. and Jiricny, J. (1996) A new class of uracil-DNA glycosylases related to human thymine-DNA glycosylase. *Nature*, **383**, 735-738.
- Georg, J. (2005) Gentechnische Production und enzymologische Charakterisierung des ExoIII-Homologs von *Methanothermobacter thermautotrophicus*. *Diplomarbeit, Georg-August-Universität Göttingen*.
- Georg, J., Schomacher, L., Chong, J.P., Majernik, A.I., Raabe, M., Urlaub, H., Muller, S., Ciirdaeva, E., Kramer, W. and Fritz, H.J. (2006) The *Methanothermobacter thermoautotrophicus* ExoIII homologue Mth212 is a DNA uridine endonuclease. *Nucleic Acids Res*, **34**, 5325-5336.
- Gorman, M.A., Morera, S. and Rothwell, D.G. (1997) The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. *The EMBO Journal*, **16**, 6548-6558.
- Grogan, D.W. (2004) Stability and repair of DNA in hyperthermophilic Archaea. *Curr Issues Mol Biol*, **6**, 137-144.
- Haushalter, K.A., Todd Stukenberg, M.W., Kirschner, M.W. and Verdine, G.L. (1999) Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors. *Curr Biol*, **9**, 174-185.
- Hendrich, B., Hardeland, U., Ng, H.H., Jiricny, J. and Bird, A. (1999) The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature*, **401**, 301-304.
- Hinks, J.A., Evans, M.C., De Miguel, Y., Sartori, A.A., Jiricny, J. and Pearl, L.H. (2002) An iron-sulfur cluster in the family 4 uracil-DNA glycosylases. *J Biol Chem*, **277**, 16936-16940.
- Hollis, T., Ichikawa, Y. and Ellenberger, T. (2000) DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, *Escherichia coli* AlkA. *Embo J*, **19**, 758-766.
- Horst, J.P.a.F.H.J. (1996) Counteracting the mutagenic effect of hydrolytic deamination of DNA 5-methylcytosine residues at high temperature: DNA mismatch N-glycosylase

- Mig.MthI of the thermophilic archaeon *Methanobacterium thermautotrophicum* THF. *Embo J*, **15**, 5459-5469.
- Hoseki, J., Okamoto, A., Masui, R., Shibata, T., Inoue, Y., Yokoyama, S. and Kuramitsu, S. (2003) Crystal structure of a family 4 uracil-DNA glycosylase from *Thermus thermophilus* HB8. *J Mol Biol*, **333**, 515-526.
- Jie Hu, A.M.a.A.R.D. (2008) A two-step nucleotide-flipping mechanism enables kinetic discrimination of DNA lesions by AGT. *PNAS*, **105**, 4615-4620.
- Jingyang Chen, F.-Y.D., David A. Case, Christopher J. Turner and JoAnne Stubbe. (2007) DNA oligonucleotides with A, T, G or C opposite an abasic site: structure and dynamics. *Nucleic Acids Res*, **36**, 253-262.
- Jones, D.H. (1994) PCR mutagenesis and recombination in vivo. *PCR Methods Appl*, **3**, S141-148.
- Kavli, B., Sundheim, O., Akbari, M., Otterlei, M., Nilsen, H., Skorpen, F., Aas, P.A., Hagen, L., Krokan, H.E. and Slupphaug, G. (2002) hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup. *J Biol Chem*, **277**, 39926-39936.
- Kelman, Z. and White, M.F. (2005) Archaeal DNA replication and repair. *Curr Opin Microbiol*, **8**, 669-676.
- Krokan, H.E., Drablos, F. and Slupphaug, G. (2002) Uracil in DNA--occurrence, consequences and repair. *Oncogene*, **21**, 8935-8948.
- Krokan, H.E., Standal, R. and Slupphaug, G. (1997) DNA glycosylases in the base excision repair of DNA. *Biochem J*, **325 (Pt 1)**, 1-16.
- Kunkel, T.A. and Wilson, S.H. (1996) DNA repair. Push and pull of base flipping. *Nature*, **384**, 25-26.
- Kyte, J. (1995) *Mechanism in Protein Chemistry*. Garland Publishing, Inc., New York and London.
- L.Schomacher. (2007) Ein neu entdeckter Weg der Reparatur hydrolytisch geschädigter DNA-Cytosinereste, etabliert im thermophilen Archaeon *Methanotermobacter thermautotrophicus* H. Dissertation, Georg-August-Universität Göttingen.
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
- Lakomek, K. (2009) Structural characterisation of the lysosomal 66,3kDa protein and of the DNA repair enzyme Mth0212 by means of X-ray crystallography, Georg-August-Universität, Göttingen.
- Lienhard, G.E. (1973) Enzymatic catalysis and transition-state theory. *Science*, **180**, 149-154.

- Lindahl, T. (1974) An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc. Natl. Acad. Sci. USA*, **71**, 3649-3653.
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.
- Lindahl, T., Ljungquist, S., Siebert, W., Hyberg, B. and Sperens, B. (1977) *DNA N-glycosidases: properties of uracil-DNA glycosylase from Escherichia coli*.
- Majernik, A.I., Jenkinson, E.R. and Chong, J.P. (2004) DNA replication in thermophiles. *Biochem Soc Trans*, **32**, 236-239.
- Mol, C.D., Arvai, A.S., Sanderson, R.J., Slupphaug, G., Kavli, B., Krokan, H.E., Mosbaugh, D.W. and Tainer, J.A. (1995a) Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. *Cell*, **82**, 701-708.
- Mol, C.D., Izumi, T., Mitra, S. and Tainer, J.A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. *Nature*, **403**, 451-456.
- Mol, C.D., Kuo, C.F., Thayer, M.M., Cunningham, R.P. and Tainer, J.A. (1995b) Structure and function of the multifunctional DNA-repair enzyme exonuclease III. *Nature*, **374**, 381-386.
- Mol, C.D., Parikh, S.S., Putnam, C.D., Lo, T.P. and Tainer, J.A. (1999) DNA repair mechanisms for the recognition and removal of damaged DNA bases. *Annu Rev Biophys Biomol Struct*, **28**, 101-128.
- Neddermann, P. and Jiricny, J. (1993) The purification of a mismatch-specific thymine-DNA glycosylase from HeLa cells. *J Biol Chem*, **268**, 21218-21224.
- Neddermann, P. and Jiricny, J. (1994) Efficient removal of uracil from G.U mispairs by the mismatch-specific thymine DNA glycosylase from HeLa cells. *Proc Natl Acad Sci U S A*, **91**, 1642-1646.
- Nickoloff, J.A.a.H., M.F. (1998) *DNA Damage and Repair, Volume I : DNA repair in Prokaryotes and Lower Eukaryotes*, Humana Press Inc., Totowa, New Jersey, USA. Humana Press Inc., Totowa, New Jersey, USA.
- 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**, 2411-2423.
- Parikh, S.S., Mol, C.D., Slupphaug, G., Bharati, S., Krokan, H.E. and Tainer, J.A. (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *Embo J*, **17**, 5214-5226.
- Pearl, L.H. (2000) Structure and function in the uracil-DNA glycosylase superfamily. *Mutation Research*, **460**, 165-181.

- Pettersen, H.S., Sundheim, O., Gilljam, K.M., Slupphaug, G., Krokan, H.E. and Kavli, B. (2007) Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms. *Nucleic Acids Res*, **35**, 3879-3892.
- Pfeifer, S. and Greiner-Stoffele, T. (2005) A recombinant exonuclease III homologue of the thermophilic archaeon *Methanothermobacter thermautotrophicus*. *DNA Repair (Amst)*, **4**, 433-444.
- Pingfang Liu, Burdzy, A. and Sowers, L.C. (2002) Substrate Recognition by Family of Uracil-DNA Glycosylases: UNG, MUG and TDG. *Chem. Research in Toxicology*, **15**, 1001-1009.
- Pingfang Liu, Theruvathu, J.A., Darwanto, A., Lao, V.V., Pascal, T., III, W.G. and Sowers, L.C. (2008) Mechanisms of base selection by the *Escherichia coli* Mismatched Uracil Glycosylase. *The J. of Biol.Chem.*, **283**, 8829-8836.
- Renos Savva, M.-H.K.B.T.P.L. (1995) The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature*, **373**, 487-493.
- Rothwell, D.G., Hang, B., Gorman, M.A., Freemont, P.S., Singer, B. and Hickson, I.D. (2000) Substitution of Asp-210 in HAP1 (APE/Ref-1) eliminates endonuclease activity but stabilises substrate binding. *Nucleic Acids Res*, **28**, 2207-2213.
- Sandigursky, M. and Franklin, W.A. (1999) Thermostable uracil-DNA glycosylase from *Thermotoga maritima* a member of a novel class of DNA repair enzymes. *Curr Biol*, **9**, 531-534.
- Sandigursky, M. and Franklin, W.A. (2000) Uracil-DNA glycosylase in the extreme thermophile *Archaeoglobus fulgidus*. *J Biol Chem*, **275**, 19146-19149.
- Sanger, F., Nicklen, S. and A.R., a.C. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, **74**, 5463-5467.
- Sartori, A.A., Fitz-Gibbon, S., Yang, H., Miller, J.H. and Jiricny, J. (2002) A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. *Embo J*, **21**, 3182-3191.
- Sartori, A.A., Schar, P., Fitz-Gibbon, S., Miller, J.H. and Jiricny, J. (2001) Biochemical characterization of uracil processing activities in the hyperthermophilic archaeon *Pyrobaculum aerophilum*. *J Biol Chem*, **276**, 29979-29986.
- Schmiedel, R., Kuettner, E.B., Keim, A., Strater, N. and Greiner-Stoffele, T. (2009) Structure and function of the abasic site specificity pocket of an AP endonuclease from *Archaeoglobus fulgidus*. *DNA Repair (Amst)*, **8**, 219-231.
- Schomacher, L., Chong, J.P., McDermott, P., Kramer, W. and Fritz, H.J. (2009) DNA uracil repair initiated by the archaeal ExoIII homologue Mth212 via direct strand incision. *Nucleic Acids Res*, **37**, 2283-2293.

- Starcuviene, V. (2001) Identification and characterisation of thermistable uracil glycosylase from the archaeon *Methanobacterium thermoautotrophicum* and the bacterium *Thermus thermophilus*, PhD thesis, Georg-August-Universität Göttingen.
- Starcuviene, V.a.H.-J.F. (2002) A novel type of uracil-DNA glycosylase mediating repair of hydrolytic DNA damage in the extremely thermophilic eubacterium *Thermus thermophilus*. *Nucleic Acids Res*, **30**, 2097-2102.
- Thayer, M.M., Ahern, H., Xing, D., Cunningham, R.P. and Tainer, J.A. (1995) Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *Embo J*, **14**, 4108-4120.
- Varshney, U. and van de Sande, J.H. (1991) Specificities and kinetics of uracil excision from uracil-containing DNA oligomers by *Escherichia coli* uracil DNA glycosylase. *Biochemistry*, **30**, 4055-4061.
- Wang, Z.a.M., D.W. (1989) Uracil-DNA glycosylase inhibitor gene of bacteriophage PBS2 encodes a binding protein specific for uracil-DNA glycosylase. *J. Biol. Chem.*, **264**, 1163-1171.
- Weston, S.A., Lahm, A. and Suck, D. (1992) X-ray structure of the DNase I-d(GGTATACC)₂ complex at 2.3 Å resolution. *J Mol Biol*, **226**, 1237-1256.
- Womble, D.D. (2000) GCG: The Wisconsin Package of sequence analysis programs. *Methods Mol Biol*, **132**, 3-22.
- Yamashita, T., Hanada, K., Iwasaki, M., Yamaguchi, H. and Ikeda, H. (1999) Illegitimate recombination induced by overproduction of DnaB helicase in *Escherichia coli*. *J Bacteriol*, **181**, 4549-4553.
- Zeikus, J.G. and Wolfe, R.S. (1972) *Methanobacterium thermoautotrophicus* sp. n., an anaerobic, autotrophic, extreme thermophile. *J Bacteriol*, **109**, 707-715.

7 Appendix

7.1 Sequences (see attached CD):

7.1.1 pET_B_001 sequence and restriction map

7.1.2 Mth212 nucleotide and amino acid sequences, restriction map

7.1.3 Mth212/D151X (X: N, A, S) mutants verification

7.1.3.1 pET_B001_Mth212/D151N sequencing chromatogram

7.1.3.2 pET_B001_Mth212D151A sequencing chromatogram

7.1.3.3 pET_B001_Mth212/D151S sequencing chromatogram

Gratitude

My profound gratitude goes to Prof. Dr. Hans-Joachim Fritz, my supervisor, for his excellent care, intense discussions and enormous support to this work. I am particularly thankful for his really infernal patience and inestimable contribution to my training as a scientist.

I appreciate PD Dr. Wilfried Krammer's kind acceptance to be the co-referee of my thesis and valuable practical and theoretical advises.

Lars Schomacher I thank for his willingness and readiness to help and especially for a good working atmosphere. I also thank him for his grate research on Mth212 and for providing the lab with *E.coli* BL_UX and BL_UXX strains.

Many thanks go to Jens Georg for his valuable ideas and detailed characterization of Mth212.

I am thankful to Dr. Kristina Lakomek, Dr. Achim Dickmanns and Prof. Dr. Ralf Ficner from Department of Structural Biology for the detailed Mth212 structure analysis and also for the constructive and interesting dispute.

Special gratitude go to all members of our research group: Ber Svetlana, Khaliuna Tseden, Dr. Anke Schürer, Dr. Blagovesta Popova, Christiane Prieß, Sabine Smolorz, Stefan Schubert and of course to Dr. Kramer's Research group, namely Christopher Ede, Bettina Huckle, Sabrina Lehmann for the grate cooperation and truly home atmosphere at the lab.

I greatly appreciate Marita Kalck's high professionalism and care.

Thanks to the staff of Göttingen Genomics Laboratory (G2L) for the high-quality DNA sequencing.

Special thanks to Agelika Löffers, Jareck Sobkowiak, Olaf Waase and Patrick Regin for the excellent work in maintaining our department's equipment.

My highest gratitude goes to my beloved parents and my sister, as well as my friends for their boundless love, support and help.

And last but not the least my love and thanks to my little son and my husband - my best team to go through the life.

Curriculum vitae

Personal details:

Name: Elena Ciirdaeva (born Jivotovskaia)

Date of birth: 04.02.1972

Place of birth: Cișhinău, Moldova

Marital status: married, one son (born in 1999)

Citizenship: Moldavian

School education:

1979-1989 Primary and Secondary education at School No.55, Cișhinău (Moldova). Grade: Matriculation certificate

Academic background:

1993-1998 Study of Biology at the Dept. of Biology and Pedology, State University of Moldova, Cișhinău (Moldova).

1997-1998 Graduate research project in the group of Prof. Dr A. Dascaluc, Institute of Plant Physiology, Academy Of Science of Moldova, Chișinău (Moldova)
Project: Isolation of saponins (ginsenosides) from callus of ginseng (*Panax ginseng*) and investigation them by HPLC.

1998 State examination (Subjects: biology and plant biochemistry)
Grade: Diploma; Specialization: Biology;
Qualifications: Biologist; Teacher of Biology and Chemistry.

1999-2002 Maternal leave

April 2003 - Feb. 2005 Fellowship, International Research School MPI, Max Planck Institute for Terrestrial Microbiology, Marburg (Germany).

From 01.06.2005 PhD project under supervision of Prof. Dr. Hans-Joachim Fritz, Dept. Molecular Genetics und Preparative Molecular Biology, Institute of Microbiology and Genetics, University of Göttingen (Germany)