

**Univerzita Karlova v Praze, Přírodovědecká fakulta**

**Katedra Fyzikální chemie**

**Charles university in Prague, Faculty of science**

**Department of Physical chemistry**

Doktorský studijní program: Fyzikální chemie

Ph.D. study program: Physical chemistry



Strukturní studie opravy abazického místa a tvorba kovalentního spojení  
opačných řetězců DNA

Structural studies of an abasic site DNA damage repair and DNA interstrand  
cross-link formation

Autoreferát disertační práce

Dissertation Thesis summary

**Mgr. Barbora Landová**

Školitel: Mrg. et Mgr. Evžen Bouřa, PhD.

Praha/Prague, 2023

## Abstrakt

Poškozením DNA se rozumí jakákoli změna nebo modifikace struktury DNA, která se odchyľuje od jejího přirozeného stavu. Abasické místo (Ap místo) je jedním z nejčastějších poškození DNA, které vzniká spontánní depurinací/depyrimidinací nebo enzymatickým odstraněním báze. Pokud se neopraví, může vést ke genetické mutaci a potenciálně způsobit onemocnění, jako je například rakovina. Pochopení mechanismu opravy DNA je zásadní pro lékařský výzkum a aplikaci.

Bakteriální MutM je glykosyláza opravující DNA, která odstraňuje poškození DNA vzniklé oxidačním stresem a zabraňuje mutacím a genomové nestabilitě. MutM patří do rodiny prokaryotických enzymů Fpg/Nei a je strukturně i funkčně podobná se svým eukaryotickým protějškům, jako jsou NEIL1-NEIL3. Zde prezentuji dvě krystalové struktury MutM z patogenní *Neisseria meningitidis*: MutM holoenzym a MutM vázaný na DNA. Volný enzym existuje v otevřené konformaci, zatímco po vazbě na DNA, dochází k podstatným strukturním změnám a přeskupení domén enzymu i ohybu DNA.

Jednou z poškození DNA opravovaných MutM je Ap místo, které, pokud není opraveno, může spontánně vést k vytvoření mezivláknového kovanetního prokřížení DNA (Ap-ICL) se sousedním adeninem na opačném vlákně DNA. Je známo, že glykosyláza NEIL3 odstraňuje Ap-ICL. Pomocí série různých oligonukleotidů jsme zkoumali rychlost tvorby, výtěžky a stabilitu Ap-ICL. Naše zjištění ukazují, jak různé báze v blízkosti Ap místa mění rychlost a výtěžek tvorby Ap-ICL *in vitro*. Na základě našich experimentálních údajů o tvorbě Ap-ICL a známé biochemii Ap místa jsme odhadli počet Ap-ICL v buňce.

## Úvod

Tato práce se zabývá DNA opravným enzymem MutM, který náleží k rodině Fpg/Nei a je zodpovědný za opravu DNA poškození zvané abazické místo.

DNA podléhá různým spontánním chemickým procesům, které mohou vést k poškození DNA a ztrátě integrity genomu. Vysoce reaktivní sloučeniny neustále napadají citlivé skupiny DNA, což vede k velkému počtu poškození. Za změny v molekule DNA, které mohou vést k jejímu poškození, jsou nejčastěji zodpovědné jak endogenní, tak exogenní sloučeniny, například ionizace nebo oxidační stres. Pokud nejsou tyto poškození opraveny, mohou vést ke genetické nestabilitě a vzniku rakoviny [1-3].

Ačkoli se zdroje mohou u mnohobuněčných i jednobuněčných organismů lišit, za hlavní činitele zodpovědné za oxidované léze DNA se považují hydroxylový radikál, singletový kyslík a peroxid vodíku [4]. Poškozeny mohou být purinové i pyrimidinové nukleotidy, což vede ke vzniku více než 100 různých lézí DNA. Aby se zabránilo škodlivým účinkům hromadícího se poškození DNA, vyvinuly se různé opravné dráhy DNA, z nichž každá se specializuje na opravu specifických typů poškození [5].

## Abazické místo

První obrannou linií proti poškozené DNA je dráha odstraňující báze (BER), která je zahájena enzymatickým nebo spontánním štěpením N-glykosidické vazby poškozené báze za vzniku abazického místa (Ap - Apurinic/Apyrimidinic) jako meziprojektu před úplnou opravou DNA. Ap místo je rozpoznáno Ap endonukleázou, což vede k přerušení jednoho řetězce DNA. Dále je místo vyplněno DNA polymerázou a ligováno DNA ligázou. Eukaryotická DNA polymeráza se vyskytuje s chybovostí 1 chyba na  $10^7$  syntetizovaných bází. Je schopna samoregulace a opravování vlastních chyb [6-8].

Ap místa mohou vznikat také spontánně ve značném množství, přičemž v jedné lidské buňce vzniká každý den přibližně 10 000 míst Ap [10,11]. V ustáleném stavu je v savcích buňkách přítomno několik tisíc Ap míst. Uracil DNA glykosylasa (UDG) je dobře prozkoumaný enzym, který se podílí na opravě odstraněné báze [12-15]. Její

primární funkcí je eliminovat uracilová rezidua, která mohou vzniknout spontánně nebo enzymatickou deaminací cytosinu [16].

V Ap místě může kromě spontánní tvorby mezivláknové příčné vazby DNA (ICL) docházet k samovolnému nebo enzymatickému štěpení. Konkrétně Ap endonukleázy štěpí fosfodiesterovou vazbu v Ap místě, čímž vzniká konec s 3'-deoxyribóza-5'-fosfátem a jednořetězcový zlom [17]. Alternativně může dojít ke spontánnímu přerušení poškození s poločasem rozpadu 8 dní. Rychlost reakce je ovlivněna aminy a zvýšeným pH vzhledem k povaze reakce, která zahrnuje katalýzu bazickou hydrolyzou.

## Formamidopyrimidin DNA glycosylasa (MutM)

MutM je bifunkční enzym charakterizovaný jednak glykosylasovou aktivitou, která je zodpovědná za odstranění báze, a jednak  $\beta,\delta$ -lyasovou aktivitou, která opravuje vzniklé Ap místo odštěpením 3' fosfodiesterové páteře a způsobuje zlom DNA [26]. Počátečním krokem v mechanismu opravy je nukleofilní atak, kdy iminová skupina Pro2 napadá kovalentní vazbu C1' ribosy poškozené báze. Tato reakce vytvoří Schiffovu bázi s ribosylovým zbytkem v DNA. Glu3 MutM hraje klíčovou roli jako donor protonu a stabilizuje meziprodukt. MutM pomocí  $\beta,\delta$ -eliminace ribosy na 3' a 5' konci tvoří vedlejší produkt 4-hydroxypenta-2,4-dien [27]. Pro2 je součástí konzervovaného N-konce katalytického motivu PELPEV proteinů MutM, ačkoli třetí a šestá aminokyselina se mohou lišit [28].

K dnešnímu dni bylo publikováno obrovské množství krystalových struktur Fpg/Nei glykosylas. Vůbec první poskytnutá struktura byla holoenzym z organismu *Bacillus stearothermophilus* (PDBID: 1EE8) [29]. Následně byla publikována struktura MutM vázaného na DNA (PDBID: 1L1T, 1L1Z, 1L2B, 1L2C, 1L2D) [28,30]. Další výzkum byl přednostně zaměřen na hledání substrátu, rozpoznávání, preferenci a význam nepárové báze pro poškozené a detaily reakčního mechanismu [27,31-34].

## Kovalentní prokřížení opačných DNA řetězců (ICL)

ICL je poškozené DNA, kde se vytváří kovalentní vazba mezi komplementárními vlákny DNA a vytváří tak bariéru, která zastavuje replikaci DNA [48]. Byla studována řada forem ICL, které se liší jak strukturním složením, tak chemickými vlastnostmi [56-59]. ICL je toxické poškození, které může vznikat jako důsledek působení endogenních i exogenních činidel. Může vznikat také spontánně, kdy volná aminoskupina na adeninu nebo guaninu na opačném vlákně DNA reaguje s aldehydovým zbytkem Ap místa. Aldehydová skupina otevřené formy ribózy na C1' vytváří N-glykosidickou vazbu s aminoskupinou báze na opačném vlákně. ICL brání replikačnímu a transkripčnímu mechanismu kovalentní vazbou mezi oběma vlákny DNA, což vyžaduje odstranění opravy drahou specializovanou na ICL.

V současné době jsou známy tři dráhy pro opravu ICL: NEIL3 glycosylasová, acetaldehydová a Fanconiho anémie (FA) [47,49]. FA opravuje cisplatinový ICL a ICL způsobené hořčičným plynem. Pokud dráha NEIL3 selže, přijde na řadu FA, která je zodpovědná za dvouřetězcový zlom, následně opravovaný homologní rekombinací (HR). Obě dráhy jsou spojeny s replikačním mechanismem DNA a při opravě ICL se vzájemně doplňují. Dráha NEIL3 přebírá odpovědnost v raných fázích replikace a je spojena s replikační vidličkou, která se zastaví před ICL. NEIL3 je specializovaná dráha zodpovědná za opravu Ap-ICL a psoralenový ICL, v tomto případě nedochází ke generování dvouřetězcových zlomů [61].

### NEIL3 glycosylasa

Bylo prokázáno, že NEIL3 může štěpit jednu z nenativních N-glykosidických vazeb ve struktuře Ap-ICL, což vede k rozpojení vazby tohoto poškození. Tyto experimenty byly provedeny v extraktech z vajíček *Xenopus laevis* a stanovily NEIL3 jako primární dráhu pro opravu Ap-ICL [47-50]. Přesné molekulární mechanismy opravy Ap-ICL a kompletní reakční mechanismus NEIL3 však zůstávají neznámé.

NEIL3 rekrutován k zastavené replikační vidličce se pravděpodobně zaměřuje na báze nacházející se v blízkosti spojnice vláken DNA v konfiguraci tvaru X nebo Y

[48,52,53]. Schopnost NEIL3 katalyticky odstraňovat psoralen a ICL odvozené od Ap lze demonstrovat na známých substrátech. Nenativní glykosidická vazba přítomná v Ap-dA ICL vykazuje podobnost s glykosidickou vazbou, která se nachází ve FapyA a FapyG. Obdobně nasycené thymínové kruhy v psoralenové ICL vykazují podobnost se substráty, jako jsou DTH a thyminoglykol.

## Cíle práce

- Struktura MutM: Pomocí proteinové krystalografie určit atomární strukturu proteinu MutM z patogenní bakterie *Neisseria meningitidis* a na tomto základě charakterizovat jeho roli při opravě DNA.
- Strukturní a funkční charakterizace komplexu MutM:DNA: Pomocí proteinové krystalografie odhalit atomární strukturu komplexu MutM s jeho substrátem. Charakterizace vazby vazebnými experimenty, určení disociačních konstant pomocí fluorescenční anisotropie a eseje založené na pohybu DNA na gelu.
- Tvorba Ap-ICL: Optimalizovat spontánní tvorbu Ap-ICL a zjistit, jak sekvence DNA v blízkosti vznikajícího poškození ovlivňuje tvorbu a stabilitu Ap-ICL.

## Použité metody

V této práci byly použity standardní metody molekulární biologie a proteinové krystalografie:

Molekulární biologie: Klonování DNA pro přípravu expresních vektorů, exprese a purifikace protein, enzymatické aktivní eseje – fluorescenční anisotropie, metoda založená na dělení a posuvu na gelu, neenzymatické eseje.

Proteinová krystalografie: Krystalizace proteinu a protein:DNA komplexu, sběr difrakčních dat na rotační anodě a synchrotronu, procesování dat a řešení struktury pomocí metody molekulárního nahrazení, upřesňování struktury.



## Výsledky

Výsledky v této práci byly publikovány ve dvou mezinárodních časopisech s impaktovým faktorem.

### Krystalové struktury a konformační změny MutM při vazbě na DNA

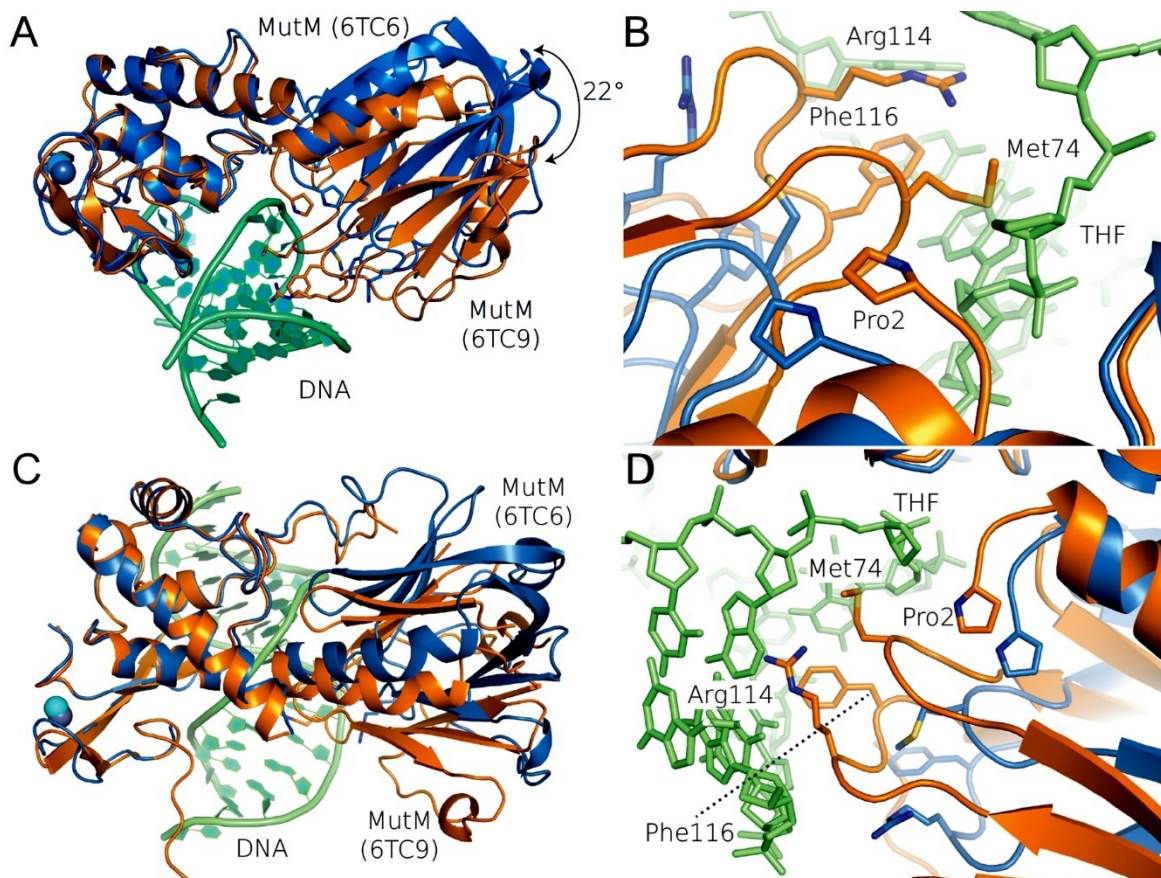
V první publikaci byl protein MutM krystalizován jako holoenzym a v komplexu s navázaným substrátem, DNA s poškozením napodobujícím Ap místo (THF), které je jeho přirozeným substrátem. Pro komplexní charakterizaci byly provedeny vazebné experimenty s využitím biofyzikálních metod fluorescenční anizotropie a pohybu DNA na gelu.

Hodnoty  $K_D$  byly vypočteny na základě měření fluorescenční anizotropie zahrnující dvouvláknovou i jednovláknovou DNA, s napodobením Ap místa THF, i bez poškození. Oligonukleotidy byly komerčně syntetizovány. Nejslabší vazbu jsme naměřili se ssDNA (bez imitace poškození), kde  $K_D$  bylo přibližně 27 nM. Disociační konstanty pro zbývající substráty, včetně dsDNA (bez poškození), ssTHF a dsTHF (s napodobením poškození), byly podobné a pohybovaly se mezi 7-9 nM.

Byla vyřešena krystalová struktura holoenzymu MutM z patogenní bakterie *Neisseria Meningitidis* (PDBID: 6TC6). Krystal z uspořádání sedící kapky difraktoval s rozlišením 2,9 Å a struktura byla řešena metodou molekulárního náhrzení s použitím výchozího modelu MutM z *E. coli* (PDBID: 1K82). Zpočátku model neposkytoval uspokojivé mapy elektronové hustoty. Pro dosažení úspěšného řešení byl model rozdělen na subdomény NTD a CTD. Tyto subdomény byly použity jako nezávislé modely pro hledání molekulárního náhrzení, přičemž byly vytipovány rozdíly v orientaci CTD a NTD mezi strukturou holoenzymu a strukturou s navázanou DNA. Oblast od Thr217 po Tyr242 vykazovala na mapě elektronové hustoty špatnou viditelnost, a proto byla z konečného modelu vyloučena. Tato oblast je typicky dobře uspořádaná u MutM vázaného na DNA.

Krystalová struktura MutM vázaného na DNA (PDBID: 6TC9) z *Neisseria Meningitidis*, obsahující THF pro napodobení Ap místa, byla také vyřešena pomocí metody molekulárního náhrzení, kde jako výchozí model byla použita struktura MutM z této studie (PDBID: 6TC6). Pro úspěšné dosažení molekulárního nahrazení byly opět NTD a CTD rozděleny a každá subdoména byla samostatně umístěna do softwaru Phaser. Výsledný model dosáhl rozlišení 2,15 Å. Vazba na MutM na DNA způsobila ohyb DNA přibližně 60°. Důležité je, že přítomnost THF nemá vliv na konformační změnu.

Vazba na DNA vyvolává výrazné konformační změny i ve strukturním uspořádání MutM, přičemž CTD i NTD procházejí značnou reorganizací. V programu PyMol jsme vypočítali celkovou střední kvadratickou odchylku (RMSD), která pro celý protein činila 2,41 Å. Superpozice jednotlivých subdomén odhalila rotaci CTD i NTD směrem k DNA o úhel 22,3° (viz obr. 1). Při superpozici NTD se navíc atom zinku z domény ZnF posunul o 11,4°.



Obrázek 1. Strukturální zarovnáníporovnání CTD MutM navázaného na DNA oranžově a volného MutM modře (A-D) ukazující konformační změny při vazbě na DNA (zeleně). Ilustrace rotačního pohybu o 22° v panelu A s pohledem shora v panelu C. Pohledy z různých úhlů (B, D) ukazují detaily pohybu aminokyselinových zbytků Met 74, Arg114 a Phe 116 směrem k navázané DNA.

Data jsou součástí následné publikace.

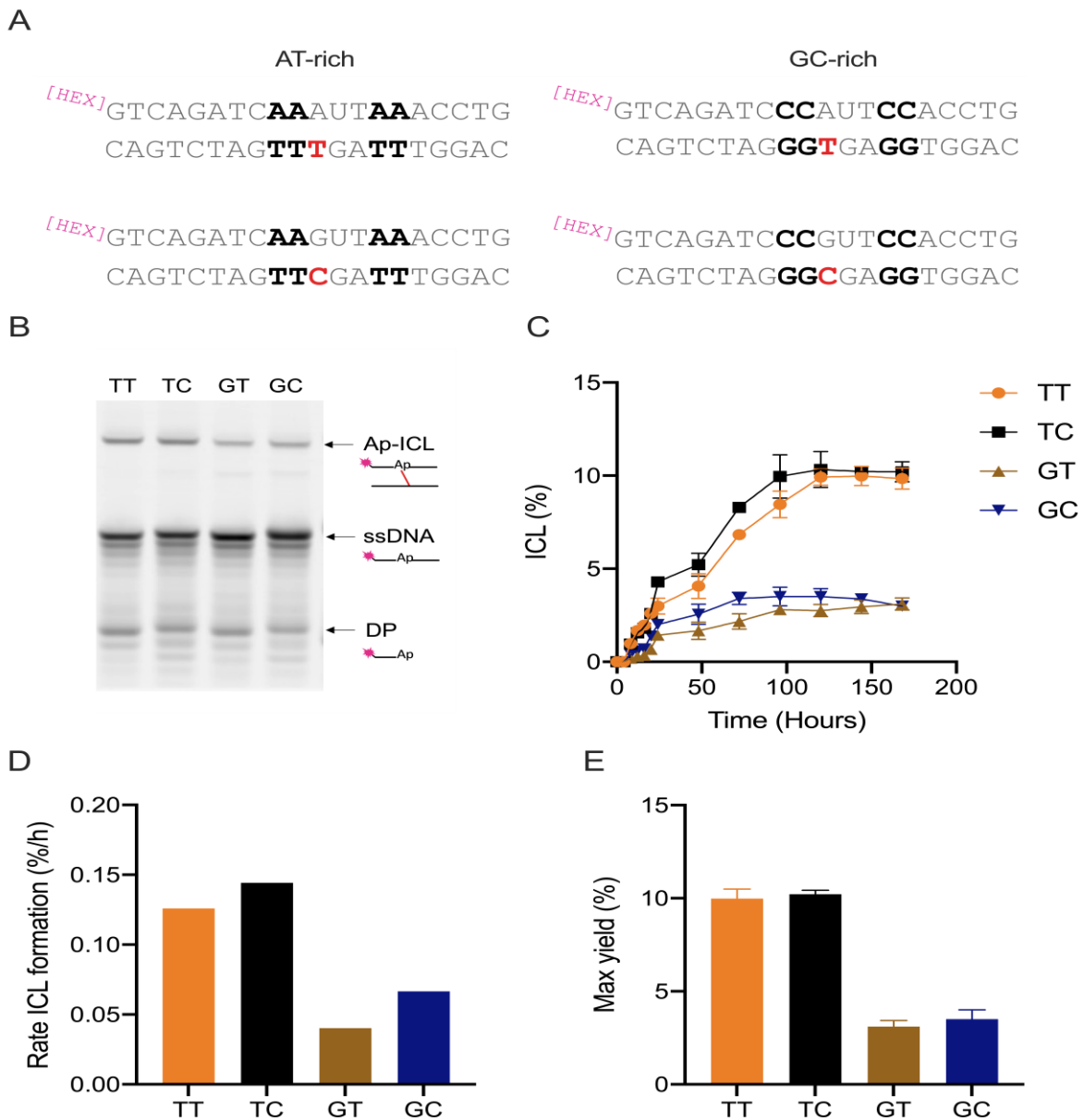
Landova, B., and Silhan, J. (2020) Conformational changes of DNA repair glycosylase MutM triggered by DNA binding, *Febs Letters* 594, 3032-3044.

## Rychlost tvorby mezivláknové příčné vazby v abasickém místě

Jak Ap místa, tak Ap-ICL se spontánně vytvářejí v molekule DNA v jakémkoli kontextu sekvence DNA. V této studii jsme provedli komplexní výzkum Ap-ICL, specifické podtřídy DNA poškození s významnými důsledky pro stabilitu genomu. Navrhli jsme sadu oligonukleotidů, u nichž jsme zpočátku měnili báze v blízkosti poškození (4. a 5. pozice), přičemž jsme ponechali fixovaný pouze adenin na protilehlém vlákně. Později jsme variabilní sekvenci rozšířili o oblasti bohaté na AT a GC. Rychlost tvorby Ap-ICL jsme měřili ve třech samostatných experimentech a v každém z nich jsme určili maximální výtěžek. Kromě toho jsme provedli kontrolní měření s hypoxantinem na 4. i 6. pozici, abychom potvrdili výhodnost námi zvolené pozice pro tvorbu kovalentní vazby Ap-ICL.

Naše údaje jasně prokázaly závislost na sekvenci v blízkosti poškození. Bylo zjištěno, že vznik poškození závisí jak na 4., tak 5. pozici, přičemž guanin v 5. poloze (naproti místu Ap) vykazoval nejvyšší potenciál pro tvorbu léze a adenin ve 4. poloze rovněž ovlivnil tvorbu. Dále jsme učinili zajímavý objev, že oblast bohatá na AT zvyšuje výtěžek i rychlost tvorby Ap-ICL (viz obrázek 2). Byl proveden experiment měření

stability tohoto poškození, kde jsme ukázali, že během 14 dní se rozpadne 30% z počátečního množství Ap-ICL.



Obrázek 2. Tvorba Ap-ICL v dsDNA obsahující oblasti bohaté na AT a GC. Schematické znázornění duplexů DNA obsahujících oblasti bohaté na AT nebo GC v blízkosti Ap místa (A). Oligonukleotidy bohaté na AT s thyminem nebo cytosinem na 4. pozici se označují TT nebo TC. Stejně tak oligonukleotidy bohaté na GC s thyminem a cytosinem na 4. pozici se označují GT nebo GC. Zastavené reakce byly separovány na 20% denaturačním PAGE gelu (B), kde DP znamená degradační produkt. (C) Reakční kinetika tvorby Ap-ICL pro oligonukleotidy bohaté na AT a GC. (D) Rychlost tvorby

*Ap-ICL a (E) maximální výtěžek vyneseny do grafu pro všechny sekvence. Chybové úsečky pocházejí ze tří nezávislých měření.*

Data jsou součástí následné publikace.

Huskova, A.\*, Landova, B.\*, Boura, E., and Silhan, J. (2022) The rate of formation and stability of abasic site interstrand crosslinks in the DNA duplex, *DNA Repair* 113.

\*S ekvivalentním přispěním obou autorů.

## Závěr

Jedním z hlavních cílů práce bylo prozkoumat závislost tvorby Ap-ICL na sekvenci nukleotidů v okolí vznikajícího poškození a zároveň objasnit mechanismus předcházející vzniku poškození.

Naše zjištění zdůrazňují důležitou roli sousední sekvence, zejména v uspořádání napodobujícím oblasti bohaté na AT nebo GC, kde jsme pozorovali nápadnou preferenci vzniku tohoto poškození ve volnějším oblastech bohatých na AT. Tato preference naznačuje větší prostorový požadavek pro vznik tohoto typu poškození. Dále jsme zjistili, že báze sousedící s poškozeným místem má také menší vliv na rychlost vzniku Ap-ICL.

Naše práce významně přispěla k pochopení tvorby Ap-ICL, což je oblast, které byla v rámci poškození a oprav nukleových kyselin věnována poměrně malá pozornost. Naše výsledky, zejména pokud jde o závislost na sekvenci a frekvenci výskytu, nabízejí cenné poznatky o složitosti tvorby Ap-ICL.

Dalším aspektem naší práce je, že jsme jako první provedli komplexní srovnání volné a DNA vázané formy proteinu MutM pocházející ze stejného organismu, a to krystalizací jak holoenzymu, tak formy vázané na DNA. Tato analýza poskytuje cenný pohled na strukturální změny, ke kterým dochází při interakci MutM s DNA, a také na to, jak tyto změny souvisejí s jeho úlohou při opravě DNA.

## Abstract

DNA damage refers to any alteration or modification in the DNA structure that deviates from its natural state. Abasic site (Ap site) is one of the most common DNA lesions resulting from spontaneous depurination/depyrimidination or enzymatic base excision. When left unrepaired it can lead to a cascade of genetic mutations, potentially causing diseases like cancer. Understanding DNA repair mechanisms is vital for medical research and applications.

Bacterial MutM is a DNA repair glycosylase, removing DNA damage generated by oxidative stress and preventing mutations and genomic instability. MutM belongs to the Fpg/Nei family of procaryotic enzymes, sharing structural and functional similarities with their eukaryotic counterparts, such as NEIL1-NEIL3. Here, I present two crystal structures of MutM from pathogenic *Neisseria meningitidis*: MutM holoenzyme and MutM bound to DNA. The free enzyme exists in an open conformation, while upon binding to DNA, both the enzyme and DNA undergo substantial structural changes and domain rearrangement.

One of the DNA lesion repaired by MutM is the Ap site, which, if not repaired, may spontaneously lead to the formation of an abasic site interstrand crosslink (Ap-ICL) with an adjacent adenine in the opposite strand. NEIL3 glycosylase is known to remove Ap-ICL. With a panel of different oligonucleotides, we investigated the rates of formation, the yields, and the stability of Ap-ICL. Our findings demonstrate how different bases in the vicinity of the AP site change crosslink formation *in vitro*. Based on our experimental data on Ap-ICL formation and known biochemistry of the Ap site we have estimated the number of Ap-ICLs within the cell.

## Introduction

DNA undergoes various spontaneous chemical reactions that could lead to DNA damage and loss of genome integrity. Moreover, highly reactive compounds constantly attack susceptible groups of DNA molecule, resulting in a vast number of DNA lesions. Both endogenous and exogenous compounds, such as ionization or oxidative stress, are the most often responsible for changes in DNA molecule that can lead to DNA lesions. If not repaired, these lesions may result in genetic instability and cancer [1-3].

Although the sources may differ for both multicellular and single-cellular organisms, the main agents responsible for oxidized DNA lesion are thought to be hydroxyl radical, singlet oxygen and hydrogen peroxide [4]. Both purine and pyrimidine nucleotides are damaged, leading to more than 100 different DNA lesions. To prevent the harmful effects of accumulating DNA damage, various DNA repair pathways have developed, each specialized in repairing specific types of lesions [5].

### Abasic site (Ap site)

The first line of defence against damaged DNA is the base excision repair pathway (BER), which is initiated by enzymatical or spontaneous cleavage of the N-glycosidic bond of the damaged base, forming an abasic site (Ap - Apurinic/Apyrimidinic) as an intermediate before the DNA is fully repaired. The Ap site is recognized by Ap endonuclease, resulting into a single-stranded break. Furthermore, it is filled by DNA polymerase and ligated by DNA ligase. Eukaryotic DNA polymerase occurs with an error rate of 1 error per  $10^7$  synthesised bases. It is capable of self-regulation and repairing its own mistakes [6-8].

Ap sites can also be generated spontaneously in significant quantities, with approximately 10,000 Ap sites formed each day in a single human cell [9,10]. Several thousand of Ap sites are present in mammalian cells at steady state. Uracil DNA glycosylase (UDG) is a well-known enzyme involved in base excision repair, and somatic hypermutation during antibody diversification is yet another source of Ap sites



[11-14]. Its primary function is to eliminate uracil residues, which can result from the spontaneous or enzymatic deamination of cytosine [15].

The Ap site, in addition to its spontaneous formation of DNA interstrand crosslink (ICL), can undergo either spontaneous or enzymatic cleavage. Specifically, Ap endonucleases cleave the phosphodiester bond at the Ap site, generating a nick with a 3'-deoxyribose-5'-phosphate and a single-strand break [16]. Alternatively, the lesion can spontaneously break with a half-life of 8 days. The reaction rate is influenced by amines and pH increase due to the nature of reaction, which involves base hydrolysis catalysis.

## Formamidopyrimidine DNA glycosylase (MutM)

MutM is bifunctional enzyme characterized by both glycosylase activity, responsible for removing the base, and its  $\beta,\delta$ -lyase activity, which processes the generated Ap site by cleaving phosphodiester backbone 3' causing a DNA break [17]. The initial step in the repair mechanism is nucleophilic attack, where the imine group of Pro2 attacks C1' of the ribose of the damaged base via a covalent bond. This reaction creates a Schiff base with ribosyl residue in DNA. Notably, Glu3 of MutM plays a crucial role as a proton donor and stabilizes the intermediate, making it relatively stable. MutM is regenerated by  $\beta,\delta$ -elimination of the ribose at the 3' and 5' ends, liberating a by-product 4-hydroxypenta-2,4-diene [18]. Pro2 is a part of conserved N-terminus catalytic motif PELPEV of MutM proteins, although the third and sixth amino acids may vary [19].

Up to this date, a vast number of crystal structures of Fpg/Nei glycosylases have been published. The very first provided architecture was of the holoenzyme from *Bacillus stearothermophilus* (PDBID: 1EE8) [20]. Afterwards, DNA-bound MutM was reported (PDBID: 1L1T, 1L1Z, 1L2B, 1L2C, 1L2D) [19,21]. Subsequent investigations preferably focused on substrate search, recognition, preference, the importance of the orphaned base to the lesion, and the details of the reaction mechanism [18,22-25].

## DNA interstrand cross-link

DNA interstrand cross-link (ICL) is a DNA lesion that forms covalent bonds between DNA strands, creating a barrier that stops DNA replication [26]. Numerous forms of ICLs have been studied, exhibiting difference in both structural composition and chemical properties [27-30]. This ICL is a toxic lesion that could originate from both endogenous and exogenous factors. Moreover, it can form spontaneously through a process where the exocyclic amino group on a base reacts with the aldehyde residue at the Ap site of adenine or guanine on the opposite DNA strand. The aldehyde group of the ring-opened ribose at C1' forms an N-glycosidic bond with an amine group of a base on the opposite strand. ICLs obstruct replication and transcription machinery by covalent bond between the two DNA strands, requiring removal by DNA repair pathway specialized for ICLs.

Currently, three pathways are now known for the repair of interstrand crosslink lesion: the NEIL3 glycosylase pathway, the acetaldehyde pathway and the Fanconi anaemia (FA pathway) [31,32]. FA it is known to repair cisplatin-ICL and nitrogen mustard-ICL. If the NEIL3 pathway fails, FA proteins are responsible for double strand breaks, which are then repaired by homologous recombination (HR). Both pathways are linked to DNA replication machinery and complement each other in repairing ICLs. The NEIL3 pathway takes charge in the early stages of replication and is coupled with replication fork, which stalls on the ICL. NEIL3 pathway is responsible for repairing Ap-ICL, psoralen-ICL, aldehyde-ICL, without generating double-strand breaks [33].

### Endonuclease VIII-like 3 (NEIL3 glycosylase)

Significantly, NEIL3 plays a key participant in Ap-ICL repair. It has been demonstrated that NEIL3 can cleave one of the non-native N-glycosidic bonds within the Ap-ICL structure, resulting in the unhooking of this lesion. These experiments were conducted in *Xenopus laevis* egg extracts and have established NEIL3 as the primary pathway for Ap-ICL repair [26,31,32,34]. However, the precise molecular mechanisms of the Ap-ICL repair and complete reaction mechanism of NEIL3 remain unknown.

When NEIL3 is recruited to a stalled replication fork, it appears to target bases located near the junction of DNA strands in X-shaped or Y-shaped DNA configuration [26,35,36]. The ability of NEIL3 to catalytically remove psoralen and Ap-derived ICL can be explained. The non-native glycosidic bond present in the Ap-dA ICL exhibits resemblance to the glycosidic bond found in FapyA and FapyG. Similarly, the saturated thymine rings in the psoralen ICL exhibit similarity to well-established substrates such as DTH and thymine glycol.

## Aims of the study

- MutM structure: Investigate the *Neisseria meningitidis* MutM structure using protein crystallization to reveal its atomic structure and understand its role in DNA repair.
- MutM:DNA structural and functional characterization: Examine MutM from *Neisseria meningitidis* while it is bound to DNA substrate, utilizing protein crystallization to reveal atomic structure. Conduct binding experiments and determine dissociation constants to understand its interaction with potential substrates using fluorescence anisotropy and gel-based assays.
- Ap-ICL formation: Optimize spontaneous production of abasic site DNA interstrand crosslink and explore how DNA sequence in the vicinity of the rising damage influences Ap-ICL formation and stability.

## Used methods

Standard methods of molecular biology and protein crystallography were used in this work.

Molecular biology: DNA cloning for preparation of expression vectors, protein expression and purification, enzymatic and non-enzymatic activity assays.

Protein crystallography: Protein and Protein-DNA complex crystallisation, diffraction data collection using both synchrotron and home source rotation anode, data processing, structure model preparation using molecular replacement method, structural refinement.

## Results

Results of this thesis are summarized in two published articles in foreign impact journals.

### Crystal structures and conformational changes of MutM upon DNA binding

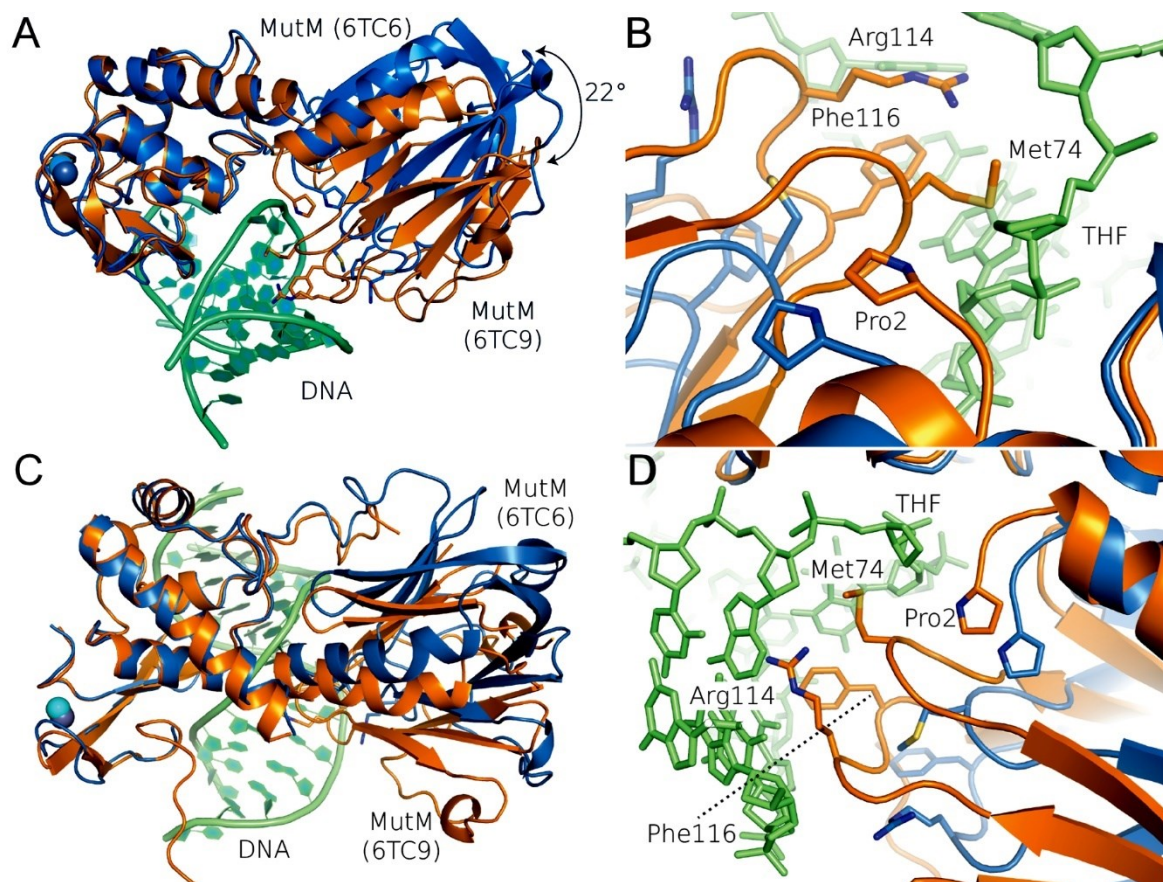
In the first publication, the MutM protein was crystalized as the holoenzyme and in complex with its bound substrate, DNA with mimicking Ap site, which is its natural substrate. To provide a comprehensive understanding, binding experiments were performed using biophysical methods fluorescent anisotropy and gel-based assays.

$K_D$  values were calculated based on fluorescence anisotropy measurements involving both double-stranded and single-stranded DNA, with and without a mimic of damaged DNA (THF). The oligonucleotides were commercially synthesized. We observed the weakest affinity for ssDNA (without mimicking damage), which was approximately 27 nM, which is still within the nanomolar range. Dissociation constants for the remaining substrates, including dsDNA (without damage), ssTHF and dsTHF (with mimicking damage), were similar, ranging between 7-9 nM.

The crystal structure of the holoenzyme MutM from the pathogenic bacteria *Neisseria Meningitidis* (PDBID: 6TC6) was solved. The crystal from the sitting drop setup diffracted to a resolution of 2.9 Å and the structure was determined by molecular replacement method using the MutM from *E. coli* (PDBID: 1K82) as the initial model. At first, the model did not produce satisfactory electron density maps. To achieve a successful solution, the model was divided into NTD and CTD subdomains. These subdomains were used as independent search models for molecular replacement, punctuating differences in the orientation of CTD and NTD between the holoenzyme and DNA-bound structures. The region from Thr217 to Tyr242 exhibited poor visibility in the electron density map and therefore was excluded from the final model. Typically, this region is well ordered when MutM is bound to DNA.

The crystal structure of DNA-bound MutM (6TC9) from *Neisseria Meningitidis*, containing THF to mimic Ap site, was also solved by molecular replacement using MutM structure from this study (PDBID: 6TC6). Again, to achieve successful molecular replacement, NTD and CTD were divided, and each subdomain individually placed in Phaser software. The resulting model reached a resolution of 2.15 Å. Upon DNA binding, MutM induced a twist of approximately 60°. Importantly, the presence of THF is not attributed for a conformational change.

The DNA binding induces significant conformational changes in structural MutM arrangement of MutM, with both CTD and NTD undergoing substantial reorganization. In PyMol we calculated overall root-mean-square deviation (RMSD) which for the entire protein was 2.41 Å. Aligning of individual subdomains revealed a rotation of both CTD and NTD towards the DNA by an angle of 22.3° (see Figure 1). Moreover, during NTD superimposition, the zinc atom from the ZnF domain moved by 11.4°.



*Figure 1. Structural alignments of CTDs of DNA bound MutM in orange and free MutM in blue (A-D) showing conformation changes upon DNA (green) binding. Illustration of the rotation movement of 22° in A panel with the top view in C panel. Views from different angles (B, D) show detail of residues Met 74, Arg114 and Phe 116 movement toward bound DNA.*

Data are included in followed publication.

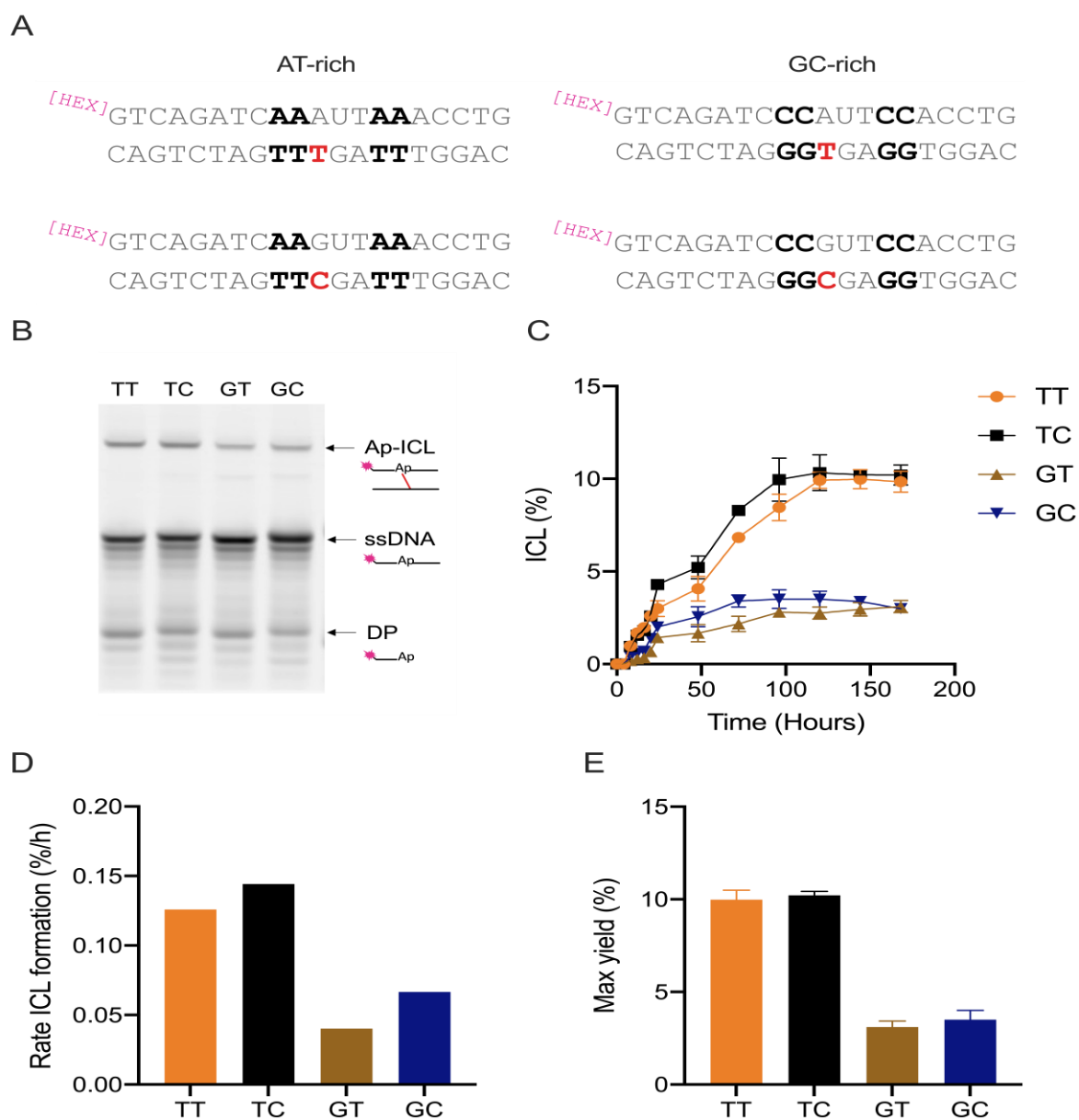
Landova, B., and Silhan, J. (2020) Conformational changes of DNA repair glycosylase MutM triggered by DNA binding, *Febs Letters* 594, 3032-3044.

## The rate of formation of abasic site interstrand crosslink

Both Ap sites and Ap-ICLs are spontaneously generated within the DNA molecule at any given DNA sequence context. In this study, we performed a comprehensive investigation of Abasic site DNA interstrand crosslink (Ap-ICL), a specific subclass of DNA lesion with significant implications for genomic stability. We designed a set of oligonucleotides where, initially, we varied the bases in the vicinity of the lesion (4<sup>th</sup> and 5<sup>th</sup> position), while keeping only adenine on opposing strand fixed. Later, we extended the variable sequence to include AT-rich and GC-rich regions. We measured the rate of the Ap-ICL formation in three separate experiments and determined the maximum yield of each of them. Additionally, we conducted control measurements with hypoxanthine at both 4<sup>th</sup> and 6<sup>th</sup> positions to confirm the preferable position for the covalent bond formation of Ap-ICL.

Our data clearly demonstrated a dependency on the sequence near the lesion. Both the 4<sup>th</sup> and 5<sup>th</sup> position were found to be crucial, with guanine at the 5<sup>th</sup> position (opposite to Ap site) showing the highest potential for lesion formation, and adenine at the 4<sup>th</sup> position being significant as well. Furthermore, we made an interesting discovery that AT-rich region increasing both yield and rate of Ap-ICL formation (see Figure 2). We also performed stability assay where we showed relative stability when 30% of the lesion was dissolved in 14 days.





*Figure 2. Formation of Ap-ICL in dsDNA containing AT-rich and GC-rich regions. Schematic representation of DNA duplexes containing AT-rich or GC-rich regions in the vicinity of the Ap site (A). AT-rich oligonucleotides with thymine or cytosine on the 4<sup>th</sup> position are named TT or TC. Likewise, GC-rich oligonucleotides with thymine and cytosine on the 4<sup>th</sup> position are named GT or GC. Stopped reaction were run on 20% denaturing PAGE gel (B), where DP stands for degradation product. (C) The reaction kinetics of Ap-ICL formation for AT and GC-rich oligonucleotides. (D) The rate of Ap-*

*ICL formation and (E) the maximum yield plotted for all sequences. Error bars comes from three independent repeats.*

Data are included in followed publication.

Huskova, A.\*, Landova, B.\*, Boura, E., and Silhan, J. (2022) The rate of formation and stability of abasic site interstrand crosslinks in the DNA duplex, *DNA Repair* 113.

\*Both authors contributed equally

## Conclusion

One of the main goals of the thesis was to investigate the dependence of Ap-ICL formation on the nucleotide sequence in the vicinity of the rising damage, while also shedding light on the mechanism prior to the damage origin.

Our findings highlight the critical role of the neighboring sequence, particularly in arrangement mimicking AT or GC-rich regions, where we observed a notable preference for this lesion to manifest in the less tightly packed AT-rich regions. This preference indicates a spatial requirement for the formation of this type of damage. Moreover, we discovered that the base adjacent to the damaged site has a slide impact on the rate of Ap-ICL formation.

Our work has significantly contributed to understanding Ap-ICL formation, an area that has received relatively limited attention within the field of nucleic acid damage and repair. Our results, particularly regarding sequence dependency and the frequency of occurrence, offer valuable insights into the complexity of Ap-ICL formation.

Another aspect of our work is that we were the first to conduct a comprehensive comparison between the bound and unbound form of the MutM protein, both originating from the same organism by crystallization both holoenzyme and DNA-bound form. This analysis provide a valuable insight into structural changes that occur when MutM interacts with DNA and also how these changes are linked to its role in DNA repair.

## Publications included in dissertation thesis

Landova, B., and Silhan, J. (2020) Conformational changes of DNA repair glycosylase MutM triggered by DNA binding, *Febs Letters* 594, 3032-3044.

Huskova, A., Landova, B., Boura, E., and Silhan, J. (2022) The rate of formation and stability of abasic site interstrand crosslinks in the DNA duplex, *DNA Repair* 113.

## Publications not included in dissertation thesis

Horova, V., Landova, B., Hodek, J., Chalupsky, K., Krafcikova, P., Chalupska, D., Duchoslav, V., Weber, J., Boura, E., and Klima, M. (2021) Localization of SARS-CoV-2 Capping Enzymes Revealed by an Antibody against the nsp10 Subunit, *Viruses-Basel* 13.

## Zkratky/Abbreviations

Ap	Apurinic/apurimidinic, abasic Apurinové/apurimidinové, abasické
Ap-ICL	Abasic site interstrand crosslink Kovalentní spojení řetězců s abazickým místem
BER	base excision repair Oprava odstraněním báze
CTD	C-terminal domain C terminální doména
DNA	Deoxyribonucleic acid Deoxyribonukleová kyselina
dsDNA	double-strand DNA dvouřetězcová DNA
dsTHF	double-strand DNA containing tetrahydrofuran dvouřetězcová DNA obsahující tetrahydrofuran
DTH	dihydrothimine dihydrothimin
FA	Fanconi anemia Fanconiho anémie
FapyA	4,6-diamino-5-formamidopyrimidine 4,6-diamino-5-formamidopyrimidin
FapyG	2,6-diamino-4-hydroxy-5-formamidopyrimidine 2,6-diamino-4-hydroxy-5-formamidopyrimidin
HEX	hexachlorofluorescein hexachlorofluorescein
HR	homologue recombination homologní rekombinace
ICL	interstrand crosslink

	kovalentní prokřížení řetězců DNA
Nei	Endonuclease VIII Endonukleasa VIII
NEIL	Endonuclease VIII-like
NTD	N-terminal domain N terminální doména
RMSD	root-mean-square deviation střední kvadratická odchylka
ssDNA	single-strand DNA jednořetězcová DNA
ssTHF	single-strand DNA containing tetrahydrofuran jednořetězcová DNA obsahující tetrahydrofuran
THF	tetrahydrofuran
UDG	Uracyl-DNA glycosylase Uracyl-DNA glykosylasa
ZnF	Zinc finger Zinečnatý prst

## Bibliography

- [1] D.E. Barnes, T. Lindahl, Repair and genetic consequences of endogenous DNA base damage in mammalian cells, *Annual Review of Genetics* 38 (2004) 445-476. 10.1146/annurev.genet.38.072902.092448.
- [2] A. Tubbs, A. Nussenzweig, Endogenous DNA Damage as a Source of Genomic Instability in Cancer, *Cell* 168 (2017). 10.1016/j.cell.2017.01.002.
- [3] T. Lindahl, INSTABILITY AND DECAY OF THE PRIMARY STRUCTURE OF DNA, *Nature* 362 (1993) 709-715. 10.1038/362709a0.
- [4] J.A. Imlay, S. Linn, DNA DAMAGE AND OXYGEN RADICAL TOXICITY, *Science* 240 (1988) 1302-1309. 10.1126/science.3287616.
- [5] S.S. Wallace, Base excision repair: A critical player in many games, *DNA Repair* 19 (2014) 14-26. 10.1016/j.dnarep.2014.03.030.
- [6] S.D. McCulloch, T.A. Kunkel, The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases, *Cell Research* 18 (2008) 148-161. 10.1038/cr.2008.4.
- [7] J.C. Fromme, G.L. Verdine, Base excision repair, in: W. Yang (Ed.) *DNA Repair and Replication 2004*, pp. 1-41.
- [8] A.L. Lu, X.H. Li, Y.S. Gu, P.M. Wright, D.Y. Chang, Repair of oxidative DNA damage - Mechanisms and functions, *Cell Biochemistry and Biophysics* 35 (2001) 141-170. 10.1385/cbb:35:2:141.
- [9] T. Lindahl, B. Nyberg, RATE OF DEPURINATION OF NATIVE DEOXYRIBONUCLEIC ACID, *Biochemistry* 11 (1972) 3610-&. 10.1021/bi00769a018.
- [10] T. Lindahl, A. Andersson, RATE OF CHAIN BREAKAGE AT APURINIC SITES IN DOUBLE-STRANDED DEOXYRIBONUCLEIC ACID, *Biochemistry* 11 (1972) 3618-+. 10.1021/bi00769a019.
- [11] C.D. Mol, A.S. Arvai, R.J. Sanderson, G. Slupphaug, B. Kavli, H.E. Krokan, D.W. Mosbaugh, J.A. Tainer, CRYSTAL-STRUCTURE OF HUMAN URACIL-DNA GLYCOSYLASE IN COMPLEX WITH A PROTEIN INHIBITOR - PROTEIN MIMICRY OF DNA, *Cell* 82 (1995) 701-708. 10.1016/0092-8674(95)90467-0.
- [12] C. Rada, G.T. Williams, H. Nilsen, D.E. Barnes, T. Lindahl, M.S. Neuberger, Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice, *Current Biology* 12 (2002) 1748-1755 Pii s0960-9822(02)01215-0. 10.1016/s0960-9822(02)01215-0.
- [13] A.L. Ross, J.E. Sale, The catalytic activity of REV1 is employed during immunoglobulin gene diversification in DT40, *Molecular Immunology* 43 (2006) 1587-1594. 10.1016/j.molimm.2005.09.017.
- [14] E.J. Steele, Mechanism of somatic hypermutation: Critical analysis of strand biased mutation signatures at A:T and G:C base pairs, *Molecular Immunology* 46 (2009) 305-320. 10.1016/j.molimm.2008.10.021.
- [15] D. Svilar, E.M. Goellner, K.H. Almeida, R.W. Sobol, Base Excision Repair and Lesion-Dependent Subpathways for Repair of Oxidative DNA Damage, *Antioxidants & Redox Signaling* 14 (2011) 2491-2507. 10.1089/ars.2010.3466.
- [16] B. Demple, T. Herman, D.S. Chen, CLONING AND EXPRESSION OF APE, THE CDNA-ENCODING THE MAJOR HUMAN APURINIC ENDONUCLEASE - DEFINITION OF A FAMILY OF DNA-REPAIR ENZYMES, *Proceedings of the National Academy of Sciences of the United States of America* 88 (1991) 11450-11454. 10.1073/pnas.88.24.11450.
- [17] D.O. Zharkov, G. Shoham, A.P. Grollman, Structural characterization of the Fpg family of DNA glycosylases, *DNA Repair* 2 (2003) 839-862. 10.1016/s1568-7864(03)00084-3.
- [18] J.C. Fromme, G.L. Verdine, DNA lesion recognition by the bacterial repair enzyme MutM, *Journal of Biological Chemistry* 278 (2003) 51543-51548. 10.1074/jbc.M307768200.
- [19] J.C. Fromme, G.L. Verdine, Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM, *Nature Structural Biology* 9 (2002) 544-552. 10.1038/nsb809.
- [20] M. Sugahara, T. Mikawa, T. Kumasaka, M. Yamamoto, R. Kato, K. Fukuyama, Y. Inoue, S. Kuramitsu, Crystal structure of a repair enzyme of oxidatively damaged DNA, MutM (Fpg), from an

- extreme thermophile, *Thermus thermophilus* HB8, *Embo Journal* 19 (2000) 3857-3869. 10.1093/emboj/19.15.3857.
- [21] R. Gilboa, D.O. Zharkov, G. Golan, A.S. Fernandes, S.E. Gerchman, E. Matz, J.H. Kycia, A.P. Grollman, G. Shoham, Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA, *Journal of Biological Chemistry* 277 (2002) 19811-19816. 10.1074/jbc.M202058200.
- [22] F. Coste, M. Ober, T. Carell, S. Boiteux, C. Zelwer, B. Castaing, Structural basis for the recognition of the FapydG lesion (2,6-diamino-4-hydroxy-5-formamidopyrimidine) by formamidopyrimidine-DNA glycosylase, *Journal of Biological Chemistry* 279 (2004) 44074-44083. 10.1074/jbc.M405928200.
- [23] Y. Qi, M.C. Spong, K. Nam, A. Banerjee, S. Jiralerspong, M. Karplus, G.L. Verdine, Encounter and extrusion of an intrahelical lesion by a DNA repair enzyme, *Nature* 462 (2009) 762-U779. 10.1038/nature08561.
- [24] Y. Qi, K. Nam, M.C. Spong, A. Banerjee, R.J. Sung, M. Zhang, M. Karplus, G.L. Verdine, Strandwise translocation of a DNA glycosylase on undamaged DNA, *Proceedings of the National Academy of Sciences of the United States of America* 109 (2012) 1086-1091. 10.1073/pnas.1111237108.
- [25] K.P. de Jesus, L. Serre, C. Zelwer, B. Castaing, Structural insights into abasic site for Fpg specific binding and catalysis: comparative high-resolution crystallographic studies of Fpg bound to various models of abasic site analogues-containing DNA, *Nucleic Acids Research* 33 (2005) 5936-5944. 10.1093/nar/gki879.
- [26] J.Q. Zhang, J.C. Walter, Mechanism and regulation of incisions during DNA interstrand cross-link repair, *DNA Repair* 19 (2014) 135-142. 10.1016/j.dnarep.2014.03.018.
- [27] D.M. Noll, T.M. Mason, P.S. Miller, Formation and repair of interstrand cross-links in DNA, *Chemical Reviews* 106 (2006) 277-301. 10.1021/cr040478b.
- [28] F. Coste, J.M. Malinge, L. Serre, W. Shepard, M. Roth, M. Leng, C. Zelwer, Crystal structure of a double-stranded DNA containing a cisplatin interstrand cross-link at 1.63 angstrom resolution: hydration at the platinated site, *Nucleic Acids Research* 27 (1999) 1837-1846. 10.1093/nar/27.8.1837.
- [29] K.M. Johnson, N.E. Price, J. Wang, M.I. Fekry, S. Dutta, D.R. Seiner, Y.S. Wang, K.S. Gates, On the Formation and Properties of Interstrand DNA-DNA Cross-Links Forged by Reaction of an Abasic Site with the Opposing Guanine Residue of 5'-Cap Sequences in Duplex DNA, *Journal of the American Chemical Society* 135 (2013) 1015-1025. 10.1021/ja308119q.
- [30] J. Nakamura, M. Nakamura, DNA-protein crosslink formation by endogenous aldehydes and AP sites, *DNA Repair* 88 (2020) 102806. 10.1016/j.dnarep.2020.102806.
- [31] M. Raschle, P. Knipscheer, M. Enoiu, T. Angelov, J.C. Sun, J.D. Griffith, T.E. Ellenberger, O.D. Scharer, J.C. Walter, Mechanism of replication-coupled DNA interstrand crosslink repair, *Cell* 134 (2008) 969-980. 10.1016/j.cell.2008.08.030.
- [32] D.R. Semlow, J.Q. Zhang, M. Budzowska, A.C. Drohat, J.C. Walter, Replication-Dependent Unhooking of DNA Interstrand Cross-Links by the NEIL3 Glycosylase, *Cell* 167 (2016) 498-+. 10.1016/j.cell.2016.09.008.
- [33] P.R. Martin, S. Couve, C. Zutterling, M.S. Albelazi, R. Groisman, B.T. Matkarimov, J.L. Parsons, R.H. Elder, M.K. Sapparbaev, The Human DNA glycosylases NEIL1 and NEIL3 Excise Psoralen-Induced DNA-DNA Cross-Links in a Four-Stranded DNA Structure, *Scientific Reports* 7 (2017) 17438. 10.1038/s41598-017-17693-4.
- [34] R.A. Wu, D.R. Semlow, A.N. Kamimae-Lanning, O.V. Kochenova, G. Chistol, M.R. Hodskinson, R. Amunugama, J.L. Sparks, M. Wang, L. Deng, C.A. Mimoso, E. Low, K.J. Patel, J.C. Walter, TRAP is a master regulator of DNA interstrand crosslink repair, *Nature* 567 (2019) 267-+. 10.1038/s41586-019-1002-0.
- [35] F.A. Ali, L. Renault, J. Gannon, H.L. Gahlon, A. Kotecha, J.C. Zhou, D. Rueda, A. Costa, Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate, *Nature Communications* 7 (2016) 10708. 10.1038/ncomms10708.
- [36] Z.N. Yuan, L. Bai, J.C. Sun, R. Georgescu, J. Liu, M.E. O'Donnell, H.L. Li, Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation, *Nature Structural & Molecular Biology* 23 (2016) 217-224. 10.1038/nsmb.3170.