

Thesis for the Master's degree in Molecular Biosciences

Main field of study in Biochemistry

# Human Neil DNA glycosylases

*Towards structure determination of a  
NEIL-DNA complex*

Anette Johansen



60 study points

Department of Molecular Biosciences

Faculty for Mathematics and Natural Sciences

UNIVERSITY OF OSLO 06/2012

© 2012 Anette Johansen

Human Neil DNA glycosylases

Towards structure determination of a NEIL-DNA complex

Anette Johansen

<http://www.duo.uio.no/>

Print: Reprosentralen, Universitet i Oslo

# Human Neil DNA glycosylases

*Towards structure determination of a NEIL-DNA  
complex*

Anette Johansen





# Acknowledgments

This thesis concludes my Master degree at the Department of Molecular Biosciences, University of Oslo. The laboratory part of this study was carried out externally at the Centre for Molecular Microbiology and Neuroscience (CMBN) and Institute of Medical Microbiology (Rikshospitalet), from January 2011 to June 2012.

Many people deserve a huge thank you for helping me accomplish this project. First of all, I would like to thank my supervisors, group leader Prof. Magnar Bjørås and Dr. Bjørn Dalhus at Rikshospitalet, and my internal supervisor Prof. Jon Nissen-Meyer at the Department of Molecular Biosciences, University of Oslo, for making this project feasible. To Bjørn, for being a great supervisor, always having time to answer my questions and for giving me invaluable guidance in the lab and in the writing process. To Magnar, for always showing interest in my results and for good advices.

I would also like to thank my colleagues at IKB for making such an inspiring and great social environment. A special thanks to Dr. Alexander Rowe for the invaluable help with the formatting and reading of this thesis, and to Mari and Pernille for good advices, help and for always making lab life so much fun.

Several other people also deserve a thank you for helping me accomplish this master project, especially *Studentorchesteret Biørneblæs* for making campus a place for more than studies, and *Nedre Blindern Flisespikkeri* for the critical reading of this thesis.

Finally, I wish to thank my family and friends for always being there for me. A special thanks to Iván, for your calm, encouragement, love and great shoulder massages. *¡Ere' el mejor!*

Oslo, June 2012

Anette Johansen



# Abstract

The repair of DNA damage is crucial for the maintenance of genome stability. In base excision repair, lesion-specific DNA glycosylases recognize and cleave a variety of base lesions. The human Nei-like (NEIL) DNA glycosylases NEIL1, NEIL2 and NEIL3 have specificity for oxidized pyrimidines, in particular when present in single-strand DNA or bubble structures. Even though their substrate specificity is known, the role of these enzymes remains unclear. Whereas the expression of NEIL1 seems to be cell-cycle dependent, NEIL2 is expressed during the entire cell-cycle and NEIL3 expression has been shown in connection with embryonic development. To date, only the 3D X-ray structure of NEIL1 has been solved.

In this thesis, truncated versions of NEIL1 have been designed for co-crystallization and cross-linking with DNA for 3D structure determination by X-ray crystallography. We show that a truncated version of NEIL1, in combination with the crucial selection of an appropriate cryo-protectant, results in a much improved diffraction of a crystal containing a NEIL1-DNA complex. Since no crystal structures of NEIL2 and NEIL3 have been reported, we have screened for crystallization conditions of the free protein of these two enzymes. In relation to this, truncated versions of NEIL3 have been designed to improve the protein stability and expression. We further show that for both NEIL1 and NEIL3, a few residues in difference between truncated versions can be essential for the stability and expression of the enzymes. So far, no crystals have been obtained for NEIL2 and NEIL3.

DNA repair enzymes are interesting drug targets in relation to cancer therapy. Therefore, potential DNA glycosylase inhibitors have been tested on NEIL2 to study their effect on the NEIL2 activity. In this thesis, we show that the same compounds that inhibit other human DNA glycosylases such as OGG1, NTH1 and NEIL1 also seem to inhibit NEIL2.



# Abbreviations

## Abbreviations

---

A	Adenine
aa	Amino acid
AAG	Alkyladenine DNA glycosylase
AP (-site)	Abasic (site)
APE	AP-endonuclease
APS	Ammonium persulfate
BER	Base excision repair
β-ME	β-mercaptoethanol
bp	Base pairs
BPB	Bromphenol blue
BSA	Bovine serum albumine
C	Cytosine
CV	Column volumes
DBS	Double strand breaks
DHT	Dihydrothymine
DHU	Dihydrouracil
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNA pol	DNA polymerase
dNTP	Deoxyribonucleotide triphosphate
dRP	Deoxyribose phosphate
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
E	Glutamic acid
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
FapyA	4,6-diamino-5-formamidopyrimidine
FapyG	2,6-diamino-4-hydroxy-5-formamidopyrimidine
Fpg	Formamidopyrimidine DNA N-glycosylase
G	Guanine
Glu	Glutamic acid
Gh	Guanidinohydantoin
His	Histidine

H2TH	Helix two turn-helix
HhH	Helix-hairpin-helix
HLR	HEAT-like repeat
HR	Homologous recombination
HTX	High throughput crystallization
IC	Inhibitory concentration
IPTG	Isopropyl- $\beta$ -D-Thiogalactopyranoside
K	Lysine
LB	Luria Bertani
Lys	Lysine
mAU	Milli absorption units
3-meA	3-methyladenine
7-meG	7-methylguanine
MES	2-(N-morpholino)ethanesulfonic acid
Met	Methionine
MGMT	O <sup>6</sup> -methylguanine-DNA methyltransferase
MMR	Mismatch repair
MOPS	3-(N-morpholino)propanesulfonic acid
MPD	3-methyl-1,5-pentanediol
mq-H <sub>2</sub> O	Milli-Q filtered and ion-exchanged water
NEB	New England Biolabs
Nei	Endonuclease VIII
NEIL	Nei-like
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
Ni-NTA	Nickel-nitrilotriacetic acid
Nth	Endonuclease III
OD	Optical density
5-OHC	5-hydroxycytosine
5-OHU	5-hydroxyuracil
8-oxoA	7,8-dihydro-8-oxoadenine
8-oxoG	7,8-dihydro-8-oxoguanine
PEG	Polyethylene glycol
PNK	Polynucleotide kinase
Pro	Proline

PUA	Phospho $\alpha,\beta$ -unsaturated aldehyde
Q	Glutamine
Ser	Serine
ssDNA	Single-stranded DNA
SSB	Single-strand break
Sp	Spiroiminodihydantoin
TEMED	Tetramethylethylenediamine
THF	(3-hydroxytetrahydrofuran-2-yl)-methylphosphate
UDG	Uracil DNA glycosylase





# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Genome integrity and DNA modifications . . . . .	1
1.1.1	Response to DNA damage . . . . .	5
1.2	DNA repair . . . . .	6
1.2.1	DNA repair mechanisms . . . . .	7
1.2.2	DNA repair and cancer treatment . . . . .	9
1.3	Base excision repair . . . . .	10
1.4	DNA glycosylases . . . . .	12
1.4.1	Structural classes . . . . .	12
1.4.2	DNA glycosylases repairing oxidative DNA base lesions . . . .	14
1.5	Human endonuclease VIII-like DNA glycosylases . . . . .	16
1.5.1	NEIL1 . . . . .	18
1.5.2	NEIL2 . . . . .	19
1.5.3	NEIL3 . . . . .	20
1.6	Protein X-ray crystallography . . . . .	22
1.7	Aims . . . . .	24
<b>2</b>	<b>Methods</b>	<b>25</b>
2.1	Generation of constructs . . . . .	25
2.1.1	Generation of NEIL1 constructs . . . . .	25
2.1.2	Generation of NEIL2 constructs . . . . .	29
2.1.3	Generation of NEIL3 constructs . . . . .	30
2.1.4	Transformation of cells . . . . .	30
2.1.5	Verification of mutants by DNA sequencing . . . . .	31
2.2	Expression tests . . . . .	32
2.2.1	Expression tests of NEIL1, NEIL2 and NEIL3 . . . . .	32
2.3	Purification of NEIL1, NEIL2 and NEIL3 . . . . .	33
2.3.1	Purification of NEIL1 and NEIL2 . . . . .	33

2.3.2	Purification of NEIL3 . . . . .	34
2.4	Generation of complexes between NEIL proteins and DNA . . . . .	36
2.4.1	Cross-linking of NEIL1 and NEIL2 with abasic DNA . . . . .	36
2.4.2	Mixing of NEIL1, NEIL2 and NEIL3 with THF-DNA . . . . .	37
2.5	Glycosylase activity studies . . . . .	38
2.5.1	Isotope $^{32}\text{P}$ labeling of DNA substrate . . . . .	38
2.5.2	Glycosylase activity assays of NEIL2 . . . . .	38
2.5.3	Inhibition of NEIL2 glycosylase activity . . . . .	39
2.6	Crystallization screening . . . . .	40
<b>3</b>	<b>Results and Discussion</b>	<b>45</b>
3.1	Generation of constructs . . . . .	48
3.1.1	Generation of NEIL1 constructs . . . . .	48
3.1.2	Generation of NEIL2 constructs . . . . .	52
3.1.3	Generation of NEIL3 constructs . . . . .	53
3.2	Expression tests . . . . .	56
3.2.1	Expression tests of NEIL1 . . . . .	56
3.2.2	Expression test of NEIL2 . . . . .	58
3.2.3	Expression tests of NEIL3 . . . . .	59
3.3	Purification of NEIL1, NEIL2 and NEIL3 . . . . .	60
3.3.1	Purification of NEIL1 . . . . .	60
3.3.2	Purification of NEIL2 . . . . .	66
3.3.3	Purification of NEIL3 . . . . .	69
3.4	Generation of complexes between NEIL proteins and DNA . . . . .	74
3.4.1	Cross-linking of NEIL1 with abasic DNA . . . . .	74
3.4.2	Cross-linking of NEIL2 with abasic DNA . . . . .	76
3.4.3	Mixing of NEIL1, NEIL2 and NEIL3 with THF-DNA . . . . .	78
3.5	Glycosylase activity studies . . . . .	81
3.5.1	Isotope $^{32}\text{P}$ labeling of DNA substrate . . . . .	81
3.5.2	Glycosylase activity assays of NEIL2 . . . . .	82
3.5.3	Inhibition of glycosylase activity . . . . .	83
3.6	Crystallization screening . . . . .	85
3.7	Data collection and processing . . . . .	87
3.8	Final conclusions and future work . . . . .	90
	<b>Bibliography</b>	<b>95</b>

# 1 Introduction

## 1.1 Genome integrity and DNA modifications

The carrier of genetic information, DNA, is a dynamic molecule that is continuously exposed to various factors that have an impact on its stability. Exogenous sources such as radiation and chemicals, and endogenous agents from cellular processes such as oxidative metabolism and replication errors, contribute to the loss of genome integrity (Hakem [2008]). Removal of these modifications is crucial to the organism since the genomic damage they may cause can lead to mutagenesis, carcinogenesis and aging (Friedberg [2003]; Lindahl [1993]). Although DNA is subject to a broad range of damaging agents, the genetic information is conserved from generation to generation, implying the presence of a control system that maintains the essential information stored in the DNA molecule. Defects in these DNA repair mechanisms have been shown to lead to hypersensitivity to DNA lesions, accumulation of damage and eventually to metabolic disorders and cancer, emphasizing the importance of a functional DNA repair system (Christmann *et al.* [2003]).

Modification of the DNA can be introduced by a variety of sources. The genome may come in contact with environmental agents such as cigarette smoke, UV light from sunlight, ionizing agents and chemicals, some of the latter being derived from drugs. Nevertheless, the majority of DNA damage in aerobic organisms is believed to be introduced by intracellular factors such as the reactive oxygen species (ROS) generated by mitochondrial respiration and from external ionizing radiation, but also by spontaneous decay and replication errors (De Bont [2004]; Lindahl [1993]; Slupphaug [2003]). UV light and ionizing radiation can induce dimer formation between bases, chemicals may attach adducts, introduce cross-links between DNA strands or produce single- or double stranded breaks. Hydrolysis of bases in nucleotides lead to non-instructive abasic sites, whereas base deamination can change the base-pairing properties. Finally, over 100 oxidative modifications have been identified,

the majority introduced by ROS, affecting both the bases and the sugar-phosphates (Ciccia & Elledge [2010]; Hoeijmakers [2001]). The consequences of these harmful modifications are diverse and include interference with cellular processes including replication, transcription and cell-cycle regulation. Bulky adducts, photoinduced dimers and cross-links are cytotoxic lesions that can stall polymerase activity and thus block replication and transcription, or they can prevent chromosome segregation, eventually leading to cell death. Nucleotide alterations due to hydrolysis, oxidation or base loss are mutagenic lesions that would lead to miscoding and mutations if left unrepaired (Hoeijmakers [2001]). Lesions that affect the DNA bases can be divided into four groups: depurination/depyrimidation, deamination, alkylation and oxidation. These are discussed in detail below.

### **Depurination / Depyrimidation**

A base is lost upon depurination or depyrimidation, producing a non-coding abasic (AP) site. The abasic sites are among the most common endogenous lesions found in DNA and they arise spontaneously due to hydrolysis of the glycosylic bond between the base and the ribose, particularly for the purine bases adenine and guanine (Lindahl [1993]; Lindahl & Nyberg [1972]), or by exposure to ROS, radiation or mutagens. In addition, DNA base repair pathways generate AP-sites. In the base excision repair pathway, the damaged base is removed by a DNA glycosylase to form an AP site as one intermediate in the repair pathway (Dalhus *et al.* [2009]). The remaining ribose moiety of the nucleotide exists in an equilibrium between the ring formed furanose and the linear aldehyde, the latter being reactive and shown to generate cytotoxic interstrand cross-links in DNA (Dutta *et al.* [2007]).

### **Deamination of DNA bases**

The spontaneous hydrolytic deamination which is enhanced by ROS and nitroso compounds of cytosine (C), adenine (A) and guanine (G) forming uracil, hypoxanthine and xanthine respectively, is highly mutagenic and will lead to mispairing if not repaired. Uracil, which has the same base pairing properties as thymine (T), produces a transition from C:G to T:A if not removed (Verri *et al.* [1992]). The resulting keto group in hypoxanthine and xanthine does not have the same binding properties as the lost amino group in adenine and guanine, and generates mutations by A:T to G:C or G:C to A:T transitions, respectively (Kow [2002]).

### Alkylated DNA base damage

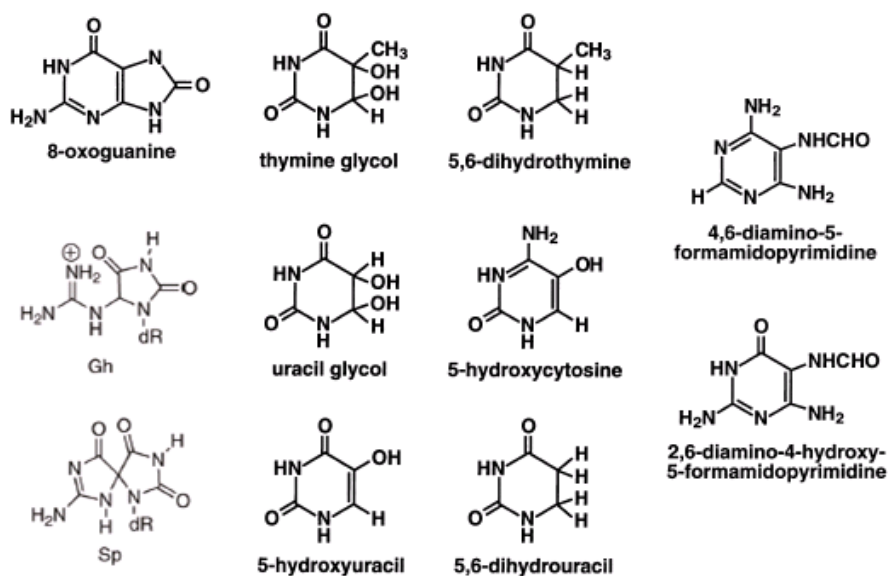
Alkylating agents are present both in the environment and intracellularly. In cells, S-adenosylmethionine (SAM) participates in both enzymatic and non-enzymatic methylation, and is mainly responsible for 7-methylation of guanine (7-meG), but also 3-methylation of adenine (3-meA) (De Bont [2004]; Sedgwick *et al.* [2007]). The random N-alkylation of purine/pyrimidine rings, and O-alkylation of the keto groups of guanine and thymine may be harmful. 7-methylguanine exhibits the same coding properties as guanine, though with a weakened N-glycosylic bond between the base and the sugar, which promotes the formation of a cytotoxic AP-site, while 3-methylguanine is cytotoxic and hinders replication (De Bont [2004]; Drabløs *et al.* [2004]). In addition, the toxic 1-methyladenine is produced in large quantities by methylating agents, and is removed by AlkB (Falnes *et al.* [2002]; Sedgwick *et al.* [2007]).

### Oxidative DNA base damage

Oxidative damage of DNA is introduced by ionizing radiation and endogenous ROS. Organisms that grow aerobically are exposed to ROS as  $O_2^{\bullet-}$ ,  $OH^{\bullet}$  and  $H_2O_2$  during normal metabolism (Evans *et al.* [2004]). Oxidation can change the chemical properties of a base and lead to erroneous base pairing or non-coding derivatives. Representative oxidized bases are illustrated in Fig. 1.1.

One of the most common forms of oxidative damage on DNA purine bases is the oxidation of guanine to 7,8-dihydro-8-oxoguanine (8-oxoG), which in its *syn* conformation will form Hogsteen base pairs with adenosine instead of cytosine. In humans, OGG1 removes the majority of 8-oxoG lesions when opposing cytosine, thus preventing the mutagenic incorporation of adenine opposite to the lesion by the DNA polymerase (Bjørås *et al.* [2002]). In addition, 8-oxoG can be readily oxidized further to guanidinohydantoin (Gh) and spiroiminodihydantoin (Sp), which frequently lead to G→T transversions (Hailer *et al.* [2005]). The oxidized guanine and adenine products 2,6-diamino-4-hydroxy-5-formamidopyrimidine (Fapy-G) and 4,6-diamino-5-formidopyrimidine (Fapy-A) with an opened imidazole ring are also abundant, Fapy-G is known to stabilize a Hogsteen base pair with adenine, and Fapy-A may block DNA synthesis (Wallace [2002]).

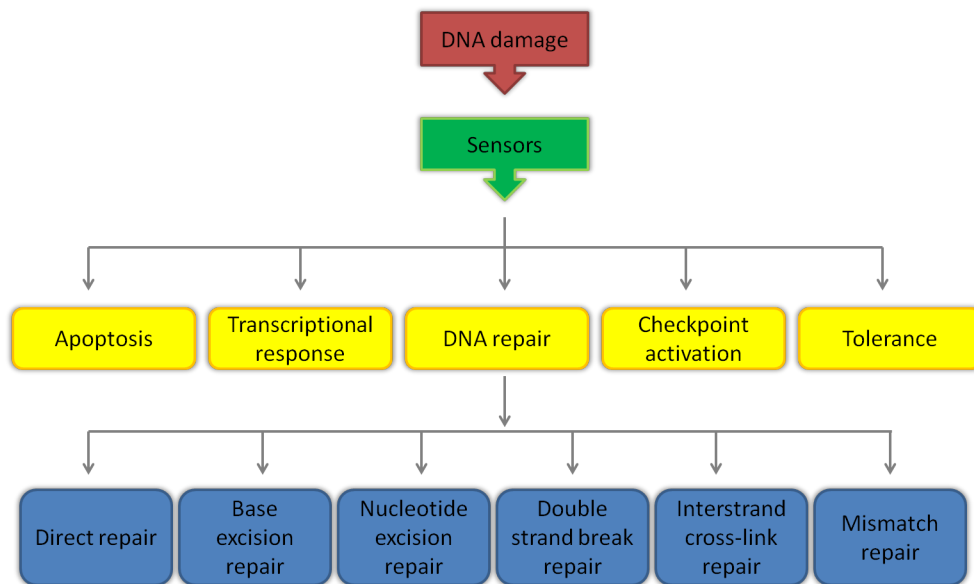
The oxidation of thymine to thymine glycol (Tg) is well studied and the Tg lesion is known to maintain Watson-Crick base pairing with adenine. However, the lesion distorts the structure of the DNA helix, and will eventually stall the polymerase (Wallace [2003]). In addition, thymine can be oxidized to dihydrothymine (DHT). Upon oxidation of cytosine bases, the produced cytosine glycol will be deaminated to form uracil glycol, or dehydrate to 5-hydroxycytosine (5-OHC), whereas 5-hydroxyuracil (5-OHU) can be formed by dehydration of uracil glycol. The *E.coli* DNA glycosylases Nth and Nei have been shown to remove both oxidative thymine and cytosine lesions (Wallace [2003]).



**Figure 1.1:** Oxidized DNA bases. Abbreviations: Gh: guanidinohydantoin; Sp: spiroiminodihydantoin. Figure modified from Wallace [2003] and Zhao et al. [2010].

### 1.1.1 Response to DNA damage

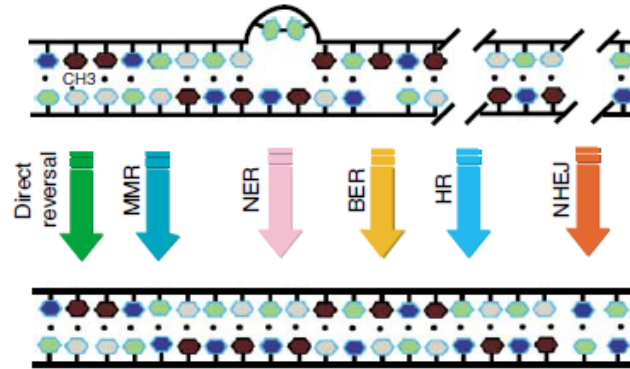
A cell must respond properly in order to maintain genomic stability and cell survival when DNA damage sensors recognize damaged DNA as distinct from undamaged DNA. The five responses that can be triggered are apoptosis, transcriptional alterations, cell cycle arrest, DNA damage repair or DNA damage tolerance (Fig. 1.2) and together they determine the cell's fate, deciding whether it should be allowed to survive, with a possible mutation, or whether programmed cell death should be activated (Madhusudan & Middleton [2005]; Sancar *et al.* [2004]). Usually, the pathways function independently from each other, although proteins specific to one response may participate in other responses. Defects in any of these responses have been shown to eventually lead to genomic instability and degeneration (Sancar *et al.* [2004]).



**Figure 1.2:** DNA damage responses and DNA repair pathways in mammalian cells. Adapted from Madhusudan & Middleton [2005].

## 1.2 DNA repair

The essential role of DNA repair in cell survival and genome integrity is demonstrated by both the highly conserved role of the repair proteins during evolution and the diseases that appear when one or several DNA damage repair mechanisms are deficient. About 130 genes have so far been identified in the human genome that participate in DNA repair (Wood *et al.* [2001]). The repair systems often overlap in their substrate recognition and repair, and internal and external acquired damage are repaired in the same manner as they often produce the same type of DNA lesions. The DNA repair systems are shown to act in connection with cellular processes like cell cycle control, transcription and apoptosis, among others (Slupphaug [2003]). The repair pathways can be divided into several categories: direct repair, mismatch repair, nucleotide excision repair, base excision repair and strand break repair (Fig. 1.3).



**Figure 1.3:** DNA damage repair systems. Abbreviations: MMR: mismatch repair; NER: nucleotide excision repair; BER: base excision repair; HR: homologous recombination; NHEJ: non-homologous end joining. Figure from Hakem [2008].



### 1.2.1 DNA repair mechanisms

#### Direct repair

The one-step reversal of DNA damage *in situ* does not require multiprotein complexes, base excision or backbone breakage, nor does it require several steps to remove the lesion. The direct repair removes mainly alkyl adducts attached by different alkylating agents and pyrimidine dimers formed due to UV-light exposure. (Hakem [2008]; Sedgwick *et al.* [2007]). Whereas photolyases, present in many species but not in humans, recognize pyrimidine dimers and photoproducts by an enzymatic and light-dependent mechanism, methyltransferases are “suicidal enzymes” that become nonfunctional after lesion transfer. In humans, the O<sup>6</sup>-methylguanine-DNA methyltransferase (MGMT), also denoted Agt, restores the damaged nucleotide by transfer of the O<sup>6</sup> methyl group from the guanine to a cysteine residue in the active site. In addition, dioxygenases such as AlkB in *E.coli* and ABH in humans have been shown to repair 1-methyladenine and 3-methylcytosine by an oxidative dealkylation mechanism (Dalhus *et al.* [2009]; Falnes *et al.* [2002]).

#### Mismatch repair

Mispaired nucleotides and insertions or deletions formed by polymerase slippage during replication are repaired by the mismatch repair system (MMR). Mismatch repair can be summarized in 4 steps: recognition of the mismatch, recruitment of repair complexes, identification of the hemi-methylated template strand and degradation of the lesion-containing daughter strand and eventually synthesis of a new strand (Hoeijmakers [2001]). In *E.coli*, the template strand is recognized by MutS, which recruits MutL. The binding activates the endonuclease activity of MutH, the newly synthesized strand is excised and UvrD helicase and several exonucleases remove up to 100 base pairs before DNA polymerase III fills the gap (Jiricny [2006]). In addition to improvement of replication fidelity, the MMR proteins are involved in the regulation of mitotic and meiotic recombination, the affinity maturation of antibodies and possibly in transcription-coupled repair. Deficiency in the mismatch repair system is known to promote cancer (Slupphaug [2003]).

## Nucleotide excision repair

The nucleotide excision repair (NER) machinery removes base lesions that may lead to distortion of the DNA helix structure, including various cross-links, pyrimidine dimers, photoproducts caused by UV radiation and bulky base adducts formed by chemical agents (Huang *et al.* [1992]; Reardon & Sancar [2003]). Substrate specificity is accomplished by over 30 different proteins that participate in lesion recognition and the excision of a 24-32 nucleotide long single-strand DNA stretch, followed by DNA polymerase strand synthesis and sealing of the nick by a ligase (Fagbemi *et al.* [2011]). Two subpathways of NER have been identified: the global genome repair recognizes the majority of lesions which are localized in nontranscribed regions of the genome, while transcription-coupled repair acts upon lesions encountered in transcribed regions (Lindahl [1999]). The importance of this repair system is illustrated in the Cockayne and xeroderma pigmentosum syndromes which cause extreme photosensitivity and predisposal to skin cancer (Laat *et al.* [1999]; Friedberg *et al.* [2004]).

## Base excision repair

Various DNA base lesions, among them the most common damage forms arising from endogenous agents, are excised by base excision repair (BER) enzymes. Lesion-specific DNA glycosylases remove the damaged base by hydrolysis of the N-glycosylic bond, leaving an apurinic/apyrimidinic (AP) site. A single stranded nick 3' to the generated AP site is produced either by the lyase activity of a bifunctional DNA glycosylase, or by an AP-endonuclease. Finally, a new nucleotide is incorporated by the DNA polymerase  $\beta$ , and a ligase seals the nick (Krokan *et al.* [1997]; Seeberg *et al.* [1995]). The BER pathway will be discussed in detail in sec. 1.3.

### Strand break repair

Single and double strand breaks (SSB and DBS, respectively) are produced by ionizing radiation and ROS by-products. In addition, SSB are produced during excision repair and DBS are normal intermediates in processes such as meiosis and V(D)J recombinations in cells of the immune system (Sancar *et al.* [2004]). In mammalian cells, the double strand breaks are repaired by either homologous recombination (HR) or non-homologous end-joining (NHEJ), the choice of which seems to be cell-cycle dependent (Slupphaug [2003]). HR occurs by the crossing-over of two adjacent DNA molecules which thereby provides a template for error-free DNA synthesis before ligation of the double-stranded nick, whereas in NHEJ the DBS are directly rejoined without any template, often in an error-prone manner leading to mutations (Hakem [2008]; Krokan *et al.* [2004]). In SSB repair, the origin of the lesion determines the initial recognition step, whereas the downstream steps occur by a common mechanism, often including enzymes that are involved in the BER pathway (Caldecott [2008]).

### 1.2.2 DNA repair and cancer treatment

Several DNA repair enzymes and pathways are targets of anticancer drugs and treatments. Anticancer drugs exhibit their function by several mechanisms, one of them being to target and introduce lesions to the DNA. Cancer cells divide rapidly and require a repair system that efficiently removes lesions in order to not stall the cell cycle. If the amount of genomic lesions becomes too high, the cell will eventually undergo programmed apoptosis instead of damage repair. It has been shown that DNA repair enzymes can decrease the efficiency of the cancer treatment. Therefore, it is hypothesized that cancer cells can be sensitized and the efficiency of radiation therapy and cytostatic drugs can be increased by selectively inhibiting DNA repair enzymes or pathways (Lord & Ashworth [2012]).

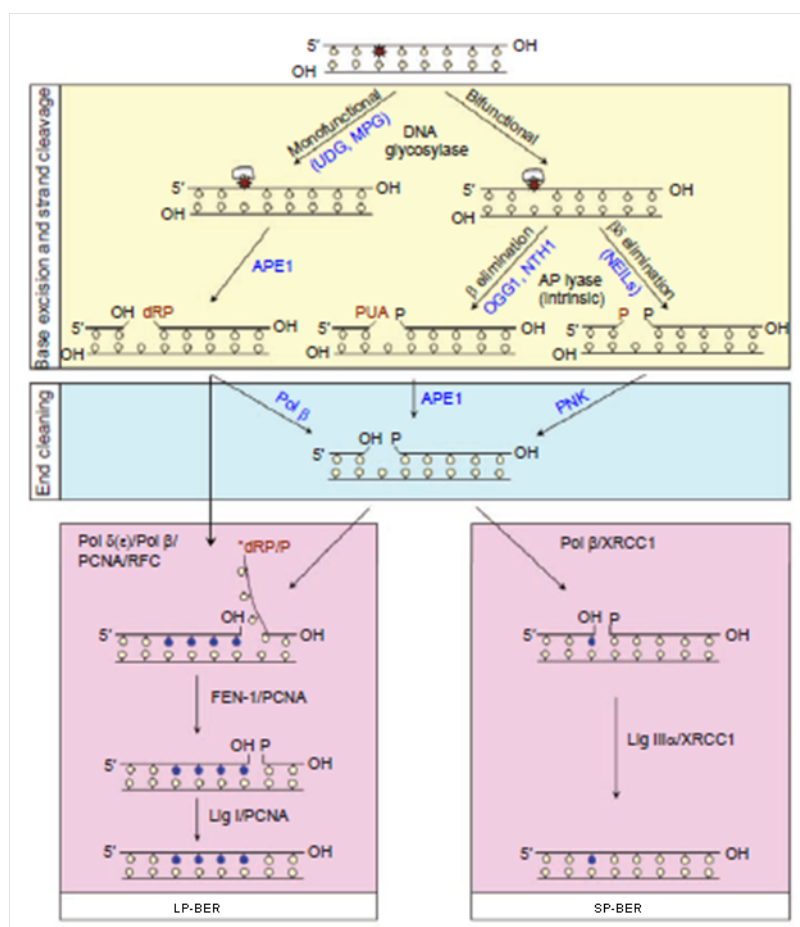
## 1.3 Base excision repair

The base excision repair pathway is a versatile and conserved pathway from prokaryotes to mammals (Hegde *et al.* [2008]). The substrate-specific DNA glycosylases recognize a variety of different lesions caused by deamination, radiation, ROS, alkylating agents and replication errors, including AP sites and single-strand breaks. In contrast to the lesions recognized and removed by NER, most of the lesions processed by the BER system do not cause major helix distortions (Krokan *et al.* [2004]). In addition to the DNA glycosylase, an AP endonuclease (APE) or an AP lyase is required for lesion repair, followed by a DNA polymerase (DNA pol) and a DNA ligase. The DNA glycosylases are classified as either mono- or bifunctional, the latter displaying an additional lyase activity. Base excision and strand cleavage occur through distinct mechanisms for the mono- and bifunctional glycosylases, eventually leading to a common end cleaving step before nucleotide synthesis, as illustrated in Fig. 1.4. The final DNA synthesis is accomplished either through a long-patch pathway or a short-patch pathway, discussed below (Schärer & Jiricny [2001]).

The monofunctional DNA glycosylases hydrolyze the N-glycosylic bond between the damaged base and the sugar with an activated water molecule, producing an abasic site. In mammals, APE1 or a AP lyase cleaves the DNA backbone 5' or 3' to the AP-site to generate a 3' OH and 5' deoxyribose phosphate (dRP) terminus which can be further processed by the lyase activity of the DNA pol  $\beta$  (Alonso *et al.* [2006]; Dalhus *et al.* [2009]). The bifunctional DNA glycosylases possess both glycosylase and AP lyase activity, the latter either through a  $\beta$ -elimination or a  $\beta\delta$ -elimination. When a base damage is recognized, an activated amine group of the enzyme attacks the N-glycosylic bond and the base is excised before the lyase activity incises the strand 3' of the AP-site. The resulting nick differs in its 3' terminal depending on the lyase mechanism. A 3' phospho  $\alpha,\beta$ -unsaturated aldehyde (3' PUA) is produced in the  $\beta$ -elimination reaction while the  $\beta\delta$ -elimination reaction creates a 3' phosphate. The latter is removed by a polynucleotide kinase (PNK) and the 3' PUA is excised by APE1 (Hazra *et al.* [2007]; Hegde *et al.* [2008]).

The BER pathway is divided into two subpathways, denoted short-patch and long patch BER depending on the number of nucleotides that is displaced by the DNA polymerase. In short-patch BER, the single nucleotide that was removed by the DNA glycosylase is re-inserted by a DNA pol  $\beta$ /XRCC1 complex before DNA ligase

III $\alpha$  seals the gap (Hegde *et al.* [2008]; Slupphaug [2003]). The long-patch BER requires many of the same factor that are used in DNA replication, including the DNA pol  $\delta$  and  $\epsilon$ , the sliding clamp PCNA and the flap endonuclease (FEN1 in humans). The DNA pol  $\delta/\epsilon$  displaces up to 10 nucleotides during the DNA synthesis to form a 5' flap extending from the DNA helix. The flap is processed by FEN1 before the gap is sealed by DNA ligase I (Slupphaug [2003]). The choice of short or long-patch BER is believed to depend on, among several factors, the concentration of available BER enzymes and the state of the processed 5' terminal moiety (Dalhus *et al.* [2009]).



**Figure 1.4:** BER subpathways. Base excision and strand cleavage is initiated by monofunctional (left) or bifunctional (right) DNA glycosylases. After end cleaning, DNA synthesis takes place by either long-patch BER (LP-BER; left panel) or short-patch-BER (SP-BER; right panel). The initial damage is represented as a star (\*). Figure adapted from Hegde *et al.* [2008].

## 1.4 DNA glycosylases

The key enzymes in the BER pathway are among the best understood enzyme classes in nucleic acid metabolism. The relatively small monomeric proteins, between 30-50 kDa, are conserved from *E.coli* to mammals and initiate the BER pathway without requiring any cofactors. The majority have a broad substrate specificity with overlapping lesion preferences, underscored by the variety of damage they recognize (Krokan *et al.* [1997]). To date, about 10 different human DNA glycosylases are known (Grin & Zharkov [2011]).

Regardless of the substrate, the DNA glycosylases have several reaction principles in common. They possess a positive DNA-binding surface where DNA interactions, primarily with the lesion-containing strand, occur through salt bridges and hydrogen bonds. For all known DNA glycosylases, a base flipping mechanism bends the damaged base out of the helix and exposes it to the active site. This accommodation of the base allows specific recognition of the lesion, thus discriminating the damaged base from a normal base. Finally, a flexible loop or a residue side chain intercalate into the DNA duplex in order to form base-stacking interactions and to prevent back-flip of the damaged base (Dalhus *et al.* [2009]).

The DNA glycosylases can be divided into five superfamilies based on their overall fold and characteristic motifs. Interestingly, DNA glycosylases that recognize similar substrates do not necessarily have the same tertiary structure, and enzymes in the same structural superfamily do not necessarily recognize the same lesions (Dalhus *et al.* [2009]). The superfamilies are discussed in the following section.

### 1.4.1 Structural classes

#### Uracil DNA glycosylase superfamily

The uracil DNA glycosylase (UDG) was the first DNA glycosylase to be discovered, and was shown to recognize uracil formed due to deamination of cytosine (Krokan *et al.* [1997]). The active site of the enzymes in the UDG superfamily has been shown to be almost entirely conserved between prokaryotes and multicellular organisms, and the typical fold consists of a single  $\alpha/\beta$  domain with 8  $\alpha$ -helices and a central four-stranded parallel twisted  $\beta$  sheet. UDG distinguishes uracil from the much alike thymine by shape complementarity in the recognition pocket, and flips the miscoding

nucleotide out of the DNA helix upon recognition (Mol *et al.* [1995]; Pearl [2000]; Savva & Pearl [1995]). In humans, the UNG DNA glycosylase recognizes and excises uracil (Krokan *et al.* [1997]).

### **Helix-hairpin-helix superfamily**

Members of the helix-hairpin-helix (HhH) superfamily have been identified and their crystal structure solved in all three kingdoms of life, and they are shown to be the most diverse superfamily. The proteins have been named after a conserved motif, the helix-hairpin-helix, involved in DNA binding. Structurally, the enzymes consist of two domains with mainly  $\alpha$  helices and the active site localized in a cleft between the domains. Enzymes that contain a  $\text{Fe}_4\text{S}_4$  cluster are classified in a subfamily of the HhH superfamily (Thayer *et al.* [1995]). Endonuclease III (Nth) and OGG1, which recognize oxidized bases, and MutY which is specific for A:8-oxoG mismatches, all possess the HhH motif, illustrating the absence of overlap between functional and structural properties of the DNA glycosylases (Zharkov [2007]).

### **Helix two turn-helix superfamily**

The N-terminal domain of enzymes in the helix two turn-helix (H2TH) superfamily of DNA glycosylases normally consists of an  $\alpha$ -helix with the catalytic residues Pro2 and Glu3, followed by a two-layered antiparallel  $\beta$ -sandwich structure, and a C-terminal domain with the conserved DNA binding H2TH motif, as well as a zinc finger motif of two antiparallel  $\beta$  strands (Dalhus *et al.* [2009]). The two domains are connected by a flexible linker, forming a positively charged DNA binding pocket located in the cleft between the two domains. The initial methionine is removed during polypeptide maturation, leaving a highly conserved N-terminal PE-helix among members of the H2TH family (Grin & Zharkov [2011]; Zharkov [2007]). Two representative DNA glycosylases from the H2TH family, formamidopyrimidine DNA N-glycosylase (Fpg) and endonuclease VIII (Nei), give an alternative name to the structural superfamily: Fpg/Nei enzymes. The human orthologs of Nei, NEIL1, NEIL2 and NEIL3 also belong to this structural family, and will be discussed in sec. 1.5.

### **Alkyladenine DNA glycosylase superfamily**

The alkyladenine DNA glycosylase (AAG) superfamily is characterized by a single domain of mixed  $\alpha/\beta$  structure with 7  $\alpha$  helices and 8  $\beta$  strands. A curved, antiparallel  $\beta$  sheet make up the protein core, and a protruding  $\beta$  hairpin can insert into the DNA helix upon lesion recognition (Lau *et al.* [1998]). AAG superfamily members have been found in both prokaryotes and mammals (Aamodt *et al.* [2004]). In humans, AAG has been shown to remove a series of alkylated bases as well as deaminated adenine (hypoxanthine) (Sedgwick *et al.* [2007]).

### **HEAT-like repeat DNA glycosylase superfamily**

Prokaryotic AlkC and AlkD DNA glycosylases, shown to remove alkylated bases, were found to belong to the recent discovered HEAT-like repeat (HLR) DNA glycosylase superfamily, without amino acid sequence similarity to any other known protein (Alseth *et al.* [2006]). Homology modeling determined the structure of the enzymes to consist of a superhelix of  $\alpha$ -helical hairpins containing 6 HLR motifs and a positive charged cleft suitable for double-stranded DNA accomodation (Dalhus *et al.* [2007]). So far, no eukaryotic equivalent has been found (Dalhus *et al.* [2009]).

## **1.4.2 DNA glycosylases repairing oxidative DNA base lesions**

The DNA glycosylases that recognize and cleave oxidized DNA bases are divided into two functional classes depending on their substrate preferences for oxidized purines or oxidized pyrimidines (Dalhus *et al.* [2009]). The functional classes do not necessary correspond to the structural classification of the DNA glycosylases. However, the DNA glycosylases that are known to act on oxidative DNA base lesions belong to either the HhH or H2TH structural superfamilies, and they all seem to possess AP/lyase activity (Hazra *et al.* [2007]). Tab.1.1 summarizes the main substrates and the structural folds of the *E.coli* and mammalian DNA glycosylases that act on oxidative DNA base damage.

In *E.coli*, three different DNA glycosylases are known to recognize and excise oxidative damaged DNA bases. Nei and Nth belongs to two different structural families, but do nevertheless excise mainly the same oxidized pyrimidines such as thymine



glycol (Tg), 5-OHU and 5-OH, though with different catalytic efficiencies (Jiang *et al.* [1997]; Wallace [2003]). Fpg has the same overall fold as Nei, but recognizes mostly oxidized purines such as 8-oxoG, FapyA and FapyG (Zharkov *et al.* [2003]).

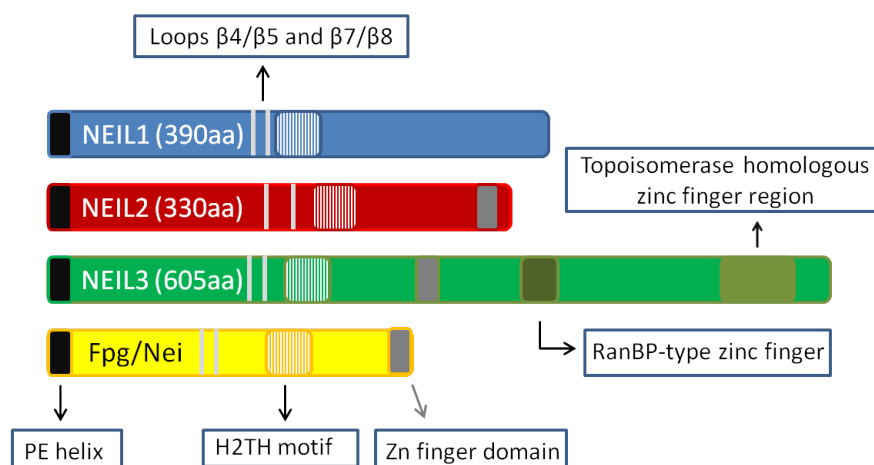
While Fpg is almost exclusively found in bacteria, eukaryotic homologs of Nei have also been discovered in vertebrates, and Nth is widely distributed in all three kingdoms of life. A functional homolog to Fpg that acts on 8-oxoG, Ogg, has been found in several species. Eukaryotic Ogg has low sequence similarity to the prokaryotic Fpg, but the orthologs found in yeast, mouse and human exhibit significant sequence homology between each other (Boiteux & Radicella [1999]). For a long time, it was thought that Nth and Ogg were the only DNA glycosylases in eukaryotes that could repair oxidized bases. However, in 2002 several groups identified eukaryotic Nei homologs based on the sequences of Fpg and Nei (Bandaru *et al.* [2002]; Hazra *et al.* [2002a]; Morland *et al.* [2002]; Takao *et al.* [2002]). The structure of the discovered proteins were shown to resemble that of Fpg and Nei, though with highest sequence similarity to the Nei protein, giving them the term Nei-like (Neil) proteins (Grin & Zharkov [2011]). The human orthologs of Nei have been the subject of study in this thesis, and will be discussed more in detail in the following section.

**Table 1.1:** *E.coli* and mammalian DNA glycosylases acting on oxidative DNA base lesions.

<i>E.coli</i>		
Enzyme	Main substrates	Structural superfamily
Nth	5-OHU, 5-OHC, Tg, DHU, DHT	HhH
Nei	Tg, DHT	H2TH
Fpg	8-oxoG, FapyA, FapyG	H2TH
<i>Mammalian</i>		
Enzyme	Main substrates	Structural superfamily
NTH1	Tg, DHU, DHT	HhH
OGG1	8-oxoG, FapyA, FapyG	HhH
NEIL1	8-oxoG, FapyA, FapyG, Tg, 5-OHU	H2TH
NEIL2	5-OHU, 5-OHC, DHU	H2TH
NEIL3	FapyA, FapyG, Sp, Gh	H2TH

## 1.5 Human endonuclease VIII-like DNA glycosylases

Three human homologs of the *E.coli* Nei DNA glycosylase were in 2002 discovered by several groups (Bandaru *et al.* [2002]; Hazra *et al.* [2002a]; Morland *et al.* [2002]; Takao *et al.* [2002]). The human endonuclease VIII-like (Nei-like; NEIL) DNA glycosylases NEIL1, NEIL2 and NEIL3 exhibited sequence homology to the bacterial Nei and Fpg enzymes, and were predicted to possess the same overall fold typical for the H2TH structural superfamily and to display glycosylase activity on oxidized DNA bases. When identified, NEIL1 and NEIL2 were characterized and shown to act on primarily oxidized pyrimidines (Hazra *et al.* [2002a]; Morland *et al.* [2002]). Recently, DNA glycosylase activity of Neil3 in mouse was shown (Li [2008]), and AP lyase activity has been detected in the human NEIL3 protein (Takao *et al.* [2009]).



**Figure 1.5:** Scheme of the location of the characteristic structural motifs in *Fpg/Nei* and the human *Nei* homologs, adapted from Grin & Zharkov [2011].

In addition to the conserved DNA binding H2TH motif, the NEIL1, NEIL2 and NEIL3 proteins display several structural motifs characteristic to the Fpg/Nei superfamily (Fig. 1.5). First, they all possess a catalytic active N-terminal residue, normally a proline, which amino group is responsible for the nucleophilic attack of the N-glycosylic bond and the formation of a Schiff base. In NEIL3, an N-terminal valine has been shown to exhibit the same nucleophilic properties (Liu *et al.* [2010]). The adjacent glutamic acid residue has also been shown to be required for the glycosylase activity (Bandaru *et al.* [2002]). Next, several residues that are located in

the short connecting loops between the  $\beta$ -strands in the N-terminal domain interact with the DNA strand. Among these a lysine residue in the  $\beta 2/\beta 3$  loop, Lys54 in NEIL1 and Lys81 in NEIL3, coordinates the 5'-phosphate of the damaged deoxynucleotide, whereas other residues in the  $\beta 4/\beta 5$  loop, such as Gly80, Met81 and Ser82 in NEIL1, intercalate with the DNA helix when the damaged base is removed and interact with the base opposite to the lesion (Zharkov [2007]). Finally, the C-terminal domain of the NEIL proteins exhibit a DNA binding zinc finger motif that interact with the major groove of DNA. The zinc finger may vary between the different NEILs, illustrated by the zincless finger present in NEIL1, and the unusual zinc finger domain in NEIL2 distinct from the Nei/Fpg zinc fingers (Das *et al.* [2004]). On the other hand, NEIL3 contains several DNA binding zinc finger domains in the extended C-terminal region in addition to the typical zinc finger domain present in the Fpg/Nei superfamily proteins (Rosenquist [2003]).

Even though the NEILs show overlapping substrate specificity with other eukaryotic BER enzymes, their mode of action in the BER pathway is not identical. After cleavage of DNA backbone by NEIL1 or NEIL2 the remaining 3'-phosphate must be removed by a PNK (polynucleotide kinase) (Hegde *et al.* [2008]). The role of the NEIL enzymes in DNA repair has been questioned. They may serve as back-up enzymes for NTH due to the common substrate specificity, and lack of these enzymes in several eukaryotic species suggests a non-essential role. Furthermore, animals deficient in NEIL enzymes do not show any specific phenotype associated with genomic instability (Grin & Zharkov [2011]). Different from other DNA glycosylases, NEIL 1 and NEIL2 recognize and process damaged DNA bases in single-stranded DNA and in bubble-DNA (Dou *et al.* [2003]). This has led to the hypothesis that the enzymes might be involved in DNA repair during transcription and replication, though no study has so far elucidated the biological significance of the activity on bubble substrates (Grin & Zharkov [2011]).

## 1.5.1 NEIL1

The 44 kDa large NEIL1 with 390 residues is by far the most studied enzyme of the three Nei homologs. When identified, NEIL1 was early shown to be active on oxidized pyrimidines (Bandaru *et al.* [2002]; Hazra *et al.* [2002a]), and to overlap with the substrate specificity of the already characterized OGG and NTH. Later, activity on oxidized purines has been reported (see below). Upon DNA base damage recognition, NEIL1 catalyzes the removal of the damaged base by a  $\beta\delta$ -elimination mechanism, leaving a 3'-phosphate-containing product as a substrate for PNK (Hegde *et al.* [2008]).

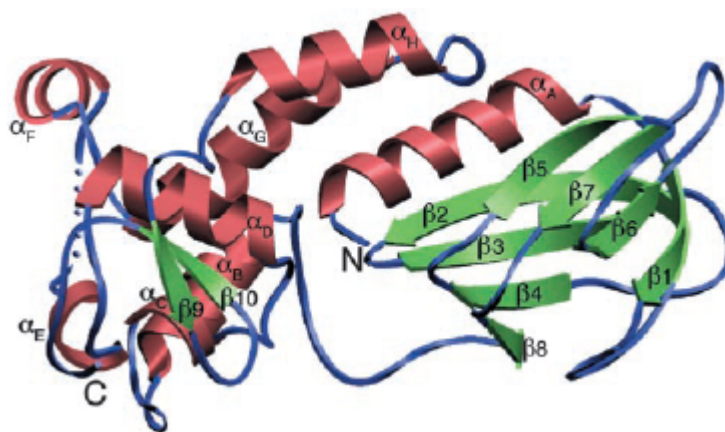
### Substrate specificity

The substrate specificity of NEIL1 is context dependent. The enzyme recognizes primarily oxidized lesions such as 5-OHU, 5-OHC and Tg, but also AP-sites and the ring-saturated products DHT and DHU. Activity on 8-oxoG has been reported in dsDNA but not in ssDNA (Dou *et al.* [2003]; Morland *et al.* [2002]), whereas other studies report low NEIL1 activity (Bandaru *et al.* [2002]; Krishnamurthy *et al.* [2008]). The oxidized guanine lesions Gh and Sp are removed from both single- and double-stranded DNA (Hailer *et al.* [2005]; Zhao *et al.* [2010]), and 8-oxoA is recognized with a cytosine in the opposite strand (Grin *et al.* [2010]). In addition, NEIL1 recognizes FapyA and FapyG (Hazra *et al.* [2002a]). Interestingly, NEIL1 has a general preference for lesions in ssDNA and bubbles rather than in duplex DNA (Dou *et al.* [2003]). In addition, studies have shown that NEIL1 is expressed mainly in S-phase, suggesting a role of the enzyme in replication or transcription repair (Hazra *et al.* [2002a]). Recently, Jaruga *et al.* [2010] reported that NEIL1 seem to be involved in the NER pathway. Hence, the role of the enzyme is not yet fully understood.

### Structure

NEIL1 has been crystallized and its structure solved for a truncated version lacking 56 C-terminal residues, denoted NEIL1 (337aa) (Doublié *et al.* [2004]), after several unsuccessful attempts of crystallizing full-length NEIL1 (Bandaru *et al.* [2004]). The solved structure is shown in Fig. 1.6. NEIL1 (337aa) contains the N-terminal domain corresponding to prokaryotic Fpg/Nei, and was shown to be even more active on Tg

than full-length NEIL1. Interestingly, no electron density beyond residue 290 was identified. Further, NEIL1 does not display any zinc finger motif in the C-terminal domain, and thus not coordinate zinc. The zincless finger is however required for the glycosylase activity, as the activity is greatly reduced when a highly conserved arginine within the motif is removed (Doublié *et al.* [2004]). To date, no crystal structure of NEIL1 in complex with DNA has been reported. Several structures of complexes between prokaryotic Nei and viral Neil and different DNA lesions have been solved. Modeling of NEIL1 in complex with DNA based on these structures reveals significant differences between viral Neil and NEIL1 in the DNA binding region (Imamura *et al.* [2009 2012]). Thus, a structural complex of NEIL1 and DNA is necessary in order to elucidate the DNA binding properties of NEIL1. This thesis focuses on the structural determination of NEIL1 in complex with DNA.



**Figure 1.6:** Crystal structure of NEIL1 as presented by Doublié *et al.* [2004]. The model comprises residues 1-290. Residues 203-207 are disorderd and shown as blue spheres.

### 1.5.2 NEIL2

Even though NEIL2, 36 kDa large with 330 residues, was characterized as a functional DNA glycosylase together with NEIL1 with preferences for oxidized pyrimidines, the exact role of the enzyme is not fully known. NEIL2 exhibits activity on 5-OHU and 5-OHC, particularly in single-stranded DNA and bubble DNA (Bandaru *et al.* [2002]; Hazra *et al.* [2002b]; Morland *et al.* [2002]). The enzyme also excises lesions such as DHT and DHU, in addition to Tg and to some extent 8-oxoG (Dou *et al.* [2003]) .

In the BER pathway, NEIL2 requires the recruitment of PNK to process the produced 5-phosphate nick, and has been shown to make stable interactions with downstream BER enzymes as DNA polymerase  $\beta$ , ligase III $\alpha$  and XRCC1 (Das *et al.* [2006]). In contrast to NEIL1 which is active mainly in S-phase of the cell-cycle, NEIL2 expression has been shown to be cell-cycle independent (Kinslow *et al.* [2010]). This year, (Mandal *et al.* [2012]) found NEIL2 to co-localize with PNK in human mitochondria (Mandal *et al.* [2012]).

To date, no crystal structure of NEIL2 has been solved. Sequence alignment with Fpg/Nei, NEIL1 and NEIL3, suggests that the NEIL2 fold is similar to its homologs, consisting of the conserved catalytic PE helix in the N-terminal domain and a typical H2TH motif. However, the zinc finger motif in the C-terminal domain of NEIL2 is not identical to the conserved zinc finger domains among the Fpg/Nei enzymes. Das *et al.* [2004] have shown that the zinc finger domain is essential for the structural integrity of the enzyme and required for activity. In this thesis, efforts have been done in crystal screening of NEIL2.

### 1.5.3 NEIL3

The third Nei homolog, NEIL3, was discovered at the same time as NEIL1 and NEIL2 (Morland *et al.* [2002]), but much less is known about its biochemical functions. At first, no glycosylase activity was detected for recombinant NEIL3, and it was hypothesized that NEIL3 required *in vivo* modifications to exhibit activity (Morland *et al.* [2002]). Recently, bifunctional DNA glycosylase activity was shown in mouse Neil3 and human NEIL3. The studies showed NEIL3 to display activity on ss AP-sites, but not on ds AP-sites, whereas mouse Neil3 recognized and cleaved Sp, Gh, FapyA and FapyG, preferentially in ssDNA or bubble structures. Furthermore, the expression of NEIL3 in *E.coli* cells deficient in Fpg, Nei and Nth improved the survival of the bacteria after ROS treatment and reduced both the spontaneous mutation frequency and the FapyG level in DNA, suggesting that NEIL3 participate in the FapyG repair *in vivo* (Liu *et al.* [2010]; Takao *et al.* [2009]).

The 68 kDa NEIL3 enzyme with 605 residues comprises a valine in position 2 instead of the absolute conserved Pro2 found in other Fpg/Nei enzymes. The primary amino group in valine serves as a nucleophile in the same manner as the secondary nucleophile of proline. Further, it consists of a long C-terminal part that contains several zinc finger domains in addition to the conserved zinc finger motif found in

Fpg/Nei enzymes. The RanBP-type zinc finger and the duplicated GRF zinc finger, the latter found in a homologous region to the DNA-topoisomerase III $\alpha$ , have been suggested to account for the difficulties in the purification of the full-length protein (Krokeide *et al.* [2009]).

A study of NEIL3 expression levels during mouse brain development showed that the enzyme is highly expressed in embryonic cells when neurogenesis begins, and that the expression decreases as development proceeds. No NEIL3 is detected in adult brain. The expression was specific in areas known to harbor neural stem cells and progenitor cells, indicating that NEIL3 expression is tightly regulated both spatially and temporally (Hildrestrand *et al.* [2009]). Another study has demonstrated the involvement of NEIL3 in stress-induced neurogenesis where NEIL3-deficient mice failed to produce neuronal progenitors and to replace damaged tissue after induced hypoxic-ischemic stroke. Furthermore, NEIL3 was shown to be the main DNA glycosylase in the removal of stress-induced Gh and Sp in proliferating cells even in the presence of NEIL1, suggesting a role in regeneration rather than in postmitotic cells (Sejersted & Hildrestrand [2011]).

## 1.6 Protein X-ray crystallography

X-ray crystallography is a widely used method for structure determination of macromolecules. The individual atoms in a molecule are typically 1-2 Å apart, so X-ray wavelengths between 0.5-1.6 Å are well suited for “imaging” the atomic structure of a molecule. Since the scattering effect from one single macromolecule is too weak, a crystal of highly ordered molecules is required. When a crystal is exposed to X-rays, the scattering from the individual molecules combines by interference from multiple copies of identical unit cells in the crystal lattice to form a diffraction pattern from which the three-dimensional structure can be calculated (Fig. 1.7) (Blow [2002]). The crystallization process, often the experimental bottleneck, consists in reducing the protein solubility under controlled conditions so that the individual protein molecules stack against each other in a highly ordered fashion. A large number of factors affect the protein crystallization, among them, the protein homogeneity and concentration, precipitants, additive, buffers, pH, temperature and the experimental setup (Kundrot [2004]).

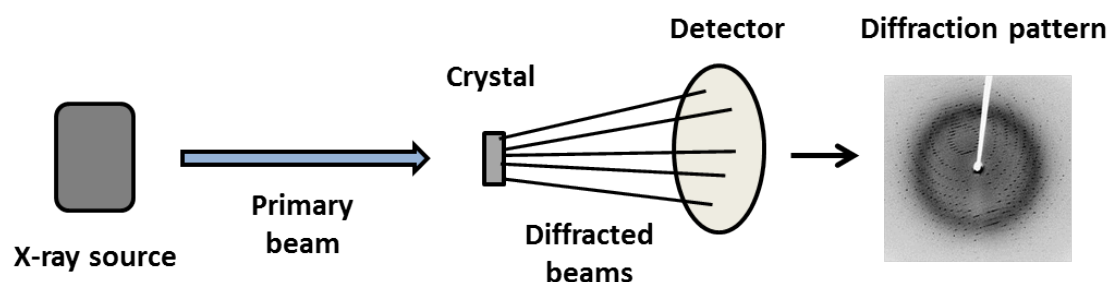
The most common techniques for crystallization screening are vapor diffusion using hanging drops or sitting drops and microbatch under oil. In vapor diffusion, protein and reagent are mixed in a droplet either on a glass plate above the reservoir solution (hanging-drop) or on a droplet platform (sitting-drop). The reagent concentration in the droplet is lower than in the reservoir solution, hence water will evaporate from the droplet until an equilibrium exists between the droplet and the reservoir. In this process, the protein will also concentrate and possibly reach supersaturation. In microbatch experiments, small volumes of protein and reagent are mixed under oil. The oil prevents rapid water evaporation, thus allowing the slow equilibration towards supersaturation of the protein (Chayen [1999]).

Free radicals are produced when crystals are exposed to X-rays. The radicals will eventually damage the crystal, resulting in badly diffracting crystals. By cooling the crystal to around 100 K using liquid nitrogen, the diffusion rate of the free radicals can be reduced. However, crystals can consist of more than 50% water. Thus, a crystal will most probably crack when frozen if it is not added to a protecting cryo-solution. Further, the water and salt from the crystallization buffer will form crystals that can interfere with the protein crystal diffraction or even split the crystal lattice of the protein crystal. Cryo-protectants are small compounds that are more flexible than water molecules, yet with high hydrogen binding capacity and polarity.



When they replace some of the water in the crystal channels, a glass-like phase is produced, preventing the breakage of the crystal when it is flash-frozen (Branden & Tooze [1999]).

Each atom in a crystal will produce scattered X-rays in all directions when the crystal is exposed to X-rays, and the positively interfering waves will give rise to a diffraction pattern. Since X-rays cannot be focused by electric or magnetic fields no lenses can collect the diffraction waves (also known as reflections) and thus conserve the phases from the diffracted waves. Thus, the phases, necessary for solving the protein structure, must be calculated from the amplitudes of the detected reflections through a series of Fourier transform calculations, sometimes even from several data sets. Once the phases are determined, an electron density map is calculated and used to build a model of the protein. During the modeling, several rounds of refinement are performed in order to reduce the differences between the calculated and experimental amplitudes. Finally, a three-dimensional structure model of the protein is obtained that can be used to investigate protein properties such as folding, active site geometry, protein-ligand interactions to mention but a few (Branden & Tooze [1999]).



**Figure 1.7:** *Schematic representation of an X-ray diffraction experiment. The intensity of the diffraction beams is measured by a detector, and can be used to calculate an electron density map.*

## 1.7 Aims

This thesis focuses on the human DNA glycosylases NEIL1, NEIL2 and NEIL3. The exact role of these enzymes in humans cells remains mostly unknown, although they have DNA glycosylase activity for oxidized bases. Whereas the crystal structure of apo NEIL1 has been solved, no structure is known for NEIL1 in complex with DNA. Neither have structures for NEIL2 nor NEIL3 been reported. A structural model will be useful in understanding the substrate specificity and catalytic mechanism of these enzymes. In addition, enzymes that participate in DNA repair can reduce the efficiency of cancer treatment. This therapeutic resistance could possibly be prevented by specific inhibition of these DNA repair enzymes. The aims of this master project have been the following:

### NEIL1

- Design suitable truncated versions of NEIL1 for co-crystallization and cross-linking with DNA and determine the 3D structure of the complex
- Design glycosylase-deficient mutants for determination of the 3D structure of a co-crystallized complex with damage-containing DNA
- Design glycosylase-deficient mutants for glycosylase/lyase activity studies

### NEIL2

- Screen for crystallization conditions and determine the 3D structure of NEIL2 alone or as a cross-linked complex with DNA
- Design glycosylase-deficient mutants for determination of the 3D structure of a co-crystallized complex with damage-containing DNA
- Investigate whether DNA glycosylase inhibitors identified for the enzymes OGG1, NTH1 and NEIL1 also inhibit the DNA glycosylase activity of NEIL2

### NEIL3

- Design truncated versions of NEIL3, optimize the purification protocol and express the protein in sufficient amounts for crystallization screening
- Determine the 3D structure of NEIL3 alone or in complex with DNA
- Design glycosylase-deficient mutants for glycosylase/lyase activity studies

## 2 Methods

### 2.1 Generation of constructs

For primer sequences, recipes and programs, refer to the Appendix.

#### 2.1.1 Generation of NEIL1 constructs

A truncated version of NEIL1, NEIL1 (337aa), was already available in a pET22b vector. The construct gives NEIL1 fused to a C-terminal hexahistidine tag (6x His-tag) following directly after NEIL1. The pET22b also encodes an ampicillin resistance gen for selection. This plasmid was used as a starting point for design of even shorter versions of NEIL1 as described in the introduction. Three different strategies were attempted as described below.

#### Restriction digestion and ligation

To make even more truncated versions of the NEIL1 enzyme, different primers were designed for PCR amplification of constructs A-D (Tab. 2.1). In addition, parallel experiments with primers containing a TEV protease cleavage site between NEIL1 and the 6x His-tag were designed. The forward and reverse primers contained recognition sites for the restriction enzymes NdeI and XhoI, respectively.

**Table 2.1: NEIL1 constructs A-D.**

Parallel	Construct
A	pET22-NEIL1-1-286aa-6xHis
B	pET22-NEIL1-1-286aa-TEV-6xHis
C	pET22-NEIL1-1-290-6xHis
D	pET22-NEIL1-1-290-TEV-6xHis

The template vector, NEIL1 (337aa), was obtained from the glycerol stock by preparation of an overnight culture, and isolation from the cells with the QIAprep Spin MiniPrep Kit (Qiagen). The pET22b plasmid-DNA was eluted in 50  $\mu$ l mq-H<sub>2</sub>O. The 4 different parallels A-D were amplified with their corresponding forward and reverse primers A-D by 1  $\mu$ l of the Pfu high-fidelity polymerase (Fermentas) in a 50  $\mu$ l PCR solution. The PCR reactions were run for 35 cycles. After addition of 10  $\mu$ l of a 6x DNA loading buffer (Fermentas) to the PCR reactions, 50  $\mu$ l of the mixtures were separated on a 0.7% agarose gel (Lonza) by agarose electrophoresis. 50 ml 0.7% agarose solution was mixed with 5  $\mu$ l SybrSafe DNA gel staining (Invitrogen), and 10  $\mu$ l  $\lambda$  DNA/EcoRI + HindIII (Fermentas) was applied as a marker. The gel was scanned and visualized on a Kodak Gel logic 200 image system with UV light. The PCR products were excised from the gel and the DNA purified with the QIAEXII Gel Extraction Kit (Qiagen), eventually eluting DNA in 30  $\mu$ l mq-H<sub>2</sub>O. 16  $\mu$ l of purified product and 32  $\mu$ l of the pET22b vector were double digested with 1 or 2  $\mu$ l of the restriction enzymes NdeI and XhoI (New England Biolabs: NEB) in NEB-buffer #4 at 37 °C for 2-3 hours. 16  $\mu$ l of each amplified PCR product were double digested a total volume of 20  $\mu$ l, while 32  $\mu$ l of the vector was incubated with the restriction enzymes in a reaction volume of 40  $\mu$ l. After incubation, the reactions were subsequently separated on a 0.7% agarose gel and DNA purified from the gel. The inserts A-D and the vector were eluted in 20  $\mu$ l and 30  $\mu$ l mq-H<sub>2</sub>O, respectively.

Ligation of the inserts into the pET22b vector was performed by mixing 12  $\mu$ l insert with 5  $\mu$ l vector together with 1  $\mu$ l T4 DNA ligase (NEB), leaving the reaction at room temperature for 1-3 days. A control reaction contained 12  $\mu$ l mq-H<sub>2</sub>O instead of insert. 50  $\mu$ l electrocompetent *E.coli* ER2566 cells were transformed with 1  $\mu$ l of the ligation reactions, and plated onto LB-agar plates with ampicillin for selection of positive mutants. For procedure, see sec. 2.1.4.

In order to verify that the ligation reaction was successful, a restriction digestion control was carried out. 10 colonies from the incubated agar plates of each of the 4 constructs were picked for inoculation of 5 ml LB-medium containing 100  $\mu$ g/ml ampicillin. After culturing at 37 °C overnight, the DNA was isolated with the MiniPrep Kit, eluting DNA in 50  $\mu$ l mq-H<sub>2</sub>O. 20  $\mu$ l of the plasmid solution were double digested with the restriction enzymes NdeI and XhoI, and 10  $\mu$ l of the solution applied on a 0.7% agarose gel for analysis.

### TOPO cloning

Cloning of the truncated constructs of NEIL1 was performed by use of the TOPO<sup>®</sup> TA Cloning<sup>®</sup> Kit (Invitrogen) as the conventional cloning of the amplified inserts into the vector, as described in the section above, failed to give any positive clones of truncated NEIL1 .

The 4 inserts A-D were amplified by PCR as previously described, though with a Taq polymerase (NEB) adding one overhanging deoxyadenosine (A) to the 3-end of the inserts. The template for the reaction was the same as for the conventional cloning described in the previous section. The supplied pCR<sup>™</sup>2.1-TOPO<sup>®</sup> vector contained one overhanging 3'deoxythymidine (T), facilitating the ligation of the amplified insert into the linear TOPO vector. The TOPO<sup>®</sup> TA Cloning<sup>®</sup> protocol will be described briefly.

4 µl of the amplified inserts were incubated with 1 µl TOPO vector for ligation at room temperature for 30 minutes in a supplied salt solution. 50 µl of the supplied chemical competent *E.coli* OneShot<sup>®</sup> DH5 $\alpha$ <sup>™</sup> TOP 10 cells were transformed with 2 µl of the ligation mix, and incubated at 37 °C overnight on LB-agar plates containing 50 mg/ml antibiotics and 40 mg/ml X-galactose. White colonies were picked for cultivation overnight in 5 ml LB-medium with antibiotics. DNA was isolated from the cells with the MiniPrep Kit. 20 µl DNA was double digested with the restriction enzymes NdeI and XhoI in a control cut reaction, and analyzed on a 0.7% agarose gel as described in the previous section. In addition to the lambda  $\lambda$  DNA/EcoRI + HindIII marker, a 1kb standard (NEB) and a 100 bp standard (NEB) were used.

Promising constructs from the control cut reaction were confirmed by DNA sequencing (see sec.2.1.5 for details), and electrocompetent *E.coli* ER2566 cells were transformed with the verified vectors (sec.2.1.4). Again, DNA was isolated from an overnight culture, either with the MiniPrep Kit, or the NucleoBond Xtra Midi Kit (Macherey-Nagel). To release the insert from the TOPO vector, 34 µl of the isolated plasmid were double digested with the restriction enzymes NdeI and XhoI in a total reaction volume of 40 µl. An empty pET22-b vector was also treated in the same manner. 40 µl of both the inserts and the empty vector were mixed with 8 µl 6x DNA loading buffer (Fermentas), and eventually separated on a 0.7% agarose gel. Isolation from the gel was carried out with the QIAEXII Gel Extraction Kit as previously described, and the inserts and the vector were eluted in 40 µl and 30 µl mq-H<sub>2</sub>O, respectively.

Finally, 12  $\mu$ l purified insert was ligated into 5  $\mu$ l of the empty pET22b vector, and electrocompetent *E.coli* ER2566 cells were used for transformation of 1  $\mu$ l of the reaction mixture. The products were eventually checked by DNA sequencing.

### QuikChange

The QuikChange site-directed mutagenesis protocol (Stratagene) was eventually followed as ligations of the amplified truncated inserts into the pET22b vector were not successful, neither with the conventional cloning or via use of the TOPO vector (see sec. 3.1.1 in Results and Discussion). To obtain truncations of the NEIL1 (337aa) construct, mutagenic primers consisting of two complementary oligomers with the necessary nucleotide mutations were designed.

The various designed NEIL1 truncations are shown in Tab. 2.2. The NEIL1 (337aa) construct was used as a template for the truncations. As the QuikChange protocol does not recommend primers to be longer than 45 bases, the mutations were introduced by the use of two primer sets, thus performing the QuikChange mutagenesis in two consecutive steps.

**Table 2.2: Truncated NEIL1 constructs designed by QuikChange. A 6x His-tag and a stop codon were introduced after residue 286, 305 or 325.**

Enzyme	Full name of construct
NEIL1 (286aa)	pET22-NEIL1-1-286aa-6xHis
NEIL1 (305aa)	pET22-NEIL1-1-305aa-6xHis
NEIL1 (325aa)	pET22-NEIL1-1-325aa-6xHis

The first three histidine residues of the new His-tag following residues 286aa, 305aa or 325aa were introduced in the first round of PCR. The reaction was run for 20 cycles, and the Pfu high-fidelity polymerase was used. The 50  $\mu$ l PCR product was digested with 1  $\mu$ l of the restriction enzyme DpnI (NEB) for 1 hour at 37 °C in order to remove methylated parental DNA. Electrocompetent *E.coli* ER2566 cells were transformed with 1  $\mu$ l of the treated PCR reaction mix, and plated onto LB-agar plates containing ampicillin as a selection marker. Plasmid DNA was eventually isolated from the cells and sequenced (refer to sec. 2.1.5 for details). Correctly mutated products were used as templates in the second PCR reaction, now using the second set of primers for introduction of the final three histidine residues in the

tag and the stop codon. The DpnI digestion, transformation of electrocompetent *E.coli* cells and verification of the constructs by DNA sequencing were repeated.

Eventually, chemical competent *E.coli* BL21 Codon Plus (DE3) RIPL cells were transformed with the verified mutated constructs according to the protocol described in sec. 2.1.4. A glycerol stock was prepared as mentioned below, and stored at -70 °C.

In addition to the truncated versions of NEIL1, the glycosylase-deficient mutants NEIL1-305 E3Q and NEIL1-286 K54Q were designed for use in co-crystallization experiments and for activity studies. The mutants were made by site-directed mutagenesis using the QuikChange protocol (Tab. 2.3). Primer sets with an E3Q or K54Q mutation, respectively, were designed and used for this purpose.

**Table 2.3: NEIL1 glycosylase-deficient mutants made by QuikChange.**

Enzyme	Full name of construct
NEIL1-305 E3Q	pET22-NEIL1-1-305aa-E3Q-6x His
NEIL1-286 K54Q	pET22-NEIL1-1-286aa-K54Q-6x His

### 2.1.2 Generation of NEIL2 constructs

Full-length NEIL2 wild-type (wt) was already available in the lab as a glycerol stock, inserted in a pET22b vector with a C-terminal 6x His-tag. A glycosylase-deficient mutant, denoted NEIL2 E3Q, was designed by introduction of a E3Q mutation (Tab. 2.4) by the QuikChange site-directed mutagenesis protocol as previously described. Transformation into *E.coli* cells, verification of mutagenesis and glycerol stock preparation were performed analogous to the NEIL1 constructs.

**Table 2.4: NEIL2 glycosylase-deficient mutant designed by QuikChange.**

Enzyme	Full name of construct
NEIL2 E3Q	pET22-NEIL2-full-length-E3Q

### 2.1.3 Generation of NEIL3 constructs

An already truncated version of NEIL3 in the expression vector pET-Duet-1 with a C-terminal 6x His-tag, denoted NEIL3 (301aa), was available in the lab. New truncations were designed by QuikChange site-directed mutagenesis (Tab. 2.5). A 6x His-tag and a stop codon were introduced after 282 or 289aa residues, completely analogous to the design of truncated NEIL1.

**Table 2.5: Truncated NEIL3 constructs designed by QuikChange. A 6x His-tag and a stop codon were introduced after residue 282 or 289.**

Enzyme	Full name of construct
NEIL3 (282aa)	pET-duet-NEIL3-1-282-6xHis
NEIL3 (289aa)	pET-duet-NEIL3-1-282-6xHis

In addition, a glycosylase-deficient mutation was introduced into the three truncated construct (Tab. 2.6), by the QuikChange protocol. Transformation into *E.coli* cells, verification of mutagenesis and glycerol stock preparation were performed as for the NEIL1 constructs.

**Table 2.6: Glycosylase-deficient NEIL3 mutants designed by QuikChange.**

Enzyme	Full name of construct
NEIL3-282 K81Q	pET-duet-NEIL3-1-282-K81Q-6xHis
NEIL3-289 K81Q	pET-duet-NEIL3-1-289aa-K81Q-6xHis
NEIL3-301 K81Q	pET-duet-NEIL3-1-301aa-K81Q-6xHis

### 2.1.4 Transformation of cells

#### Transformation of electrocompetent cells

50  $\mu$ l of electrocompetent *E.coli* ER2566 cells were thawed on ice, and transformed with 1  $\mu$ l plasmid by electroporation using 10 pulses of 100  $\mu$ s at 2000 V. Keeping the transformed cells on ice, 950  $\mu$ l of SOC-medium was added to the cells, and the cells were cultured at 37 °C for approximately 40 minutes with shaking. Small aliquots of the cell suspension were spread onto LB-agar plates with 50  $\mu$ g/ml ampicillin. The plates were incubated overnight at 37 °C or at room temperature for 2-3 days.



### **Transformation of chemical competent cells**

15-25 µl of chemical competent *E.coli* BL21 Codon Plus (DE3) RIPL cells (Agilent Technologies) were thawed on ice before addition of 1 µl plasmid. The cells were subsequently heat-shocked at 42 °C for 15 seconds, left on ice for 2 minutes, and incubated for approximately 1 hour at 37 °C with shaking in 1 ml SOC-medium. The cell culture was plated onto LB-agar plates supplemented with 50 µg/ml ampicillin, and incubated overnight at 37 °C or at room temperature for 2-3 days.

### **Preparation of glycerol stocks**

A single colony from a freshly made agar plate was selected for inoculation of 5 ml LB-medium containing 100 µg/ml ampicillin, and cultured overnight at 37 °C. A glycerol stock of the construct was prepared by mixing 1 ml cell culture with 0.5 ml 60% glycerol before storage at -70 °C.

### **2.1.5 Verification of mutants by DNA sequencing**

To verify design of correct mutants and truncations during and after the cloning procedures, 5-10 colonies from freshly made LB-agar plates containing transformed cells carrying the mutated or truncated plasmid were inoculated in 5 ml LB-medium. The cell cultures were grown overnight at 37 °C with shaking. After isolation of plasmid-DNA from the overnight cultures, 150-300 ng DNA was mixed with 1 µl of a 3.5 pmol forward or reverse primer in two separate tubes. After addition of mq-H<sub>2</sub>O to a final volume of 15 µl, sequencing was carried out at the in-house sequencing facility. The primers were specific to the sequence of interest, see the Appendix for details.

## 2.2 Expression tests

### 2.2.1 Expression tests of NEIL1, NEIL2 and NEIL3

To prepare samples for the expression tests of the NEIL1 and NEIL2 proteins, pre-autoclaved LB-medium containing 0.5 M sorbitol and 2.5 mM betaine was inoculated with 10 ml/l overnight culture from a glycerol stock with the construct of interest. For the expression tests of NEIL3 constructs, pre-autoclaved LB-medium was used. Ampicillin was added to a final concentration of 100 µg/ml, and the cultures incubated at 37 °C with shaking at 180 rpm until the cell density reached OD<sub>600</sub> ~ 0.5. The temperature was then lowered to 16 °C. After 1 hour, 1 ml of the cell culture was collected before the protein expression was induced by addition of 0.25 mM IPTG. The cells were cultured overnight, and a 1 ml sample was collected at the end of the induction period.

In addition, an alternative expression test was performed for NEIL1 and NEIL3. Here, the cell cultures were allowed to grow until the cell density reached OD<sub>600</sub> ~ 0.7. Then, the protein expression was induced by addition of 0.25 mM IPTG at 37 °C, and the cells incubated for another 4 hours. A 1 ml sample was collected before induction, and another 1 ml sample 4 hours after induction.

The cell pellets from the 1 ml samples before and after induction were resuspended in 150 µl protein crack buffer. After resuspension and vortexing, the samples were sonicated for 10 seconds and heated to 95 °C for 15 minutes. Finally, 10 µl of the uninduced sample, and 7.5 µl, 5 µl and 2.5 µl of the sample after induction were analyzed on a 12% NuPage gel (Invitrogen) with 5 µl SeeBlue Plus 2 standard (Invitrogen). The proteins on the gel were visualized with Coomassie Brilliant Blue protein stain.

## 2.3 Purification of NEIL1, NEIL2 and NEIL3

To screen for crystallization conditions, large quantities of protein is needed. NEIL1, NEIL2 and NEIL3 protein were therefore purified in large amounts for this purpose in accordance with the following protocols.

### 2.3.1 Purification of NEIL1 and NEIL2

For simplicity, in this section NEIL1, NEIL2 and their mutants will be denoted NEIL as they are all purified in the same way. For buffer details, see the Appendix. The proteins were kept on ice or at 4 °C during the whole purification process.

Pre-autoclaved LB-medium containing sorbitol and betaine were inoculated with 10 ml/l overnight culture made from a glycerol stock of the desired construct. After adding ampicillin (100 µg/ml), the cells were grown at 37 °C with shaking at 180 rpm, until the OD<sub>600</sub> reached ~ 0.5. The temperature was lowered to 16 °C, and the protein expression was induced by adding 0.25 mM IPTG after 1 hour. Then the cells were left to grow overnight. The day after, the cells were harvested by centrifugation at 5000 g for 20 minutes. Protein purification was done from 6 or 12 liters cell culture.

Cell pellets were resuspended in 15 ml cold sonication buffer per liter cell culture, and sonicated on ice for 3 x 30 seconds with an amplitude of 60. The cell debris was centrifuged at 27 000 g for 20 minutes, and the supernatant mixed with Ni-NTA agarose (Qiagen; 1 ml suspension per liter cell culture) by tilting for 20-30 minutes. The mix was applied to an Econo column (BioRad), and the flow-through collected before washing of the column with 15 column volumes (CV) of the sonication buffer. A 50 mM imidazole buffer was added in a volume of 8 CV, followed by elution of NEIL proteins with 5 CV of a 300 mM imidazole buffer. 3-5 fractions of 8-10 ml each were collected from the 50 mM imidazole elution, and 3-5 fractions of 5 ml each were collected from the 300 mM imidazole elution. 15 µl of all the fractions were mixed with 5 µl 4x NuPage loading buffer (Invitrogen) and heated to 70 °C for 10 minutes, before separation on a 12% NuPage gel. Fractions containing NEIL proteins were pooled.

Before further purification, the pooled fractions were desalted by dialysis against a low salt buffer A. After 2 hours, the solution was applied to a pre-equilibrated HiTrap

SPXL column (GE Healthcare) using an Äkta FPLC system (GE Healthcare), and NEIL protein was separated from other proteins in the solution by a salt gradient elution. The column was first washed with 4 CV of the low salt buffer A, before the high salt buffer A concentration (2 M NaCl) was reached after 12 CV. Fractions of 1.5 ml were collected, and eventually separated on a 12% NuPage gel together with the collected flow-through and wash fractions. Fractions with pure NEIL were pooled.

### 2.3.2 Purification of NEIL3

In this section, NEIL3 refers to all the different NEIL3 variants. For buffers details, see the Appendix. All purification steps were carried out by keeping the protein on ice or at 4 °C at all times.

The NEIL3 protein is expressed with a C-terminal 6x His-tag in a pET-Duet-1 vector which encodes an ampicillin resistance gene for selection. LB-medium containing 100 µg/ml ampicillin was inoculated with 10 ml/l overnight culture prepared from a glycerol stock of the desired NEIL3 construct. The cells were incubated at 37 °C with shaking at 180 rpm, until the OD<sub>600</sub> reached ~ 0.5. The temperature was decreased to 16 °C, and the protein expression was induced 1 hour later by adding 0.25 mM IPTG. The cells were cultured overnight, and harvested the next day as mentioned in the previous sections. The purification was done from 12 liters of cell culture. The protocol for the first Ni-NTA agarose purification step was essentially the same as for NEIL1 and NEIL2, except that the batch method was used instead of an Econo column. After mixing of protein and Ni-NTA agarose on a tilt board for 20-30 minutes, the suspension was centrifuged for 10 minutes. The supernatant was removed, and the agarose washed by tilting with 10 CV cold sonication buffer. Centrifugation was repeated, and the wash fraction removed. The same procedure was repeated with both the 50 mM and 300 mM imidazole buffers. The suspension was washed with 8 CV of the 50 mM imidazole buffer, and NEIL3 eluted from the agarose with 2 CV of the 300 mM imidazole buffer. 15 µl of all fractions were mixed with 5 µl 4x NuPage buffer and heated to 70 °C for 10 minutes before separation on a 12% NuPage gel.

Fractions containing NEIL3 were pooled, and concentrated to a volume of 0.5-1 ml in Amicon Ultrafree tubes (Millipore) with a 10K-cut-off filter. Immediately after filtering, the protein was applied to a pre-equilibrated Superdex 75 size-exclusion column (GE Healthcare) using an Äkta FPLC system, and NEIL3 was separated from other remaining proteins by gel filtration. The column was washed with 1 CV of a gel filtration buffer, and fractions of 0.5 ml collected. Pure fractions containing NEIL3 were pooled.

## 2.4 Generation of complexes between NEIL proteins and DNA

In order to co-crystallize NEIL protein with DNA, purified NEIL protein was cross-linked or mixed with DNA before the crystallization screening.

### 2.4.1 Cross-linking of NEIL1 and NEIL2 with abasic DNA

In this section, NEIL1, NEIL2 and their different mutants will be denoted NEIL for simplicity.

Abasic site-containing DNA (AP-DNA) was prepared by first annealing two complementary single-stranded DNA strands (11 or 13mer) at 80 °C for 2 minutes, followed by cooling at room temperature. In the middle of the oligo, a U:G base pair was located. 5 mM duplex DNA and 70  $\mu$ l UDG (uracil DNA glycosylase; NEB) were incubated in a 10 x UDG reaction buffer (NEB) in a total volume of 1 ml at 37 °C for 2 hours, generating the abasic (AP) site by removal of the uracil base.

Already purified NEIL was dialysed against a cross-linking buffer at 4 °C for 2 hours. When necessary, the protein was concentrated to a volume of 6-7 ml. The 6-7 ml of NEIL protein, 1 ml AP-DNA solution and 50 mM NaBH<sub>4</sub> were added to a cross-linking buffer in a final volume of 10 ml, and subsequently incubated at 37 °C for 30 minutes. The cross-linking reaction was stopped by adding 20% glucose and incubated at room temperature for another 30 minutes. The different combinations of NEIL protein and AP-DNA lengths are shown in Tab. 2.7.

**Table 2.7: Name of enzyme and number of nucleotides in AP-DNA used in the cross-linking experiments.**

Enzyme	Number of nucleotides in AP-DNA strand
NEIL1 (286aa)	11
NEIL1 (305aa)	11
NEIL1 (305aa)	13
NEIL2 (full-length)	13

The cross-linked protein solution was sterile filtered, and applied to a pre-equilibrated Resource S column (GE Healthcare) for ion exchange chromatography using an Äkta FPLC system. The elution was carried out for 10 CV, reaching a final concentration of 60% of a high salt buffer B (2M NaCl). The column was washed with 4 CV of a low salt buffer B containing 50 mM NaCl, and both the flow-through and the wash fraction were collected. The eluate was collected in fractions of 1 ml. Eventually, 15 µl of all fractions were mixed with 5 µl 4x NuPage buffer, and heated to 70 °C for 10 minutes before separation on a 12% NuPage gel. Fractions with pure, cross-linked NEIL protein were pooled and concentrated to a final concentration of 8-10 mg/ml. The protein concentrations were determined by the Bradford assay (BioRad).

### 2.4.2 Mixing of NEIL1, NEIL2 and NEIL3 with THF-DNA

In this section, the term NEIL is used for both NEIL1, NEIL2 and NEIL3.

An 11mer single-stranded DNA with a tetrahydrofuran (THF) nucleotide serving as an AP-site analogue, was annealed with its complementary strand by heating to 80 °C for 2 minutes before cooling at room temperature. The resulting double stranded THF-DNA was mixed with purified NEIL protein with a concentration in the range 8-14 mg/ml in a molar ratio 1.2:1, and incubated on ice for 30 minutes. For NEIL1, the protein was mixed with different concentrations of glycerol (up to 50%) before mixing with DNA. The different combinations of NEIL protein and the opposite nucleotide in the complementary strand used in the experiments are shown in Tab. 2.8.

**Table 2.8: Used enzymes and type of nucleotide in the complementary strand opposite to the THF-DNA. All strands contained 11 nucleotides.**

Enzyme	Nucleotide in complementary strand
NEIL1 (305aa)	Adenine
NEIL1 (305aa)	Thymine
NEIL2 (full-length)	Adenine
NEIL3 (301aa)	Cytosine

## 2.5 Glycosylase activity studies

It has been shown that DNA repair enzymes can decrease the efficiency of cancer treatment, making them interesting drug targets. Therefore, possible DNA glycosylase inhibitors already available in the lab were tested on NEIL2. The used oligo sequences and the buffer recipes can be found in the Appendix.

### 2.5.1 Isotope $^{32}\text{P}$ labeling of DNA substrate

A 5-hydroxyuracil-containing DNA oligo (5-OHU) was mixed with 1  $\mu\text{l}$  of the isotopic labeled  $\gamma\text{-}^{32}\text{P}\text{-ATP}$  (Amersham Biosciences) in a 20  $\mu\text{l}$  reaction mix together with 1  $\mu\text{l}$  T4 polynucleotide kinase (PNK; NEB). The reaction mix was incubated at 37 °C for 30 minutes. After 5 minutes of incubation at 80 °C, 2  $\mu\text{l}$  of the labeled oligo was kept as a single strand control, and mixed with 2  $\mu\text{l}$  of a DNA loading buffer. A complementary oligo was added to the remaining solution to give a 5-OHU:G double-stranded substrate. The mixture was further incubated at 90 °C for 2 minutes, followed by 2 minutes of incubation at room temperature, and finally on ice for 5 minutes.

4  $\mu\text{l}$  of the DNA loading buffer were mixed with the double stranded sample, and the single-stranded and double-stranded samples were purified on a 20% native gel. After purification, the gel was kept in a storage phosphor screen (GE Healthcare) for 5 minutes, and subsequently scanned with a Typhoon 9410 Variable Mode Imager (GE Healthcare). The band corresponding to the double stranded substrate was excised from the gel and dissolved in 300  $\mu\text{l}$  mq- $\text{H}_2\text{O}$  overnight at 4 °C.

### 2.5.2 Glycosylase activity assays of NEIL2

The glycosylase activity assays were performed with protein that had been stored on ice, and for samples containing 50% glycerol stored at -20 °C. Since the compounds with potential inhibitor properties for DNA glycosylase activity (see sec. 2.5.3) were all dissolved in DMSO, a series of activity assays with various amounts of NEIL2 were carried out with and without DMSO in order to monitor the effect of DMSO on activity. All assays and dilutions were executed in two separate parallels.

The NEIL2 protein, with or without glycerol, was diluted in 90  $\mu\text{l}$  protein dilution buffer. 7 different protein concentrations between 150 ng/ $\mu\text{l}$  and 0.15 pg/ $\mu\text{l}$  were



used in the activity assays. 3.5  $\mu\text{l}$  mq-H<sub>2</sub>O and 0.5  $\mu\text{l}$  DMSO was mixed before addition of 1  $\mu\text{l}$  NEIL2. The negative control was prepared by replacing NEIL2 in one sample with mq-H<sub>2</sub>O. A dilution of the isotopic labeled 5-OHU substrate was made by mixing 20  $\mu\text{l}$  of the substrate with 380  $\mu\text{l}$  mq-H<sub>2</sub>O. A mastermix was prepared by adding 1  $\mu\text{l}$  of the diluted 5-OHU:G substrate to 2  $\mu\text{l}$  of a reaction buffer, and addition of mq-H<sub>2</sub>O to a total volume of 5  $\mu\text{l}$ . The mastermix was added to the protein samples, and the mixtures incubated at 37 °C for 20 minutes. After incubation at 37 °C for 20 minutes, 10  $\mu\text{l}$  of a stop solution was added to all samples, and the samples heated to 95 °C for 5 minutes. Finally, the samples and the negative control were analyzed on a 20% denaturing gel. The gels were placed on a 3M paper, vacuum dried at 80 °C for 1 hour and subsequently transferred to a Storage Phosphor Screen (GE Healthcare). The gels were scanned after 1-3 days on a Typhoon 9410 Variable Mode Imager, and the glycosylase activity quantified by the ImageQuant TL version 2003.02 program (Amersham Biosciences).

### 2.5.3 Inhibition of NEIL2 glycosylase activity

The compounds A-P with potential inhibitor properties for the glycosylase activity of NEIL2 were investigated by incubation of protein with the different compounds (dissolved in DMSO). It is important to perform these experiments in the linear range of activity, thus NEIL2 concentrations to give 20-30% activity on 5-OHU:G were used. All assays and dilutions were carried out in two separate parallels.

A protein mix consisting of 1  $\mu\text{l}$  diluted protein and 3.5  $\mu\text{l}$  mq-H<sub>2</sub>O was prepared. For the negative control, NEIL2 was exchanged for mq-H<sub>2</sub>O, and 0.5  $\mu\text{l}$  DMSO was added to both the negative and positive control samples. A mastermix was prepared by adding 1  $\mu\text{l}$  diluted substrate to 2  $\mu\text{l}$  reaction buffer and 2  $\mu\text{l}$  mq-H<sub>2</sub>O. 0.5  $\mu\text{l}$  of the 10 mM compound A-P, dissolved in DMSO, was added to 5  $\mu\text{l}$  protein mix in each parallel A-P. 5  $\mu\text{l}$  mastermix was added, and the samples immediately incubated at 37 °C for 20 minutes. As for the titration experiments, 10  $\mu\text{l}$  of a stop solution was added to each sample before heated to 95 °C for 5 minutes and eventually analyzed on a 20% denaturing gel. For the remaining protocol, see the previous section.

## 2.6 Crystallization screening

The NEIL1, NEIL2 and NEIL3 proteins were concentrated to 3.5-15 mg/ml for screening of crystallization conditions that would give well-ordered crystals that diffract to high resolution necessary for solving the structure. Several methods and crystallization screens were used, as indicated in Tab. 2.9, Tab. 2.10 and Tab. 2.11.

### Crystal screening

In the microbatch method, equal amounts of protein and crystallization buffer was combined in one solution and placed under paraffin oil. Using a Oryx6 robot (Douglas Instruments), 0.3  $\mu$ l crystallization buffer was dispensed together with 0.3  $\mu$ l protein sample into the wells of a Vapour Batch Plate (Douglas Instruments). The plates were covered with paraffin oil, and stored at 4 °C or room temperature ( $\sim$  23 °C).

Sitting-drop experiments were as well performed with the Oryx6 robot. Here, 0.33  $\mu$ l crystallization buffer from a pre-dispensed reservoir and 0.67  $\mu$ l protein were mixed to a droplet and dispensed into the sitting-drop position on MRC 2 Well Crystallization plates (Swissci). The plates were sealed with Crystal Clear Sealing Tape (Hampton) and stored at 4 °C. In addition, fractions of NEIL2 protein were shipped to the High-Throughput Crystallization (HTX) facility at the ESRF synchrotron in Grenoble for sitting drop screening at 4 °C. There, only 0.1  $\mu$ l protein sample was required for each crystallization condition.

The manual screening performed by the hanging-drop method was executed at room temperature or at 4 °C. Drops consisting of 1  $\mu$ l crystallization buffer and 1  $\mu$ l protein sample were prepared on silica-coated cover glass (Hampton). The glass plate was sealed over a well of a VDX plate (Hampton) with 500  $\mu$ l crystallization buffer. Both commercial screens and a home-made kit were used in the screening. The conditions for the home-made screen are listed in Tab. 2.12. All screens were stored at 4 °C. In order to obtain larger crystals, micro-seeding was attempted for NEIL1 (305aa) in complex with AP-DNA. Droplets containing small crystals to be used in seeding were transferred to a tube containing the corresponding crystallization buffer and small beads. The small crystals were crushed into tiny fragments by vortexing. The seeds were transferred to already prepared drops in a hanging drop screen at 4 °C

## 2.6 Crystallization screening

using a seeding tool (Hampton). The thin fiber was streaked through the drops at different time intervals (Tab. 2.13).

**Table 2.9: Screening conditions for NEIL1 (305aa) in complex with 11mer AP-DNA.**

Method	Temperature	Protein concentration	Crystallization screen
Sitting drop	23 °C	8-10 mg/ml	Sigma Basic + Extension (Sigma-Aldrich) ProPlex (Molecular Dimensions)
Hanging drop	23 °C	8-10 mg/ml	Kit 1 <sup>a</sup>

<sup>a</sup> For screening conditions in Kit1 refer to table 1.12

**Table 2.10: Screening conditions for NEIL2 (full-length) without DNA and cross-linked to DNA.**

NEIL2			
Method	Temperature	Protein concentration	Crystallization screen
Sitting drop	4 °C	8.5 mg/ml	Crystal Screen Lite & PEG/Ion (Hampton) MembFac & Natrix (Hampton) Index Screen (Hampton) The Classics (Qiagen/Nextal) JCSG (Qiagen/Nextal) Wizard I + II (Emerald BioSystems)
Sitting drop	23 °C	8.5 mg/ml	Sigma Basic + Extension (Sigma-Aldrich) ProPlex (Molecular Dimensions) Alternative Precipitation
NEIL2 cross-linked to 13mer DNA			
Method	Temperature	Protein concentration	Crystallization screen
Sitting drop	23 °C	8 mg/ml	Sigma Basic + Extension (Sigma-Aldrich)

**Table 2.11: Screening conditions for crystallization of NEIL3 (301aa) and NEIL3 (289aa).**

<b>NEIL3 (301aa)</b>			
<b>Method</b>	<b>Temperature</b>	<b>Protein concentration</b>	<b>Crystallization screen</b>
Microbatch	23 °C	8-15 mg/ml	JCSG+ (Qiagen) Index (Hampton) Wizard I + II (Emerald Biosystems)
Hanging drop	4 °C	3.5-11 mg/ml	Index (Hampton) Wizard I + II (Emerald Biosystems)
<b>NEIL3 (289aa)</b>			
<b>Method</b>	<b>Temperature</b>	<b>Protein concentration</b>	<b>Crystallization screen</b>
Hanging drop	4 °C	7-13 mg/ml	Wizard I + II (Emerald Biosystems)

**Table 2.12: Home-made kit 1 screen. Combinations of PEG 3350 and sodium tartrate in wells A1-A6 and B1-B6.**

<b>Sodium-tartrate</b>	<b>PEG 3350</b>					
	<b>10%</b>	<b>12%</b>	<b>14%</b>	<b>16%</b>	<b>18%</b>	<b>20%</b>
<b>150 mM</b>	A1	A2	A3	A4	A5	A6
<b>200 mM</b>	B1	B2	B3	B4	B5	B6

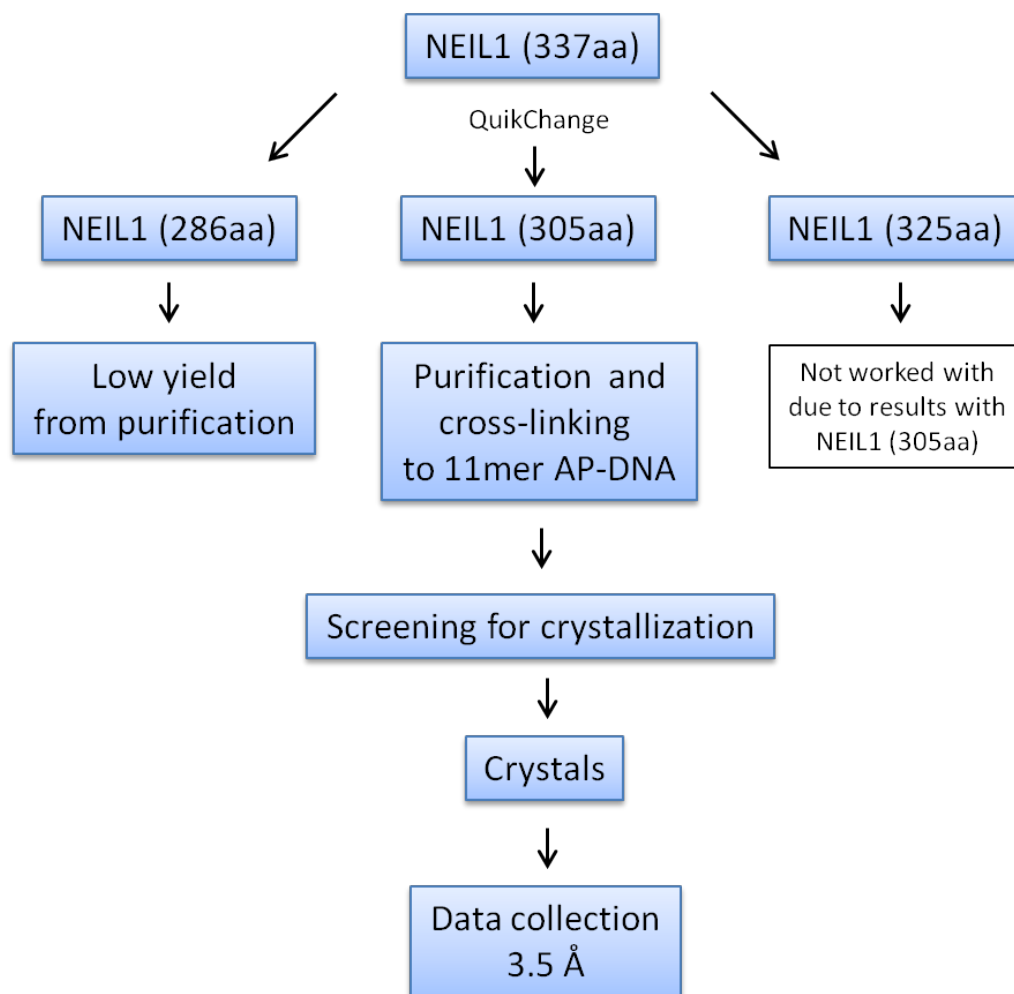
**Table 2.13: Seeding of NEIL1 (305aa) cross-linked to AP-DNA. Combination of screening conditions and seeding time intervals in wells A1-A6 to H1-H6.**

Crystallization conditions	Time interval for seeding (hours)						
	1 - 2 - 4 - 6 - 8 - 16						
10% PEG 3350 150 mM sodium tartrate	A1	A2	A2	A3	A4	A5	A6
10% PEG 3350 200 mM sodium tartrate	B1	B2	B2	B3	B4	B5	B6
12% PEG 3350 150 mM sodium tartrate	C1	C2	C2	C3	C4	C5	C6
12% PEG 3350 200 mM sodium tartrate	D1	D2	D2	D3	D4	D5	D6
14% PEG 3350 150 mM sodium tartrate	E1	E2	E2	E3	E4	E5	E6
14% PEG 3350 200 mM sodium tartrate	F1	F2	F2	F3	F4	F5	F6
16% PEG 3350 150 mM sodium tartrate	G1	G2	G2	G3	G4	G5	G6
16% PEG 3350 200 mM sodium tartrate	H1	H2	H2	H3	H4	H5	H6

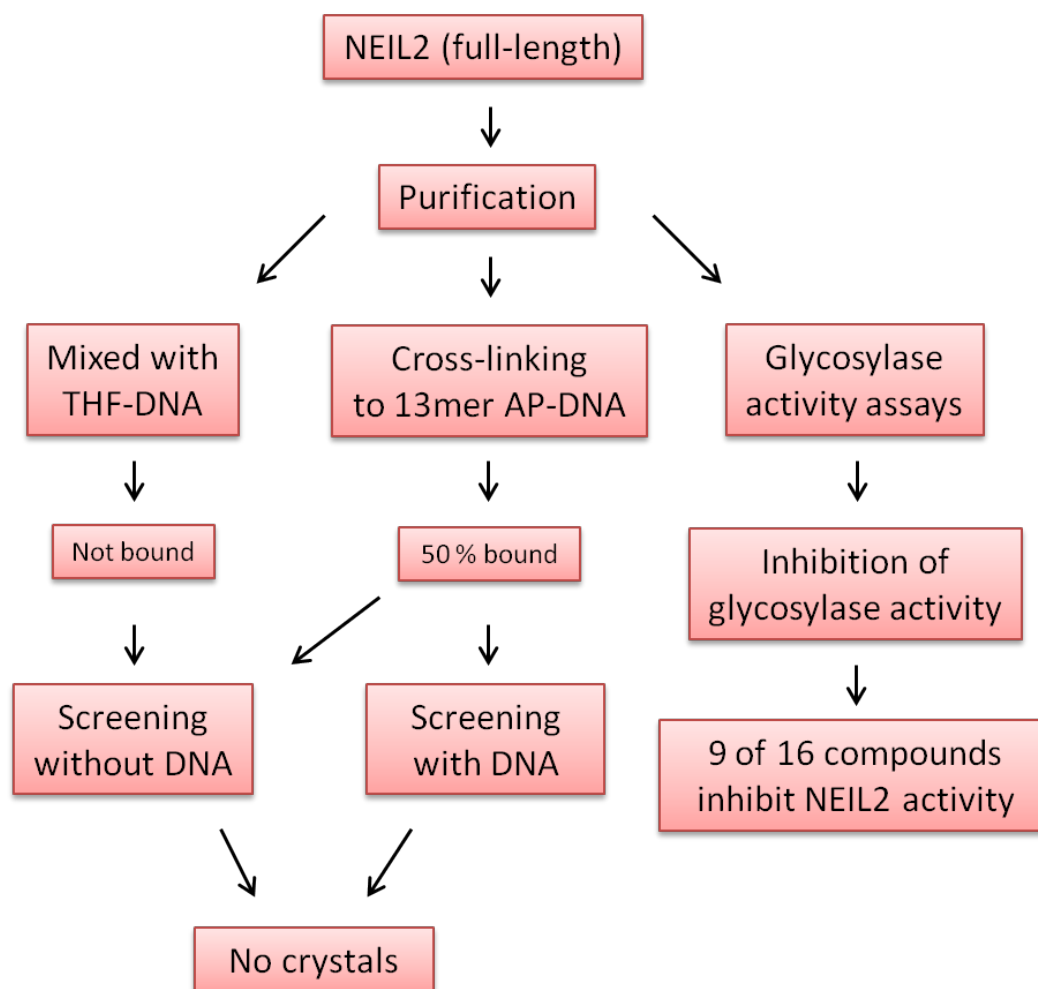
### **Data collection, processing and structure determination and refinement**

Crystals of NEIL1 (305aa) cross-linked with DNA were flash-frozen in liquid nitrogen after a short soak in various cryo-protectant solutions. Diffraction data for the crystals were collected at beamline BL14.1 at Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung (BESSY) in Berlin. The diffraction data were processed and integrated with iMosflm and scaled with Scala in CCP4i (Leslie & Powell [2007]). The Phaser program (McCoy *et al.* [2007]) was used to solve the structures by molecular replacement using NEIL1 as a search model (pdb code: 1TDH; Doubl   *et al.* [2004]). Free R-values were calculated from 5% of all the reflections for monitoring the refinement. Rigid body refinement and restrained refinement of the solved structures were performed with Refmac4 in CCP4i, and the obtained electron density maps ( $2F_o - F_c$  and  $F_o - F_c$ ) were manually inspected in Coot and side chains were checked and adjusted if needed. Due to the low resolution of the data, no water molecules were added to the model. No DNA was added to the model either, as the diffraction density showing the location of the DNA was too undefined to fit nucleotides with certainty.

### 3 Results and Discussion

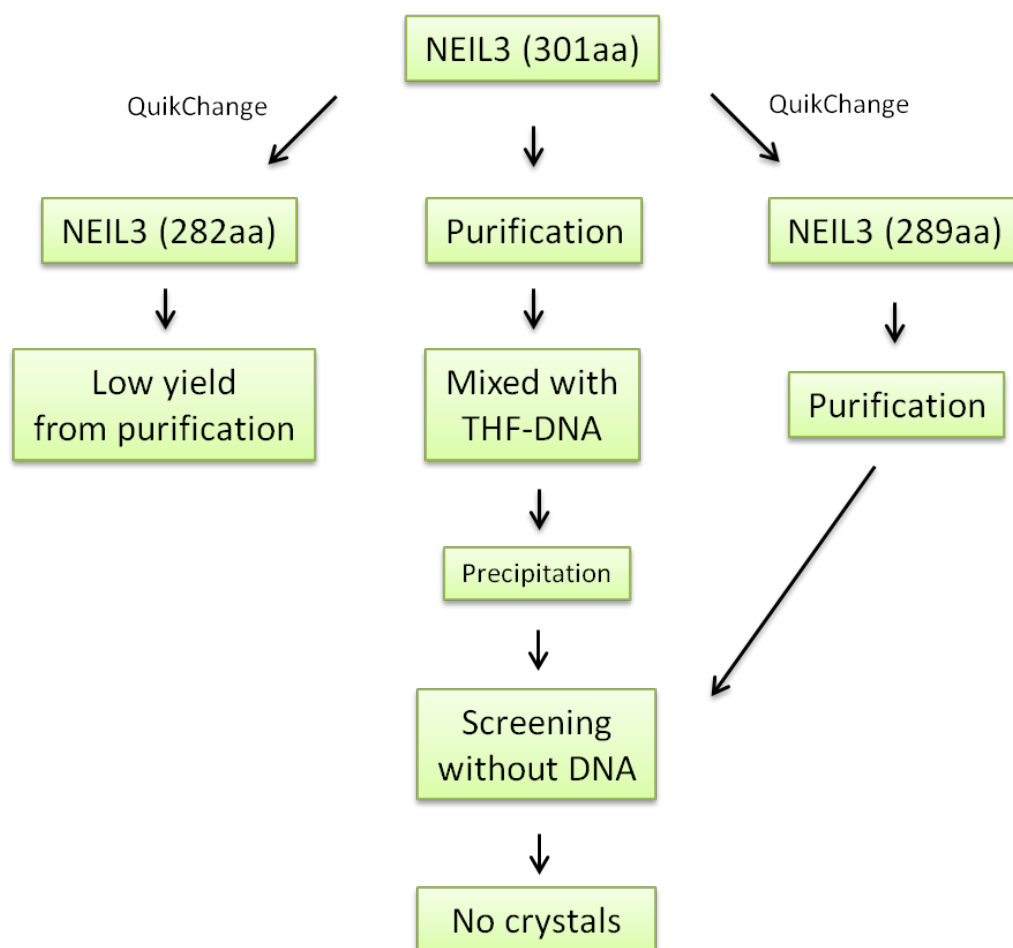


**Figure 3.1:** Flow chart of studied NEIL1 truncations. Three NEIL1 protein constructs were designed by QuikChange, of which one was purified, cross-linked to DNA and obtained crystals from that diffracted to 3.5 Å.



**Figure 3.2:** Flow chart of full-length NEIL2 studies. NEIL2 was purified for crystallization and screened both with and without DNA. In addition, potential DNA glycosylase inhibiting compounds were tested.





**Figure 3.3:** Flow chart of studied *NEIL3* truncations. *NEIL3* (301aa) protein was purified and screened for crystallization without DNA. In addition, two *NEIL3* protein constructs were designed by QuikChange, of which one was purified in high enough yield for crystallization screening.

Figures 3.1, 3.2 and 3.3 summarize the experimental outline for *NEIL1*, *NEIL2* and *NEIL3*, respectively. The designed glycosylase-deficient mutants were not studied and are therefore not included in the flow-charts.

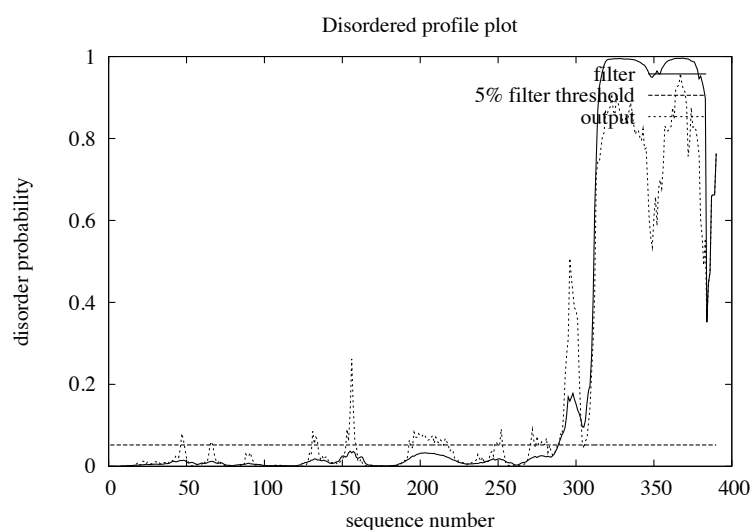
## 3.1 Generation of constructs

### 3.1.1 Generation of NEIL1 constructs

#### Generation of truncated NEIL1 constructs

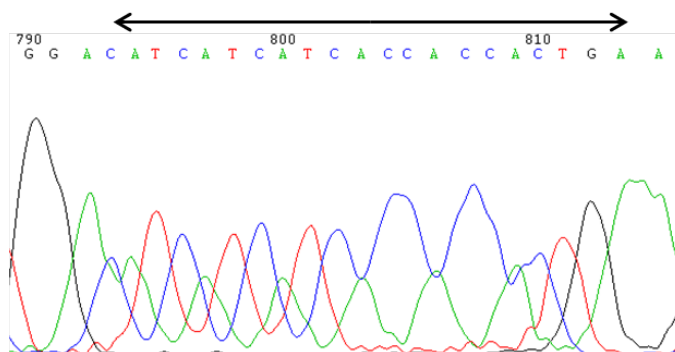
The in-house NEIL1 truncated protein with 337aa amino acids (337aa) has been crystallized in complex with DNA in the lab, but all tested crystals diffracted poorly. Several attempts have been made to optimize the crystallization conditions, however, no improvements in terms of diffraction were obtained. This truncated version of NEIL1 corresponds to the one that has been crystallized without DNA (Doublié *et al.* [2004]). A disordered profile analysis of wild-type NEIL1, shown in Fig. 3.4, suggested that the approximately 100 last residues in the sequence are highly flexible. Such flexible regions could possibly prevent optimal crystallization to occur, and are in general not well suited for crystallization. NEIL1 is known to contain a longer C-terminal sequence than its prokaryotic homologs Fpg and Nei and sequence alignments of human NEIL1 with *E.coli* Fpg and Nei show that the first 281aa residues of NEIL1 correspond to the full length of the Fpg/Nei enzymes (Morland *et al.* [2002]). One of the aims of the current project is thus to design and test even further truncated forms of the 337aa version of NEIL1 in order to hopefully obtain crystals that diffract to a higher resolution. Hence, three truncated forms of NEIL1 with 286, 305 or 325 residues were designed.

The design of truncated NEIL1 versions was not successful with neither conventional cloning of amplified inserts directly into the pET22b expression vector or via use of a TOPO vector. In both cases, amplification of the designed truncations was achieved, but the ligation of the truncated inserts into the pET22b vector failed. All LB-agar plates streaked with ER2566 cells transformed with the ligation reactions or the control reaction contained approximately the same amount of colonies, suggesting that no ligation of the insert had occurred. The reason for the failure of the ligation is unclear, since the restriction sites had been verified by DNA sequencing, and separate control cut reactions with the restriction enzymes had been performed. The activity of the used T4 DNA ligase was also confirmed by other experiments in the lab.

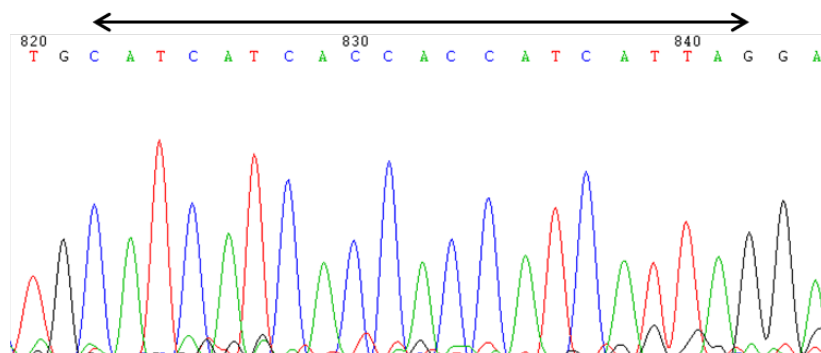


**Figure 3.4:** *Disordered profile plot of full-length NEIL1 (390aa). The C-terminal domain is predicted to be disordered with high probability.*

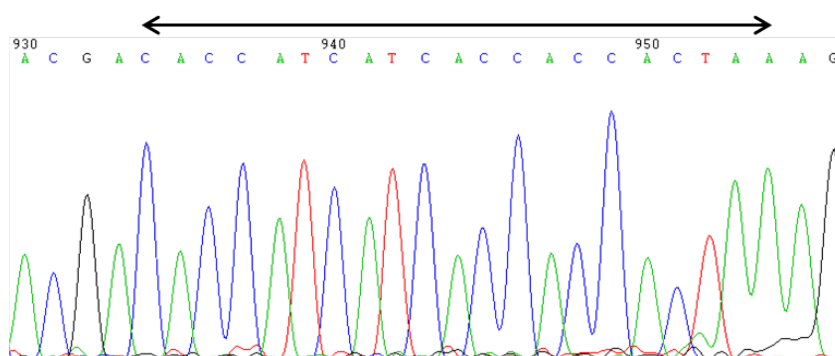
Eventually, truncated versions of NEIL1 were obtained by use of two consecutive rounds of QuikChange site-specific mutagenesis. The in-house NEIL1 (337aa) construct was used as a template for all the truncations. The first three histidine residues of the new His-tag following residues 286aa, 305aa or 325aa were introduced by the first round of PCR, and the methylated DNA template was digested with Dpn I before *E.coli* ER2566 cells were transformed with the PCR reaction and plated onto LB-agar plates. Five colonies were selected for preparation of overnight cultures followed by plasmid-DNA isolation and sequencing. After verification of a successful mutagenesis, the isolated DNA was used as a the template for the second PCR reaction to introduce the final three histidine residues and the stop codon. Finally, verified constructs were transformed into *E.coli* BL21 Codon Plus (DE3) RIPL cells. Glycerol stocks of the constructs were prepared and stored at -70 °C. Sections of the chromatograms of the confirmed mutations are shown in Fig.3.5, Fig. 3.6 and Fig. 3.7. The primer sequences can be found in the Appendix.



**Figure 3.5:** Chromatogram of sequenced NEIL1 (286aa). The introduced 6x His-tag (CAT and CAC) and the stop codon (TGA) are indicated by the arrow.



**Figure 3.6:** Chromatogram of sequenced NEIL1 (305aa). The introduced 6x His-tag (CAT and CAC) and the stop codon (TAG) are indicated by the arrow.

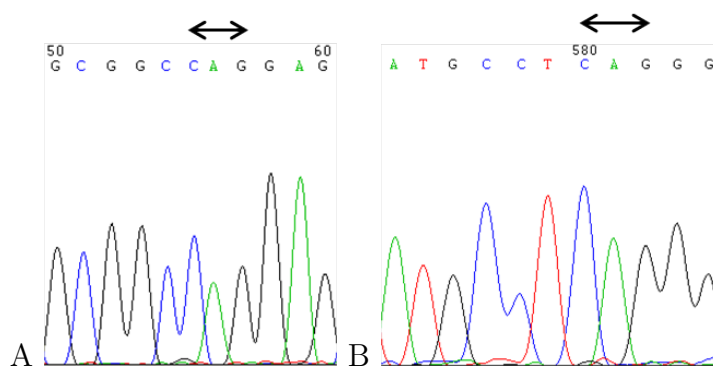


**Figure 3.7:** Chromatogram of sequenced NEIL1 (325aa). The introduced 6x His-tag (CAT and CAC) and the stop codon (TAA) are indicated by the arrow.

### Generation of glycosylase-deficient NEIL1 constructs

The N-terminal PE helix of NEIL1 is known to be the catalytic active site of the enzyme. The amino group of Pro2 acts as a nucleophile by cleaving the N-glycosylic bond and forms a Schiff base intermediate, whereas the carboxylic group of Glu3 interacts with the oxygen atom in the ribose moiety of the damaged nucleotide. Both residues have been shown to be required for the glycosylase and lyase activities as site-directed mutation of these amino acids inactivates the enzyme (Bandaru *et al.* [2002]; Dou *et al.* [2003]). Further, the Lys54 residue coordinates the 5-phosphate of the damaged nucleotide (Wallace [2003]). It has recently been shown that both Pro2 and Lys54 may alternate as catalytic nucleophiles in the glycosylase/AP lyase reaction (Erik S. Vik, unpublished). However, the mechanistic details remain indefinite.

After successful design of the truncated NEIL1 constructs, two glycosylase deficient mutants, NEIL1-286 K54Q and NEIL1-305 E3Q, were designed with the QuikChange method. The NEIL1-286 K54Q mutant was made for glycosylase/lyase activity studies in order to elucidate the role of Lys54 in catalysis, whereas the NEIL1-305 E3Q mutant was designed for use in co-crystallization experiments with damaged DNA. NEIL1 (286aa) and NEIL1 (305aa) were used as templates for the design of NEIL1-286 K54Q and NEIL1-305 E3Q, respectively. Since only one residue was mutated in these constructs, one set of mutagenic complementary primers was sufficient to obtain the mutants. The verified mutations are shown in the sequencing chromatograms in Fig. 3.8.

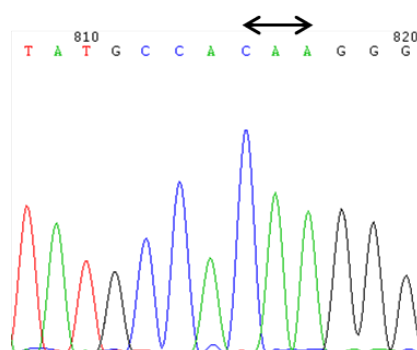


**Figure 3.8:** Chromatogram of sequenced A: NEIL1-286 K54Q and B: NEIL1-305 E3Q. The arrows indicate the altered amino acid (A: Lys = AAG → Gln = CAG. B: Glu = GAG → Gln = CAG).

### 3.1.2 Generation of NEIL2 constructs

NEIL2 has the same overall fold as the NEIL1 and the homologous *E.coli* Fpg/Nei enzymes, and possesses DNA glycosylase/lyase activity with the N-terminal PE-helix being the active site. In addition, the conserved H2TH DNA binding motif is present in NEIL2, but the zinc finger domain is different from the conserved zinc finger domain found in the other Fpg/Nei enzymes. Normally, the Fpg/Nei enzymes contain a zinc finger with four cysteines, whereas in NEIL2 the second cysteine in the zinc finger domain is replaced by a histidine. Mutation of the zinc finger residues inactivates the DNA binding activity (Das *et al.* [2004]). NEIL2 excises 5-hydroxyuracil (5-OHU), 5-hydroxycytosine (5-OHC) and dihydrouracil (DHU). In contrast to NEIL1, NEIL2 has low activity for thymine glycol (Tg) and 8-oxoG (Hazra *et al.* [2002b]). Site-directed mutation of the Glu3 residue in the catalytic active PE-helix has been shown to abolish the glycosylase activity of the Fpg/Nei enzymes (Wallace [2003]). Whereas structures exist for both viral and human NEIL1, no structure is known for NEIL2 (Doublié *et al.* [2004]; Imamura *et al.* [2009]). A lack of glycosylase activity can be utilized in co-crystallization with DNA containing a damaged base. A structure of NEIL2 in complex with DNA will allow a detailed analysis of protein-DNA interactions, with emphasis on the substrate preference.

An E3Q glycosylase-deficient mutant of NEIL2, NEIL2 E3Q, was therefore made by the QuikChange method in the same manner as for the NEIL1-305 E3Q mutant. Successful mutagenesis was confirmed by sequencing (Fig. 3.9).



**Figure 3.9:** Chromatogram of sequenced NEIL2 E3Q. The arrow indicates the altered amino acid Glu = GAA → Gln = CAA).

### 3.1.3 Generation of NEIL3 constructs

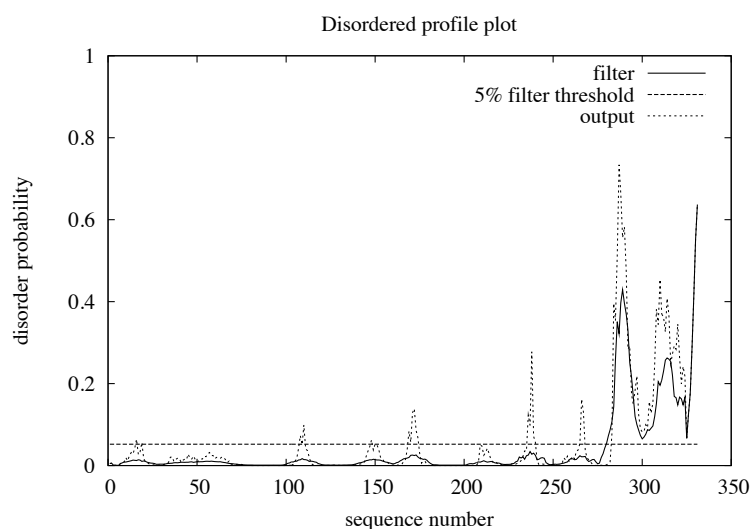
#### Generation of truncated NEIL3 constructs

The functions of the third Nei homolog, NEIL3, still remain unclear and only recently NEIL3 DNA glycosylase activity was shown. In single-stranded and bubble DNA, NEIL3 recognizes and cleaves the oxidized products of 8-oxoG, guanidinohydantoin (Gh) and spiroiminodihydantoin (Sp), in addition to FapyA, FapyG and ss AP-sites (Li [2008]; Takao *et al.* [2009]). The Pro2 in the catalytic active PE-helix in other Fpg/Nei enzymes is exchanged for a Val2 residue in NEIL3 which is shown to display the same catalytic activity as the Pro2 residues (Takao *et al.* [2009]). In addition to the typical Fpg/Nei N-terminal domain and zinc finger domain, NEIL3 contains a long C-terminal part with several zinc finger binding domains not found in other DNA glycosylases (Morland *et al.* [2002]). Removal of the Fpg/Nei-like zinc finger domain abolishes the AP lyase activity, suggesting that as for NEIL2, the zinc finger binding domain is necessary for DNA binding activity.

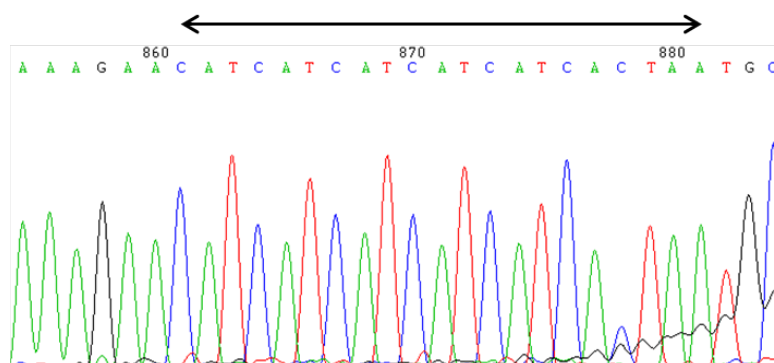
Purification of full-length NEIL3 (605aa) has been attempted without success. Removal of the RanBP-type zinc finger and the duplicated GRF zinc finger in the flexible C-terminal part of the protein, resulting in a 301 residue long NEIL3 version, did however allow purification of the protein (Krokeide *et al.* [2009]). The truncated NEIL3 (301aa) has been subject to screening for crystallization in the lab without any positive outcome (Medya Salah, Master thesis, IMBV/UiO 2010). A disordered profile analysis, shown in Fig. 3.10, indicated that the C-terminal part of the NEIL3 (301aa) structure could be flexible, and thus possibly preventing optimal crystallization. Sequence alignments of NEIL3 and the prokaryotic homologs Fpg and Nei show the first 282aa residues of NEIL3 to correspond to the Fpg/Nei enzymes (Morland *et al.* [2002]). NEIL3 constructs with C-terminal truncations corresponding to the lengths of the Fpg/Nei sequences were therefore made.

The in-house construct of truncated NEIL3 with 301aa amino acids (301aa) construct was used as a template for QuikChange site-directed mutagenesis to design the new 282aa and 289aa truncations of NEIL3, NEIL3 (282aa) and NEIL3 (289aa), respectively. Similar to the NEIL1 protocol for the generation of truncated constructs, two set of mutagenic primers were designed and the new His-tag and stop codon were introduced by two consecutive rounds of PCR.

The verification of successful mutagenesis by sequencing is shown in Fig. 3.11 and Fig. 3.12. *E.coli* BL21 Codon Plus (DE3) RIPL cells were eventually transformed with the mutated plasmids, and glycerol stocks with the mutants were prepared and stored at -70 °C.

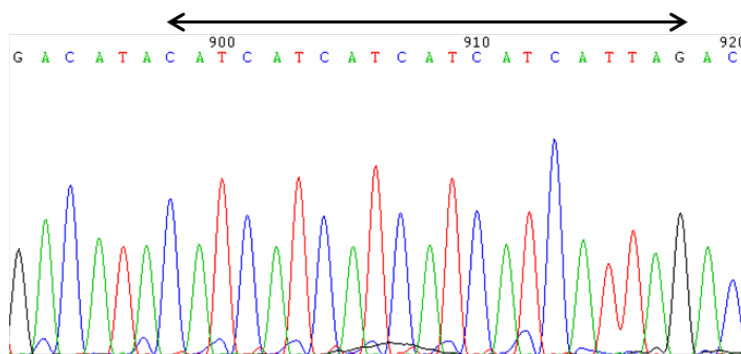


**Figure 3.10:** *Disopred analysis of NEIL3, showing that the C-terminal may be flexible.*



**Figure 3.11:** Chromatogram of sequenced NEIL3 (282aa). The introduced 6x His-tag (CAT and CAC) and stop codon (TAA) following residue 282aa are indicated by the arrow.

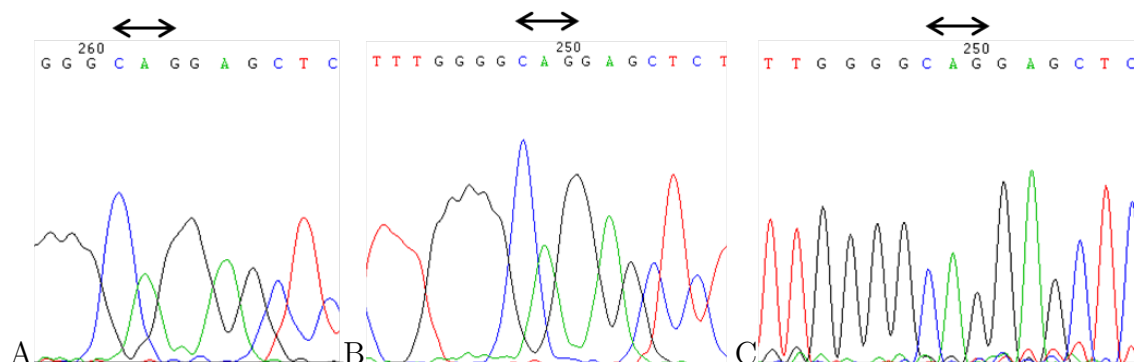




**Figure 3.12:** Chromatogram of sequenced NEIL3 (289aa). The arrow indicates the introduced 6x His-tag (CAT) and the stop codon (TAG) after residue 289aa.

### Generation of glycosylase-deficient NEIL3 constructs

NEIL3 has DNA glycosylase and AP lyase activity for oxidized derivatives of guanine, including guanidinohydantoin (Gh) and spiroiminodihydantoin (Sp) (Liu *et al.* [2010]). In order to test the role of the catalytic lysine Lys81 in damage recognition and base removal, two putative glycosylase deficient mutants, NEIL3-282 K81Q and NEIL3-289 K81Q, were designed by the QuikChange method. The NEIL3-282 K81Q and NEIL3-289 K81Q mutants were made for glycosylase/lyase activity studies in order to elucidate the biochemical functions of the altered residues. NEIL3 (282aa) and NEIL3 (289aa), respectively, were used as templates for the mutagenesis. Since only one residue was mutated in these constructs, one set of complementary primers was sufficient to obtain the mutants. The verified mutations by sequencing are shown in Fig. 3.13.



**Figure 3.13:** Chromatogram of the sequenced A: NEIL3-282 K81Q; B: NEIL3-289 K81Q; and C: NEIL3-301 K81Q. The arrows indicate the altered amino acid (Lys = AAG → Gln = CAG).

## 3.2 Expression tests

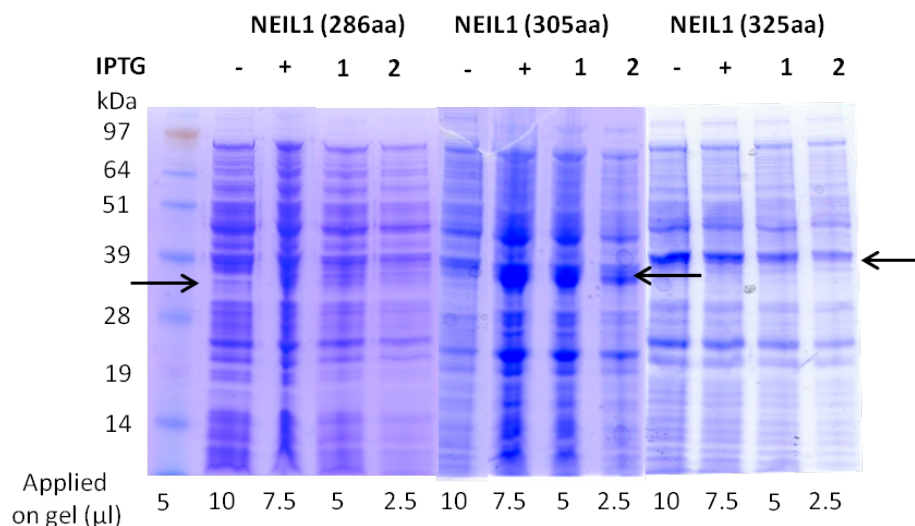
### 3.2.1 Expression tests of NEIL1

Overnight cultures from the glycerol stocks that contained the truncated NEIL1 (286aa, 305aa and 325aa) proteins were used for inoculation of larger cultures with LB-medium containing sorbitol, betaine and ampicillin. When the cell density reached an  $OD_{600} \sim 0.5$ , the temperature was lowered to 16 °C and 1 ml of the uninduced cell culture was collected before the protein expression was induced by adding 0.25 mM IPTG. The next day, a sample from the induced cell culture was, together with the uninduced sample, analyzed on a denaturing gel. In addition, an alternative expression at 37 °C for 4 hours was tested for the NEIL1 (286aa) protein. The cell culture was allowed to reach a cell density of  $OD_{600} \sim 0.7$ , and an uninduced sample was removed before protein expression was induced by adding 0.25 mM IPTG at 37 °C. After 4 hours, an induced sample was collected for analysis on a denaturing gel. For both protocols, the expression level was measured by mixing the uninduced and the induced sample with a protein crack buffer before analyzing the samples on a NuPage gel.

No expression of the NEIL1 (286aa) protein was observed, neither at 16 °C overnight nor at 37 °C. No expression was however visible (Fig. 3.14). The alternative expression method employed at 37 °C for 4 hours did not improve the protein expression (result not shown).

The NEIL1 (305aa) protein, on the other hand, was expressed to some extent. The results from the expression test are shown in Fig. 3.14. The band corresponding to NEIL1 (305aa) is stronger in the induced culture than in the uninduced culture, compared to other bands in the samples. Even though the expression of NEIL1 (305aa) was not very high, it was still higher than for the more truncated NEIL1 (286aa) construct. Somehow, the 19 residues in difference between the constructs are of importance for protein expression.

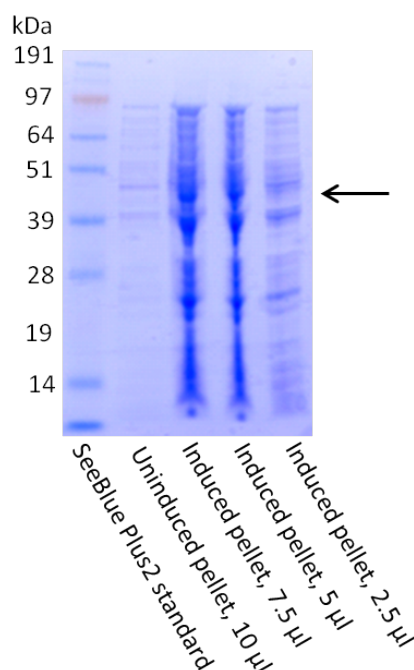
Surprisingly, the expression test for NEIL1 (325aa) at 16 °C, shown in Fig. 3.14, revealed that the protein expression was lower for this construct than for the 10 residue shorter NEIL1 (305aa) construct. The NEIL1 (325aa) protein is expected to appear around 37kDa. A protein band of the expected size is visible, however, the same band is detected in the sample from the uninduced cell culture and is probably an expressed bacterial protein.



**Figure 3.14:** Expression tests of NEIL1 constructs at 16 °C. NEIL1 (286aa) is shown in the left panel, NEIL1 (305aa) in the middle and NEIL1 (325aa) in the right panel. The arrows indicate where NEIL1 protein expression was expected to be visible, whereas the dilutions of induced cell culture are denoted “1” and “2”.

### 3.2.2 Expression test of NEIL2

Full-length NEIL2 (wt) was expressed in a pET22 vector, and an overnight culture was prepared from the glycerol stock. One liter LB-medium containing sorbitol, betaine and ampicillin was inoculated with 10 ml of the overnight culture. The protein expression was induced with 0.25mM IPTG at 16 °C similar to the NEIL1 constructs. Figure Fig.3.15 shows that the NEIL2 protein is hardly expressed at all; only a narrow band corresponding to the size of NEIL2 is visible. Similar to the analysis of NEIL1 expression, three thin bands at approximately 25, 39 and 45 kDa are visible, suggesting the bands to be bacterial proteins. However, purification of NEIL2 from a 6 liter culture is still possible with quite good yield.

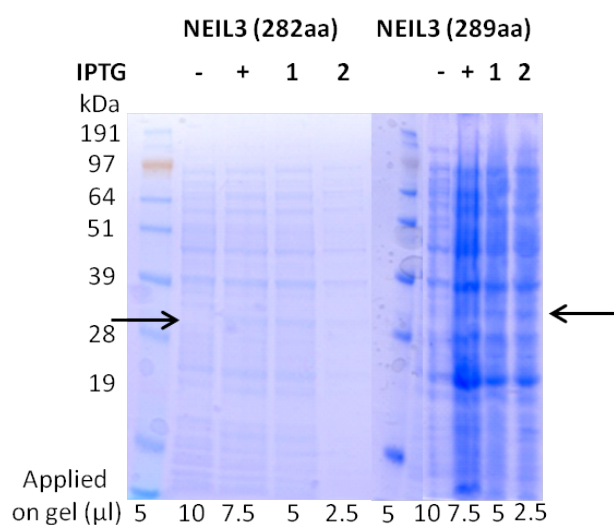


**Figure 3.15:** *Expression test of NEIL2 at 16 °C. The arrow indicates where NEIL2 expression was expected to be seen.*

### 3.2.3 Expression tests of NEIL3

Overnight cultures with the NEIL3 (282aa and 289aa) constructs were prepared from their respective glycerol stocks, and the expression tests were performed in LB-medium containing ampicillin, allowing the cells to grow until the cell density was  $OD_{600} \sim 0.5$ . The protein expression was induced by adding 0.25mM IPTG 1 hour after the temperature was decreased to 16 °C, and the cell cultures were incubated overnight. A sample for the analysis was removed before induction, and the next day.

The expression tests of the two truncated constructs showed hardly any induction at all (Fig. 3.16). Similar to the NEIL1 (286aa) construct, an alternative expression test at 37 °C was performed for the NEIL3 (282aa) construct, without any increase of the protein expression (not shown). Interestingly, it was possible to purify the NEIL3 (289aa) construct to quite a good yield, but not the 7 residue shorter NEIL3 (282aa) construct. The detection of the three bands at approximately 25, 39 and 45 kDa in the NEIL1, NEIL2 and NEIL2 expression tests indicates that the bands correspond to bacterial proteins.



**Figure 3.16:** Expression tests of NEIL3 proteins at 16 °C. Left panel: NEIL3 (282aa); right panel: NEIL3 (289aa). The arrows indicate the expected size of expressed protein, whereas the dilutions of the induced samples are denoted “1” and “2”.

### 3.3 Purification of NEIL1, NEIL2 and NEIL3

Although some of the induction tests suggested low expression of the different versions of NEIL1, NEIL2 and NEIL3, as shown in the previous section, each protein was purified from IPTG-induced cell cultures to test if enough protein could still be obtained for crystallization.

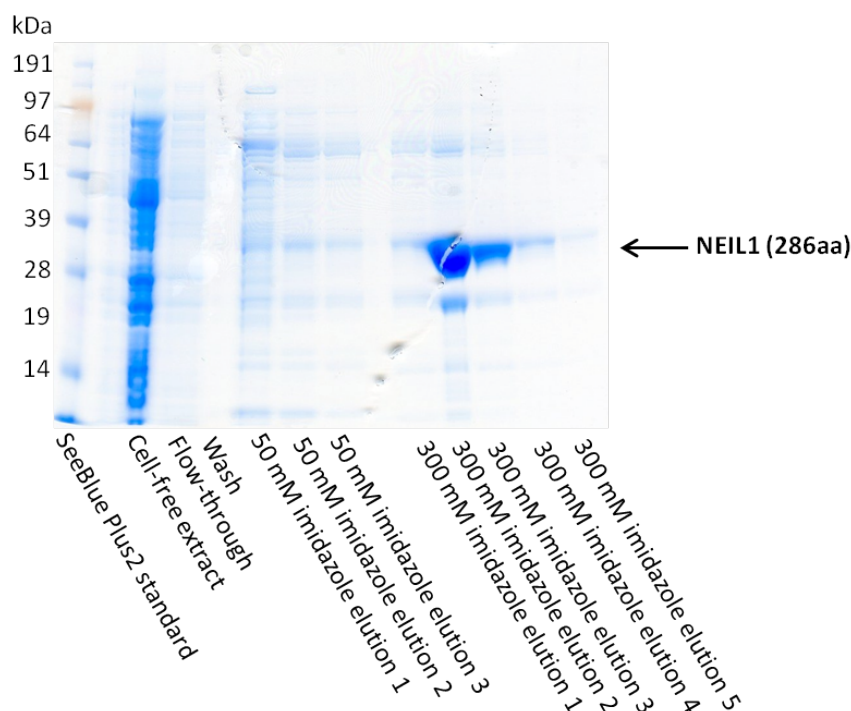
#### 3.3.1 Purification of NEIL1

An already established purification protocol for the original truncation of NEIL1, NEIL1 (337aa), was followed for the purification of the new truncated versions of NEIL1. Briefly, a glycerol stock of the different NEIL1 proteins was used to inoculate LB-medium containing sorbitol, betaine and ampicillin. The protein expression was induced at 16 °C by adding 0.25mM IPTG and the cells were harvested the next day by centrifugation. Cell pellets were resuspended in a cold sonication buffer and lysed by ultrasonication to disrupt the cell membrane. Eventually, cell debris was removed by centrifugation and the supernatant was mixed with already equilibrated Ni-NTA agarose. The His-tag present on the proteins' C-terminal has a high affinity for the Ni<sup>2+</sup> ions, allowing the protein to bind strongly to the Ni-NTA agarose while bacterial proteins bind to a lower extent or do not bind at all. After mixing on a tilt board for 20-30 minutes, the suspension was applied to an Econo column, and the flow-through collected. The column was washed with 10 mM and 50 mM imidazole buffers, before elution of NEIL1 protein with a 300 mM imidazole buffer. All fractions were collected.

None of the proteins were sufficiently pure after this first purification step, and required further purification. Fractions rich in NEIL1 were pooled and desalted by dialysis, and applied to a HiTrap SP XL cation exchange column for elution using a salt gradient. NEIL1 is expected to bind to the negatively charged column material due to the positive charges in proximity to the DNA-binding region on the protein surface. By increasing the salt concentration, the salt ions will bind to the column material and to the protein surface. Since the surfaces of different proteins are likely to be distinctly charged, a steady increase in salt ions via a gradient will lead to elution of the different proteins from the column at different salt concentrations. This may allow a good separation of NEIL1 from the remaining proteins in the protein extract.

#### Purification of NEIL1 (286aa)

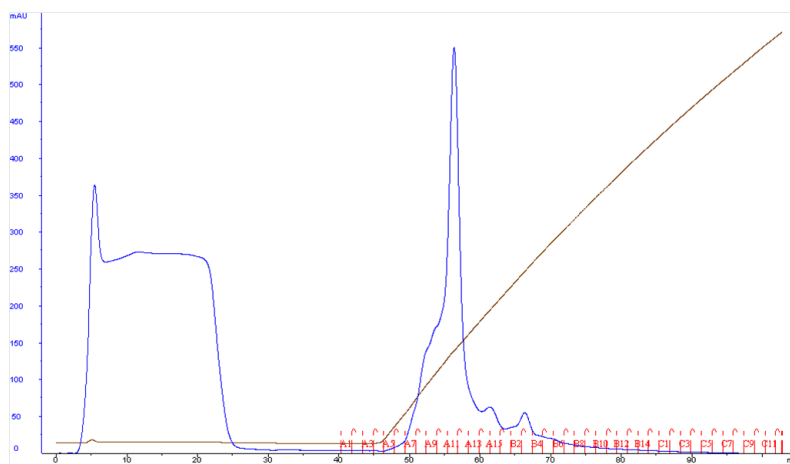
Even though the expression test of NEIL1 (286aa) did not show any strong induction of the protein, purification with Ni-NTA agarose was carried out from 12 liter cell culture. The NEIL1 (286aa) protein was actually expressed to some extent, allowing it to be purified (Fig. 3.17). Most of the protein eluted at 300 mM imidazole, although some NEIL1 was also observed in the 50 mM imidazole fractions.



**Figure 3.17:** Purification of NEIL1 (286aa) with Ni-NTA agarose, analyzed on a 12% NuPage gel.

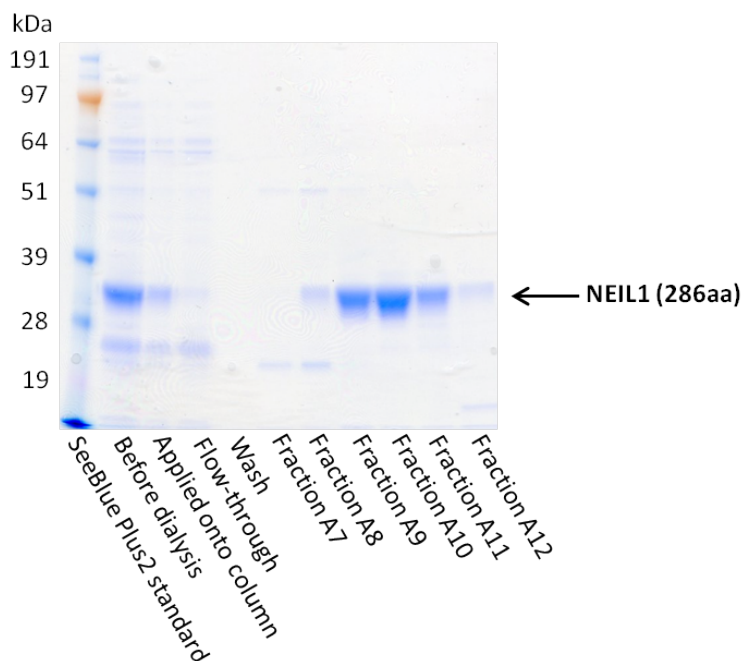
The NEIL1 (286aa) protein was not stable in the imidazole buffer and precipitated. The rapid precipitation was a challenge since a high yield is required for crystallization screening. Compared to the original truncated version of NEIL1, for which the purification protocol had been established, the NEIL1 (286aa) protein exhibited a very different behavior. The protein expression was much lower, and the protein was much less stable than NEIL1 (337aa).

To avoid precipitation, different approaches were attempted. First, the entire purification protocol had to be accomplished in one day since the protein was unstable in the 300 mM imidazole buffer; leaving the samples in the 300 mM imidazole buffer overnight lead to heavy precipitation. To decrease the time spent on the first Ni-NTA purification step, the batch method (explained in section 2.3.2) was employed. The highest yields were achieved when the batch method was used. The dialysis step, which also led to precipitation, was replaced by rapid dilution of the protein in a low salt buffer. This allowed faster application of the protein onto the HiTrap SP column for further purification without any loss of protein due to a time consuming dialysis. Fortunately, the NEIL1 (286aa) protein was stable for several days in the elution buffer used in the ion exchange step. The chromatogram of the purification is included in Fig. 3.18, and the subsequent gel analysis of selected fractions is shown in Fig. 3.19.



**Figure 3.18:** *Chromatogram of the NEIL1 (286aa) purification by ion exchange chromatography with a HiTrap SPXL column. Blue graph: elution profile at 280nm; brown line: salt gradient. Absorbance is measured in milli absorbance units (mAU), and the x-axis indicates elution buffer volume in ml. Fractions are shown in red on the x-axis.*



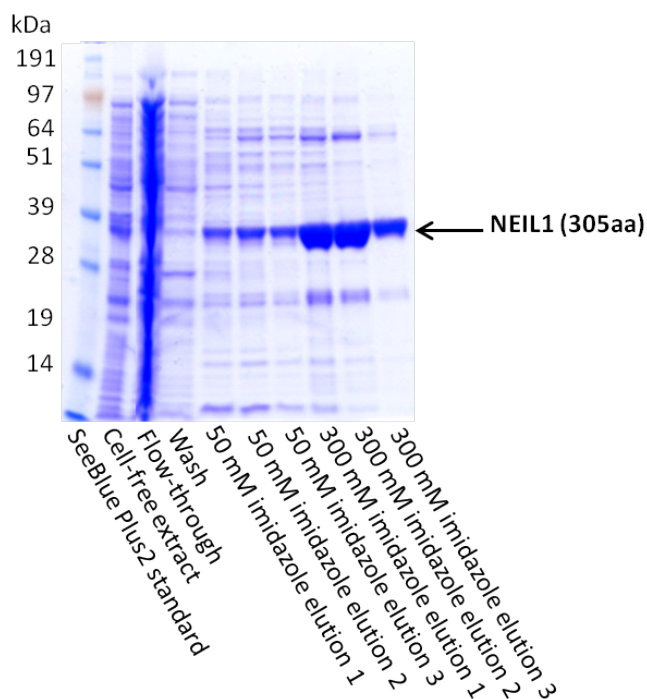


**Figure 3.19:** Analysis on a 12% NuPage gel of fractions A7-A12 from the salt gradient elution of NEIL1 (286aa).

#### Purification of NEIL1 (305aa)

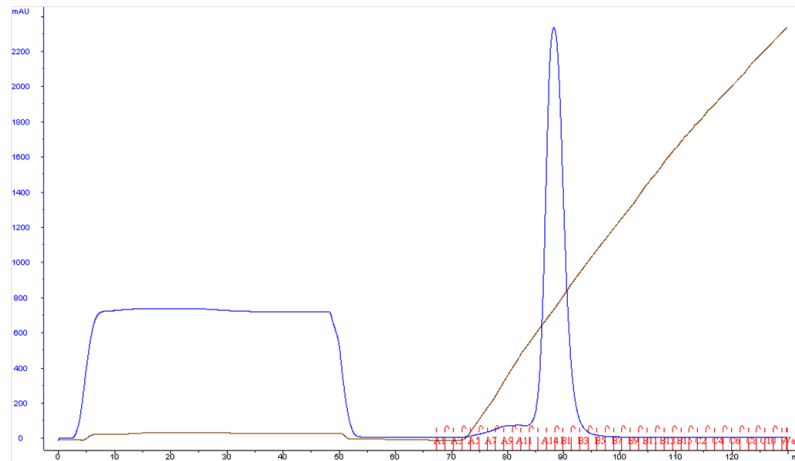
Since the cross-linking of NEIL1 (286aa) protein with DNA was unsuccessful and the yield not very high (see sec. 3.4.1) a slightly longer protein, NEIL1 (305aa), was cloned and expressed. The expression test did not show very high expression of the protein, but still higher than for the NEIL1 (286aa) construct.

The yield from purification with Ni-NTA agarose (Fig. 3.20) from 6 liter cell culture was much higher than what could be expected from the qualitative expression test. Most of the protein eluted in fractions with 300 mM imidazole, but also to a limited extent with 50 mM imidazole. For further purification, both the 50 mM and 300 mM imidazole fractions were selected.

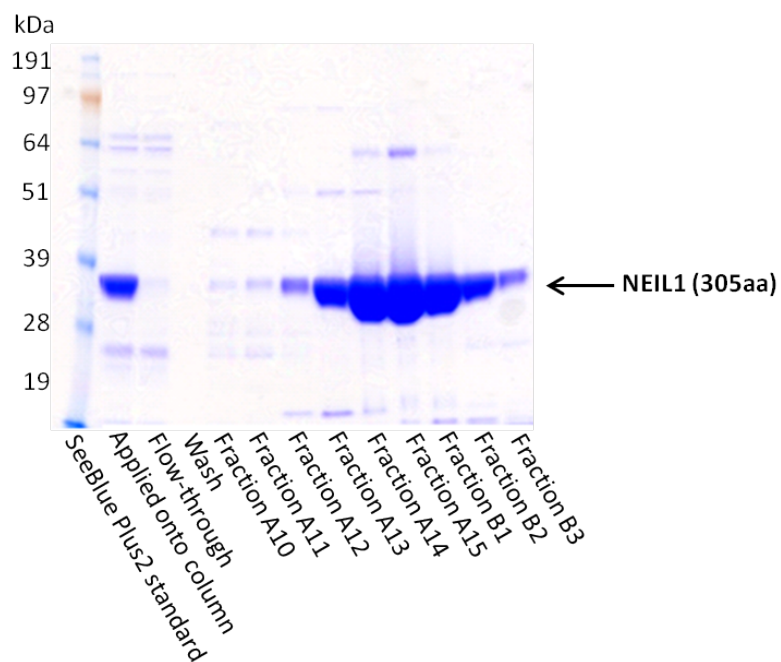


**Figure 3.20:** Analysis on a 12% NuPage gel of the purification of NEIL1 (305aa) with Ni-NTA agarose.

After dialysis, the protein solution was applied to a HiTrap SP XL cation exchange column for salt gradient elution (Fig. 3.21), and selected fractions were analyzed on a NuPage gel (Fig. 3.22). The resulting fractions contained a lot of pure protein, but also a weak band corresponding to a protein with higher molecular weight, possibly a multimer of NEIL1. The band was not analyzed further due to the very weak intensity compared to the NEIL1 (305aa) band. The expression of the NEIL1 (305aa) protein was not very high, but still observable. The high yield from the Ni-NTA agarose purification could suggest that the His-tag of the protein was more accessible than for the NEIL1 (286aa) protein, for instance by being localized on the protein surface enhancing contact between the protein and the Ni-NTA agarose. Unlike the NEIL1 (286aa) protein, the NEIL1 (305aa) protein was stable in the imidazole buffer overnight, and did not precipitate during dialysis. Purification of NEIL1 (325aa) was not carried out. The NEIL1 (305aa) protein is closer in sequence length to the Fpg/Nei enzymes, and was therefore prioritized.



**Figure 3.21:** Chromatogram of the purification of NEIL1 (305aa) on a HiTrap SPXL ion exchange column. Blue graph: elution profile at 280nm; brown line: salt gradient. Absorbance is measured in milli absorbance units (mAU), and the x-axis shows the elution buffer volume in ml. Fractions are indicated in red on the x-axis.

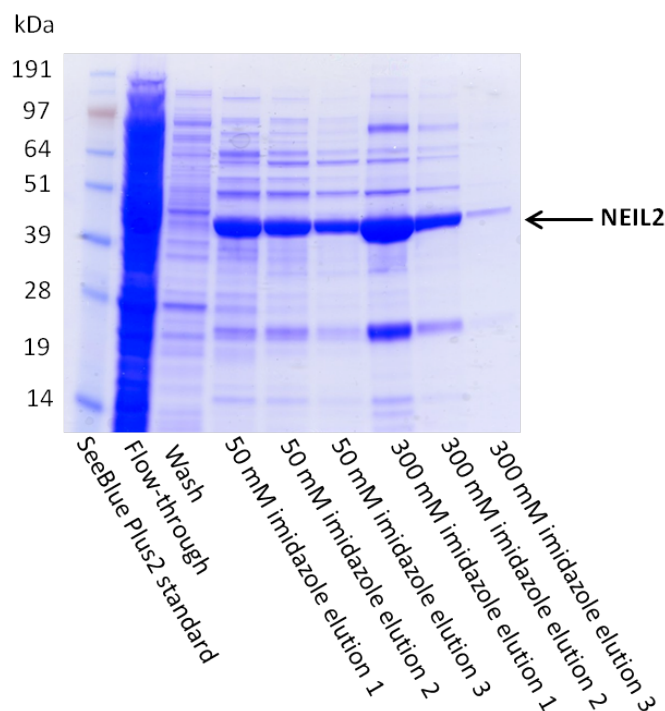


**Figure 3.22:** Analysis on a 12% NuPage gel of fractions A10-B3 from the salt gradient elution of NEIL1 (305aa).

### 3.3.2 Purification of NEIL2

NEIL2 was purified in high quantities for screening of crystallization conditions with and without DNA, and for glycosylase activity studies. A purification protocol was already established in the lab, similar to the NEIL1 protocol. The protein was purified in its full-length version, including a C-terminal 6x His-tag. The pET22b vector with the NEIL2 construct was available as a glycerol stock. Shortly, an overnight culture prepared from the glycerol stock was used to inoculate 6 liter LB-medium containing sorbitol, betaine and ampicillin. The protein expression was induced with 0.25 mM IPTG, and the cells cultured overnight at 16 °C. After harvesting the cells by centrifugation, the cell pellets were resuspended in a cold sonication buffer. Cells were lysed by ultrasonication, the cell debris was removed by centrifugation and the supernatant mixed with Ni-NTA agarose. As for the NEIL1 proteins, His-tagged NEIL2 protein was bound to the Ni-NTA agarose, while the majority of the remaining bacterial proteins did not bind.

The Ni-NTA agarose was washed with a 50 mM imidazole buffer to remove weakly bound proteins, leading to elution of some NEIL2 protein as well (see Fig. 3.23). Finally, the majority of NEIL2 protein was eluted with a 300 mM imidazole buffer. The NEIL2 fractions were not very pure, and contained some other protein bands. All the fractions from the 50 mM and 300 mM imidazole buffer elutions were pooled and desalted by dialysis for salt gradient elution on a HiTrap SP XL or a Resource S cation exchange column.



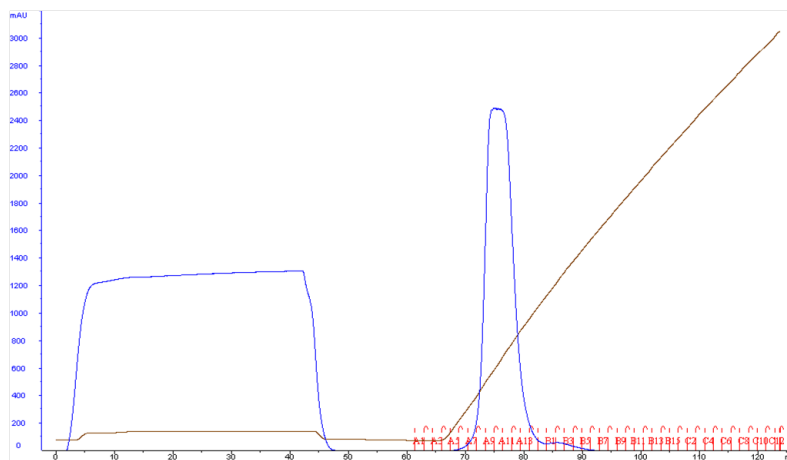
**Figure 3.23:** Analysis on a 12% NuPage gel of the NEIL2 purification with Ni-NTA agarose.

A chromatogram of the purification of NEIL2 from a HiTrap column is included in Fig. 3.24, and the subsequent gel analysis of selected fractions is presented in Fig. 3.25. A small amount of the NEIL2 protein did not bind to the column and was lost in the flow-through, but the total yield was still reasonably high. Some other bands were visible after the ion exchange chromatography, suggesting that another purification step would be required. Noteworthy, purification with a more sensitive Resource S column did not improve the purity.

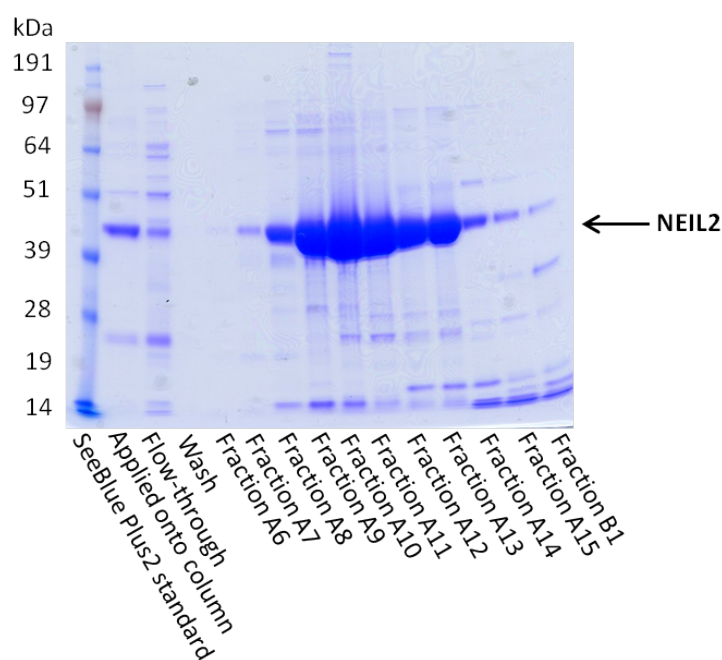
Gel filtration with a Superdex 75 size-exclusion column was also attempted in order to improve the purity of the NEIL2 protein. The protocol was the same as for the NEIL3 gel filtration protocol (sec. 3.3.3). The purity did not increase significantly, and moreover, NEIL2 precipitated when stored overnight in the gel filtration buffer A containing 100 mM NaCl.

Even though the expression test did not show any substantial induction of NEIL2, the yield from 6 liter of cell culture was high, and enabled cross-linking experiments and crystallization screening of the protein to be carried out. Some other protein bands were visible after PAGE analysis of the last purification step, but the frac-

tions contained mostly NEIL2. Moreover, after cross-linking of NEIL2 with DNA additional purification steps are required and the purity is expected to improve.



**Figure 3.24:** Chromatogram of the salt gradient elution of NEIL2 on a HiTrap SPXL column. Blue graph: absorbance at 280nm; brown line: salt gradient. Absorbance is measured in milli absorbance units (mAU), the x-axis shows the elution buffer volume in ml. Fractions are indicated in red on the x-axis.



**Figure 3.25:** NEIL2 fractions A6-B1 from ion exchange chromatography analyzed on a 12% NuPage gel.

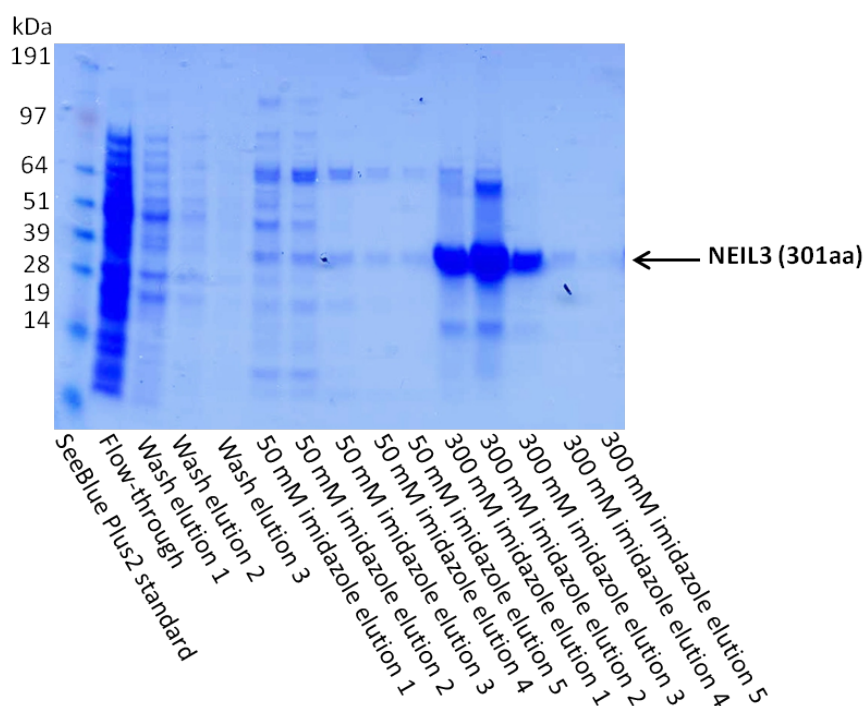
#### 3.3.3 Purification of NEIL3

The purification process of NEIL3 is known to be cumbersome from previous experience in the lab and from the literature Krokeide *et al.* [2009]. An already established protocol in the lab was followed in order to obtain a high enough yield of the protein for crystallization screening, that is, ideally a concentration between 8-10 mg/ml. The protocol was followed as explained in (sec.2.3.2). All the NEIL3 proteins were expressed with a C-terminal 6x His-tag in a pET-Duet-1 vector containing ampicillin resistance. In brief, an overnight culture was prepared from a glycerol stock of the desired NEIL3 version and was used to inoculate 12 liter LB-medium containing ampicillin. The cells were incubated at 37 °C, before protein expression was induced at 16 °C and the cells cultured overnight. Harvesting of the cells was done by centrifugation, and the cell pellets were resuspended in a cold sonication buffer. After ultrasonication and disruption of the cell membrane, the cell debris was removed by centrifugation. The remaining supernatant was mixed with Ni-NTA agarose, the suspension applied to an Econo column and the flow-through containing unbound proteins collected. The histidine residues in the C-terminal NEIL3 His-tag bound to the column while other weakly bound proteins were washed off of the column with the sonication buffer and the 50 mM imidazole buffer. Eventually, a 300 mM imidazole buffer was used to elute the NEIL3 protein.

At this point, precipitation was already visible. To decrease the precipitation, various approaches were tested. By keeping the protein on ice from the moment of elution, that is, store the tube on ice from the first eluted drop, it was possible to prevent some precipitation. Since NEIL3 was not stable in the imidazole buffer, a buffer change was required. From experience in the lab, NEIL3 was known to precipitate during dialysis. Thus, a gel filtration step was applied to change the buffer in a more gentle manner, and at the same time allowing a final polishing step. This required concentration of the protein fractions by centrifugation to a final volume of less than 1 ml. During this step, the protein continued to precipitate. To decrease the centrifugation time, the batch method (explained in sec.2.3.2) was used for the affinity chromatography purification. This method allowed the NEIL3 protein to be eluted in a small volume of the imidazole buffer, and thus reduce the centrifugation time by 50% relative to the previous protocol. The concentrated protein was immediately applied to a pre-equilibrated Superdex 75 size-exclusion column. After elution, the protein was stable in the gel filtration buffer for several days.

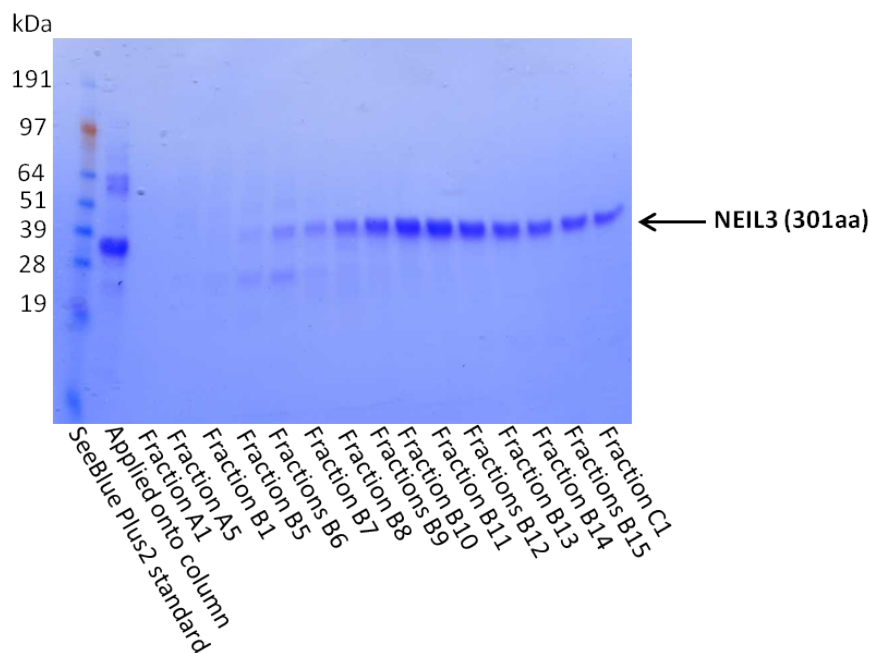
### Purification of NEIL3 (301aa)

The truncated NEIL3 (301aa) protein was purified with Ni-NTA agarose on an Econo column or with the batch method. The gel analysis of the Ni-NTA purification on an Econo column is shown in Fig. 3.26. Some NEIL3 (301aa) protein elution was observed with 50 mM imidazole, however most of the NEIL3 protein eluted with 300 mM imidazole in quite pure fractions. Due to low stability in the imidazole buffer, the fractions were immediately concentrated to a final volume of 0.5-1 ml, and 0.5 ml fractions were applied to the Superdex 75 size-exclusion column. Selected fractions were analyzed on a 12% NuPage gel, presented in Fig. 3.27. Purified NEIL3 (301aa) was incubated with lesion-containing DNA and eventually used for screening of crystallization conditions.



**Figure 3.26:** A 12% NuPage analysis of the NEIL3 (301aa) purification from Ni-NTA agarose.

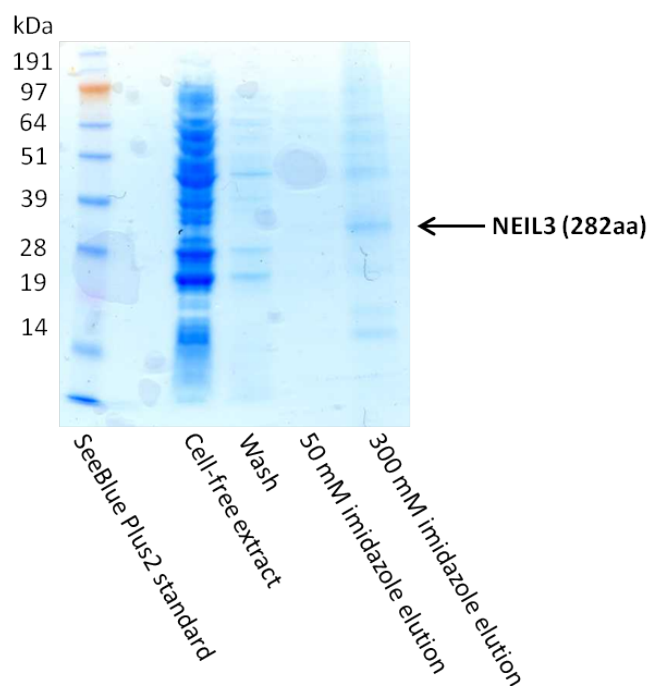




**Figure 3.27:** Analysis of selected fractions from the purification of NEIL3 (301aa) by gel filtration with a Superdex 75 column.

#### Purification of NEIL3 (282aa)

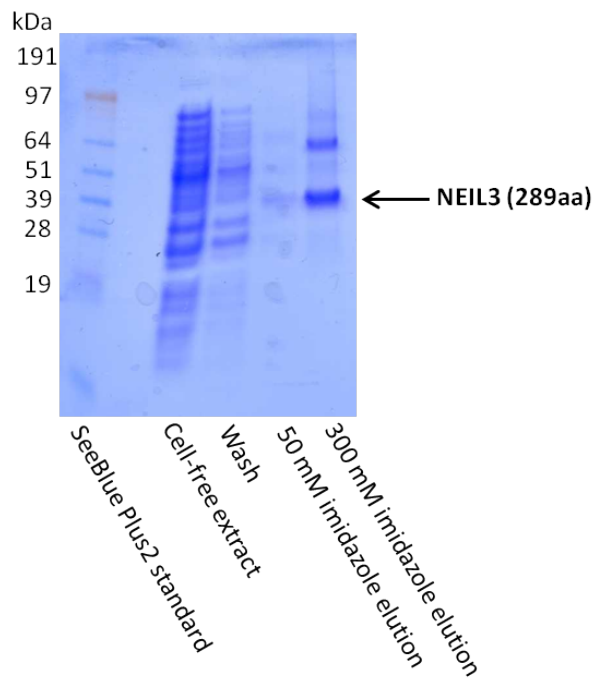
The even shorter NEIL3 (282aa) protein corresponds to the length of the bacterial orthologs Fpg/Nei. Unfortunately, this truncated version was even more unstable than the NEIL3 (301aa) truncation, and the protein expression was very low. Purification was carried out once from 12 liter cell culture, with almost no yield after the affinity chromatography step with Ni-NTA agarose (Fig. 3.28). The small amount that was collected precipitated rapidly. Gel filtration was attempted, without any great success. Since high protein concentrations are required in order to screen for crystallization conditions, purification was instead tried from the slightly longer NEIL3 (289aa).



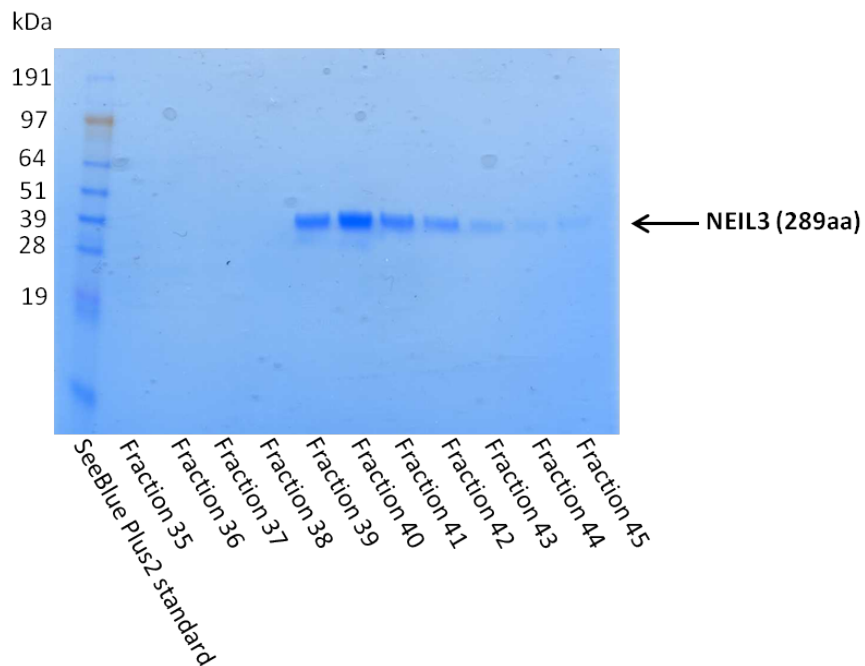
**Figure 3.28:** *Fractions from the purification of NEIL3 (282aa) from Ni-NTA agarose analyzed on a 12% NuPage gel.*

### Purification of NEIL3 (289aa)

The NEIL3 (289aa) protein was purified from 12 liters of cell culture following the same protocol as for NEIL1 (301aa). For the affinity chromatography step, the batch method was employed (Fig. 3.29). The achieved protein yield after the final gel filtration step (Fig. 3.30) was lower than that for the original truncated NEIL3 (301aa) protein, but still higher than for the 7 residue shorter NEIL3 (282aa) truncated version. The pooled NEIL3 (289aa) protein was stable in the gel filtration buffer, and some screening for crystallization conditions was carried out. Even though the protein stability did not improve from that of NEIL3 (301aa), it was nevertheless more stable than the NEIL3 (282aa) protein. Hopefully, the protein yield can be increased by further adjustments of the established protocol. This has however not been attempted due to lack of time before the submission of this thesis.



**Figure 3.29:** *NEIL3 (289aa) purification with Ni-NTA agarose analyzed on a 12% NuPage gel.*



**Figure 3.30:** *NEIL3 (289aa) purified by gel filtration. Selected fractions were analyzed on a 12% NuPage gel.*

## 3.4 Generation of complexes between NEIL proteins and DNA

In order to obtain crystals of NEIL protein together with damaged DNA, NEIL was subjected to damaged-containing oligos as explained in sec. 2.4.1 before screening for crystallization conditions. Two different approaches were attempted: cross-linking of NEIL and abasic DNA (AP-DNA), and mixing of NEIL and DNA containing a tetrahydrofuran (THF) nucleotide (THF-DNA). For buffer recipes and primer sequences, refer to the Appendix.

### 3.4.1 Cross-linking of NEIL1 with abasic DNA

The already established cross-linking protocol for NEIL1 was followed for cross-linking of NEIL1 protein and DNA, and eventually screening for crystallization conditions. A duplex DNA strand containing one deoxyuridine nucleotide was incubated with UDG (Uracil DNA glycosylase) in order to remove the uracil base to give a proper abasic (AP) site in the middle of the strand. The NEIL protein was desalted by dialysis to avoid any possible decrease of the protein's affinity for the DNA by shielding of positive charges on the protein surface due to high salt concentration. The cross-linking reaction was carried out by incubation of protein and DNA with NaBH<sub>4</sub> at 37 °C for 30 minutes, and the reaction quenched with 20% glucose. The reductant NaBH<sub>4</sub> stabilizes the formed Schiff base intermediate due to the  $\beta$ -elimination mechanism, and thus traps the covalent complex between NEIL and damaged DNA (Zharkov *et al.* [2002]). To avoid any precipitate on the column upon ion exchange chromatography, the cross-linking solution was sterile filtered to remove particles from precipitated protein formed during the cross-linking reaction at 37 °C.

#### Cross-linking of NEIL1 (286aa) with AP-DNA

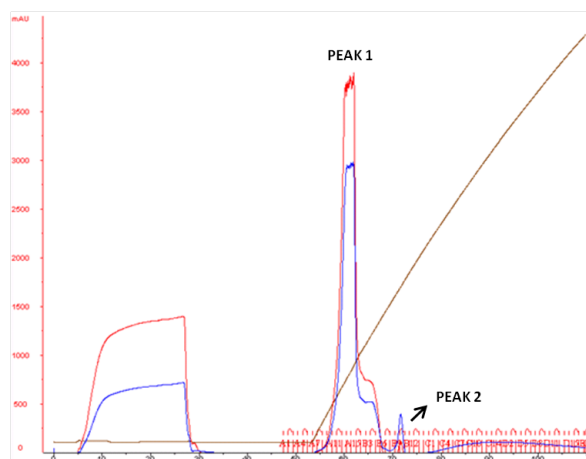
Cross-linking of NEIL1 (286aa) was carried out by incubating the protein with an 11mer AP-DNA with guanine opposite the abasic site. During ion exchange chromatography with a Resource S column only one form of protein bound to the column. Low yield made it challenging to see if any protein had reacted with DNA

(data not shown). Since a high yield is required for crystallization screening, cross-linking to DNA was instead tried with the NEIL1 (305aa) version.

#### **Cross-linking of NEIL1 (305aa) with AP-DNA**

The NEIL1 (305aa) protein was purified in a high quantity and could easily be cross-linked with AP-DNA and further purified for screening for crystallization. NEIL1 (305aa) was cross-linked to an 11mer or 13mer AP-DNA, both oligos containing a guanine opposite to the generated AP-site. Cross-linked protein was separated from unbound protein by ion exchange chromatography, shown in Fig. 3.31. High absorbance at 260nm, corresponding to DNA, was detected in the flow-through, indicating that excess DNA did not bind to the column. A peak with high absorbance at both 260nm and 280nm, peak 1, was detected followed by a small peak 2 at 280nm. Absorbance at 280nm is mainly due protein, indicating that peak 1 consisted of both DNA and protein. Since free DNA should not bind to the negatively charged material of the cation exchange column, the presence of DNA absorbance in peak 1 suggests that the DNA is bound to the protein. Further, NEIL1 protein in complex with negatively charged DNA should retain less on the anion material of the column and elute earlier than uncomplexed NEIL1. Analysis on a NuPage gel (Fig. 3.32) supported the appearance of a cross-linked complex in the fractions corresponding to peak 1. For these fractions, a small shift of a few kDa in size was visible compared to the control sample of unbound protein and the protein present in peak 2. Most of the protein was cross-linked to DNA, and the purest fractions were pooled and concentrated to 8-9 mg/ml for screening.

Cross-linking with 13mer AP-DNA was not successful (data not shown). All the protein eluted as one peak without any detection of absorbance at 260nm corresponding to DNA, and no shift in protein size was visible after separation on a denaturing gel. The reasons for the lack of cross-linking in this case are unclear, but could be due to lack of AP-site processing by the UDG, unreactive NaBH<sub>4</sub>, loss of enzyme activity, or that the enzyme somehow did not exhibit activity on the 13mer DNA. Since cross-linking with the 11mer DNA was successful, the experiment was not repeated with the 13mer DNA.



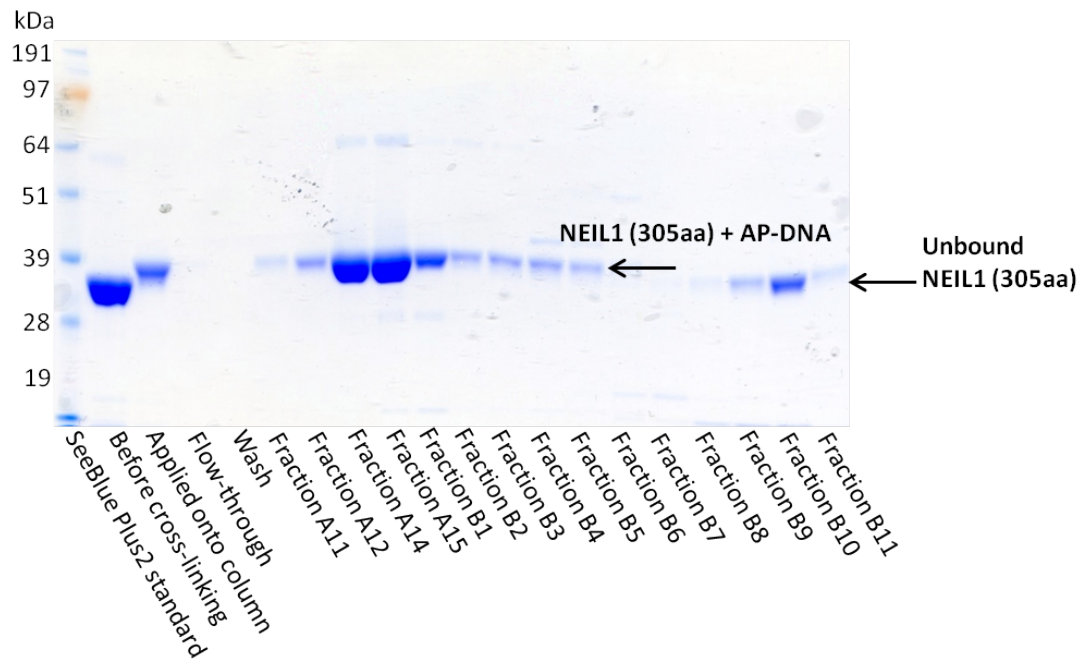
**Figure 3.31:** Chromatogram of the salt gradient elution on a Resource S column of NEIL1 (305aa) cross-linked to 11mer AP-DNA. Red graph: absorbance profile at 260nm; blue graph: absorbance at 280nm; brown line: salt gradient. Absorbance is measured in milli absorbance units (mAU), the x-axis shows the elution buffer volume in ml. Fractions are indicated in red on the x-axis.

### 3.4.2 Cross-linking of NEIL2 with abasic DNA

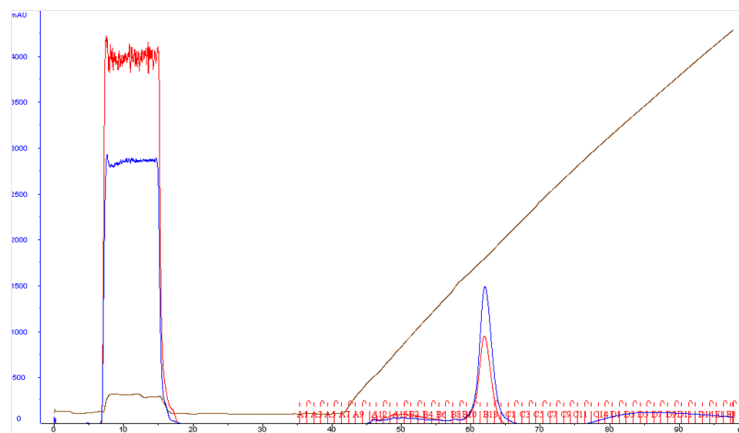
Full-length NEIL2 was cross-linked to a 13mer AP-DNA with guanine opposite the abasic site. It turned out that NEIL2 was not stable in the cross-linking buffer A with 50 mM NaCl, thus a higher concentration had to be used and NEIL2 was dialysed against a cross-linking buffer B containing 100 mM NaCl.

NEIL2 was incubated with the generated AP-DNA oligo and purified by ion exchange chromatography. One peak was detected on the chromatogram (Fig. 3.33), though without high DNA absorbance.

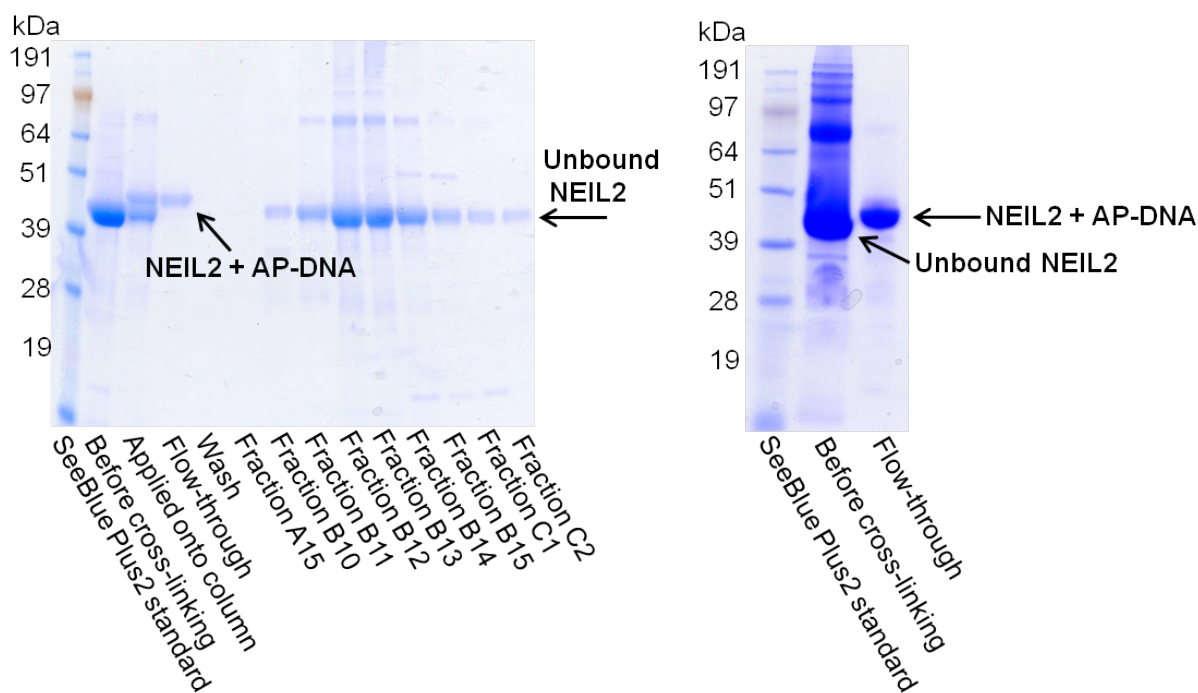
Analysis on a denaturing gel showed that only unliganded protein had bound to the Resource S column, and that the cross-linked NEIL2 protein did not bind to the column but had been collected in the flow-through (Fig. 3.34). The analyzed fraction of the cross-linking reaction revealed that less than 50 % of NEIL2 seemed to be cross-linked. The flow-through was thus concentrated and analyzed on a denaturing gel together with unbound NEIL2, showing that the flow-through contained pure, cross-linked NEIL2 (Fig. 3.34).



**Figure 3.32:** A 12% NuPage analysis of fractions A11-B11 of the salt gradient elution of NEIL1 (305aa) cross-linked to 11mer AP-DNA. Lane 2: sample of uncomplexed NEIL1 (305aa); lane 3: cross-linked NEIL1 (305aa) applied onto Resource S column.



**Figure 3.33:** Chromatogram of the salt gradient elution on a Resource S column of NEIL2 cross-linked to 13mer AP-DNA. Red graph: absorbance profile at 260nm; blue graph: absorbance at 280nm; brown line: salt gradient. Absorbance is measured in milli absorbance units (mAU), the x-axis shows the elution buffer volume in ml. Fractions are indicated in red on the x-axis, and the flow through was collected for the first 20 ml.



**Figure 3.34:** A 12% NuPage analysis of selected fractions from the salt gradient elution of NEIL2 cross-linked to 13mer AP-DNA. Left: analysis of collected flow-through and selected fractions; right: analysis of concentrated flow-through. For both gels: lane 2: sample of uncomplexed NEIL2.

### 3.4.3 Mixing of NEIL1, NEIL2 and NEIL3 with THF-DNA

In another approach to obtain crystals with NEIL and DNA, the NEIL protein was incubated with an 11mer duplex DNA with a tetrahydrofuran (THF) nucleotide serving as an AP-site analogue, abbreviated THF-DNA. The THF analogue will not react with the protein, but may form a heterodimer complex, as with the viral Neil1 ortholog MvNeil (Imamura *et al.* [2009]). The protocol is explained in sec. 2.4.2.

#### NEIL1

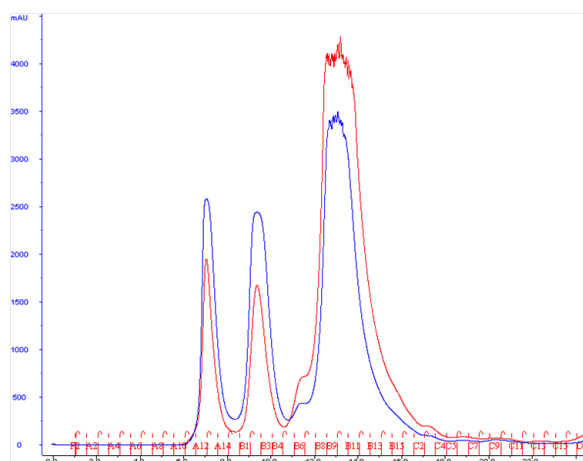
In order to increase the chances of obtaining crystals of NEIL1 in complex with DNA, NEIL1 was also mixed with an 11mer THF-DNA, which will not form a covalent complex with the protein. An adenine (A) or a thymine (T) was located opposite to the THF nucleotide in the complementary strand. After generation of the double stranded DNA oligo, DNA was incubated on ice with purified protein. NEIL1 (305aa) had a concentration of around 8 mg/ml. In some of the experiments, the



protein solution contained 50% glycerol. However, upon addition of DNA the protein precipitated, regardless of both the base in the opposite strand and presence of glycerol. Screening for crystallization conditions with THF-AP-DNA was therefore not accomplished.

## NEIL2

NEIL2 was also mixed with an 11mer THF-DNA, containing an adenine opposite to the AP site. The protein was concentrated to 14.5 mg/ml, and incubated with duplex DNA on ice. In contrast to NEIL1, NEIL2 did not precipitate. In order to separate unbound protein from the DNA-protein complex the solution was thus applied to a Superdex 75 column for gel filtration. A gel filtration buffer B containing the salt concentration at which NEIL2 eluted from the ion exchange column (550 mM NaCl), was used. The chromatogram from the size-exclusion chromatography is shown in Fig. 3.35.



**Figure 3.35:** *Chromatogram of the purification of NEIL2 and THF-DNA by gel filtration with a Superdex 75 column. Red graph: elution profile at 260nm; blue graph: elution profile at 280nm. Absorbance is measured in milli absorbance units (mAU), and the x-axis shows the elution buffer volume in ml. Fractions are indicated in red on the x-axis.*

Interestingly, two separate peaks with high absorbance for protein were present. Analysis on a denaturing gel showed the molecular weight of the proteins in the corresponding fractions to be identical (data not shown). This could indicate that the fractions contained a monomer and a multimer of NEIL2, the latter presumably being separated to monomers under the denaturing PAGE conditions. The peak with high absorbance at 260nm, corresponding to a high DNA content, eluted last from the column, suggesting that DNA had not bound to the protein. As the smallest components in a solution are expected to elute last in size-exclusion chromatography, the last peak would correspond to free DNA, and not to a DNA-protein complex. Separation of the fractions on a NuPage gel confirmed that the last peak was protein-free (data not shown).

The lack of binding could be due to high salt content in the buffer. NEIL2 eluted from the ion exchange column at a high salt concentration, thus incubation with DNA and the following purification was performed with the gel filtration buffer B containing 550 mM NaCl. As for the cross-linking experiments, a high salt concentration could prevent DNA-binding. Executing the experiment with a lower salt concentration could possibly lead to DNA-binding. Unfortunately, this is not an option since NEIL2 is not stable at low salt concentrations. One option is to introduce the DNA at an earlier step in the purification protocol, and then exchange the high-salt buffer with a low-salt buffer by dialysis before gel filtration with Superdex 75. This strategy will work if NEIL2 is more stable when bound to DNA.

### NEIL3

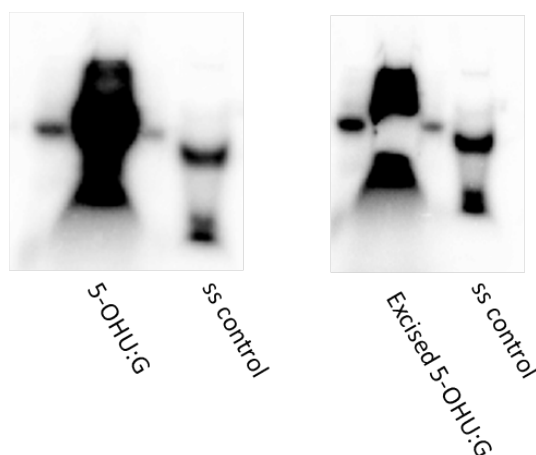
Mixing of NEIL3 and THF-DNA was attempted instead of the cross-linking experiment because of the poor stability of the protein and the low yield. The cross-linking reaction is a much rougher treatment than mixing with DNA to form a heterodimeric complex. The 11mer THF-AP-DNA contained a cytosine opposite to the AP-site. NEIL3 (301aa), stored in a gel filtration buffer, was concentrated to 10 mg/ml and mixed with the duplex DNA on ice. Precipitation was visible immediately. Possibly, addition of glycerol, lower protein concentration or choice of other oligos could prevent protein precipitation. The experiment was however not repeated, nor attempted for the NEIL3 (289aa) protein due to the low protein yield of that form.

## 3.5 Glycosylase activity studies

Inhibition of DNA repair enzymes is of interest since they have been shown to decrease the efficiency of cancer treatment (Lord & Ashworth [2012]). Several compounds which potentially inhibit DNA glycosylases have been identified by virtual docking with OGG1 (Mari Ytre-Arne, Master thesis, IMBV/UiO 2010). The most promising compounds have been tested in glycosylase cleavage assays with OGG1 and NEIL1, showing that the compounds that inhibited OGG1 also inhibited NEIL1 (Mari Ytre-Arne, unpublished). Thus, the compounds were also tested on NEIL2.

### 3.5.1 Isotope $^{32}\text{P}$ labeling of DNA substrate

A 40mer uracil-containing strand was labeled with  $\gamma\text{-}^{32}\text{P}\text{-ATP}$  by incubation with a T4 polynucleotide kinase (PNK) at 37 °C. A single-strand control was removed after kinase inactivation at 80 °C and the complementary strand, with a guanine opposite to the uracil, was added to the labeled strand and incubated further at 90 °C, then room temperature and finally on ice. The sample and the single-strand control were mixed with DNA loading buffer and separated on a native gel. Eventually, the gel was transferred to a phosphor storage screen, and the gel scanned after 5 minutes. The band corresponding to double-stranded DNA was cut out (Fig. 3.36), and the gel scanned again to confirm the incision of the band corresponding to double-stranded DNA. The gel piece was transferred to a tube and mixed with  $\text{mq-H}_2\text{O}$  in order to extract the DNA from the gel.



**Figure 3.36:** A scanned 20% native gel showing double-stranded 5-OHU:G before and after excision of the band. Single-stranded DNA was used as a control.

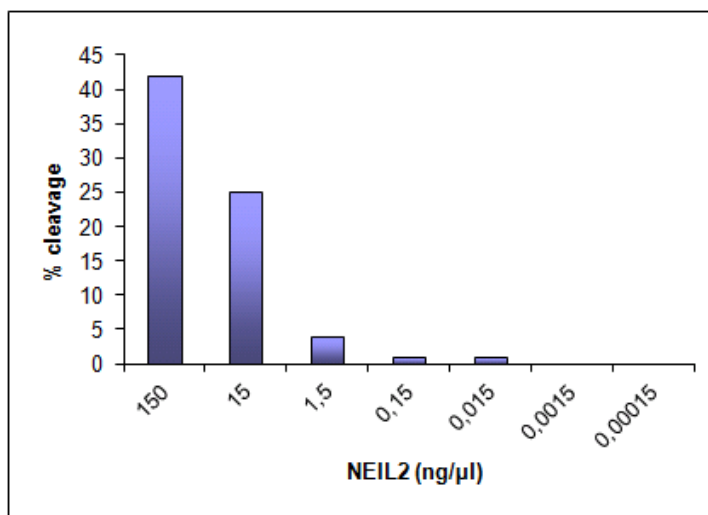
### 3.5.2 Glycosylase activity assays of NEIL2

From the literature, NEIL2 is known to exhibit glycosylase activity against 5-OHU:G and to be a bifunctional glycosylase that cleaves the phosphodiester bond in the damaged strand (Hazra *et al.* [2002b]). In order to carry out the glycosylase inhibition assays with a protein concentration within the linear range of enzyme activity, the cleavage reaction should correspond to 20-40% cleavage of the damaged substrate.

NEIL2 was thus diluted in a protein dilution buffer to 7 different concentrations ranging from 150 ng/ $\mu$ l to 0.15 pg/ $\mu$ l. A negative control was prepared by exchanging enzyme with mq-H<sub>2</sub>O. The enzyme was mixed with labeled substrate and incubated at 37 °C. A stop solution was added and the mix heated to 95 °C, followed by separation on a denaturing gel. The gel was stored in a phosphor storage screen, before scanning and activity measurement the next day.

The enzyme was stored on ice, or at -20 °C in 50% glycerol. Since the potential inhibitor compounds were dissolved in DMSO, titration assays with NEIL2 were performed with and without DMSO in order to detect any activity differences because of the presence of DMSO.

The experiments showed that glycerol is required for NEIL2 glycosylase activity, as no activity could be detected for the samples without glycerol (data not shown). Further, DMSO did not alter the enzyme activity (not shown). The quantified results from the titration experiments with NEIL2 are illustrated in Fig. 3.37, suggesting that a concentration of 15 ng/ $\mu$ l NEIL2 will give 25% cleavage activity. The value from the negative control was used as a baseline and subtracted for the activity measurements.

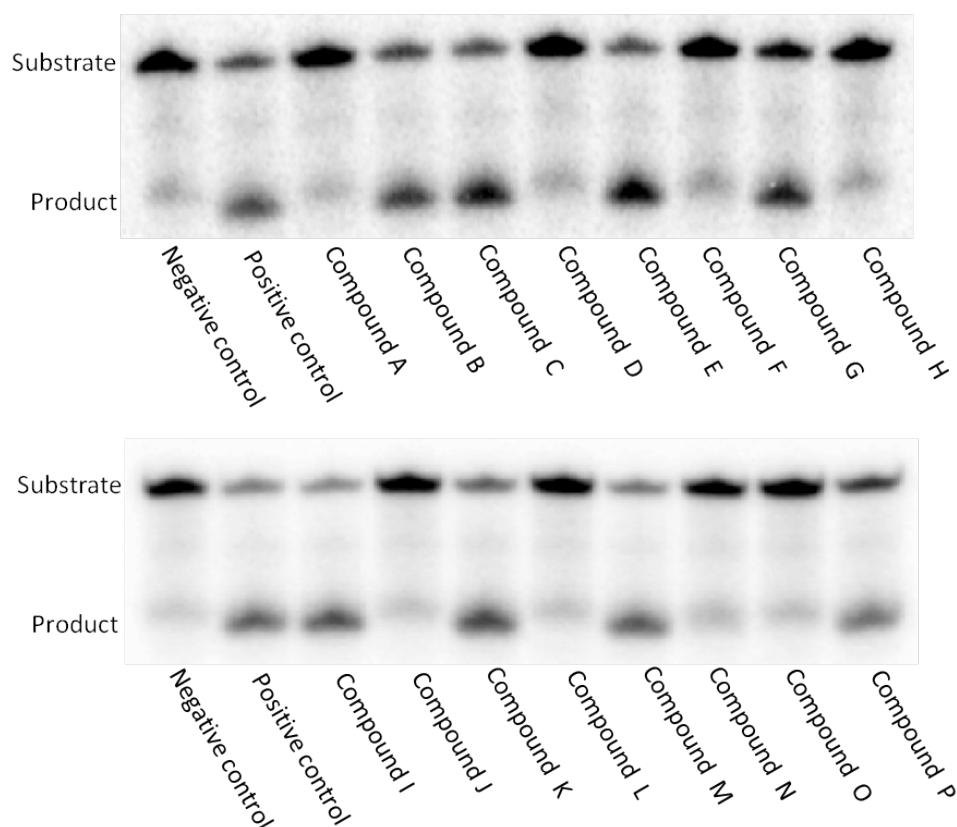


**Figure 3.37:** *Titration of NEIL2 activity on 5-OHU:G. Y-axis: % cleavage; x-axis: NEIL2 concentration in ng/ $\mu$ l.*

#### 3.5.3 Inhibition of glycosylase activity

Compounds with inhibitor potential on NEIL2 glycosylase activity were tested qualitatively. NEIL2 enzyme with a concentration of 15 ng/ $\mu$ l, corresponding to 25% activity on the 5-OHU:G substrate, was used. The negative control contained mq-H<sub>2</sub>O instead of enzyme. Both the positive and negative control were supplemented with DMSO. Enzyme was mixed with the inhibitor compound, added substrate and incubated at 37 °C. After addition of the stop solution and heating to 95 °C, the samples were separated on a denaturing gel. The gels were transferred to a phosphor storage screen and scanned the following day.

The results in Fig. 3.38 suggest that 9 of the 16 compounds, inhibit the NEIL2 glycosylase activity. These compounds are the same that have been shown to inhibit OGG1 activity (Mari Ytre-Arne, unpublished). It should be noted that the used inhibitor compound concentrations are rather high in order to detect a positive effect, and that further titration assays with the promising compounds must be done to determine the  $IC_{50}$  values.



**Figure 3.38:** *Inhibition of NEIL2 by potential the DNA glycosylase inhibitor compounds A-P. Substrate: 5-OHU; product: cleaved DNA.*

## 3.6 Crystallization screening

### Screening of NEIL1

The NEIL1 (305aa) protein was cross-linked to an 11mer AP-DNA, separated from unbound protein and concentrated to 8-8.5 mg/ml. Screening was performed with commercial kits by sitting drop using the Oryx6 robot at room temperature. Screens were stored at 4 °C, but no crystal formation was found for the tested conditions after 10 weeks.

For the NEIL1 (337aa) protein in complex with an 11mer AP-DNA, crystallization conditions with PEG 3350 existed (sec.2.6). NEIL1 (305aa) protein was therefore screened by hanging drop at room temperature and stored at 4 °C. After 1 day, small transparent needles appeared in the wells with 10 and 12% PEG 3350. No differences in trends were observed for the conditions containing 150 mM or 200 mM sodium tartrate. After 1 week, the crystals had grown bigger, and smaller needles had appeared in the wells with up to 20% PEG 3350. Representative crystals are shown in Fig. 3.39. The crystals were transferred to a cryo-protectant solution and flash frozen in liquid nitrogen. The utilized cryo-protectants are listed in Tab.3.1.



**Figure 3.39:** *Crystals of NEIL1 (305aa) cross-linked to 11mer AP-DNA.*

**Table 3.1: Cryo-protectants used for freezing of NEIL1 (305aa) + AP-DNA crystals. MPD: 3-methyl-1,5-pentanediol.**

Cryo-protectant
20% ethylene glycol
20% PEG400
20% glucose
20% MPD

### Screening of NEIL2

NEIL2 protein (without DNA) was shipped for sitting-drop screening at 4 °C at the High-Throughput Crystallization facility at the ESRF synchrotron in Grenoble. Unfortunately, no crystals have formed in these experiments. Screening of NEIL2 without DNA or in complex with 13mer AP-DNA was carried out home by sitting drop with the Oryx6 robot. The screens were stored at both 4 °C and at room temperature. So far, no crystallization conditions with a positive outcome have been discovered.

### Screening of NEIL3

Screening of crystallization conditions with commercial kits for NEIL3 was carried out by vapor batch with the Oryx6 robot, and by hanging drop at 4 °C. All screens were stored at 4 °C. No crystallization conditions have been found.



## 3.7 Data collection and processing

Crystals of NEIL1 (305aa) in complex with an 11mer AP-DNA were transferred to a cryo-protectant solution and flash frozen in liquid nitrogen before data collection at the BESSY synchrotron in Berlin. Diffraction data was collected at 100 K. Crystals that had formed in 18% PEG 3350 diffracted to the best resolution (see sec. 2.6 for conditions). While collecting diffraction data, it became apparent that the cryo-protectant was crucial for the quality of the data. Of 30 crystals, data sets were recorded for three crystals, all of them soaked in 20% ethylene glycol. The crystals diffracted to 3.5-4 Å, and the data set with highest resolution was further processed and refined. The published structure of apo NEIL1 (1TDH; Doublé *et al.* [2004]) was used as a search model to solve the structure from our data of the cross-linked complex by molecular replacement (Tab. 3.2).

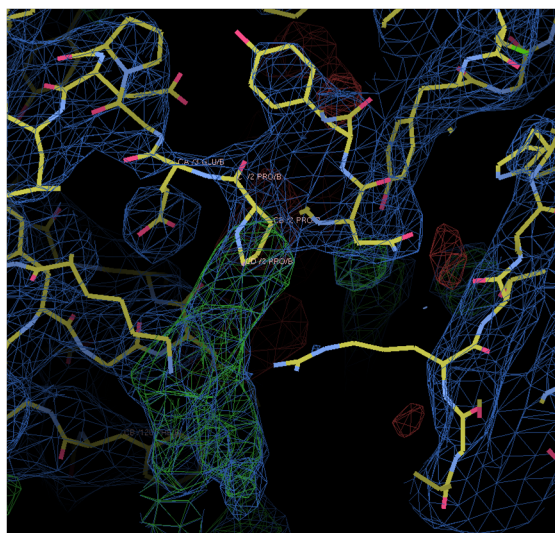
**Table 3.2: Crystal data, data collection and refinement statistics of NEIL1 (305aa) with AP-DNA.**

Crystal data			
Space group	P3 <sub>1</sub> 21		
a, b, c (Å)	94.46/94.46/269.99 90/90/120		
Data collection			
X-ray source	BESSY BL 14.1	Resolution range (Outer shell)	48.10-3.50 (3.58-3.50)
Wavelength	0.9797 Å	Completeness (%) <sup>a</sup>	99.8/100
Oscillation range	1.0 degree	Redundancy <sup>a</sup>	4.7/4.8
Exposure time	20 seconds	I/σ(I) <sup>a</sup>	6.8/2.1
Number of pictures	70	R <sub>merge</sub> (%) <sup>a</sup>	0.184/0.774
Refinement statistics			
R <sub>fac</sub>	0.320		
R <sub>free</sub>	0.396		

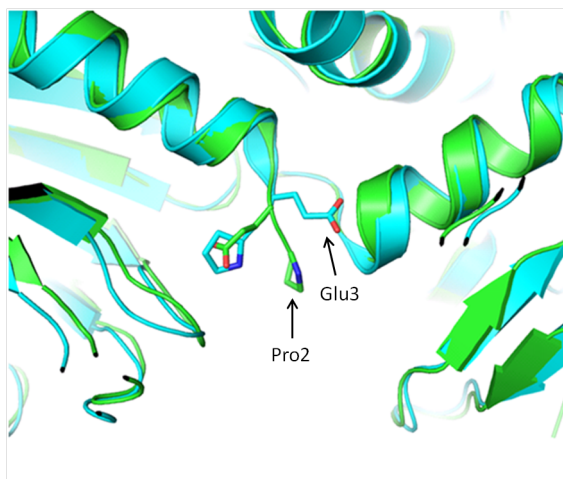
<sup>a</sup>Value before blackslash: Overall; after blackslash: Outershell

The protein was refined using Phaser. After molecular replacement, the model was improved by rigid body refinement followed by restrained refinement. The quality and the resolution of the data set did not allow any water molecules to be added to the structure. Positive electron density for DNA in the active site was visible, however, apparently only for ssDNA although the cross-linking experiment was performed with dsDNA. The quality of the residual electron density was not high enough to reliably build in DNA, but the current model represents a huge improvement compared to the crystals of the original truncated NEIL1 (337aa) which only diffracted to around 10 Å.

Interestingly, the catalytic active Pro2 and Glu3 residues seem to have switched location. Due to the cross-linking reaction, Pro2 is covalently bound to the ribose, hence should it be in close proximity to the positive electron density corresponding to DNA ((Fig. 3.40) and Fig. 3.41). A flip of the N-terminal tail of the protein has not been previously reported or even anticipated, neither for free NEIL1 (Doubl   *et al.* [2004]), the viral Neil1 in complex with DNA (Imamura *et al.* [2009 2012]) nor in the prokaryotic E.coli Fpg- or Nei-DNA complexes (Coste *et al.* [2004]; Zharkov *et al.* [2002]). Whether the residue shift is an unnatural static crystal conformation or an actual part of the mechanism, remains to be shown.



**Figure 3.40:** *Electron density maps showing the active site of NEIL1 (305aa). Positive electron density (green) is visible for DNA (not modeled). The positions of the Pro2 and Glu3 residues were switched since Pro2 is cross-linked to the DNA and must be closest to the DNA (see also Fig. 3.41).*



**Figure 3.41:** *Close-up of active site. The structure of NEIL1 (305aa) cross-linked to DNA (shown in green) is superimposed to apo NEIL1 (cyan; pdb code: 1TDH), showing that the N-terminal Pro2 and Glu3 residues are interchanged in the DNA-bound structure.*

## 3.8 Final conclusions and future work

### NEIL1

The published structure of apo NEIL1 (Doubl   *et al.* [2004]) was finally solved for a 337 residue long version of the protein, after multiple trials with the full-length form and various truncated forms. The C-terminal part of NEIL1 is structurally disordered from around residue 300 and onward. Removal of the full disordered part lead to crystals with low quality, and inclusion of part of the unstructured tail (residues 300–337) was essential for obtaining well-diffracting crystals. However, no density was visible after residue 290. Previously, crystallization of this variant of NEIL1 (337aa) cross-linked to AP-DNA has been extensively carried out in our lab, resulting in crystals with bad diffraction (less than 10 Å resolution). Several strategies have been tried to improve the quality of the crystals, including variations in the length of DNA, using additives in the crystallization buffer, varying cryo-protectants and applying crystal humidity control and crystal annealing without success. Therefore, this project has focused on even shorter versions of NEIL1.

Three other truncated versions of NEIL1 were designed: NEIL1 (286aa), NEIL1 (305aa) and NEIL1 (325aa). NEIL1 (286aa) corresponds to the Fpg/Nei sequence length and was difficult to express and purify to the necessary yield for crystallization. Of the NEIL1 (305aa) and NEIL1 (325aa) versions, the 305 residue long protein was closest in sequence to the Fpg/Nei enzymes and was hence prioritized for expression and purification. Surprisingly, the protein yield was much higher than expected from the expression test. NEIL1 (305aa) was successfully cross-linked to an 11mer AP-DNA and crystallized with the same conditions as for the NEIL1 (337aa) protein. Under data collection, we learned that the choice of cryo-protectant was crucial for the diffraction of this complex. Only crystals soaked in 20% ethylene glycol diffracted to a resolution better than 5 Å, with the best data set diffracting to 3.5 Å. Unfortunately, the resolution and the corresponding electron density map was not good enough to build a model of the cross-linked DNA, but positive electron density for ssDNA was however visible. The protein was crystallized with double-stranded DNA, yet the density suggests that only single-stranded DNA is present in the crystal. Interestingly, the N-terminal part of the catalytic active PE helix seems to have flipped relative to the apo protein structure. If the observed flip is a result of an unnatural, static intermediate induced by crystal contacts or an essential part

of the DNA cleavage mechanism, remains to be solved.

Although no final model of NEIL1 in complex with DNA could be built and refined during this project, it has been shown that by truncating the protein and selecting the optimal cryo-protectant, the diffraction can be considerably improved. Work is in progress to optimize the cross-linking protocol to retain dsDNA in the protein-DNA complex. The cross-linking reaction at 37 °C has been replaced with a similar reaction at 4 °C to possibly prevent separation of the two strands in the short 11mer DNA. Cross-linking with larger DNA oligomers (13mer, 15mer) should also increase chances of binding of double-stranded DNA. Further, the protocol for cryo-protectant soaking should be improved by varying the ethylene glycol concentrations and by adding the cryo-protectant directly to the crystallization buffer when screening.

Recently, the structure of a glycosylase-deficient mutant of Neil1 from the mimivirus, MvNeil1, was solved in complex with a damage-containing nucleotide (Imamura *et al.* [2009 2012]). By mutation of the Glu3 residue to a glutamine (Q), the damaged nucleotide will not be excised and the interactions between the recognized base and the active site can be investigated. The viral MvNeil1 and human NEIL1 do have the same overall fold as they belong to the same structural family. However, in the DNA binding regions and in surface exposed loops participating in DNA binding, the two orthologs exhibit substantial sequence and structural differences. Therefore, a human NEIL1 in complex with DNA is required to explain the substrate specificity of the human NEIL1. In this thesis, an E3Q glycosylase-deficient version of NEIL1 (305aa) has been designed for this purpose. This mutant will be used for co-crystallization (without cross-linking) with DNA containing oxidized bases known to be recognized by NEIL1, like thymine glycol and 5-hydroxycytosine.

The Lys54 residue in NEIL1 has recently been shown to alternate with the Pro2 residue in the AP lyase reaction (Erik S. Vik, unpublished). Hence, a second mutant, NEIL1-286 K54Q, has been designed to investigate the function of the Lys54 in catalysis.

## NEIL2

NEIL2 has been purified and screened for crystallization in its full-length as its structure is predicted to be well-ordered. To date, no structure of NEIL2 has been solved. In this project, full-length NEIL2 has been expressed and purified in a high yield, although the expression test did not show high protein expression. Screening for crystallization has been carried out for the apo enzyme and for NEIL2 cross-linked to an 13mer AP-DNA. So far, no crystals have been observed and screening should continue both with and without DNA. Similar to NEIL1, a glycosylase-deficient mutant of NEIL2, NEIL2 E3Q, has been designed for crystallization with damaged-containing DNA and investigations of the protein-DNA interactions.

Potential inhibitors of DNA glycosylase activity have been tested on NEIL2 in this thesis. The compounds with potential inhibiting properties have previously been tested on OGG1 and NEIL1 (Mari Ytre-Arne, unpublished). So far, the compounds that inhibit OGG1 and NEIL1 also seem to inhibit NEIL2 activity. The results are however preliminary as high concentrations of the inhibitors were used. Thus, further studies with lower inhibitor concentrations must be carried before any conclusions can be made and the results compared with the corresponding OGG1 and NEIL1 data.

## NEIL3

Compared to the homologs NEIL1 and NEIL2, not much is known about NEIL3. Only recently, DNA glycosylase activity for the oxidized purines Gh and Sp was reported(Liu *et al.* [2010]). No structure of NEIL3 has been solved yet. The full-length NEIL3 contains several C-terminal zinc finger domains which are not typical of the Fpg/Nei family, and their presence has been suggested to complicate the purification of the protein (Krokeide *et al.* [2009]). In addition, the C-terminal part is predicted to be flexible. The starting point of this project was a truncated version of the NEIL3 (605aa) protein: NEIL1 (301aa), without the flexible C-terminal part. A protocol of the NEIL1 (301aa) version was already established in the lab for high yield purification (Medya Salah, Master thesis, IMBV/UiO, 2010). It should be noted that compared to the yield from purification of NEIL1 and NEIL2, the yield from the NEIL3 purification is much lower. The low stability of the protein and following precipitation has been particularly challenging during purification.

NEIL3 (301aa) was purified and incubated with THF-DNA, leading to heavy precipitation of the protein. Screening was therefore carried out for the apo NEIL3 (301aa) protein, without any result. Thus, two new truncated versions of NEIL3 were designed: NEIL3 (282aa) and NEIL3 (289aa), the former corresponding to the Fpg/Nei enzyme sequence length. None of the new protein versions were expressed to a high extent in the expression tests. Purification was nevertheless attempted, resulting in a very low yield of unstable NEIL3 (282aa) and a higher yield of NEIL3 (289aa). Interestingly, the 7 last residues in the C-terminal part had a great impact on protein expression and stability. Several modifications of the purification protocol were tested, but the NEIL3 (282aa) protein continued to precipitate, finally resulting in a too low yield for crystallization screening. Since no crystal structure of NEIL3 is known, also a structure without DNA is of high interest. NEIL3 (289aa) was therefore only screened without DNA due to the relatively low yield. The low stability of the NEIL3 enzymes requires continuous storage at 4 °C. The in-house Oryx6 robot is however operated at room temperature, impeding fast automatic screening with less protein than possible by the manual hanging-drop screening. Ideally, further screening should be carried out with a robot at 4 °C, as for the NEIL2 screening. So far, no NEIL3 crystals have formed.

In addition, mutants of the three truncated NEIL3 proteins were designed. The NEIL3-301 K81Q, NEIL3-282aa K81Q and NEIL3-289aa K81Q proteins can be used in activity experiments to investigate the role of Lys81 in catalysis.

The expression of NEIL3 has been shown to be increased in cancer tissues compared to normal tissues (Hildrestrand *et al.* [2009]). The potential DNA glycosylase inhibitors tested on OGG1, NEIL1 and NEIL2 should therefore also be used in studies with NEIL3.





# Bibliography

- Aamodt, R. M., Falnes, P. Ø., Johansen, R. F., Seeberg, E., & Bjørås, M. 2004. The *Bacillus subtilis* counterpart of the mammalian 3-methyladenine DNA glycosylase has hypoxanthine and 1,N(6)-ethenoadenine as preferred substrates. *The Journal of Biological Chemistry*, **279**(14), 13601–6.
- Alonso, A., Terrados, G., Picher, A. J., Giraldo, R., Blanco, L., & Larraga, V. 2006. An intrinsic 5'-deoxyribose-5-phosphate lyase activity in DNA polymerase beta from *Leishmania infantum* supports a role in DNA repair. *DNA Repair*, **5**(1), 89–101.
- Alseth, I., Rognes, T., Lindbäck, T., Solberg, I., Robertsen, K., Kristiansen, K. I., Mainieri, D., Lillehagen, L., Kolstø, A., & Bjørås, M. 2006. A new protein superfamily includes two novel 3-methyladenine DNA glycosylases from *Bacillus cereus*, AlkC and AlkD. *Molecular Microbiology*, **59**(5), 1602–9.
- Bandaru, V., Sunkara, S., Wallace, S. S., & Bond, J. P. 2002. A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to *Escherichia coli* endonuclease VIII. *DNA Repair*, **1**(7), 517–29.
- Bandaru, V., Cooper, W., Wallace, S. S., & Doublié, S. 2004. Overproduction, crystallization and preliminary crystallographic analysis of a novel human DNA-repair enzyme that recognizes oxidative DNA damage. *Biological Crystallography*, **60**(6), 1142–4.
- Bjørås, M., Seeberg, E., Luna, L., Pearl, L. H., & Barrett, T. E. 2002. Reciprocal “ Flipping ” Underlies Substrate Recognition and Catalytic Activation by the Human 8-Oxo-Guanine DNA Glycosylase. *Journal of Molecular Biology*, **317**, 171–177.
- Blow, David. 2002. *Outline of Crystallography for Biologists*. Oxford University Press.

- Boiteux, S., & Radicella, J. P. 1999. Base excision repair of 8-hydroxyguanine protects DNA from endogenous oxidative stress. *Biochimie*, **81**(1-2), 59–67.
- Branden, C., & Tooze, J. 1999. *Introduction to Protein Structure*.
- Caldecott, K. W. 2008. Single-strand break repair and genetic disease. *Nature Reviews. Genetics*, **9**(8), 619–31.
- Chayen, N. E. 1999. Recent advances in methodology for the crystallization of biological macromolecules. *Journal of Crystal Growth*, **198-199**(Mar.), 649–655.
- Christmann, M., Tomicic, M. T., Roos, W. P., & Kaina, B. 2003. Mechanisms of human DNA repair: an update. *Toxicology*, **193**(1-2), 3–34.
- Ciccia, A., & Elledge, S. J. 2010. The DNA damage response: making it safe to play with knives. *Molecular Cell*, **40**(2), 179–204.
- Coste, F., Ober, M., Carell, T., Boiteux, S., Zelwer, C., & Castaing, B. 2004. Structural basis for the recognition of the FapydG lesion (2,6-diamino-4-hydroxy-5-formamidopyrimidine) by formamidopyrimidine-DNA glycosylase. *The Journal of Biological Chemistry*, **279**(42), 44074–83.
- Dalhus, B., Helle, I. H., Backe, P. H., Alseth, I., Rognes, T., Bjørås, M., & Laerdahl, J. K. 2007. Structural insight into repair of alkylated DNA by a new superfamily of DNA glycosylases comprising HEAT-like repeats. *Nucleic Acids Research*, **35**(7), 2451–9.
- Dalhus, B., Laerdahl, J. K., Backe, P. H., & Bjørås, M. 2009. DNA base repair–recognition and initiation of catalysis. *FEMS Microbiology Reviews*, **33**(6), 1044–78.
- Das, A., Rajagopalan, L., Mathura, V. S., Rigby, S. J., Mitra, S., & Hazra, T. K. 2004. Identification of a zinc finger domain in the human NEIL2 (Nei-like-2) protein. *The Journal of Biological Chemistry*, **279**(45), 47132–8.
- Das, A., Wiederhold, L., Leppard, J. B., Kedar, P., Prasad, R., Wang, H., Boldogh, I., Karimi-Busheri, F., Weinfeld, M., Tomkinson, A. E., Wilson, S. H., Mitra, S., & Hazra, T. K. 2006. NEIL2-initiated, APE-independent repair of oxidized bases in DNA: Evidence for a repair complex in human cells. *DNA Repair*, **5**(12), 1439–48.
- De Bont, R. 2004. Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis*, **19**(3), 169–185.

- Dou, H., Mitra, S., & Hazra, T. K. 2003. Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *The Journal of Biological Chemistry*, **278**(50), 49679–84.
- Doubl  , S., Bandaru, V., Bond, J. P., & Wallace, S. S. 2004. The crystal structure of human endonuclease VIII-like 1 (NEIL1) reveals a zincless finger motif required for glycosylase activity. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(28), 10284–9.
- Drabl  s, F., Feyzi, E., Aas, P. A., Vaagb  , C., Kavli, B., Bratlie, M., Pe  a Diaz, J., Otterlei, M., Slupphaug, G., & Krokan, H. E. 2004. Alkylation damage in DNA and RNA—repair mechanisms and medical significance. *DNA repair*, **3**(11), 1389–407.
- Dutta, S., Chowdhury, G., & Gates, K. S. 2007. Interstrand cross-links generated by abasic sites in duplex DNA. *Journal of the American Chemical Society*, **129**(7), 1852–3.
- Evans, M. D., Dizdaroglu, M., & Cooke, M. S. 2004. Oxidative DNA damage and disease: induction, repair and significance. *Mutation Research*, **567**(1), 1–61.
- Fagbemi, A. F., Orelli, B., & Sch  rer, O. D. 2011. Regulation of endonuclease activity in human nucleotide excision repair. *DNA Repair*, **10**(7), 722–9.
- Falnes, P.   ., Johansen, R. F., & Seeberg, E. 2002. AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature*, **419**(6903), 178–82.
- Friedberg, E. C. 2003. DNA damage and repair. *Nature*, **421**(January), 436–440.
- Friedberg, E. C., McDaniel, L. D., & Schultz, R. 2004. The role of endogenous and exogenous DNA damage and mutagenesis. *Current Opinion in Genetics & Development*, **14**(1), 5–10.
- Grin, I. R., & Zharkov, D. O. 2011. Eukaryotic endonuclease VIII-Like proteins: New components of the base excision DNA repair system. *Biochemistry (Moscow)*, **76**(1), 80–93.
- Grin, I. R., Dianov, G. L., & Zharkov, D. O. 2010. The role of mammalian NEIL1 protein in the repair of 8-oxo-7,8-dihydroadenine in DNA. *FEBS letters*, **584**(8), 1553–7.

- Hailer, M. K., Slade, P. G., Martin, B. D., Rosenquist, T., & Sugden, K. D. 2005. Recognition of the oxidized lesions spiroiminodihydantoin and guanidinohydantoin in DNA by the mammalian base excision repair glycosylases NEIL1 and NEIL2. *DNA Repair*, **4**(1), 41–50.
- Hakem, R. 2008. DNA-damage repair; the good, the bad, and the ugly. *The EMBO Journal*, **27**(4), 589–605.
- Hazra, T. K., Izumi, T., Boldogh, I., Imhoff, B., Kow, Y. W., Jaruga, P., Dizdaroglu, M., & Mitra, S. 2002a. Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **99**(6), 3523–8.
- Hazra, T. K., Kow, Y. W., Hatahet, Z., Imhoff, B., Boldogh, I., Mookapati, S. K., Mitra, S., & Izumi, T. 2002b. Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. *The Journal of Biological Chemistry*, **277**(34), 30417–20.
- Hazra, Tapas K, Das, Aditi, Das, Soumita, Choudhury, Sujata, Kow, Yoke W, & Roy, Rabindra. 2007. Oxidative DNA damage repair in mammalian cells: a new perspective. *DNA repair*, **6**(4), 470–80.
- Hegde, M. L., Hazra, T. K., & Mitra, S. 2008. Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Research*, **18**(1), 27–47.
- Hildrestrand, G. A., Neurauter, C. G., Diep, D. B., Castellanos, C. G., Krauss, S., Bjørås, M., & Luna, L. 2009. Expression patterns of Neil3 during embryonic brain development and neoplasia. *BMC Neuroscience*, **10**(Jan.), 45.
- Hoeijmakers, JHJ. 2001. Genome maintenance mechanisms for preventing cancer. *Nature*, **411**, 366–374.
- Huang, J. C., Svoboda, D. L., Reardon, J. T., & Sancar, A. 1992. Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proceedings of the National Academy of Sciences of the United States of America*, **89**(8), 3664–8.
- Imamura, K., Wallace, S. S., & Doublié, S. 2009. Structural characterization of a

- viral NEIL1 ortholog unliganded and bound to abasic site-containing DNA. *The Journal of Biological Chemistry*, **284**(38), 26174–83.
- Imamura, K., Averill, A., Wallace, S. S., & Doublié, S. 2012. Structural characterization of viral ortholog of human DNA glycosylase NEIL1 bound to thymine glycol or 5-hydroxyuracil-containing DNA. *The Journal of Biological Chemistry*, **287**(6), 4288–98.
- Jaruga, P., Xiao, Y., Vartanian, V., Lloyd, R. S., & Dizdaroglu, M. 2010. Evidence for the involvement of DNA repair enzyme NEIL1 in nucleotide excision repair of (5'R)- and (5'S)-8,5'-cyclo-2'-deoxyadenosines. *Biochemistry*, **49**(6), 1053–5.
- Jiang, D, Hatahet, Z, Melamede, R J, Kow, Y W, & Wallace, S S. 1997. Characterization of Escherichia coli endonuclease VIII. *The Journal of Biological Chemistry*, **272**(51), 32230–9.
- Jiricny, J. 2006. The multifaceted mismatch-repair system. *Nature Reviews. Molecular Cell Biology*, **7**(5), 335–46.
- Kinslow, C. J., El-Zein, R., Rondelli, C. M., Hill, C. E., Wickliffe, J. K., & Abdel-Rahman, S. Z. 2010. Regulatory regions responsive to oxidative stress in the promoter of the human DNA glycosylase gene NEIL2. *Mutagenesis*, **25**(2), 171–7.
- Kow, Y. W. 2002. Repair of deaminated bases in DNA. *Free Radical Biology and Medicine*, **33**(7), 886–893.
- Krishnamurthy, N., Zhao, X., Burrows, C. J., & David, S. S. 2008. Superior Removal of Hydantoin Lesions Relative to Other Oxidized Bases by the Human DNA Glycosylase hNEIL1. *Biochemistry*, **47**, 7137–7146.
- Krokan, H. E., Standal, R., & Slupphaug, G. 1997. DNA glycosylases in the base excision repair of DNA. *Biochemical Journal*, **325**, 1–16.
- Krokan, H. E., Kavli, B., & Slupphaug, G. 2004. Novel aspects of macromolecular repair and relationship to human disease. *Journal of Molecular Medicine (Berlin, Germany)*, **82**(5), 280–97.
- Krokeide, S. Z., Bolstad, N., Laerdahl, J. K., Bjørås, M., & Luna, L. 2009. Expression and purification of NEIL3, a human DNA glycosylase homolog. *Protein Expression and Purification*, **65**(2), 160–4.
- Kundrot, C E. 2004. Which strategy for a protein crystallization project? *Cellular and Molecular Life Sciences*, **61**(5), 525–36.

- Laat, W. L., Jaspers, N. G. J., & Hoeijmakers, J. H. J. 1999. Molecular mechanism of nucleotide excision repair. *Genes & Development*, **13**, 768–785.
- Lau, Y, Schärer, O D, Samson, L, Verdine, G L, & Ellenberger, T. 1998. Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. *Cell*, **95**(2), 249–58.
- Leslie, A. G. W, & Powell, H. R. 2007. *Evolving Methods for Macromolecular Crystallography*.
- Li, Guo-Mi. 2008. Mechanisms and functions of DNA mismatch repair. *Cell research*, **18**(1), 85–98.
- Lindahl & Nyberg. 1972. Rate of depurination of native deoxyribonucleic acid. *Biochemistry*, **11**(19), 3610–3618.
- Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.
- Lindahl, T. 1999. Quality Control by DNA Repair. *Science*, **286**(5446), 1897–1905.
- Liu, M., Bandaru, V., Bond, J. P., Jaruga, P., Zhao, X., Christov, P. P., Burrows, C. J., Rizzo, C. J., Dizdaroglu, M., & Wallace, S. S. 2010. The mouse ortholog of NEIL3 is a functional DNA glycosylase in vitro and in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(11), 4925–30.
- Lord, C. J., & Ashworth, A. 2012. The DNA damage response and cancer therapy. *Nature*, **481**(7381), 287–94.
- Madhusudan, S., & Middleton, M. R. 2005. The emerging role of DNA repair proteins as predictive, prognostic and therapeutic targets in cancer. *Cancer Treatment Reviews*, **31**(8), 603–17.
- Mandal, S. M., Hegde, M. L., Chatterjee, A., Hegde, P. M., Szczesny, B., Banerjee, D., Boldogh, I., Gao, R., Falkenberg, M., Gustafsson, C. M., Sarkar, P. S., & Hazra, T. K. 2012. Role of human DNA glycosylase Nei-like 2 (NEIL2) and single strand break repair protein polynucleotide kinase 3'-phosphatase in maintenance of mitochondrial genome. *The Journal of Biological Chemistry*, **287**(4), 2819–29.
- McCoy, A. J, W., Grosse-Kunstleve R., Adams, P. D., Winn, M. D, Storoni, L. C., & Read, R. J. 2007. Phaser Crystallographic Software. *Journal of Applied Crystallography*, **40**, 658–674.

- Mol, C D, Arvai, a S, Slupphaug, G, Kavli, B, Alseth, I, Krokan, H E, & Tainer, J a. 1995. Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell*, **80**(6), 869–78.
- Morland, I., Rolseth, V., Luna, L., Rognes, T., Bjørås, M., & Seeberg, E. 2002. Human DNA glycosylases of the bacterial Fpg / MutM superfamily : an alternative pathway for the repair of 8-oxoguanine and other oxidation products in DNA. *Structure and Sequence analysis. Nucleic Acids Research*, **30**(22), 4926–4936.
- Pearl, L H. 2000. Structure and function in the uracil-DNA glycosylase superfamily. *Mutation Research*, **460**(3-4), 165–81.
- Reardon, J. T., & Sancar, A. 2003. Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes & Development*, **17**(20), 2539–51.
- Rosenquist, T. 2003. The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. *DNA Repair*, **2**(5), 581–591.
- Sancar, A., Lindsey-Boltz, L., Unsal-Kaçmaz, K., & Linn, S. 2004. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annual Review of Biochemistry*, **73**(Jan.), 39–85.
- Savva, R., & Pearl, L. H. 1995. The structural basic of specific base excision repair by uracil-DNA glycosylase. *Nature*, **373**, 487–493.
- Schärer, O. D., & Jiricny, J. 2001. Recent progress in the biology, chemistry and structural biology of DNA glycosylases. *BioEssays*, **23**(3), 270–81.
- Sedgwick, B., Bates, P., Paik, J., Jacobs, S. C., & Lindahl, T. 2007. Repair of alkylated DNA: recent advances. *DNA Repair*, **6**(4), 429–42.
- Seeberg, E, Eide, L, & Bjørås, M. 1995. The base excision repair pathway. *Trends in Biochemical Science*, **20**(10), 391–397.
- Sejrsted, Y., & Hildrestrand, G. A. 2011. Endonuclease VIII-like 3 (Neil3) DNA glycosylase promotes neurogenesis induced by hypoxia-ischemia. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(46), 18802–7.
- Slupphaug, G. 2003. The interacting pathways for prevention and repair of oxidative DNA damage. *Mutation Research*, **531**(1-2), 231–251.
- Takao, M., Kanno, S., Kobayashi, K., Zhang, Q., Yonei, S., van der Horst, G., & Yasui, A. 2002. A back-up glycosylase in Nth1 knock-out mice is a functional Nei

- (endonuclease VIII) homologue. *The Journal of Biological Chemistry*, **277**(44), 42205–13.
- Takao, M., Oohata, Y., Kitadokoro, K., Kobayashi, K., Iwai, S., Yasui, A., Yonei, S., & Zhang, Q. 2009. Human Nei-like protein NEIL3 has AP lyase activity specific for single-stranded DNA and confers oxidative stress resistance in *Escherichia coli* mutant. *Genes to cells : devoted to molecular & cellular mechanisms*, **14**(2), 261–70.
- Thayer, M M, Ahern, H, Xing, D, Cunningham, R P, & Tainer, J a. 1995. Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *The EMBO journal*, **14**(16), 4108–20.
- Verri, A, Mazzarello, P, Spadari, S, & Focher, F. 1992. Uracil-DNA glycosylases preferentially excise mispaired uracil. *The Biochemical Journal*, **287**(Nov.), 1007–10.
- Wallace, S. S. 2002. Biological consequences of free radical-damaged DNA bases. *Free Radical Biology and Medicine*, **33**(1), 1–14.
- Wallace, S. S. 2003. The enigma of endonuclease VIII. *DNA Repair*, **2**(5), 441–453.
- Wood, R D, Mitchell, M, Sgouros, J, & Lindahl, T. 2001. Human DNA repair genes. *Science*, **291**(5507), 1284–9.
- Zhao, X., Krishnamurthy, N., Burrows, C. J., & David, S. S. 2010. Mutation versus repair: NEIL1 removal of hydantoin lesions in single-stranded, bulge, bubble, and duplex DNA contexts. *Biochemistry*, **49**(8), 1658–66.
- Zharkov, D. O. 2007. Structure and conformational dynamics of base excision repair DNA glycosylases. *Molecular Biology*, **41**(5), 702–716.
- Zharkov, D. O., Golan, G., Gilboa, R., Fernandes, A. S., Gerchman, S: E., Kycia, J. H., Rieger, R., Grollman, A. P., & Shoham, G. 2002. Structural analysis of an *Escherichia coli* endonuclease VIII covalent reaction intermediate. *The EMBO Journal*, **21**(4), 789–800.
- Zharkov, D. O., Shoham, G., & Grollman, A. P. 2003. Structural characterization of the Fpg family of DNA glycosylases. *DNA Repair*, **2**(8), 839–862.



# Appendix

## Section A - Materials

**Table A - Chemicals**

Chemical	Manufacturer
1 kb DNA ladder	New England Biosciences
100 bp DNA ladder	New England Biosciences
Acetic Acid	Merck
Ammonium persulfate (APS)	Merck
Ampicillin	Merck
Bacto Agar	Difco
Bacto Trypone	Difco
Bacto Yeast Extract	Difco
Betaine hydrochloride	Sigma
$\beta$ -mercaptoethanol	Sigma
Bio-Rad Protein Assay	Bio-Rad
Bovine serum albumin (BSA, 100x)	New England Biosciences
Calcium chloride	Merck
Coomassie Blue	Amersham Biosciences
Dimethyl sulfoxide (DMSO)	Sigma
Dithiothreitol (DTT)	Sigma
DNA loading buffer (6x)	Fermentas
dNTP mix	Fermentas
<i>DpnI</i> restriction enzyme	New England Biosciences
Ethanol (Absolut Prima)	Arcus
Ethylenediaminetetraacetic acid (EDTA)	Sigma
Ethylene glycol	Sigma
D - (+) - Glucose	Sigma

D - (+) - Sorbitol	VWR
Disodium hydrogen phosphate ( $\text{Na}_2\text{HPO}_4$ )	Sigma
Glycerol	Merck
HEPES	Sigma
Imidazole	Merck
Isopropyl- $\beta$ -D-Thiogalactopyranoside (IPTG)	Sigma
Kanamycin	Sigma
Lambda DNA/EcoRI+HindIII Marker	Fermentas
Long Ranger (50%)	Lonza
LB Broth, Miller (Luria-Bertani)	Difco
Magnesium chloride ( $\text{MgCl}_2$ )	Merck
Magnesium sulphate ( $\text{MgSO}_4$ )	Sigma
Methanol	Merck
3-methyl-1,5-pentanediol (MPD)	Sigma
3-(N-morpholino)ethanesulfonic acid (MES)	Sigma
3-(N-morpholino) propanesulfonic acid (MOPS)	Invitrogen
<i>Nde</i> I restriction enzyme	New England Biosciences
Ni-NTA agarose	Qiagen
NuPage loading buffer (4x)	Invitrogen
Pfu reaction buffer (10x)	Stratagene
<i>Pfu Turbo</i> polymerase	Fermentas
Polyethylene glycol (PEG) 400	Sigma
Polyethylene glycol (PEG) 3350	Sigma
Potassium chloride (KCl)	Merck
SeaKem <sup>®</sup> GTG <sup>®</sup> Agarose	Lonza
SeeBlue Plus2 standard	Invitrogen
Sodium acetate (NaAc)	Merck
Sodium borhydride ( $\text{NaBH}_4$ )	Sigma
Sodium natrate ( $\text{C}_4\text{H}_4\text{Na}_2\text{O}_6$ )	Sigma
Sodium chloride	Merck
Sodium dihydrogen phosphate ( $\text{NaH}_2\text{PO}_4$ )	Merck
Sodium dodecyl sulphate (SDS)	Bio-Rad
SYBR <sup>®</sup> Safe DNA gel stain (100x)	Invitrogen
T4 ligase buffer (10x)	New England Biosciences
T4 ligase	New England Biosciences

Taq polymerase	New England Biosciences
Taurin	Sigma
Tetramethylethylenediamine (TEMED)	Sigma
Tris(hydroxymethyl)aminomethane (Tris)	Sigma
Uracil DNA glycosylase (UDG)	New England Biosciences
UDG reaction buffer (10x)	New England Biosciences
Urea	Duchefa Biochemie
<i>Xho</i> I restriction enzyme	New England Biosciences

**Table B - Chromatographic Column materials**

<b>Material</b>	<b>Manufacturer</b>
HiTrap SP XL column	GE Healthcare
Resource S 6 ml column	GE Healthcare
Superdex 75 R10/300 Column	GE Healthcare

**Table C - Hardware**

<b>Hardware</b>	<b>Manufacturer</b>
Avanti Centrifuge J-26 XP	Beckman Coulter
Biofuge Pico	Heraeus Instruments
ECM 830 Electro Square Porator	BTX
Eppendorf's biophotometer	Eppendorf
Eppendorf's Thermomixer comfort	Eppendorf
Gel logic 200 image system	Kodak
InnOva 4230	New Brunswick Scientific
Oryx6 Robot	Douglas Instruments
Peltier Thermal Cycler (PTC) - 200	MJ Research
Power Ease 500	Invitrogen
Typhoon 9419 Variable Mode Imager	GE Healthcare
Äkta purifier system	GE Healthcare

**Table D - Software**

<b>Software</b>	<b>Manufacturer</b>
CCP4 package	Technelysium Pty Ltd
Chromas Lite	
Coot	
ImageQuant TL Version 2003.02	Amersham Biosciences
iMosflm	Douglas Instruments
WaspRunW	

**Table E - Equipment**

<b>Equipment</b>	<b>Manufacturer</b>
Amicon Ultrafree tubes	Millipore
Crystal Clear Sealing Tape	Hampton
Cuvettes	Sarstedt
Electroporation Cuvettes Plus	BTX
MRC 2 Well Crystallization Plates	Swissci
PCR Uvette	Eppendorf
Storage Phosphor Screen	GE Healthcare
Vapour Batch Plates	Douglas Instruments

**Table F - Kits**

<b>Kit</b>	<b>Manufacturer</b>
NucleoBond (Xtra) Midi Kit	Macherey-Nagel
TOPO TA Cloning Kit	Invitrogen
QIAEXII Gel Extraction Kit	Qiagen
QIAprep Spin Miniprep Kit	Qiagen

**Table G - Crystallization Kits**

<b>Crystallization kit</b>	<b>Manufacturer</b>
Alternative Precipitation	Molecular Dimensions
Basic + Extension	Sigma
Index	Hampton
JCSG+	Hampton
ProPlex	Molecular Dimensions
Wizard I + II	Emerald Biosciences

## Section B - PCR mixtures and PCR programs

### PCR of NEIL1 inserts A-D

PCR mixture for amplification of linear NEIL1 inserts A-D			PCR program for amplification of linear NEIL1 inserts A-D		
1	μl	pET22b-NEIL1-1-337 template	94	°C	2 minutes
5	μl	10 x Pfu reaction buffer	94	°C	30 seconds
5	μl	3.5 pmol forward primer A-D	55	°C	30 seconds
			x 35		
5	μl	3.5 pmol revers primer A-D	68	°C	2 minutes
5	μl	10mM dNTP mix	72	°C	1 minute
1	μl	DMSO	4	°C	for ever
1	μl	MgSO <sub>4</sub>			
1	μl	Pfu polymerase			
26	μl	mq-H <sub>2</sub> O			
50	μl				

### PCR for TOPO cloning

PCR mixture for amplification of linear NEIL1 inserts A-D			PCR program for amplification of linear NEIL1 inserts A-D		
1	μl	pET22b-NEIL1-1-337 template	94	°C	2 minutes
5	μl	10 x Taq reaction buffer	94	°C	30 seconds
5	μl	3.5 pmol forward primer A-D	55	°C	30 seconds
			x 35		
5	μl	3.5 pmol revers primer A-D	68	°C	2 minutes
0.5	μl	50mM dNTP mix	72	°C	30 minutes
1	μl	DMSO	4	°C	for ever
1	μl	50mM MgCl <sub>2</sub>			
1	μl	Taq polymerase			
29.5	μl	mq-H <sub>2</sub> O			
50	μl				

**QuikChange of all NEIL1, NEIL2 and NEIL3 constructs**

PCR mixture for QuikChange of NEIL1, NEIL2 and NEIL3 constructs			PCR program for QuikChange of NEIL1, NEIL2 and NEIL3 constructs		
1	µl	Template	95	°C	2 minutes
5	µl	10 x Pfu reaction buffer	95	°C	30 seconds
5	µl	3.5 pmol forward primer	55	°C	30 seconds
5	µl	3.5 pmol revers primer	68	°C	8 minutes
5	µl	10mM dNTP mix	72	°C	7 minutes
1	µl	DMSO	4	°C	for ever
1	µl	Pfu polymerase			
26	µl	mq-H <sub>2</sub> O			
50	µl				

*Template for NEIL1 (286aa), NEIL1 (305aa) and NEIL1 (325aa): pET22b-NEIL1 (337)*

*Template for NEIL1 (286aaK54Q): pET-22b-NEIL1 (286aa)*

*Template for NEIL1 (305aaE3Q): pET22b-NEIL1 (305aa)*

*Template for NEIL2(E3Q): pET-22b-NEIL2*

*Template for NEIL3 (282aa) and NEIL3 (289aa): pETDuet-NEIL3 (301aa)*

*Template for NEIL3 (282aaK81Q): pETDuet-NEIL3 (282aa)*

*Template for NEIL4 (289aaK81Q): pETDuet-NEIL3 (289aa)*

*For primers, refer to section F.*

## Section C - Restriction- and ligation reactions

### Double digest reactions

Double digest reaction for NEIL1 inserts A-D			Double digest reaction for pET22b cloning vector		
16	μl	PCR product A-D	32	μl	pET22b
2	μl	NEB buffer #4	4	μl	NEB buffer #4
1	μl	Nde I	2	μl	Nde I
1	μl	Xho I	2	μl	Xho I
<hr/>			<hr/>		
20	μl		40	μl	

*The restriction reactions were incubated at 37 °C for 2-3 hours.*

### Ligation reactions

Ligation reaction of NEIL1 inserts A-D and pET22b cloning vector			Control ligation reaction of pET22b cloning vector		
12	μl	NEIL1 A-D insert	5	μl	pET22b vector
5	μl	pET22b vector	2	μl	10 x ligation buffer
2	μl	10 x ligation buffer	1	μl	T4 DNA ligase
1	μl	T4 DNA ligase	12	μl	mq-H <sub>2</sub> O
<hr/>			<hr/>		
20	μl		20	μl	

*The ligation reactions were incubated at room temperature for 1-3 days.*

### Test cutting

Test cutting of NEIL1 A-D		
20	μl	plasmid miniprep A-D
2.5	μl	NEB buffer #4
1	μl	Nde I
1	μl	Xho I
<hr/>		
24.5	μl	

*The reactions were incubated at  
37 °C for 1-3 hours.*

### TOPO cloning reaction

TOPO cloning reaction		
4	μl	PCR product
1	μl	salt solution (supplied with kit)
1	μl	TOPO vector
<hr/>		
6	μl	

*The reaction was incubated at  
room temperature for 30 minutes*



## Section D - Solutions, buffers and gels

### Running buffers for agarose- and PAGE gels, staining and destaining solutions

1 x TAE			1 x MOPS		
4.84	g	Tris base	8.38	g	MOPS
1.14	g	Glacial acetic acid	1.36	g	NaAc
0.74	g	EDTA * 2H <sub>2</sub> O	2	ml	0.5 M EDTA pH 7.2
mq-H <sub>2</sub> O to a final volume of 1 L			mq-H <sub>2</sub> O to a final volume of 1 L		
Coomassie Blue staining solution			Destaining solution		
40	%	Methanol	40	%	Methanol
10	%	Acetic acid	10	%	Acetic acid
0.1	%	Coomassie Blue	4	%	Glycerol

### Protein crack buffer

Protein crack buffer		
2	%	SDS
5	%	β-ME (14.6M)
10	%	Glycerol
60	mM	Tris pH 6.8
1.4	mg/ml	BPB

### Buffers used in NEIL1 and NEIL2 protein purification and cross-linking

Sonication buffer			50 mM imidazole buffer		
10	mM	Imidazole	50	mM	Imidazole
50	mM	Na <sub>2</sub> HPO <sub>4</sub> /Na <sub>2</sub> HPO <sub>4</sub> pH 8	50	mM	Na <sub>2</sub> HPO <sub>4</sub> /Na <sub>2</sub> HPO <sub>4</sub> pH 8
300	mM	NaCl	300	mM	NaCl
10	mM	β-ME	10	mM	β-ME
300 mM imidazole buffer			Low salt buffer A		
300	mM	Imidazole	50	mM	NaCl
50	mM	Na <sub>2</sub> HPO <sub>4</sub> /Na <sub>2</sub> HPO <sub>4</sub> pH 8	20	mM	Tris pH 7
300	mM	NaCl	10	mM	β-ME
10	mM	β-ME			

High salt buffer A			Cross-linking buffer for NEIL1		
2	M	NaCl	25	mM	Na <sub>2</sub> HPO <sub>4</sub> /NaH <sub>2</sub> PO <sub>4</sub> pH 6.8
20	mM	Tris pH 7	50	mM	NaCl
10	mM	β-ME	1	mM	EDTA

Cross-linking buffer A for NEIL2			Cross-linking buffer B for NEIL2		
25	mM	Na <sub>2</sub> HPO <sub>4</sub> /NaH <sub>2</sub> PO <sub>4</sub> pH 6.8	25	mM	Na <sub>2</sub> HPO <sub>4</sub> /NaH <sub>2</sub> PO <sub>4</sub> pH 6.8
50	mM	NaCl	100	mM	NaCl
1	mM	EDTA	1	mM	EDTA

Gel filtration buffer A for NEIL2			Gel filtration buffer B for NEIL2		
100	mM	NaCl	550	mM	NaCl
10	mM	MES pH 6.5	10	mM	MES pH 6.5
10	mM	β-ME	10	mM	β-ME

### Buffers used in NEIL3 protein purification

Sonication buffer			50 mM imidazole buffer		
10	mM	Imidazole	50	mM	Imidazole
50	mM	Tris pH 8	50	mM	Tris pH 8
300	mM	NaCl	300	mM	NaCl
10	mM	β-ME	10	mM	β-ME

300 mM imidazole buffer			Gel filtration buffer		
300	mM	Imidazole	20	mM	MES pH 6.5
50	mM	Tris pH 8	50	mM	NaCl
300	mM	NaCl	10	mM	β-ME
10	mM	β-ME			

**Buffers and gels used in NEIL2 glycosylase assays**

Protein dilution buffer			Reaction mixture for $^{32}\text{P}$ labeling of substrate	
1	mM	EDTA	2	$\mu\text{l}$ 10 x PNK buffer
25	mM	HEPES pH 7.9	1	$\mu\text{l}$ T4 PNK
1	mM	DTT	1	$\mu\text{l}$ 3.5 pmol oligo
0.1	mg/ml	BSA	1	$\mu\text{l}$ $\gamma$ - $^{32}\text{P}$ -ATP
15	%	Glycerol	15	$\mu\text{l}$ mq-H <sub>2</sub> O
			20	$\mu\text{l}$
20% native TBE PAGE gel			20% denaturing Taurin PAGE gel	
0.5	ml	5x TBE buffer	2.7	g Urea
2	ml	Long Ranger	1.4	ml mq-H <sub>2</sub> O
2.5	ml	mq-H <sub>2</sub> O	2.5	ml Long Ranger
17.5	$\mu\text{l}$	10% APS	0.3	ml 20x Taurin
1.75	$\mu\text{l}$	TEMED	30	$\mu\text{l}$ 10% APS
			3	$\mu\text{l}$ TEMED
0.5x TBE				
5.4	g	Tris base		
2.75	g	Boric acid		
20	ml	EDTA pH 8.0 (0.5M)		
mq-H <sub>2</sub> O to a final volume of 1 L				

**Media**

LB-medium			LB-medium with sorbitol and betaine		
25	g	Broth-Miller	25	g	Broth-Miller
1	L	mq-H <sub>2</sub> O	91.1	g	D - (+) - Sorbitol
<i>The medium was autoclaved</i>			0.384	g	Betaine hydrochloride
			1	L	mq-H <sub>2</sub> O
			<i>The medium was autoclaved</i>		
SOC-medium					
20	g	Bacto Typtone			
5	g	Bacto Yeast Extract			
2	ml	NaCl (5M)			
2.5	ml	KCl			
10	ml	MgSO <sub>4</sub>			
10	ml	MgCl <sub>2</sub>			
20	ml	Glucose (1M)			
mq-H <sub>2</sub> O to a final volume of 1 L					
<i>The medium was pH adjusted and autoclaved</i>					
<i>before glucose was added</i>					

## **Section E - Nucleotides used for generation of complexes between NEIL and DNA**

### **Nucleotides used in cross-linking experiments**

#### **11mer DNA**

Uracil containing strand: 5' - GCTAC-U-GATCG - 3'

Complementary strand: 5' - CGATC-G-GTAGC - 3'

#### **13mer DNA**

Uracil containing strand: 5' -GGCTAC-U-GATCGG - 3'

Complementary strand: 5' - CCGATC-G-GTAGCC - 3'

### **THF-nucleotides**

THF-containing strand: 5' - GCTAC-THF-GATCG - 3'

Complementary strand with opposite adenine: 5' - CGATC-A-GTAGC - 3'

Complementary strand with opposite cytosine: 5' - CGATC-C-GTAGC - 3'

Complementary strand with opposite thymine: 5' - CGATC-T-GTAGC - 3'

## Section F - Primer sequences

### NEIL1 primers

#### Primers used for amplification of NEIL1 inserts A-D

Forward primer (NdeI restriction site underlined):

5' - GCAGCTCGGGGCCCTCAGGCATATGAATTAT - 3'

Reverse primer A (Xho I restriction site underlined):

5' - TATCTCGAGGGGTGCCAACGGTCCAGGATCC - 3'

Reverse primer B (Xho I and TEV restriction sites underlined and in bold, respectively):

5' - TATCTCGAG**GCCCTGAAAATAAAGATTCTC**  
CCCGGGTGCCAACGGTCCAGGATCCCCC - 3'

Reverse primer C (Xho I restriction site underlined):

5' - TATCTCGAGTCCAGGATCCCCCTGGAACC - 3'

Reverse primer D (Xho I and TEV restriction sites underlined and in bold, respectively): 5' - TATCTCGAG**GCCCTGAAAATAAAGATTCTC**  
TCCAGGATCCCCCTGGAACCAGATGGTACG - 3'

#### QuikChange-primers for NEIL1 (286aa)

Forward #1 (mutated sequence underlined): 5' - GGTTCCAGGGGGATCCTGGA  
CATCATCATCCCAAAGGGCGCAAGTCC - 3'

Forward #2 (mutated sequence underlined): 5' - GGATCCTGGACATCATCAT  
CACCACCACTGAAAGTCCCGCAAAAA GAAATCC - 3'

#### QuikChange-primers for NEIL1 (305aa)

Forward #1 (mutated sequence underlined): 5' - CCAAGGCCACACAGCTG  
CATCATCACGACAGAGTGGAGGACGC - 3'

Forward #2 (mutated sequence underlined): 5' - CACAGCTGCATCATCAC  
CACCATCATTAGGACGCTTTGCCTCCAAGC - 3'

#### **QuikChange-primers for NEIL1 (325aa)**

Forward #1 (mutated sequence underlined): 5' - CCCCTTCCAGGACACGA  
CACCATCACAGAGACCTTCCTAGAGG - 3'

Forward #2 (mutated sequence underlined): 5' - GGACACGACACCATCAT  
CACCACCACTAAAGAGGACTGCAACCCAGC - 3'

#### **QuikChange-primers for NEIL1-286 K54Q**

Forward (mutated sequence underlined):

5' - TCAGCTTCAGCCCGCGGCCAGGAGCTGCGCCTGATACTG - 3'

#### **QuikChange-primers for NEIL1-305 E3Q**

Forward (mutated sequence underlined):

5' - GGAGATATACATATGCCTCAGGGCCCCGAGCTCCACCTGG - 3'

#### **Primers for sequencing**

Forward A: 5' - GGTCAAACACTTACTCC - 3'

Forward B: 5' - TAATACGACTCACTATAGGG - 3

Reverse A: 5' - GGTACTCCTGCAAGACACAGG - 3'

Reverse B: 5' - GCTAGTTATTGCTCAGCGG - 3

#### **NEIL2 primers**

#### **QuikChange-primers for NEIL1-305 E3Q**

Forward (mutated sequence underlined):

5' - GGAGATATACATATGCCCAAGGGCCGTTGGGTGGAGG - 3'

### **Primers for sequencing**

Forward: 5' - TAATACGACTCACTATAGGG - 3'

Reverse A: 5' - GCTAGTTATTGCTCAGCGG - 3'

Reverse B: 5' - CTTTCCATGGAC - 3'

### **NEIL3 primers**

#### **QuikChange-primers for NEIL3 (282aa)**

Forward #1 (mutated sequence underlined): 5' - CCTCACTGTCAAAAAGAA  
CATCATCATCATGTTGACATATGCAAGC - 3'

Forward #2 (mutated sequence underlined): 5' - CCTCACTGTCAAAAAGAA-  
CATCATCATCAT  
CATCACTAATGCAAGCTACCGACTAG - 3'

#### **QuikChange-primers for NEIL3 (289aa)**

Forward #1 (mutated sequence underlined): 5' - CCTCAACATGTTGACATA  
CATCATCATCCGACTAGAAATACTATAATCAGTTGG - 3'

Forward #2 (mutated sequence underlined): 5' - GTTGACATACATCATCAT  
CATCATCATTAGACTATAATCAGTTGGCACCAC - 3'

#### **QuikChange-primers for K81Q mutation of NEIL3 (301aa), NEIL3 (289aa) and NEIL3 (282aa)**

Forward (mutated sequence underlined):

5' - GGCGTGGAAGCTTTGGGGCAGGAGCTCTTTATGTACTTTGG - 3'

### **Primers for sequencing**

Forward: 5' - ATTATGCGGCCGTGTACAA - 3'

Reverse: 5' - GCTAGTTATTGCTCAGCG - 3'



## Section G - Nucleic acid sequences

### NEIL1 (1173nt)

ATGCCTGAGGGCCCCGAGCTGCACCTGGCCAGCCAGTTTGTGAATGAGGCCTGCAGGGCGCTGGTGTTCG  
GCGGCTGCGTGGAGAAGTCCTCTGTACGCCGCAACCCTGAGGTGCCCTTTGAGAGCAGTGCCCTACCGCAT  
CTCAGCTTCAGCCCGCGGCAAGGAGCTGCGCCTGATACTGAGCCCTCTGCCTGGGGCCCAGCCCCAACAG  
GAGCCACTGGCCCTGGTCTTCCGCTTCGGCATGTCCGGCTCTTTTCAGCTGGTGCCCCGCGAGGAGCTGC  
CACGCCATGCCACCTGCGCTTTTACACGGCCCCGCTGGCCCCCGGCTCGCCCTATGTTTCGTGGACAT  
CCGCCGGTTTCGGCCGCTGGGACCTTGGGGGAAAGTGGCAGCCGGGCGCGGGCCCTGTGTCTTGCAGGAG  
TACCAGCAGTTCAGGGAGAATGTGCTACGAAACCTAGCGGATAAGGCCTTTGACCGGCCCATCTGCGAGG  
CCCTCCTGGACCAGAGGTTCTTCAATGGCATTGGCAACTATCTGCGGGCAGAGATCCTGTACCGGCTGAA  
GATCCCCCCTTTGAGAAGGCCCGCTCGGTCTGGAGGCCCTGCAGCAGCACAGGCCGAGCCCGGAGCTG  
ACCCTGAGCCAGAAGATAAGGACCAAGCTGCAGAATCCAGACCTGCTGGAGCTATGTCACTCAGTGCCCA  
AGGAAGTGGTCCAGTTGGGGGGCAAAGGCTACGGGTCAGAGAGCGGGGAGGAGGACTTTGCTGCCTTTTCG  
AGCCTGGCTGCGCTGCTATGGCATGCCAGGCATGAGCTCCCTGCAGGACCGGCATGGCCGTACCATCTGG  
TTCCAGGGGGATCCTGGACCGTTGGCACCCAAAGGGCGCAAGTCCCGCAAAAAGAAATCCAAGGCCACAC  
AGCTGAGTCCTGAGGACAGAGTGGAGGACGCTTTGCCTCCAAGCAAGGCCCTTCCAGGACACGAAGGGC  
AAAGAGAGACCTTCCTAAGAGGACTGCAACCCAGCGGCCTGAGGGGACCAGCCTCCAGCAGGACCCAGAA  
GCTCCACAGTGCCCAAGAAGGGGAGGAGGAAGGGGCGACAGGCAGCCTCTGGCCACTGCAGACCCCGGA  
AGGTCAAGGCTGACATCCCATCCTTGGAACCAGAGGGGACCTCAGCCTCTTAG

### NEIL2 (999nt)

ATGCCAGAAGGGCCGTTGGTGAGGAAATTTACCATTTGGTCTCCCCCTTTGTGGGTGAGCAGGTGGTCA  
AGACAGGGGGCAGCAGTAAGAAGCTACAGCCCGCCAGCCTGCAGTCTCTGTGGCTCCAGGACACCCAGGT  
CCATGGAAAGAAATTATTCCCTTAGATTTGATCTAGATGAAGAAATGGGGCCCCCTGGCAGCAGCCCAACA  
CCAGAGCCTCCACAAAAAGAAAGTGCAGAAGGAAGGGGCTGCGGACCCAAAGCAGGTGCGGGAGCCCAGCG  
GGCAGAAGACCCCTTGATGGATCCTCACGGTCTGCAGAGCTCGTCCCCCAGGGCGAGGATGATTCTGAGTA  
TTTGAGAGAGAGACGCCCCGTCAGGAGATGCTGGGAGGTGGCTGCGTGTGAGCTTTGGTTTGTGTTGGCAGC  
GTTTGGGTGAACGATTTCTCCAGAGCCAAGAAAGCCAACAAGAGGGGGGACTGGAGGGACCCCTCCCCGA  
GGTTGGTCTGCACTTTGGTGGTGGTGGCTTCCTGGCATTTTATAATTGTCAGTTGTCTTGGAGCTCTTC  
CCCAGTGGTCACACCCACCTGTGACATCCTGTCTGAGAAGTTCCATCGAGGACAAGCCTTAGAAGCTCTA  
GGCCAGGCTCAGCCTGTCTGCTATACACTGCTGGACCAGAGATACTTCTCAGGGCTAGGGAACATCATT  
AGAATGAAGCCTTGACAGAGCTGGGATCCATCCCTTTCTCTCGGTTTCAGTCCTGAGTGCCCTCGCGTCG  
GGAGGTCTGGTGGATCACGTGGTGGAGTTCAGTACAGCCTGGCTGCAGGGCAAGTTCCAAGGCAGACCG  
CAGCACACACAGGTCTACCAGAAAGAACAGTGCCCTGCTGGCCACCAGGTGATGAAGGAGGCGTTTGGGC  
CCGAAGATGGGTTACAGAGGCTCACCTGGTGGTGCCCGCAGTGCCAGCCCCAGTTGTCAGAGGAGCCAGA  
GCAGTGCCAGTTCTCCTAA

**NEIL3 (1818nt)**

ATGGTGGAAGGACCAGGCTGTACTCTGAATGGAGAGAAGATTCGCGCGCGGGTGCTCCCGGGCCAGGCGG  
TGACCGGCGTGCGGGGAAGCGCTCTGCGGAGTCTGCAGGGCCGCGCCTTGCGGCTCGCAGCCTCCACGGT  
TGTGGTCTCCCCGCAGGCTGCTGCACTGAATAATGATTCCAGCCAGAATGTCTTGAGCCTGTTTAATGGA  
TATGTTTACAGTGGCGTGGAACCTTTGGGGAAGGAGCTCTTTATGTACTTTGGACCAAAAGCTTTACGGA  
TTCATTTTCGGAATGAAAGGCTTCATCATGATTAATCCACTTGAGTATAAATATAAAAATGGAGCTTCTCC  
TGTTTTTGAAGTGCAGCTCACCAAAGATTTGATTTGTTTTCTTTGACTCATCAGTAGAACTCAGAAACTCA  
ATGGAAAGCCAACAGAGAATAAGAATGATGAAAGAATTAGATGTATGTTACCTGAATTTAGTTTCTTGA  
GAGCAGAAAAGTGAAGTTAAAAAACAGAAAGCCGGATGCTAGGTGATGTGCTAATGGATCAGAACGTATT  
GCCTGGAGTAGGGAACATCATCAAAAATGAAGCTCTCTTTGACAGTGGTCTCCACCCAGCTGTTAAAGTT  
TGTCAATTAACAGATGAACAGATCCATCACCTCATGAAAATGATACGTGATTTTCAGCATTCTCTTTTACA  
GGTGCCGTAAAGCAGGACTTGCTCTCTCTAAACACTATAAGGTTTACAAGCGTCCTAATTGTGGTCAGTG  
CCACTGCAGAATAACTGTGTGCCGCTTTGGGGACAATAACAGAATGACATATTTCTGTCTCTCACTGTCAA  
AAAGAAAATCCTCAACATGTTGACATATGCAAGCTACCGACTAGAAATACTATAATCAGTTGGACATCTA  
GCAGGGTGGATCATGTTATGGACTCCGTGGCTCGGAAGTCGGAAGAGCACTGGACCTGTGTGGTGTGTAC  
TTAATCAATAAGCCCTCTTCTAAGGCATGTGATGCTTGCTTGACCTCAAGGCCTATTGATTCAGTGCTC  
AAGAGTGAAGAAAATTCTACTGTCTTTAGCCACTTAATGAAGTACCCGTGTAATACTTTTGGAAAACCTC  
ATACAGAAGTCAAGATCAACAGAAAACTGCATTTGGAACTACAACTCTTGTCTTGACTGATTTTAGCAA  
TAAATCCAGTACTTTGGAAAAGAAAAACAAAGCAAAACCAGATACTAGATGAGGAGTTTCAAAACTCTCCT  
CCTGCTAGTGTTTGTGTTGAATGATATACAGCACCCCTCCAAGAAGACAACAAACGATATAACTCAACCAT  
CCAGCAAAGTAAACATATCACCTACAATCAGTTCAGAACTAAATTATTTAGTCCAGCACATAAAAAACC  
GAAAACAGCCCAATACTCATCACCCAGAGCTTAAAAGCTGCAACCCCTGGATATTCTAACAGTGAACCTCAA  
ATTAATATGACAGATGGCCCTCGTACCTTAAATCCTGACAGCCCTCGCTGCAGTAAACACAACCGCCTCT  
GCATTCTCCGAGTTGTGGGGAAGGATGGGGAAAACAAGGGCAGGCAGTTTTATGCCTGTCTCTACCTAG  
AGAAGCACAATGTGGATTTTTTTGAATGGGCAGATTTGTCTTCCCATTCTGCAACCATGGCAAGCGTTCC  
ACCATGAAAACAGTATTGAAGATTGGACCTAACAATGGAAAGAATTTTTTTGTGTGTCTCTTGGGAAGG  
AAAAACAATGCAATTTTTTCCAGTGGGCAGAAAATGGGCCAGGAATAAAAAATTATTCTGATGCTAA

## Section H - Amino acid sequences

### NEIL1 (full-length)

Residues 286, 305 and 325 are marked in red. K54 is marked in green, and E3 is marked in blue.

MP**E**GP~~EL~~H~~LA~~SQFVNEACRALVFGGCV~~EK~~SSVSRNPEVVPFESSAYRISA  
 SARG**K**ELRLILSPLPGAQPQ**Q**EPALVFRFGMSGFQLVPREELPRHAHL  
 RFYTAPPGPRLALCFVDIRRFGRWDLGGKWQPGRGPCVLQ**EY**QQFREN  
 VLRNLADKAFDRPICEALLDQ**R**FFNGIGNYLRAEILYRLKIPPF**EK**ARSVL  
 EALQQHRPSP**EL**TL**SQ**KIRTKLQNPDLLELCHSV**PKE**VVQLGGKGYGSES  
 GEEDFAAFRAWLRCYGM**P**GMS**LQ**DRHGRTIWFQGD**P**PLAPKGRKSR  
 KKKSKAT**Q**LSPEDRVEDALPPSKAPSRT**RA**KRDLPKRTATQRPEGTSLQ  
 QDPEAPTVPKKGRRKGRQAASGHCRPRKVKADIPSLEPEGTSAS-

### NEIL2 (full-length)

E3 is marked in blue.

MP**E**GPLVRKFHHLVSPFVGQQVVK**TG**GSSKKLQPASLQSLWLQDTQVHG  
 KKLFLRFDLDEEMGPPGSSPTPEPPQKEVQKEGAADPKQVGEP**SG**QKTLD  
 GSSRSAELVPQGEDDSEYLERDAPAGDAGRWLRSFGLFGSVVWNDF**SRA**  
 KKANKRGDWRDPSPRLVLHFGGGGFLAFYNCQLSWSSSPVVTPTCDILSEK  
 FHRGQALEALGQAQPV**CY**TL**LD**QRYFSGLGNIKNEALYRAGIHPLSLG**SVL**  
 SASRREVLVDHVVEFSTAWLQ**GK**FQGR**PQ**HTQVYQKEQCPAGHQVMKEA  
 FGPEDGLQRLTWWCPQCQPQLSEEPEQCQFSLE-

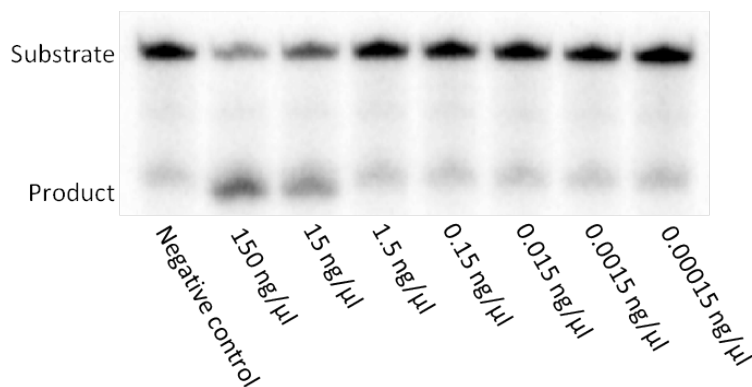
### NEIL3 (full-length)

Residues 282 and 289 are marked in red. K81 is marked in green.

MVEGPGCTLNGEKIRARVLP**GQ**AVTGVRGSALRS**LQ**GRALRLAASTVVVSPQ  
 AAALNNDSSQNVLSLFNGYVYSGVETL**GK**ELFMYFGPKALRIHF**GM**KGFIMI  
 NPLEYKYKNGASPVLEVQLTKDLICFFDSSVELRNSMESQQRIRMMKELDVC  
 SPEFSFLRAESEVKKQKGRMLGDV**MDQ**NVLPVG**NI**KNEALFDSGLHPAVK  
 VCQLTDEQIHHLMKMIRDF**SIL**FYRCRKAGLALS**KHY**KVYKRPNC**GQ**CHCRIT  
 VCRFGDNNRMTYFCPHCQ**KEN**PQHVD**ICK**LPTRNTIISWTSSRVDHVMDSVAR  
 KSEEHWT**CV**VCTLINKPSSKACD**ACL**TSRPIDSVLKSEENSTVFSHLMKYPCNTF  
 GKPHTEVKINRKTA**FG**TTTLVLTDFSNKSSTLERKTKQN**QIL**DEEFQNSPPASVC  
 LNDIQHPSKKTNDITQPSKVNISPTISSES**KLF**SPA**HK**PKTAQYSSPELKSCNP  
 GYSNSELQINMTDGPRTLNPDSPRCSKHNR**LCIL**RVVGKDGENKGRQFYACPLP  
 REAQCGFFEWADLSF**PC**NHKGKRSTMKTVLKIGPNNGKNFFVCPLGKEKQCNFF  
 QWAENGPGIKIIPGC

## Section I - Gels, primers and compounds A-P from glycosylase assays

### Titration of NEIL2 cleavage activity



*Scanned gel from titration experiments with NEIL2 in concentrations between 150 ng/μl and 0.00015 ng/μl. Substrate: 5-OHU:C.*

### 5-hydroxyuracil containing oligonucleotide

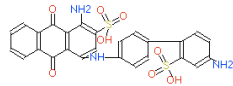
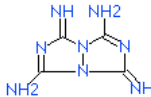
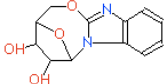
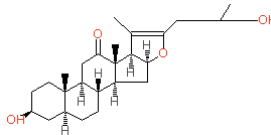
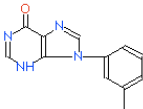
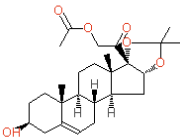
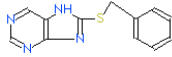
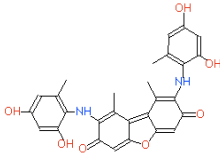
Lesion containing strand:

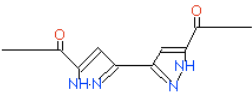
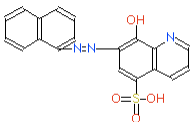
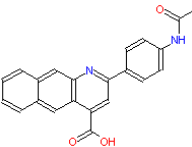
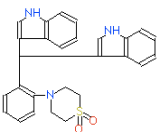
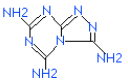
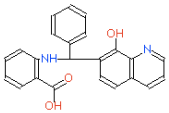
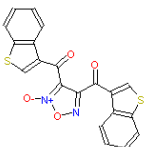
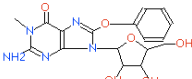
5' - GCATGCCTGCACGG-U-CATGGCCAGATCCCCGGGTACCGAG - 3'

Complementary strand:

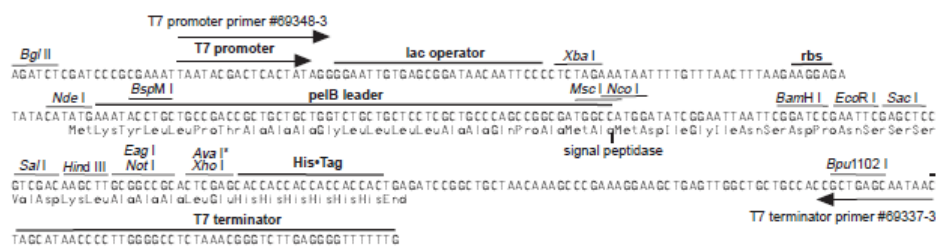
5' - CTCGGTACCCGGGGATCTGGCCATG-G-CCGTGCAGGCATGC - 3'

## Structure of compounds A-P

Compound letter	Size (Da)	Structure
A	566	
B	166	
C	248	
D	431	
E	226	
F	447	
G	242	
H	500.51	

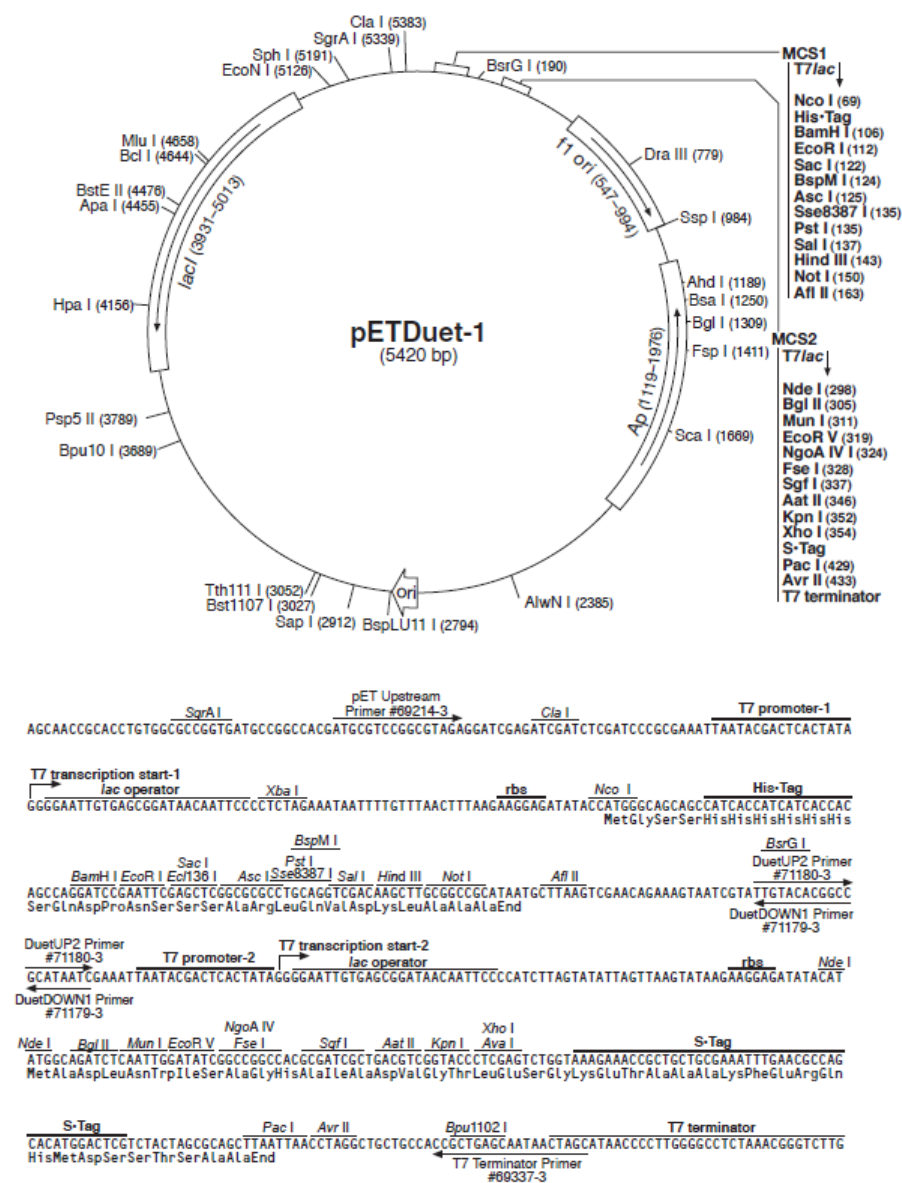
Compound letter	Size (Da)	Structure
I	246	
J	379	
K	356	
L	470	
M	166	
N	370	
O	406.43	
P	403	

pET-22b(+) sequence landmarks



125

## pET-Duet-1 vector map



pETDuet-1 cloning/expression regions