# Expression and Characterization of Recombinant Human NEIL3 from *Escherichia coli*

## MUSTAFA SALEM ALBELAZI

*Biomedical Research Centre, School of Environment and Life Sciences,*

**University of Salford, Salford. UK**

Submitted in partial fulfilment of the requirements for the Degree of

Doctor of Philosophy, 2019

# Table of Contents

## List of Tables

# List of Figures

# Acknowledgements

# List of Abbreviations:

| | |
|---|---|
| 5-OHU | 5-hydroxyuracil |
| 8-oxoG | 8-oxo-7,8-dihydroguanine |
| AP | Apurinic/apyrimidinic |
| APE1 | AP-endonuclease 1 |
| BER | Base excision repair |
| C-terminal | Carboxyl terminal |
| DHT | Dihydrothymine |
| DHU | Dihydrouracil |
| DMSO | Dimethyl sulfoxide |
| dRP | 5′-deoxyribose phosphate |
| FapyA | 4,6-diamino-5-formamidopyrimidine |
| FapyG | 2,6-diamino-4-hydoxy-5-formamidopyrimidine |
| FEN-1 | Flap endonuclease-1 |
| FPLC | Fast protein liquid chromatography |
| G | Guanine |
| Gh | Guanidinohydantoin |
| H2TH | Helix-two-turn-helix |
| HhH | Helix-hairpin-helix |
| hNEIL3 | Human NEIL3 |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| LB | Lysogeny broth |
| MmNEIL3 | Mouse NEIL3 |
| MMR | Mismatch repair |

| | |
|---|---|
| MUTYH | Human MutY homolog |
| Nei | Endonuclease VIII |
| NEIL | Endonuclease VIII-like |
| NER | Nucleotide excision repair |
| NHEJ | Non-homologous end joining |
| N-terminal | Amino terminal |
| NTH1 | Endonuclease III-like 1 |
| Nth | Endonuclease III |
| OGG1 | 8-oxoguanine DNA glycosylase |
| PAGE | Polyacrylamide gel electrophoresis |
| PCNA | Proliferating cell nuclear antigen |
| PCR | Polymerase chain reaction |
| PMSF | Phenylmethylsulfonyl fluoride |
| Polβ | DNA polymerase β |
| rpm | Revolutions per minute |
| ROS | Reactive oxygen species |
| SDS | Sodium dodecyl sulphate |
| SNP | Single nucleotide polymorphism |
| Sp | Spiroiminodihydantoin |
| TBE | Tris-borate-EDTA |
| TGS | SDS-PAGE running buffer |
| UDG | Uracil DNA glycosylase |
| TG | Thymine glycol |
| U | Uracil |
| WTB | Western blot transfer buffer |

# Abstract

The DNA glycosylase NEIL3 is one of a family of proteins that release oxidized bases from DNA, thereby initiating base excision repair. NEIL3 gene expression is normally tightly regulated and expressed only in certain rapidly dividing cells, however, human NEIL3 (hNEIL3) is also highly expressed in cells from metastatic tumours. In this project, full length hNEIL3 and truncations of the cDNA have been cloned into the pETDuet-2 expression vector for subsequent expression and purification. Each of the hNEIL3 truncations codes for the N-terminal Fpg/Nei and H2tH domains essential for DNA glycosylase activity, but lack at least one of four C-terminal motifs unique to NEIL3. Using pETDuet-2, active truncated and full length hNEIL3 have been overexpressed in *Escherichia coli* and purified. Enzyme assays were performed using oligonucleotide substrates containing one of three oxidised bases, thymine glycol (TG), 5-hydroxyuracil (5-OHU) or 8-oxoguanine (8-oxoG). Results indicate that the recombinant hNEIL3 proteins are active on single-stranded DNA (ssDNA) substrates containing 5-OHU and TG but show only weak activity on 8-oxoG. In contrast, in double-stranded DNA (dsDNA), activity on 5-OHU:G and TG:A was only weakly observed with nominal activity on 8-oxoG:C. The oxidised bases were then placed at three different sites in a model replication fork, at position -4 in ssDNA, at the fork junction (the last nucleotide in the dsDNA) and position +4 in dsDNA. For TG, hNEIL3 incised the fork substrate through β-elimination at all positions and exhibited greater activity at the +4 position than in the equivalent dsDNA substrate. For 5-OHU, a similar level of β-elimination activity was observed at the -4 and fork junction positions, but activity at the +4 dsDNA position was predominantly by β,δ-elimination when the full length hNEIL3 protein was used. Again, only weak activity was observed on 8-oxoG fork substrates and a similar incision pattern was observed to that for 5-OHU. Therefore, the results indicate that hNEIL3 acts as a bifunctional DNA glycosylase, either through β- or β,δ-elimination depending on the substrate lesion. For the first time it is shown that the C-terminal domains can influence the bifunctional DNA glycosylase activity of hNEIL3. These studies will help to define the biochemical function of this unique protein in both normal and malignant human cells.

# 1 INTRODUCTION

## 1.1 DNA damage repair

DNA damage can be caused by either endogenous sources generated during cell metabolism including hydrolysis, alkylation, oxidation or exogenous sources such as ultra violet radiation UV, ionizing radiation, and chemical agents. Failure to repair DNA damage could lead to loss of genomic integrity leading to apoptosis or cancer, however organisms have developed mechanisms of DNA repair to counter DNA damage (Hakem, 2008).

DNA nucleotides are continually modified by hydrolysis and oxidation reactions. These types of reaction can result in many types of DNA damage. For example, spontaneous hydrolytic reactions can lead to depurination and deamination, the former resulting in an apurinic site in the DNA molecule and the latter resulting in the conversion of cytosine to uracil or adenine to hypoxanthine. Oxidation can lead to chemically altered bases and single- and double-strand strand breaks in the DNA molecule.

UV radiation can cause covalent cross-linking of adjacent pyrimidines. This can lead to the formation of a thymine dimer that interferes with DNA replication. Such alterations in nucleotide structure affect mainly one DNA strand at a given site. Ionizing radiation, however, can lead to double-strand breaks, which is more serious (Liu, *et al.,* 2006).

Oxidative DNA damage caused by reactive oxygen species (ROS) which are present in all aerobic cells causing oxidative DNA damage e.g. hydrogen peroxide, superoxide and hydroxyl radical ($\cdot$OH) which produce a variety of oxidation products by attacking DNA at several positions. Oxidative damage causes single and double strand breaks. Examples of these products are 5-formyluracil, 2-oxoadenine and 8-oxoguanine and their level is dramatically induced by ionizing radiation of water (radiolysis) leading to hydroxyl radical formation (Naegeli, 1997).

DNA damage encountered by the cellular mechanisms of DNA repair to maintain genomic integrity. In response to the variety of DNA lesions, several DNA repair pathways have evolved in all organisms. The important DNA repair pathways for mammalian cells are direct repair, base excision repair (BER), mismatch repair (MMR), nucleotide excision repair (NER) and double-strand break repair (DSB), which includes non-homologous end-joining and homologous recombination. Also, the same mechanisms are used by cancer cells to survive treatment with genotoxic insults, *e.g.* ionizing radiation and chemotherapy (Naegeli, 1997).

## 1.2   Oxidative DNA damage

Oxidation of nucleic acids can be a cofactor in the carcinogenic process (Poulsen et al., 2005). The most comprehensively examined product of ROS attack on the DNA molecule is the oxidation of guanine to 8-oxo-7,8-dihydroguanine (8-oxoG), particularly the mutagenic potential of this oxidised base lesion (Cadet et al., 2002). For example, methylene blue and UVA treatment of cells produces singlet oxygen molecules that generate 8-oxoG in DNA that is mutagenic because of its tendency to mispair with adenine during DNA replication. Thus, pre-mutagenic 8-oxoG is the most dominant product of guanine oxidation (Figure 1.1) causing G-to-T transversion mutations (Henderson et al., 2002; Hsu et al., 2004). 8-oxoG has been linked with cancer risk and its cellular effects have been studied extensively, indicating a role in carcinogenesis (Hegde *et al.,* 2012). 8-oxoG can be further oxidised producing spiroiminodihydantoin (Sp) and guanidinohydantoin (Gh) in DNA (Niles *et al.,* 2001; Jena and Mishra, 2012).

The accumulation of 8-oxoG in DNA can be a major contributing factor in carcinogenesis and can be used as a cancer risk biomarker as the occurrence of this oxidative pre-mutagenic lesion in high concentrations in DNA leads to greater percentages of mutation accumulation over time, increasing the risk of cancer. These studies also linked life style with concentration levels of DNA oxidation products including 8-oxoG (McAuley-Hecht *et al.,* 1994).
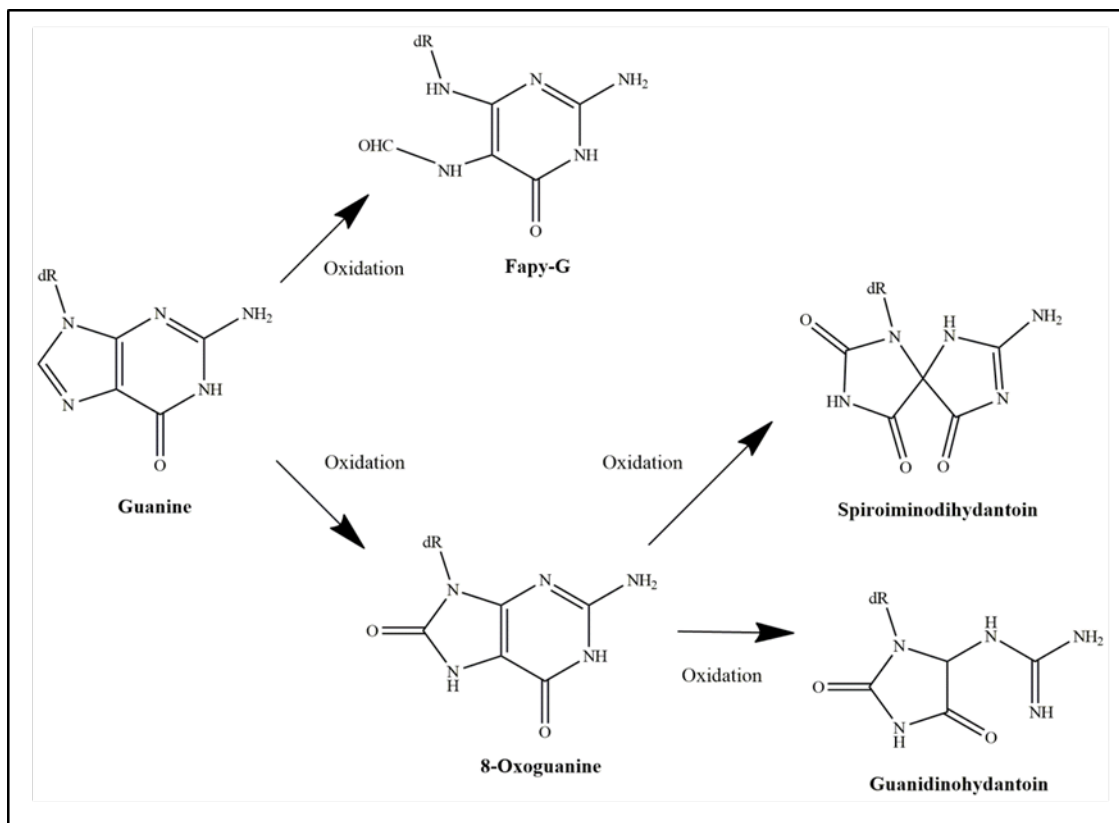
**Figure 1.1 Guanine base oxidation leading to formation of 8-oxoG and FapyG**

Guanine oxidation producing FapyG and 8-oxoG which is further oxidized to form spiroiminodihydantoin (Sp) and guanidinohydantoin (Gh). (Adapted from Thiviyanathan *et al.,* 2008)

Formamidopyrimidines (Fapy) are another mutagenic form of DNA oxidative damage caused by attack of ROS leading to the formation of FapyG (Figure 1.1) and FapyA (Jena and Mishra, 2013).

Cytosine can be attacked by hydroxyl radicals (•OH) and hydroperoxides or peroxyl radicals resulting in cytosine glycol (Figure 1.2**.**). Cytosine glycol can then either be deaminated to form uracil glycol, which is readily dehydrated to produce 5-hydroxyuracil (5-OHU), or dehydrated to form 5-hydroxycytosine (5-OHC) that can then be deaminated to form 5-OHU (Thiviyanathan *et al.,* 2008).
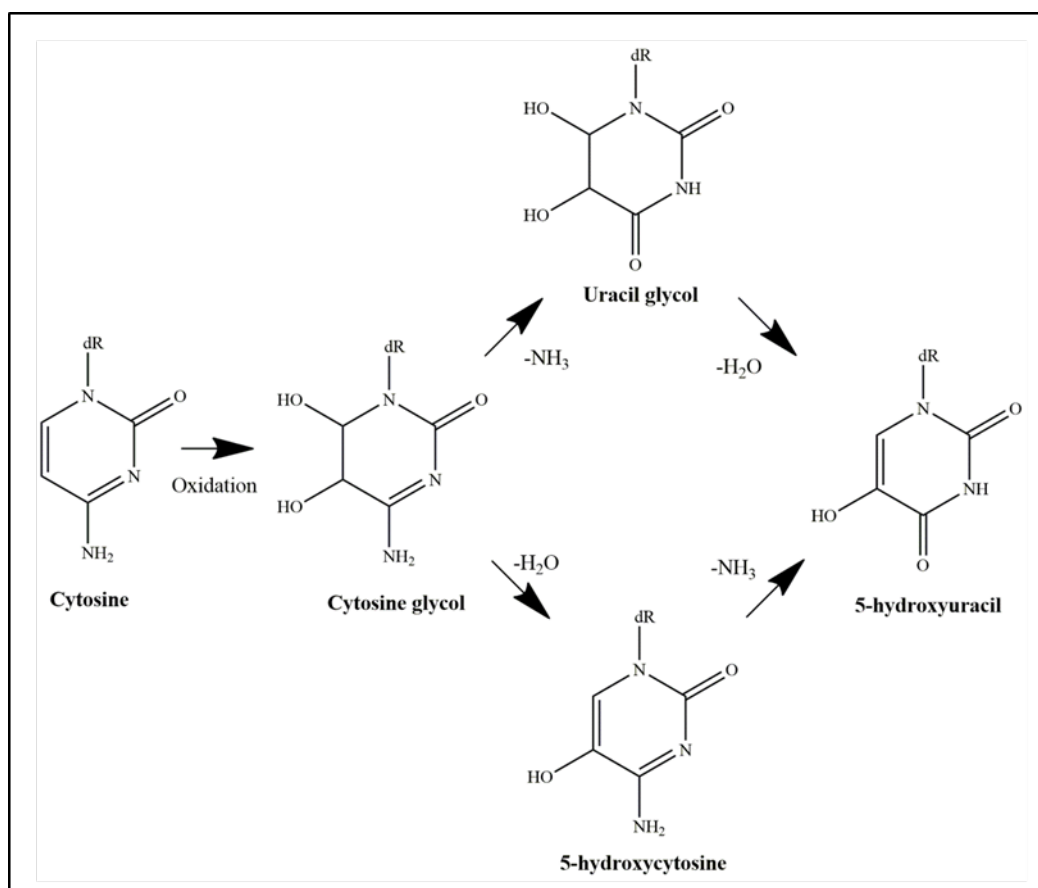
**Figure 1.2 Cytosine oxidation producing cytosine glycol followed by deamination-dehydration or dehydration-deamination forming uracil glycol, 5-OHC AND 5-OHU.**
(Adapted from Thiviyanathan *et al.,* 2008)

Thymine can also be oxidised to form a 5,6 radical adduct then a 5,6 OH-OOH intermediate by $O_2$ addition, then the latter dehydrated to produce thymine glycol (TG). Oxidation and ionising radiation of thymine cause TG formation in DNA which has a stalling effect on DNA replication (Yoon *et al.,* 2010). Thus unrepaired, accumulated oxidative damage may block replicative DNA polymerases causing replication fork stalling leading to double-strand breaks and cell apoptosis (Aller *et al.,* 2007).
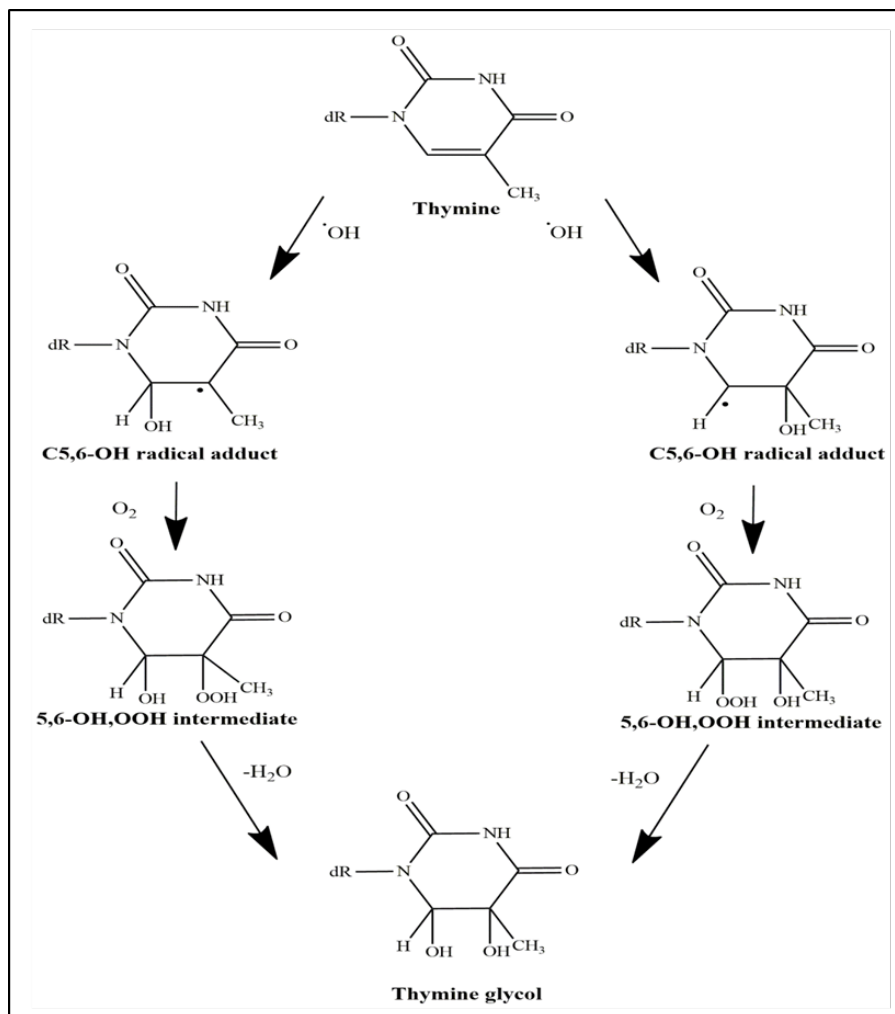
**Figure 1.3 Thymine oxidation leading to formation of C5,6-OH radical adduct transforming to unstable intermediate by O₂ addition then dehydrated to produce thymine glycol.**
(Adapted from Cadet and Wagner, 2013).

## 1.3 Base excision repair

Base excision repair (BER) repairs most of the endogenous lesions such as oxidized bases and AP-sites as well as DNA single-strand breaks (Hegde *et al.,* 2008). Studies have shown that the process is present in all organisms, form bacteria to eukaryotes including mammals. BER is an exceptional mechanism among the excision repair processes as individual base lesions are recognised and removed by specific DNA glycosylases.

In addition to DNA base oxidation, ROS attacks the deoxyribose in DNA generating single- and double-strand breaks (sugar fragments of 3′ phosphates) and BER is also used to repair the resulting single-strand breaks. The 3′ blocking groups, including 3′-phosphate, 3′-phosphoglycolaldethyde, and 3′-phosphoglycolate are also removed by enzymes in BER. The 5′ terminus could consist either of phosphate or 5′-deoxyribosephosphate depending on the DNA glycosylase that has acted on the damaged base (Demple *et al.,* 2002).

DNA glycosylase initiate BER by recognizing and removing the damaged base by hydrolysing the N-glycosylic bond leaving either an abasic site (monofunctional DNA glycosylase), or a single-strand break (bifunctional DNA glycosylase). For the former, the apurinic/apyrimidinic (AP) site is then cleaved by AP-endonuclease-1 (APE1) giving a 5′-deoxyribophosphate (dRp) end that is removed by the dRPase activity DNA polymerase β (Polβ). The resultant single nucleotide gap is filled with the appropriate complementary nucleotide by Polβ and the nick is sealed by the action of DNA ligase III/XRCC1 to complete the process (Prasad *et al.,* 2002).

If the damaged base is recognized by a bifunctional DNA glycosylase (DNA glycosylase/AP lyase), such as endonuclease III (NTH1) and 8-oxoguanine DNA glycosylase (OGG1), the AP-site is cleaved by β-elimination creating a 3′ terminal sugar phosphate (3′-dRp) which is removed by APE1 leaving a gap with a 3′ OH end (Figure 1.4; Fortini and Dogliotti, 2007).

The DNA glycosylase, endonuclease VIII (Nei)-like proteins, NEIL1 and NEIL2 are also bifunctional, but they act by a different mechanism independent of APE1 as they remove the damaged base and cleave the AP-site by β/δ elimination generating a 3′ phosphate end which is then removed by polynucleotide kinase (PNK) (Wiederhold *et al.,* 2004).

## 1.3.1 Short patch BER

BER is divided into two pathways, short- and long-patch BER. While the first steps are the same, the gap-filling steps diverge. In short-patch BER, only one nucleotide is replaced and Polβ carries out both the steps required to produce a 5′ phosphate group and DNA synthesis (Figure 1.4).
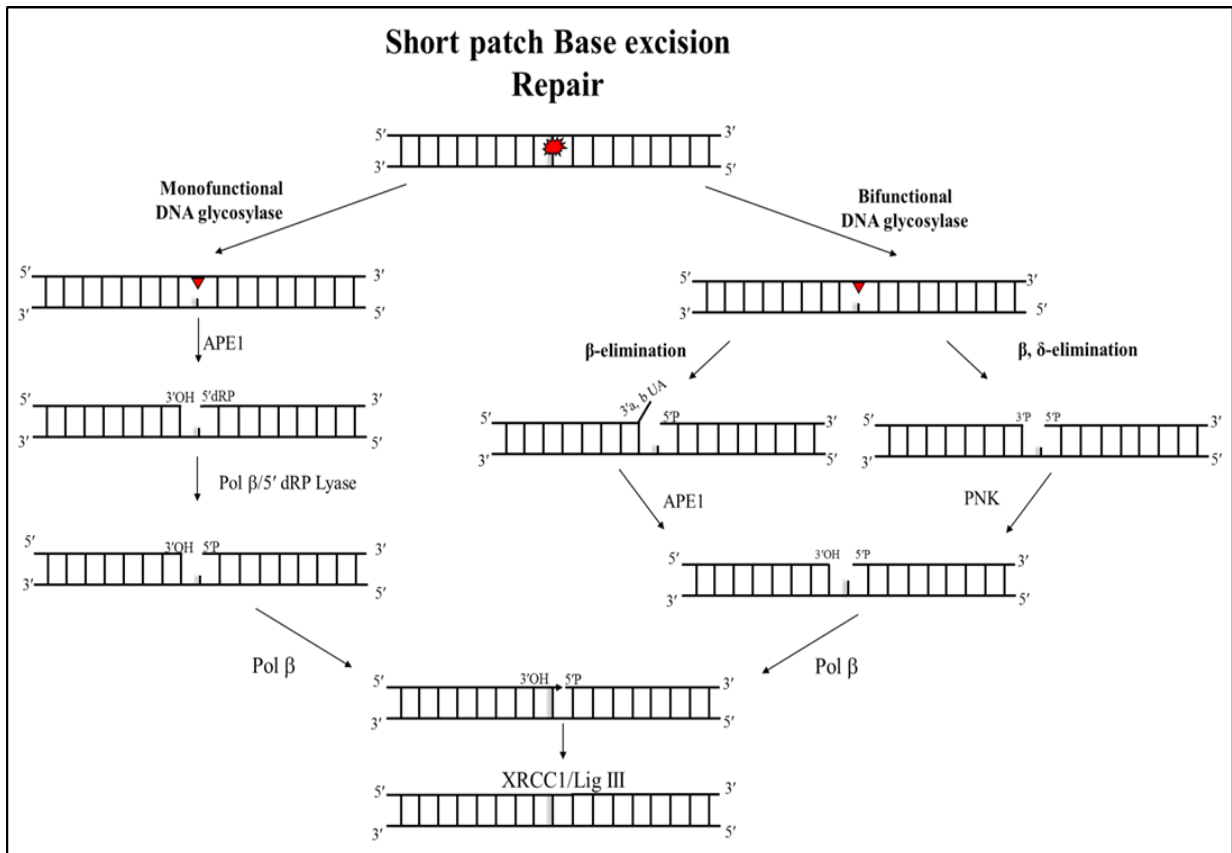
**Figure 1.4  Short patch BER**

Damaged bases in DNA are recognized by a DNA glycosylase (monofunctional or bifunctional). DNA glycosylases hydrolyse the N-glycosylic bond between the modified base and sugar backbone generating an AP-site that is then cleaved by APE1 to generate a 5′-dRp that is removed by Polβ. This is followed by gap filling by Polβ and finally DNA ligase III seals the nick and completes the process (Adapted from Wallace 2012).

## 1.3.2 Long patch BER

BER has been shown to undergo long patch repair *in vivo*, however, the choice of short- or long patch is still unclear (Figure 1.5; (Krokan, and Bjoras. 2013)). The first nucleotide is processed by Polβ as in short patch BER, then a replicative DNA polymerase (*e.g.* DNA polymerase δ or ε) replaces the following 6 to 13 nucleotides. Replication factor (RF)-C is also involved in long patch BER to recruit the sliding clamp, proliferating cell nuclear antigen (PCNA) to the damage site of DNA. This provides a platform for DNA polymerase and flap endonuclease-1 (FEN-1), which removes the displaced oligonucleotide fragment. Finally, the gap is sealed by DNA ligase I (Figure 1.5; Cappelli *et al.*, 1997; Sleeth *et al.*, 2004).
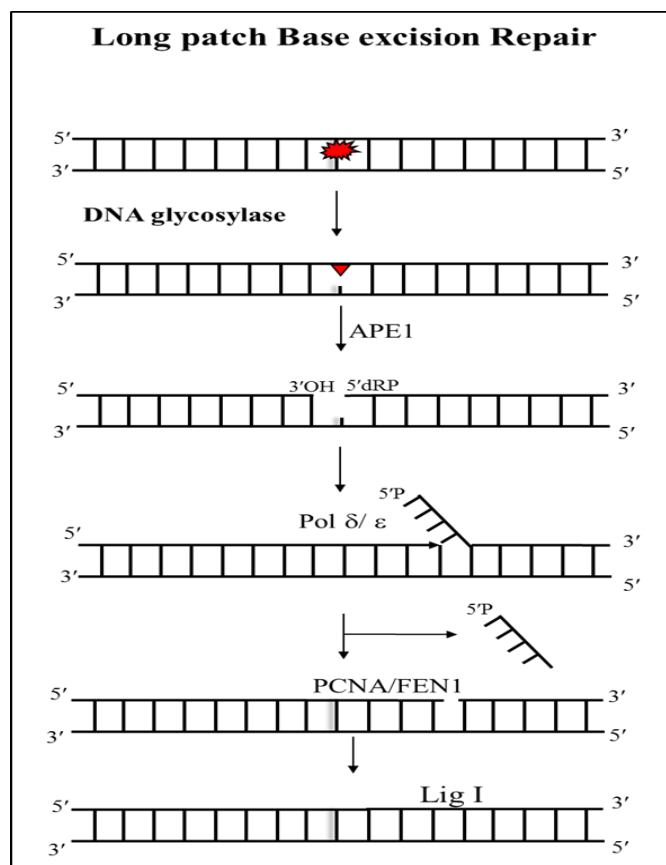


**Figure 1.5 Long patch BER**

A DNA glycosylase cleaves the glycosylic bond and the AP-site is incised by APE1 on the 5′-side. Thereafter, DNA polymerases δ or ε replace the strand with (RF)-C and PCNA in

displacement synthesis. The resulting oligonucleotide flap is removed by FEN-1 and the nick sealed by DNA ligase I.

## 1.4    DNA glycosylases for oxidized bases

The first DNA glycosylase discovered was uracil DNA glycosylase (UDG) in *E. coli* (Lindahl, 1974). As cytosine deamination generates uracil and a G-U base pair is potentially mutagenic, the author looked for activities that could be used to identify uracil in DNA. Subsequently, other bacteria, yeast, plants, mammalian cells and mitochondria-specific UDGs were also characterized with similar enzyme activities (Hegde *et al.*, 2008). In contrast to the relaxed specificity of the most DNA glycosylases for various substrates, UDG is specific for uracil in DNA as uracil fits tightly into the catalytic pocket of UDG. Mutational studies of UDG and crystal structures have revealed the structural basis of substrate specificity in preference to similar structures, such as C or T in the DNA. Because the uracil-binding pocket is narrow and deep and located in the positively charged groove of the enzyme, the DNA helix flips out to allow binding of DNA-uracil to the enzyme (Fromme and Verdine, 2003).

Generally, a DNA glycosylase recognises only abnormal bases as substrates in DNA, including uracil. However, this is not always the case and at least three mammalian DNA glycosylases remove normal bases from DNA (TDG [thymine DNA glycosylase], MBD4 [methyl-CpG-binding domain protein 4] and MUTYH) (Fromme and Verdine, 2002). The MutY gene was first described in *E. coli*; MutY deficiency induces spontaneous A-T, C-G transversion mutations (Fromme and Verdine, 2003). It mainly excises normal bases such as A & G from A-G and A (G): 8-oxoG base pair mismatches. Therefore, the gene MutY and its mammalian homolog MUTYH are basically essential, because they protect against oxidative base damage

- induced mutation by removing the base (A) in the nascent strand opposite 8-oxoG during DNA replication (Hegde *et al.,* 2008).

While uracil DNA glycosylase belongs to the UDG superfamily, a group of monofunctional DNA glycosylases, DNA glycosylases acting on oxidised base are divided into two families, either with a helix hairpin helix (HhH) domain or a helix 2-turn helix domain, and they are all bifunctional DNA glycosylases (Friedman and Stivers., 2010). Therefore, while the Nth (endonuclease III) superfamily has an HhH DNA binding domain, the Fpg/Nei (formamidopyrimidine-DNA glycosylase/endonuclease VIII) family members have a conserved H2TH motif. However, while their protein architecture may be different, some members of these families have common substrates (Hitomi *et al., 2007*).

The Nth superfamily consists of *E. coli* endonuclease III (Nth), *E. coli* MutY, bacterial and yeast AlkA, yeast Ntg1 and Ntg2, mammalian NTH1, eukaryotic MUTYH and 8-oxoguanine glycosylase (OGG1). MutY, Nth and their eukaryotic homologs have an HhH structure tailed by a loop containing mainly glycine, proline and valine residues, and a conserved catalytic aspartate residue (McCullough *et al.,* 1999; Prakash *et al.,* 2012). Conversely, MutM (Fpg) and Nei have an H2TH motif and a zinc finger or a zincless finger motif for DNA binding (Hegde *et al.,* 2008). The Fpg/Nei family consists of the bacterial proteins Fpg and Nei and their eukaryotic orthologs, Nei-like proteins NEIL1, NEIL2 and NEIL3.

On the other hand, according to previous studies on *E. coli*, the oxidized base-specific DNA glycosylases were classified according to their tertiary structures into two group-families, AP lyase reaction activity and active site characteristics. The internal lyase residue is utilized by the Nth family as the active sits of the β elimination reaction, which in turn generates a 3′ phospho α,β-unsaturated aldehyde (3′ PUA) at the strand break. On the other hand, the other

members of *E. coli* Fpg family involving Nei (endonuclease VIII) catalyse β/δ elimination at the AP site by utilizing the N-terminal proline or valine as the nucleophile and remove the residual deoxyribose producing a 3′ phosphate terminus at the DNA strand break (Drohat and Maiti, 2014)

## 1.5 Monofunctional DNA glycosylases in mammalian cells

Uracil DNA glycosylases (UDG) superfamily proteins are all monofunctional DNA glycosylases that mainly recognize and remove uracil from DNA. Several UDGs has been expressed and purified: single-strand selective monofunctional uracil DNA glycosylase 1 (SMUG1), nuclear uracil DNA glycosylase (UNG2) and mitochondrial uracil DNA glycosylase (UNG1). Additionally, the mismatch DNA glycosylases MBD4 and TDG are also members of this family (Wyatt, 2013). The catalytic domain of UDG is characterized by an α/β fold structure and UDGs show a varied substrate preference. Thus, while SMUG1 shows a preference for 5-formyluracil and 5-hydroxymethyluracil, SMUG1 and UNG2 have been found to be active on 5-OHU and 5,6-dihydroxyuracil, while SMUG1, UNG1 and UNG2 show activity on 5-flurouracil (Krokan *et al.,* 2013).

Thymine DNA glycosylase is a mismatch DNA glycosylase. TDG is active on uracil and thymine mispaired with guanine and with preference to thymine. TDG is also active on several DNA lesions such as 5-carboxylcytosine, 5-formylcytosine with specificity for 5-halogenated pyrimidines *e.g*. 5-fluorouracil (Krokan *et al.,* 2013; Morgan *et al.,* 2007). MBD4 also acts on halogenated uracil e.g. 5-fluorouracil and 5-bromouracil. MBD4 which is an HhH family member, recognizes and removes thymine and uracil mispaired with guanine at CpG and deaminated CpG sites (Sjolund *et al.,* 2012).

MUTYH is a monofunctional DNA glycosylase, also a member of the HhH superfamily, acts on adenine mispaired with 8-oxoG. MUTYH is a homolog of the bacterial MutY protein and

they both remove adenine mispaired with FapyG (Greenberg, 2012). MUTYH has also been shown to have activity on DNA containing 2-hydroxyadenine paired with guanine or adenine. Also, mutated MUTYH is associated with colon cancer syndromes such as familial adenomatous polyposis (Markkanen *et al.,* 2013; Palles *et al.,* 2013; Wallace, 2014).

The bacterial AlkA and Tag and mammalian alkyladenine DNA glycosylase (AAG) are functional homologs but have no structural similarity (Bjelland and Seeberg, 1996). AAG is a monofunctional DNA glycosylase that shows activity on a range of alkylated bases, including 3-methyladenine, 3-methylguanine and 7-methylguanine, as well as deaminated adenine (hypoxanthine) and the lipid peroxidation product 1,$N^6$-ethenoadenine (Lee *et al.,* 2009).

## 1.6    Bifunctional DNA glycosylases

The mammalian cell DNA glycosylases, OGG1, NTH1, NEIL1, NEIL2 and NEIL3 are responsible for the excision of a range of oxidised bases from DNA. In addition to this function, these bifunctional DNA glycosylases have an AP-lyase activity that results in strand incision, 3′ to the AP-site that results from base excision.

In the case of OGG1, the AP-lyase activity is often weak, therefore, the main product of OGG1-catalysed cleavage is an undamaged AP site (Hegde *et al*., 2008). There are many differences found in the specific amino acid residues that recognise the damaged bases in DNA. Based on these differences, one is considered a unifying mechanism of DNA glycosylases activity and that is the extra helical flipping of the damaged deoxynucleotide into a lesion-specific recognition pocket. According to previous studies, all DNA glycosylases bind to the minor groove at the damage site causing a twist in the DNA molecule and the base lesion flips out of the DNA major groove (Slupphaug *et al.,* 1996).

As is evident from the discussion so far, DNA glycosylases can recognise a specific range of quite diverse substrate bases. Thus, the form of base damage is specific to each DNA glycosylase and only bases that can be accommodated in the binding pocket after base flipping will be recognised and excised. Therefore, the plasticity of the catalytic pocket plays a critical role, as it allows the fit of various substrate bases and allows a range of damaged bases to be excised by one DNA glycosylase (Hegde *et al.,* 2008).

### 1.6.1 *E. coli* Fpg protein

Formamidopyrimidine DNA glycosylase (Fpg/MutM), consists of 263 amino acids with molecular weight of 30.2 kDa. Fpg is a bifunctional DNA glycosylase with glycosylase and AP-lyase activity acting through β,δ-elimination. It possesses four DNA binding cysteine-zinc finger motifs and its active site has been located in its N-terminal domain. The first substrate that Fpg was reported to be active on was 2,6-diamino-4-hydroxy-N7-methyl-5-formamidopyrimidine (Me-FapyGua; Chetsanga and Lindahl., 1979). The activity of Fpg has been extensively investigated using various methods and it has been shown to be active on 8-oxoG, 5-hydroxycytosine (5-OHU), 5-OHU, FapyG and FapyA as well as the further oxidative product of 8-oxoG, Sp, in dsDNA. Studies have shown that 8-oxoG, FapyG and FapyA are the major physiological substrates of *E. coli* Fpg (Guo *et al.,* 2010).

### 1.6.2 *E. coli* Nei protein (Endonuclease VIII)
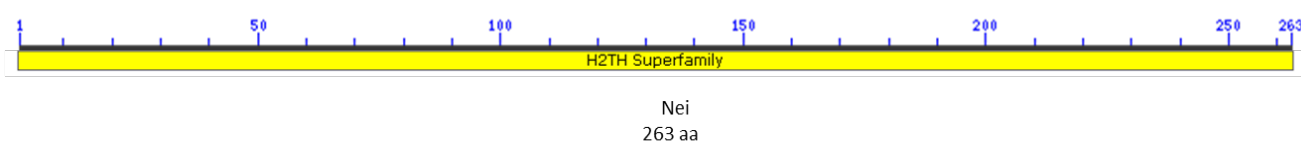


**Figure 1.6 Endonuclease VIII (Nei) protein structure**
Nei protein possesses one domain (H2TH) and consists of 263 amino acids.

Endonuclease VIII (Nei) DNA glycosylase has been discovered in *E. coli* with molecular size of 29.7 consisting of 263 amino acids (Figure 1.6), The amino acid sequence and protein structure of Nei show homology to bacterial Fpg glycosylase (Melamede *et al.,* 1994). Nei has been shown to have DNA glycosylase and AP-lyase activity on double-stranded DNA containing modified pyrimidines acting through β,δ-elimination. The substrate preference of

Nei has been investigated and the protein shows excision activity on 8-oxoG, TG, urea, β-ureidoisobutyric acid, uracil glycol, 5-OHU, 5-OHC in DNA containing single modified bases (Boiteux *et al.,* 1992).

DNA with multiple lesions was also investigated and Nei has shown activity on pyrimidine derived DNA lesions such as 5-hydroxy-5-methyl-hydantoin (5-OH-5-MeHyd), 5-OH-6-HUracil and **5**-hydroxy-**6**-hydrocytosine as well as 8-oxoG:G and 8oxo-G:A in double-stranded DNA with lower activity on 8-oxoG:C. Nei also showed efficient activity on multiple FapyA DNA lesions, but no activity on similar DNA containing FapyG (Hazra *et al.,* 2000; Wiederholt *et al.,* 2005).

## 1.7   Mammalian bifunctional DNA glycosylases

In mammalian cells there are five oxidized base specific DNA glycosylases. Previous studies determined that NTH1 and OGG1 are both members of the Nth family. NTH1 prefers oxidized pyrimidines as substrates, while OGG1 is primarily responsible for the removal of 8-oxoG and ring opened guanine, *i.e.* Fapy-G (Dalhus *et al.,* 2009). More recently, two other human DNA glycosylases that are part of the Fpg/Nei family were discovered by searching human protein databases and they have been termed NEIL1 and NEIL2 respectively (Hazra *et al.,* 2002). It has been mentioned that these enzymes are bifunctional glycosylases with broad substrate range. These were followed in the same year by the identification of a third paralog in human cells, subsequently termed NEIL3 (Morland *et al.,* 2002; Liu *et al.,* 2012). Amongst these three NEIL homologs, the N-terminal proline is present in NEIL1 and NEIL2, while NEIL3 possess a valine residue that acts as the active site nucleophile (Serre *et al.,* 2002). Previous studies showed that both NEIL1 and NEIL2 enzymes prefer modified pyrimidine substrates and are closer to Nei than Fpg.

### 1.7.1 OGG1 protein

OGG1 DNA glycosylase excises 8-oxoG from double-stranded DNA by cleaving the N-glycosylic bond and incises the phosphodiester backbone on the 3′ side of the AP-site with its AP-lyase activity through β-elimination, resulting in a single-stranded break at the site of the oxidized base. Two human OGG1 proteins has been discovered (α-hOGG1, molecular weight of 38.8 kDa 345 amino acids and β-hOGG1, molecular weight of 47.2 kDa 424 amino acids). OGG1 is the major DNA glycosylase involved in removing 8-oxoG from eukaryotic cells (Wallace *et al.,* 2012).

Guanine is the base most susceptible to oxidative damage and 8-oxoG is quatitatively one of the principal oxidised bases found in DNA. 8-oxoG is a pre-mutagenic lesion as it can base pair with adenine as well as the canonical cytosine. However, OGG1 knockout mice showed no, or only a slight increase in cancer frequency despite the increase of damage accumulation and mutation (Arai *et al.,* 2006). However, a SNP variant of OGG1 (OGG1-S326C) has been detected in 23% to 41% of the Japanese and Caucasian population (Hung *et al.,* 2005) and this SNP has been associated with various diseases and syndromes including types of cancer such as kidney, head and neck, bladder, colorectal and lung cancers (Zhou *et al.,* 2015).

### 1.7.2 NTH1 protein

Homologous to bacterial Nth, the mammalian NTH1 DNA glycosylase is an HhH superfamily member, consists of 312 amino acids and has a molecular weight of 34.4 kDa. NTH1 is a bifunctional DNA glycosylase with glycosylase-AP-lyse activity acting through β-elimination. NTH1 has overlapping substrate specificity with NEIL1 and recognizes and removes 5-OHU, 5-OHC, urea and TG in DNA containing single and multiple DNA lesions. It has also been shown that FapyA is a major preferred substrate of NTH1 where NTH1 null mice showed accumulations of FapyA DNA damage (Hu *et al.,* 2005; Chan *et al.,* 2009). A SNP variant of

NTH1 (hNTH1-Asp239Tyr) has been found in 6.2% of the global population. The mutant variant has been purified and shows no DNA glycosylase activity which suggests that humans carrying this variant may be more susceptible to carcinogenesis. Also, it has been shown that overexpression of NTH1 is detected in around half of the cases of colorectal cancers in disease-free survival, suggesting that NTH1 plays a major role in maintaining genome integrity in these patients (Koketsu *et al.,* 2004).

## 1.7.3 NEIL1
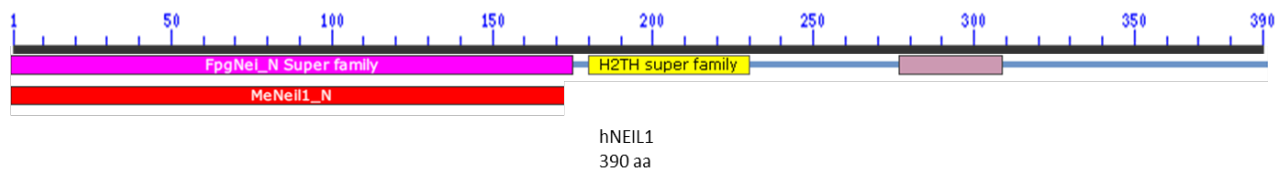


**Figure 1.7 Human NEIL1 protein possesses FpgNei superfamily, H2TH and Nei-DNA binding domains and consists of 390 amino acids.**

NEIL1 is a human homolog of the Nei bacterial protein (Nei-like), that has been identified in human cells (Hazra, *et al.,* 2002; Morland *et al.,* 2002). NEIL1 is one of the Fpg/Nei family of DNA glycosylases with more similarity of amino acid sequence to Nei than Fpg, The molecular size of NEIL1 is 43.7 kDa and 390 amino acids (Figure 1.7). The expression of NEIL1 is cell-cycle dependent and it is principally expressed during S phase and therefore it is suggested to have a role in the repair of DNA damage during DNA replication. NEIL1 DNA glycosylase acts through a β/δ-elimination mode of action utilizing proline 2 at the N-terminus (Bandaru, *et al.,* 2002; Hazra *et al.,* 2002). As evidence of its importance in safeguarding genetic integrity, NEIL1 is found in the nucleus and mitochondria (Hu *et al.,* 2005). The highest expression levels

of NEIL1 is detected in the thymus, pancreas and liver with lower expression in brain tissue (Morland *et al.,* 2002; Izumi *et al.,* 2003). The structure of NEIL1 has been studied by X-ray crystallography where it showed a lack of the conserved zinc-finger motif that is replaced by a zinc-less finger motif (Doublie *et al.,* 2003). NEIL1 has proved to be active on Me-FapyG and oxidised pyrimidines in an oligonucleotide substrate with a single oxidative lesion. However, while no activity was observed on 8-oxoG:C the further oxidation products of 8-oxoG, Sp and Gh were substrates. NEIL1 was tested on multiple DNA lesions and proved to be active on FapyG and FapyA while it showed less activity on TG and 5-OH-5-MeHyd (Jaruga *et al.,* 2004; Hazra *et al.,* 2002).

NEIL1 knockout mice and NEIL1 knockdown cells show accumulation of FapyG and FapyA but no 8-oxoG confirming the above *in vitro* studies and demonstrating NEIL1 preference to FapyG and FapyA. A study showed that FapyG is an overlapping substrate for NEIL1 and OGG1 when NEIL1 and OGG1 knockout mice show accumulation of FapyG, but 8-oxoG is accumulated only in OGG1 knockout mice confirming that NEIL1 does not recognise 8-oxoG containing DNA lesion *in vivo* (Liu et al., 2013). Also, it has been shown that NEIL1 knockout mice developed hepatocellular and pulmonary cancer in late stage of their life (Chan *et al.,* 2009). NEIL1 shows activity on lesions in quadruplex DNA structures and telomeric DNA context with preference for specific lesions such as Gh and no activity on TG or 8-oxoG confirming previous results (Zhou *et al.,* 2013 and 2015). Interestingly, NEIL1 also proved to be active on DNA crosslinks induced by psoralen (Couve-Privat *et al.,* 2007; Couve *et al.,* 2009; Martin *et al.,* 2017).

Numerous polymorphic alternates has been detected and purified such as NEIL1-Asp252Asn, NEIL1-Cys136Arg, NEIL1-Ser82Cys, NEIL1-Pro208Ser, NEIL1-DGlu28 and NEIL1-Gly83Asp, according to the National Institute of Environmental and Health Sciences Environ-

mental Genome Project (Roy *et al.,* 2007). NEIL1 SNP variants have been identified in a group of people representing the US population, where it has been found that SNPs occurred in 1% of that group. Some of these variants have been found in cancer patients such as NEIL1-Pro208Ser, which was found in patients of colorectal carcinoma and NEIL1-Gly83Asp, found in patients with cholangiocarcinoma, suggesting that people carrying those NEIL1 variants could be at higher risk of developing cancer (Cadet *et al*., 2010; Sampath *et al.,* 2012). Furthermore, several types of NEIL1 SNPs have been identified as cancer associated variants (Dallosso *et al.,* 2008; Forsbring *et al.,* 2009). Although NEIL1 is one of the DNA glycosylases, it is believed that NEIL1 is involved in other DNA repair mechanisms such as nucleotide excision repair (NER). The evidence for this has been presented by a research group judging the accumulation of 8,5′-cyclopurine-2′-deoxynucleoside lesions in NEIL1 knockout mice where these lesions are repaired by NER not BER and suggested that NEIL1 may play a role in NER (Jaruga *et al.,* 2010). The role of NEIL1 in NER is suggested as a result of NEIL1 protein interaction with the NER protein complex (Hegde *et al.,* 2015).

## 1.7.4 NEIL2
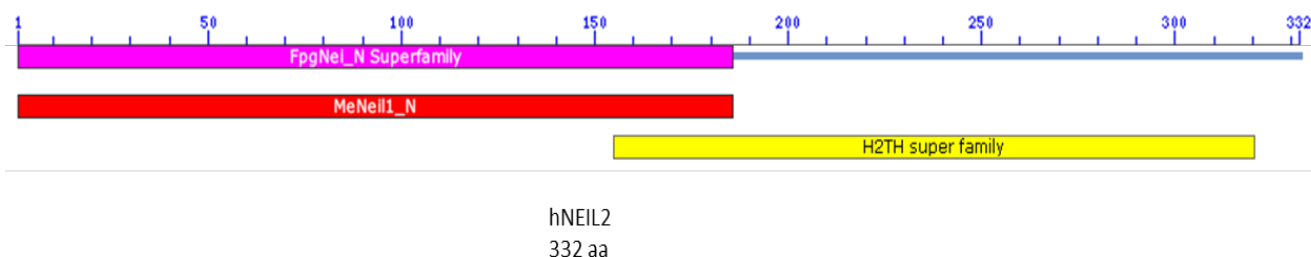


**Figure 1.8 Human NEIL2 protein possesses FpgNei superfamily and H2TH super family domains and consists of 332 amino acids.**

NEIL2 is the smallest member of the Nei-like DNA glycosylase family and the human protein consists of 332 amino acids and is a 36.8 kDa protein (Figure 1.8). NEIL2 has DNA glycosylase/AP-lyase activity acting through β,δ-elimination (Hazra, *et al*., 2002; Dou *et al*., 2003). Although NEIL2 does not share sequences homology with NEIL1, it possesses the same N-terminal proline 2 at the catalytic active site as NEIL1. It possesses the conserved H2TH and the zinc finger domains of *E. coli* Fpg/Nei and NEIL3. NEIL2 has been detected in both the nucleus and mitochondria. However, NEIL2 does not show a cell-cycle dependent expression pattern and therefore it has been suggested that NEIL2 acts in transcription-coupled DNA repair, further evidenced by the substrate preference of NEIL2 for single-stranded and bubble DNA structures which are created during transcription (Banerjee *et al*., 2011).

NEIL2 shows efficient excision activity on 5OHU:G in double-stranded oligonucleotide substrates (Hailer et al., 2005). It showed lower activity when incubated with 5,6-diHU and 5-OH-C and a very weak activity when incubated with TG or 8-oxoG containing substrates. However, NEIL2 also showed activity on the further oxidative products of 8-oxoG (Sp and Gh) in double-stranded DNA oligonucleotide substrates (Hailer *et al*., 2005). Due to the fact that NEIL2 is highly unstable and quickly loses its activity even after being stored at low

temperature and when incubated at 37°C, no data has been showed for the activity of NEIL2 on oligonucleotide substrate with multiple oxidative lesions (Hazra *et al.,* 2002).

NEIL2 knockout mice show accumulation of oxidative DNA lesions with significant sensitivity to inflammation. The increase of oxidative DNA damage was remarked as age dependent DNA damage, whilst genetic instability and telomere loss has also been detected in embryonic fibroblasts of these knockout NEIL2 mice (Chakraborty *et al.,* 2015).

Multiple polymorphic variants of NEIL2 have been detected in different types of cancers: NEIL2-Arg257Leu and NEIL2-Arg103Gln were detected in lung cancer and have been overexpressed and purified (Broderick *et al.,* 2006). NEIL2-Arg257Leu is specifically expressed in lung cancer cells and showed less activity and a significant increase in DNA lesion accumulation in comparison with NEIL2 wild type, suggesting that humans carrying this type of NEIL2 allele are more susceptible to developing lung cancer (Broderick *et al.,* 2006; Dey *et al.,* 2012).

## 1.8 NEIL3
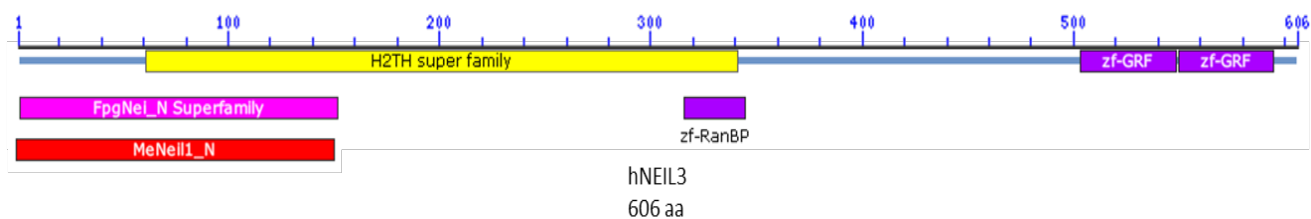


**Figure 1.9 Human NEIL3 protein possesses FpgNei superfamily, H2TH, zinc-finger (zf-RanBP) and two zf-GRF zinc finger domains and consists of 606 amino acids.**

NEIL3 is the largest member of the NEIL family in mammalian cells consisting of 606 amino acids in human cells (Morland *et al.,* 2002). The protein differs from the other NEIL paralogs in two main respects, firstly the canonical N-terminal proline is replaced by valine and NEIL3 has an extended C-terminal tail that contains multiple zinc finger domains of as yet unknown function (Liu *et al.,* 2002).

Studies found that insect cell lysates overexpressing NEIL3 can excise methyl FapyG (Morland *et al.*, 2002) and a weak AP lyase activity has been observed on single-stranded DNA substrates (Takao *et al.,* 2010). A study on full-length mouse NEIL3 and an N-terminal DNA glycosylase domain truncation indicated a strong substrate preference for oxidized purines and pyrimidines in single-stranded DNA (Liu *et al.,* 2010). The single-stranded DNA lesion preference of NEIL3 is due to its lack of two void-filling residues that stabilize the DNA duplex and interact with the opposite strand (Wallace *et al.,* 2013).

NEIL3 is expressed only in highly proliferating cells, such as those in testes, thymus and in cancer cell lines. Although studies have shown that NEIL3 has substrate specificity for oxidized pyrimidines and purines, no excision activity has been observed for 8-oxoG (Wilson & Bohr, 2007), although in common with the other NEIL paralogs it does have a major role in removing the hydantoin products of 8-oxoG oxidation, Sp and Gh (Liu *et al.,* 2010; Klattenhoff, *et al.,*

2017). However, NEIL3 also has been shown to possess excision activity on various DNA lesions in ssDNA such as, 5-OHU, 5-OHC, 5-OH-5MeHyd and TG (Qi *et al.,* 2009).

The NEIL3 protein structure features an N-terminal conserved domain and a long C-terminal tail (Figure 1.9) that distinguishes NEIL3 from NEIL1 and NEIL2. The NEIL3 protein contains the Fpg/Nei catalytic domain at the N-terminus, with the replacement of proline by valine at position 2. Additionally NEIL3 contains the H2TH DNA binding domain of the Fpg/Nei superfamily, a zinc finger motif (zf-RanBP) and tandem Zn-GRF domains at the C-terminal (Figure 1.9). The function of the C-terminal domains of NEIL3 is still unknown and is subject of continued research (Krokeide *et al.,* 2009; Krokeide *et al.,* 2013).

G-quadruplex DNA structures have been found in telomeric DNA and the promoters of several genes in eukaryotic chromosomes (Zhou *et al.,* 2013). The G-quadruplex DNA structure has been found to be susceptible to oxidative DNA damage producing, 8-oxoG, Sp, Gh and TG lesions (Zhou et al., 2013). Therefore, the activity of DNA glycosylases has been tested on G-quadruplex substrates (Zhou *et al.,* 2015). Thus, a recombinant NEIL3 protein from *Mus musculus* showed excision activity on TG in quadruplex DNA while NEIL1 from the same source exhibited a preference for Sp and Gh lesions, suggesting that NEIL3 and NEIL1 shared a DNA repair role for damaged bases in telomeres and G-quadruplex DNA (Zhou *et al*., 2013). Therefore, NEIl3 has a preference *in vitro* for single-stranded DNA substrates and quadruplex DNA structures in comparison to double-stranded DNA (Krokeide *et al*., 2009; Liu *et al*., 2010, 2012; Zhou *et al.*, 2013). This substrate preference has been suggested to be one possible function of the extended C-terminal domains of NEIL3 (Liu *et al.,* 2010).

The expression of NEIL3 in highly proliferating cells during the S-and G2-phase of the cell cycle makes NEIL3 an interesting target to investigate its role in the repair of oxidative damage in telomeres. Telomeres are DNA structures covering the end of mammalian chromosomes and are guanine rich DNA structures and thus more susceptible to oxidative damage by ROS. The

telomeric repeat proteins TRF1, TRF2 and POT1 act as a shield to protect (shelter) telomeres from the DNA damage response. (Palm and de Lange, 2008). Because telomeres are more sensitive to oxidative damage *in vitro* or *in vivo* due to G-rich components (Hewitt *et al*., 2012; Wang *et al.,* 2010), unrepaired DNA lesions in telomeres may affect the binding of the protecting telomeric proteins and lead to telomere defects that can trigger apoptosis or more significantly are linked to some types of cancer such as, colon, prostate and breast in which chromosomal instability is characterised (Chin *et al.,* 2004; Meeker *et al.,* 2002; Rudolph *et al.,* 2001).

An extensive study by Zhou *et al.* (2017) showed that NEIL3 has a crucial role in safeguarding telomere integrity against oxidative damage in highly proliferating cells, The study shows evidence that knockdown of NEIL3 in those cells leads to telomeric dysfunction, DNA bridges and ultimately cell apoptosis, suggesting a mechanism of action where NEIL3 is recruited to the telomeres utilizing its C-terminal domain which interacts with the TRFH domain of the telomeric protein TRF1 during the upregulation of NEIL3 in the cell cycle (late S/G2 phase). NEIL3 recognizes oxidative damage in telomeres and initiates BER by recruiting the long-patch BER proteins, APE1, FEN1 and PCNA. The interaction between NEIL3 and TRF1 can reduce double-strand breaks and thus prevent telomere fusion (Zhou *et al.,* 2017). Protein - protein interactions were also detected between NEIL3 and FEN1 and PCNA, while no interaction was observed between NEIL3 and Polβ, suggesting that long-patch BER is the preferred mode of BER in telomeric regions (Zhou *et al.,* 2017). When NEIL3 interaction with APE1 was tested, it was found that both a truncated version of NEIL3, consisting of the Fpg/Nei and H2TH domains, and the full-length NEIL3 protein interacted with APE1. However, interestingly, the full-length protein showed stronger interactions in comparison with the truncated NEIL3 protein, suggesting a protein-protein interaction function for the C-terminal domain of NEIL3 (Zhou *et al.,* 2017). In support of these findings, NEIL3 knockdown cells show characteristics

of telomeric dysfunction leading to mitotic defects in highly proliferating cell types, where NEIL3 is overexpressed. Thus, depletion of NEIL3 may impair telomere integrity ultimately leading to cell apoptosis, suggesting that NEIL3 could be an interesting target for novel anticancer drugs (Zhou *et al.,* 2017).

The catalytic domain of NEIL3 possesses a flexible DNA binding pocket that can accommodate a wide range of oxidized bases. This feature is utilized when NEIL3 recognizes and initiates BER in different shapes of telomeres. (Liu *et al.,* 2010, 2013).

In a study by Massaad *et al.* (2016), the effect of a naturally occurring human mutation in the N-terminal DNA glycosylase domain of NEIL3 was described. The D132V missense mutation ablated DNA glycosylase activity and in some instances led to increased autoimmunity, frequent infections and impaired B cell function in the affected individuals. In the same article, studies on NEIL3 knockout mice showed a significant increase in cell death by apoptosis of T and B cells from the spleen and also germinal centre B cells in comparison to wildtype mice. This led the authors to conclude that a lack of NEIL3 can lead to increased lymphocyte cell death and a predisposition to autoimmunity, indicating a role for NEIL3 in immune function (Massaad *et al.,* 2016).

### 1.8.1 NEIL3 in biology

To investigate the role of NEIL3 in neurogenesis throughout brain development in the embryo, the NEIL3 expression pattern was followed during brain development in mice (Hildrestrand *et al.,* 2009). Using quantitative PCR, the study showed elevated levels of NEIL3 expression corresponding to the start of neurogenesis in the developing mouse embryo. I*n situ* hybridization confirmed that NEIL3 expression was specifically in areas where neural stem and progenitor cells are found. NEIL3 expression levels declined as the brain developed and on

reaching adulthood, NEIL3 was undetectable by quantitative PCR (Hildrestrand *et al.,* 2009). However, using NEIL3 knockout mice, it was reported that NEIL3 is essential for the maintenance of neurogenesis in adult mice that NEIL3 is required for the repair of oxidative DNA damage in neural stem/progenitor cells to prevent age-associated neurodegeneration and deterioration of cognitive function (Regnell, *et al.,* 2012). In relation to this, NEIL3 has also been shown to be involved in the recovery of neurogenesis following hypoxia-ischemia in a mouse model (Sejersted *et al*., 2011).

Skarpengland *et al.* (2015) used a nested case-control study to investigate the role of several BER proteins in the risk of myocardial infarction as increased ROS and oxidative DNA damage are known to be contributing factors in atherosclerosis. Of the BER genes studied, only a SNP in NEIL3 (rs12645561TT) was associated with an increased risk of myocardial infarction. Thus, the authors concluded that decreased NEIL3 expression or activity, may result in increased atherogenesis and increased risk of heart disease (Skarpengland *et al.,* 2015).

## 1.9   DNA crosslinks

Inter-strand crosslinks (ICLs) accrue in cells due to the formation of covalent bonds between bases on the two strands of the DNA molecule. ICLs can either be induced by natural products of metabolism or by exogenous chemical agents, most notably those used in cancer chemotherapy Most common ICL - inducing chemotherapeutic agents are classified into four main groups, (i) alkylating agents, (ii) platinum – based agents such as cisplatin, carboplatin and oxaliplatin, (iii) nitrogen mustards such as cyclophosphamide and chlorambucil, and (iv) other compounds such as mitomycin C and psoralen (Deans and West, 2011).

ICL - inducing agents act mostly in similar manner utilizing two active leaving groups in order to activate the drug, the bifunctional platinum and nitrogen mustards react with guanine or

adenine bases in the two strands of the double helix at the N7 position by consecutive replacement of two chloride ions with two water molecules (Figure 1.10).
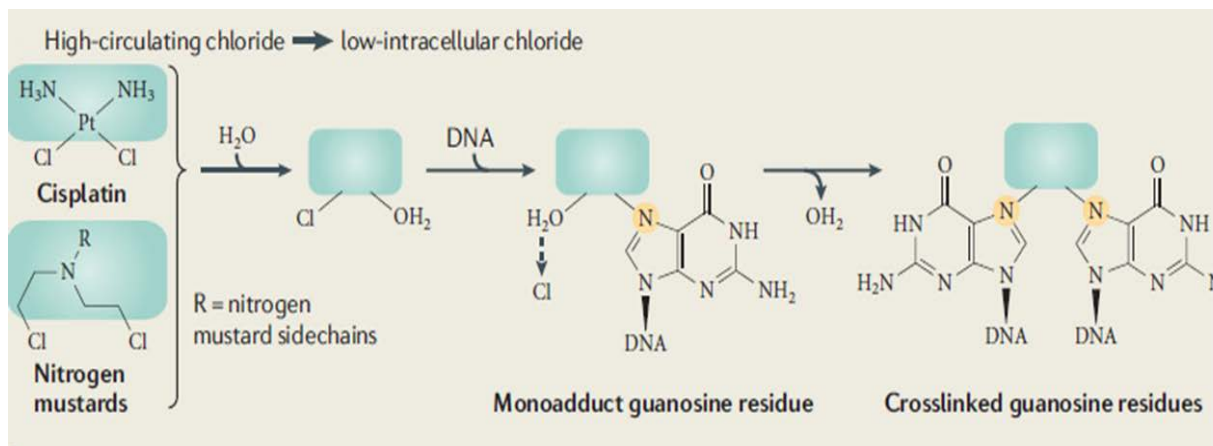


**Figure 1.10 Cisplatin and Nitrogen mustard inducing ICLs DNA lesions (taken from Deans and West, 2011)**

Psoralen, on the other hand, reacts through photon mediated cycloaddition, resulting in cross-linked thymidine nucleotides, while photon mediated cycloreduction of mitomycin C causes crosslinking between guanine bases on the opposite strands. Psoralen is widely used in the treatment of the skin condition, psoriasis, in conjunction with UVA treatment (Figure 1.11; Deans and West, 2011).
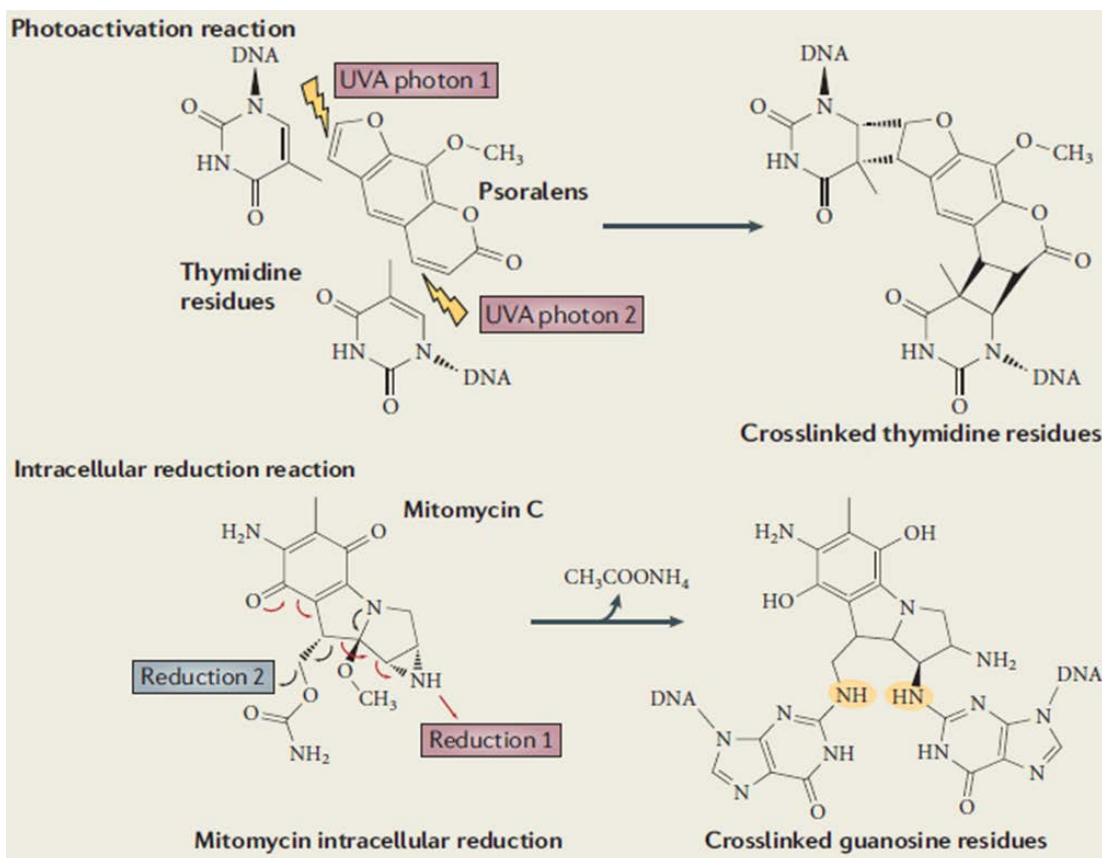
**Figure 1.11 Photoactivation reaction and intracellular reduction reaction effect on thymine and guanine generating crosslinked thymine and crosslinked guanines respectively (taken from Deans and West, 2011)**

Chemical agents introducing ICLs are used as anticancer treatments, due to the toxicity of the ICL that prevent transcription, and stall DNA replication of cells and if unrepaired could trigger cell death (Martin *et al.,* 2017).

In mammalian cells, ICLs are repaired by DNA repair machineries the incising strand at either side of the ICLs leaving a gap which filled by homologous recombination then the repair is completed by nucleotide excision repair (NER). While in eukaryotic multiple DNA repair mechanisms are involved; NER, homologous recombination, BER and the Fanconi anaemia proteins, the efficiency of the repair in order to maintain the genome integrity is dependent on

retrieving the missing information (DNA sequences) from an undamaged homologous chromosome (Deans and West, 2011).

### 1.9.1 NEIL3 in DNA crosslink repair

Human NEIL3 proved to be active on DNA crosslinks induced by psoralen on both three and four-stranded DNA structures leaving unhooked DNA products (Martin *et al.,* 2017). NEIL3 acts on both DNA strands of the four stranded ICL while Nei and NEIL1 cleave psoralen four stranded crosslinks DNA substrates leaving two double-stranded DNA with nick. NEIL3 cleave the psoralen ICL three and four-stranded DNA structures with no resultant single-strand breaks. The three and four-stranded ICL DNA can be formed in vivo during replication from replication forks where NEIL3 is upregulated and expressed in highly proliferating cells suggesting a role for NEIL3 in the repair of replication fork (Martin *et al.,* 2017). The study shows evidence that Nei, NEIL1 and NEIL3 repair three and four-stranded ICL DNA structures induced by psoralen by unhooking the substrates, hydrolysing the glycosylic bond between the oxidized base and the sugar backbone of the DNA with no resultant double-strand breaks (Martin *et al.,* 2017). The three and four-stranded ICL DNA structures can be formed during DNA replication, supporting the suggestion that Nei-like DNA glycosylases, Nei, NEIL1 and NEIL3 have a role in DNA repair at the replication fork, and furthermore, that this function is conserved from *E. coli* to human cells.

## 1.10 hNEIL3 Expression vectors

### 1.10.1 pET30a-ORF6:

The pET30a-ORF6 vector was constructed by Guo et al (2009) in an attempt to express the Rv3297 gene (MtuNei2) in *E. coli*. Previous attempts to express the Rv3297 gene in *E. coli*, based on the theory that the rare codons are the major factor in insufficient protein expression, were not successful to produce a significant amount of the protein either by augmenting tRNAs for rare codons or amplifying the gene with preferred codons for *E. coli* (Guo *et al.,* 2009).

Instead, the study proved that the formation of secondary structures surrounding the translation initiation region was the main factor affecting translation and protein expression. The vector proved efficienct for expressing proteins from G/C rich organisms and carry a proline residue at the N-terminus (Guo *et al.,* 2009).

The vector is called bicistronic because it possesses two cistrons, an upstream cistron (ORF6) which is 74 nucleotides long and reduces the local secondary structures in the mRNA which surround the translation initiation region allowing efficient translation initiation, and a downstream cistron that accommodates the target gene. The coupled translation mechanism has been anticipated in two patterns, one is that the secondary structures which inhibit translation of the downstream cistron is disturbed by the ribosome translating the upstream cistron resulting in other ribosomes ability to reach the ribosome binding site of the downstream cistron thus starting the translation of the coupled downstream gene. The other pattern is that the ribosome translating the upstream cistron is able to re-initiate translation for the downstream cistron.

## 1.10.2 pETDuet-2 bicistronic expression vector

In NEIL3 protein expression in *E. coli,* the removal of N-formylmethionine is required for the catalytic activity of the protein. The process is divided into two steps; one is the release the formyl group by a peptide deformylase which is coupled to translation, the other is N-terminal methionine cleavage by an amino peptidase (Liu *et al,* 2012).

N-terminal methionine processing is dependent on the amino acids in the second and third position of the target protein. It has been shown that valine at the penultimate (second position) and glutamic acid residues at the antepenultimate (third position) in the translated protein results in incomplete processing of the N-terminal methionine in *E. coli* (Frottin *et al.,* 2006). An engineered EcoMap (EcoMapY168A) from pBSMap (Y168A)-cGSTM produced promising results of processing the N-terminal methionine in proteins with valine residue at penultimate position (Liu *et al.,* 2012).

## 1.10.3 Construction of the bicistronic vector pETDuet-2-T7EcoMap-ORF6

Liu and colleagues (2012) modified the pETDuet-1 expression vector (Novagen) by amplifying a mutated version of the *E. coli* amino peptidase, EcoMapY168A, then the PCR product was digested and ligated into the pETDuet-1 using NcoI and SalI restriction enzymes resulting pETDuet-1-T7 EcoMapY168A (Liu *et al.*, 2012). Then the leader sequence ORF6 and hexa-histidine tag from previous pET30a-ORF6 vector were cloned into the vector using NdeI and StyI cut in the T7 terminator. The target protein mouse NEIL3 was cloned into the NdeI and XhoI cut in ORF6 leader sequence resulting of pETDuet-2-T7EcoMap-ORF6-NEIL3 expression vector (Figure 1.12). The expression of active recombinant MmNeil3 and N-terminal hNEIL3 using the pETDuet-2-EcoMap-ORF6 expression vector was successful with DNA glycosylase/lyase activity when was subjected to substrates containing Sp and thymine glycol (TG), or an AP site in single- and double-stranded DNA (Liu *et al.*, 2012).
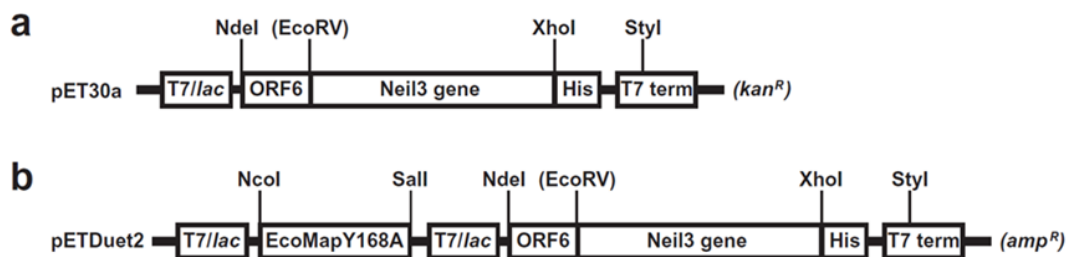


**Figure 1.12 Map of the pET30a and pETDuet-2 expression vectors containing murine NEIL3.** The figure shows the position of the ORF6 and EcoMap Y168A, the His-Tag and the restriction sites that were used to construct the vector (taken from Liu *et al,* 2012).

## 1.11   Project aims and objectives

The biochemical properties of NEIL3 are still an enigma and due to the fact that the protein is highly expressed in cancer cells, testing the function of the protein domains to determine which of the domains is responsible for the protein activity could lead to therapeutic significance. In this regard, studies have demonstrated that cells from mice were lacking NEIL3 show sensitivity to oxidative genotoxins and the DNA cross-link inducing agent, cisplatin (Rolseth, *et al.,* 2013).

In this project expression and purification of full-length and four truncated versions of hNEIL3 was undertaken to test the effect of the five protein domains of human NEIL3 (hNEIL3) on DNA damage substrates to further characterize the biochemical properties of hNEIL3. Thus, expression and purification of hNEIL3 full-length and four truncated versions was attempted to test the effect of the different protein domains on DNA glycosylase / AP lyase activity (Figure 1.13).
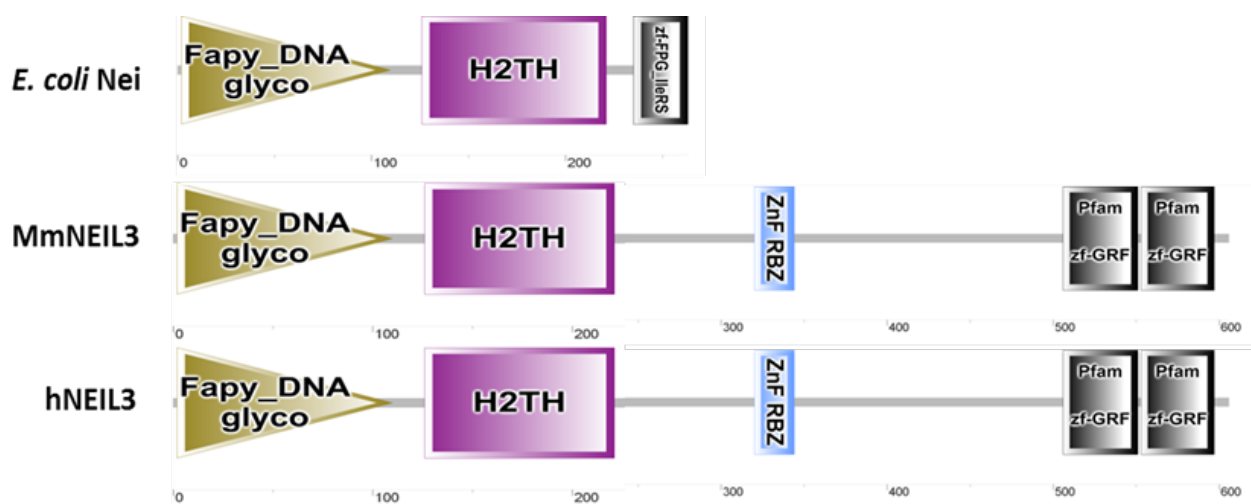


**Figure 1.13 Conserved protein domains of *E. coli* Nei, *M. musculus* NEIL3 (MmNEIL3) and hNEIL3.**
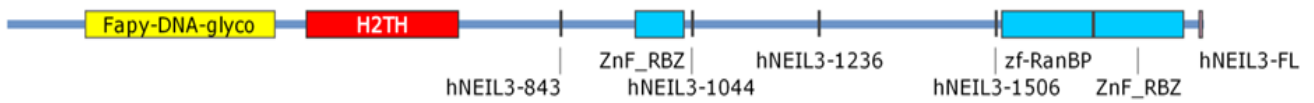
**Figure 1.14 Illustration of hNEIL3 protein shows the site of the different truncations in relation to the conserved protein domains of hNEIL3.**
The figure shows the hNEIL3 protein domains and their positions. Numbers refer to the nucleotide position in the cDNA (CDS).

The first stage of the project entailed that full-length hNEIL3 and four truncated versions (Figure 1.14) were cloned in to pET30b-ORF6. Thereafter, the ORF6-hNEIL3 (or truncated hNEIL3 cDNA) was released from the vector and introduced to the pETDuet-2 bicistronic expression vector and this used to transform *E. coli* cells and tested for protein expression. The resulting His-tagged proteins were purified by chelation and ion-exchange chromatography and biochemical enzyme assays performed to determine the effect, if any, of the C-terminal domains on DNA glycosylase / AP lyase activity.

# 2    Materials

**Table 2.1 Materials used in molecular cloning experiments**

| Material | Description |
|---|---|
| **Tris-Borate-EDTA (TBE)** | 5x TBE was prepared by dissolving 54 g of Tris base, 27.5 g of boric acid and 20 ml of 0.5 M EDTA pH8.0 in 1 L of $dH_2O$. The 5x TBE stock solutions was diluted 10 fold to 0.5 x prior to use. |
| **Agarose Gel Electrophoresis Loading Buffer** | 4 g of sucrose, 25 mg of bromophenol blue (0.25% w/v), and 2.4 ml of 0.5 M EDTA pH 8.0 was mixed and made up to 10 ml with $dH_2O$. |
| **DNA Size Marker** | 1 μl of Hyperladder 1kb (Bioline) was used to determine DNA band sizes for all agarose electrophoresis experiments. |
| **GelRed** | 4 μl of GelRed (Biotium) was added to 100 ml of agarose gel solution at approximately 50ºC. |
| **Lysogeny broth (LB)** | 5 g of LB-broth powder (Sigma-Aldrich) was added to 200 ml of $dH_2O$ followed by autoclaving at 121°C for 45 min. |
| **LB with agar** | LB with agar (Sigma-Aldrich) was prepared by adding 7 g of LB-agar powder to 200 ml of $dH_2O$ followed by autoclaving at 121°C for 45 min. |

## 2.1 Protein Expression and Purification Materials

**Table 2.2 Materials used in Recombinant Protein Expression, Purification and Analysis.**

| Material | Description |
|---|---|
| Acrylamide | Acrylamide (30% acrylamide, 0.8% bis) obtained ready for use (Bio-Rad). |
| Ammonium Persulfate (APS) | A 10% (w/v) stock solution of APS was prepared by dissolving 1 g of APS in 10 ml dH$_2$O. Aliquots were stored at -20°C. |
| Tetramethylethylenediamine (TEMED) | TEMED was obtained from Sigma-Aldrich. |
| SDS-PAGE Loading Buffer (3x) | The stock was prepared by mixing, 2.4 ml of 1M Tris-HCl pH 6.8, 3 ml 20% (w/v) SDS, 3 ml, 100% (v/v) glycerol, 1.6 μl 2-mercaptoethanol and 6 mg bromophenol blue and the final volume adjusted to 10 ml with dH$_2$O and stored at 4°C. |
| SDS-PAGE Running Buffer 10x | The 10x stock was prepared by mixing 30.2 g Tris base, 10 g of SDS and 144 g of glycine in about 900 ml of H$_2$O. The pH was adjusted to pH8.3 if required and the volume adjusted to 1 L with dH$_2$O. The 10x stock was diluted to 1x buffer just before use. |
| SDS-PAGE stacking gel buffer | 0.5M Tris-HCl pH6.8, 10% (w/v) SDS |
| SDS-PAGE separating gel buffer | 1.5M Tris–HCl pH 8.8, 10% (w/v) SDS |
| Western Blot Transfer Buffer (WTB) | 10x western blot transfer buffer (WTB) was prepared by dissolving 144 g glycine and 30.2 g Tris base and the final volume made up to 1 L with dH$_2$O. A 1x working buffer was prepared by adding 100 ml of 10x TG to 200 ml of methanol and made up to 1 L with dH$_2$O and stored at 4°C. |

| Phosphate Buffered Saline (PBS) Stock Solution (10x) | 10x PBS was prepared by adding 80 g NaCl, 2 g KCl, 7.62 g $Na_2HPO_4$ and 0.77 g $KH_2PO_4$ to 800 ml of $dH_2O$. Once fully dissolved, the pH was adjusted to pH7.4 with concentrated HCl and the total volume made up to 1 L with $dH_2O$. |
|---|---|
| PBS-Tween-20 | The working solution of PBS-Tween-20 contained 100 ml of 10x PBS and 1 ml of Tween-20 made up to 1 L with $dH_2O$. |
| Blocking buffer | Blocking buffer was prepared using Odyssey blocker solution (LI-COR) diluted 1:1 in 1x PBS. |
| Antibody dilution buffer | The antibody dilution buffer was prepared using Odyssey blocker solution (LI-COR) diluted 1:1 in 1x PBS containing 0.1% Tween-20. |
| Protein molecular weight (MW) Standards | Precision plus protein All Blue Pre-stained Protein Standards (Bio-Rad) were used to determine the molecular weight of polypeptides following SDS-PAGE and western blotting. |
| IPTG (Isopropyl β-D-1-thiogalactopyranoside) Stock Solution | IPTG stock solution was prepared to a final concentration of 100 mM by dissolving 120 mg of IPTG powder (Promega) in 5 ml $dH_2O$ then filter sterilized and stored at -20°C. |
| Bacterial pellet Lysis Buffer | Bacterial lysis buffer was prepared by mixing 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 5% (v/v) glycerol and 5 mM imidazole then $dH_2O$ to a final volume of 500 ml. To each 30 ml, 30 µl of protease inhibitor cocktail; (leupeptin, pepstatin a, chemostatin, and aprotinin, Thermo Fisher Scientific) and 100 µl of 100 mM Phenylmethane sulphonyl fluoride (PMSF; Sigma-Aldrich) was added just before use. |

| FPLC His-Trap purification buffers | **Buffer A** consisted of: 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 5% (v/v) glycerol, 5 mM imidazole and 0.1 mM PMSF. |
| --- | --- |
| | **Buffer B** consisted of: 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 5% (v/v) glycerol, 500 mM imidazole and 0.1 mM PMSF. |
| FPLC Mono S purification buffers | **Buffer A** consisted of: 50 mM Tris-HCl pH 8.0, 50 mM KCl, 1 mM EDTA pH8.0, 5% (v/v) glycerol, 1 mM dithiothreitol (DTT) (1:1000 dilution from 1 M stock) and 0.1 mM PMSF. |
| | **Buffer B** consisted of: 50 mM Tris- HCl pH 8.0, 1 M KCl, 1 mM EDTA pH8.0, 5% (v/v) glycerol, 1 mM DTT and 0.1 mM PMSF. |

## 2.2 Activity Assay Materials

**Table 2.3 Materials used in Recombinant Protein Activity Assays**

| Material | Description |
|---|---|
| **Denaturing - PAGE sample loading buffer** | 95% (v/v) formamide, 0.02% (w/v) xylene cyanol, and 0.02% (w/v) bromophenol blue. |
| **Activity assay reaction buffer** | The activity assays were carried out using 300 nM of purified proteins and 5 nM of the substrate oligonucleotides in a total volume of 10 µL of reaction mixture containing, 50 mM Tris-HCl pH 7.8, 50 mM KCl, 10 mM $MgCl_2$, 0.5 mM EDTA pH8.0, 1.5 mM DTT, 8.5% (v/v) glycerol, and 100 µg/ml bovine serum albumin (BSA). |
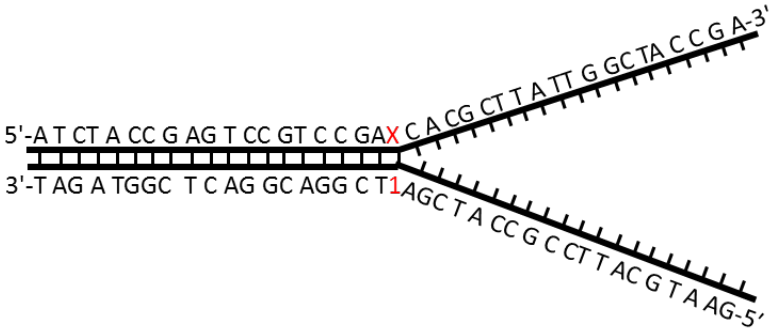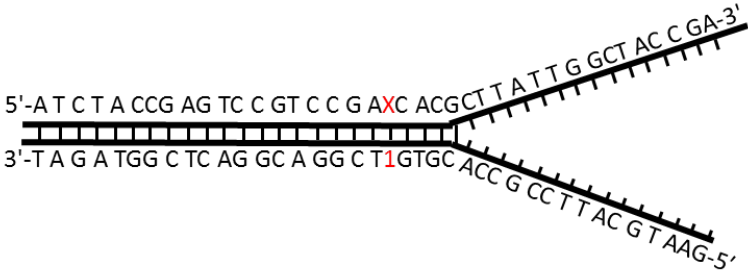
### 2.2.1 Oligonucleotide Substrates

The sequences and design of the oligonucleotides used in the activity assays is given in Table 2.4. The oligonucleotides containing the modified base were fluorescently labelled with 5′-IRDye700 (8-oxoG, TG, U) or 5′-Alexa Fluor 680 (5-OHU) and were purchased from IDT Technologies. Double-stranded oligonucleotides and mock replication fork substrates were produced by adding a 1.5-fold excess of the complementary oligonucleotide, heating at 95°C for 5 min and then left to cool at room temperature.

**Table 2.4 Oligonucleotide Substrate Design and Sequences.**

| Substrate | Oligonucleotide sequence and construct design |
|---|---|
| Single-stranded DNA substrate | 5'-ATCTACCGAGTCCGTCCGA**X**CACGCTTATTGGCTACCGA-3'<br><br>5'-A T CT A CC G AG T CC GT C C GA**X** C AC G C T TA T T GG C TA C C G A-3'<br><br>X denotes the position of the base modification (5-OHU / 8-oxoG / TG). |
| Double-stranded DNA substrate | 5′-ATCTACCGAGTCCGTCCGA**X**CACGCTTATTGGCTACCGA-3′<br>3′-TAGATGGCTCAGGCAGGCT**1**GTGCGAATAACCGATGGCT-5′<br><br>5'-A T CT A CC G AG T CC GT C C GA **X**C A C GC T T AT  T GGC T A CC G A-3'<br>3'-T AG A TGGC  T C AG GC AGG C T **1**T G C GA AT A A C CG AT GG C T-5'<br><br>X denotes the position of the base modification (5-OHU / 8-oxoG / TG), while '1' represents the appropriate complementary base opposite the oxidised base. |
| Fork -4 DNA substrate | 5′ -ATCTACCGAGTCCGTCCGA**X**CACGCTTATTGGCTACCGA-3'<br>3′-TAGATGGCTCAGGCAGCTATAGCTACCGCCTTACGTAAG-5′<br><br><br><br>X denotes the position of the base modification, 5-OHU, 8-oxoG or TG. |

| | |
|---|---|
| Fork DNA substrate | 5′-ATCTACCGAGTCCGTCCGA**X**CACGCTTATTGGCTACCGA-3′<br><br>3′-TAGATGGCTCAGGCAGGCT**1**AGCTACCGCCTTACGTAAG -5′<br><br>5'-A T CT A CC G AG T CC GT C C GA**X** C A CG CTT A TT G GCTA C C G A-3'<br>3'-T AG A TGGC T C AG GC AGG C T**1**AGC T A CC G C CT T AC G T A AG-5′<br><br>X denotes the position of the base modification (5-OHU / 8-oxoG / TG), while '1' represents the appropriate complementary base opposite the oxidised base. |
| Fork +4 DNA substrate | 5′-ATCTACCGAGTCCGTCCGA**X**CACGCTTATTGGCTACCGA-3′<br>3′-TAGATGGCTCAGGCAGGCT**1**GTGCACCGCCTTACGTAAG-5′<br><br>5'-A T C T A CCG AG TC C GT C C G A**X**C ACG CTT A TT G GCT AC C GA-3'<br>3'-T A G A TGG C TC AG GC A GG C T**1**GTGC ACC G CC T T AC G T AAG-5′<br><br>X denotes the position of the base modification (5-OHU / 8-oxoG / TG), while '1' represents the appropriate complementary base opposite the oxidised base. |

# 3 Methods

## 3.1 Cloning Methods

### 3.1.1 Bacterial Culture

For both LB and LB with agar, the selection antibiotic was added after autoclaving and cooling to approximately 50°C. The LB-agar was poured into 10 cm petri dishes, left to set and the plates stored at 4°C until needed, while the LB was stored at room temperature. *E. coli* bacterial overnight cultures were prepared in 5 ml of LB in universal tubes and 5 μl of ampicillin (100 mg/ml) or 5 μl of kanamycin (30 mg/ml) was added where appropriate, just before the addition of the *E. coli* cells. The tubes were incubated overnight (16 - 18 h) at 37°C in a shaking incubator at 220 rpm.

### 3.1.2 Plasmid DNA Extraction and Quantification

Plasmid DNA extractions were carried out as described in the Isolate II Plasmid mini kit protocol (Bioline). Three millilitres of overnight culture were centrifuged for 30 s at 11,000 *xg,* the supernatant discarded and the pellet re-suspended in 250μl re-suspension Buffer P with vortexing. Then 250 μl of Lysis Buffer P2 was added and mixed by inversion and incubated for 5 min at room temperature. Then 300 μl of Neutralization Buffer P3 was added, mixed thoroughly by inversion and centrifuged for 5 min at 11,000 *xg*. An Isolate II Plasmid Mini Spin Column was placed in a 2 ml collection tube and the clarified supernatant was pipetted into the column and centrifuged for 1 min at 11,000 *xg*. The flow-through was discarded and 500 μl wash buffer PW1 was added to the column to wash the silica membrane and centrifuged for 1 min at 11,000 *xg* after which the flow-through was discarded and 600 μl of wash buffer PW2 was added and the column centrifuged again for 1 min at 11,000 *xg*. The flow-through was discarded and the column dried by centrifuging for 2 min at 11,000 *xg*. Finally, the column was placed into a fresh 1.5 ml collection tube and 50 μl of Elution Buffer P was added onto the

column, incubated at room temperature for 1 min and then centrifuged for 1 min at 11,000 *xg*. The flow-through was collected and the plasmid DNA concentration measured using a NanoDrop 2000 instrument (Thermo Scientific) and stored at -20°C.

### 3.1.3 Agarose Gel Preparation

Agarose gels were prepared by dissolving 0.8 g of agarose (Fisher BioReagents) in 100 ml of 0.5x TBE by heating in a microwave and then the mixture left to cool to approximately 50°C and 4 μl of GelRed (Biotium) was added. The liquid agarose was then poured into a gel tray, a well-forming comb placed in position and the gel left at room temperature to set. The gel was then placed in an electrophoresis tank containing 0.5x TBE and the comb removed.

### 3.1.4 Polymerase Chain Reaction (PCR)

PCR was used to amplify the target hNEIL3 cDNA from pCMV6-AC-hNEIL3 (OriGene) for use in molecular cloning. Details of the PCR primers used are given in Table 2.2 with the XhoI site and three flanking nucleotides underlined.

**Table 3.1 PCR Primers for hNEIL3.**

| Primer | Primer sequence | *Tm* | *Tm* (XhoI site included) | CG% |
|---|---|---|---|---|
| **hNEIL3 Fwd.** | G GTG GAA GGA CCA GGC TGT ACT CTG AAT | 73.2°C | — | 53.6 |
| **hNEIL3 843 bp Rev** | <u>CCG CTC GAG</u> TTT TTG ACA GTG AGG ACA GAA ATA TGT CAT TCT GT | 72.1°C | 79.8°C | 34.3 |
| **hNEIL3 1044 bp Rev** | <u>CCG CTC GAG</u> TGA ATC AAT AGG CCT TGA GGT CAA GC | 70.7°C | 80.7°C | 46.2 |
| **hNEIL3 1236 bp Rev** | <u>CCG CTC GAG</u> ATC TAG TAT CTG GTT TTG CTT TGT TTT TCT TTC CAA AG | 71.9°C | 81.1°C | 31.6 |

| | | | | |
|---|---|---|---|---|
| **hNEIL3 1506 bp Rev** | CCG CTC GAG AGG ATT TAA GGT ACG AGG GCC ATC TGT | 70.4°C | 81.5°C | 48.1 |
| **hNEIL3 Full length Rev** | CCG CTC GAG GCA TCC AGG AAT AAT TTT TAT TCC TGG CC | 71.9°C | 82.2°C | 41.4 |
| **Sense ORF6-hNEIL3 Fwd.** | TATGAAAATCGAAGCAGGTAA ACTGGTACAGAAGG | 67.1°C | — | 40 |

PCR reaction mixtures for hNEIL3-full length and reaction conditions were as follows:

<u>PCR reaction mix for full length hNEIL3 cDNA</u>

| **Reagent** | **Amount** |
|---|---|
| 1 ng/µl Template DNA | 1.5 µl |
| 5x Reaction buffer | 10.0 µl |
| 10 mM dNTPs | 1.0 µl |
| 10 µM Forward primer | 2.5 µl |
| 10 µM Reverse primer | 2.5 µl |
| DMSO | 1.5 µl |
| Phusion DNA polymerase (NEB) | 0.5 µl |
| dH$_2$O | 30.5 µl |
| **Total** | **50.0 µl** |

<u>PCR Cycle Conditions</u>

95°C      30 s

95°C      10 s

76°C      30 s     30 cycles

72°C      30 s

72°C      10 min

4°C      Hold

For preparation of the four truncated cDNAs, the reaction mixtures and PCR reaction conditions were as follows:

PCR reaction mix for truncated hNEIL3 cDNAs

| Reagent | Amount |
| --- | --- |
| 10 ng/µl template DNA | 1.5 µl |
| 5x Reaction buffer | 10.0 µl |
| 10 mM dNTPs | 1.0 µl |
| 10 µM Forward primer | 2.5 µl |
| 10 µM Reverse primer | 2.5 µl |
| Phusion DNA polymerase | 0.5 µl |
| dH$_2$O | 32.0 µl |
| Total | 50.0 µl |

PCR Cycle Conditions

| | | |
| --- | --- | --- |
| 95°C | 30 s | |
| 95°C | 10 s | |
| 76°C | 30 s | 30 cycles |
| 72°C | 30 s | |
| 72°C | 10 min | |
| 4°C | Hold | |

### 3.1.5  Restriction Endonuclease Digests

All restriction endonuclease digests were carried out using 300 ng of plasmid DNA, 10 units of restriction enzyme and the appropriate buffer according to the manufacturer's recommendation (New England Biolabs).

**Table 3.2 Restriction Endonuclease Digest Reaction Mixture**

| Component | volume |
|---|---|
| Plasmid cDNA (~300 ng) | 5.0 µl |
| 10x NEB Buffer | 2.0 µl |
| Enzyme (20 units/µl) | 0.5 µl |
| dH$_2$O | 12.5 µl |
| Total | 20.0 µl |

Reactions were carried out at 37°C for 1 h.

### 3.1.6  DNA Ligation

All DNA ligation reactions were carried out using T4 DNA ligase (New England Biolabs). The protocol was as follows:

The amount of insert DNA required was calculated using the following formula to achieve a 3:1 insert/vector ratio:

$$X \text{ ng insert} = \frac{(y \text{ bp insert}) (z \text{ ng vector})}{\text{bp of vector}}$$

In general, 50 ng of vector was used in the DNA ligation reactions.

**Table 3.3 Reaction mix for ligation of hNEIL3 inserts into (pET30b-ORF6 and pETDuet-2) using T4 DNA ligase.**

| Reagent | Volume |
|---|---|
| 10x T4 DNA ligase Buffer | 2µl |
| Vector DNA (50 ng) | x µl |
| Insert DNA | x µl |
| Nuclease-free water | Up to 20 µl |
| Total | 20 µl |

Reactions were carried out at 16°C for approximately 18 h.

## 3.1.7 Mung Bean Nuclease

Blunting of the plasmid DNA fragment by degrading the 5′ overhang resulting from NdeI digestion was carried out using mung bean nuclease (New England Biolabs). The reaction mixture contained 5 µl of 1x mung bean nuclease reaction buffer and 1 µl (10 units) of mung bean nuclease. The reaction was incubated at 30°C for 30 min.

## 3.1.8 Clean-up and Purification from an Agarose Gel of DNA Fragments.

Digested plasmids and inserts were cleaned up using an Isolate II PCR and Gel Kit (Bioline). The sample was prepared by mixing 1 volume of sample with 2 volumes of Binding Buffer CB. Then an Isolate II PCR and Gel Column were placed into a 2 ml collection tube and the sample loaded into the column and centrifuged for 30 s at 11,000 *xg*. The flow-through was discarded and the silica membrane of the column was washed by adding 700 µl Wash Buffer CW to the column and centrifuged for another 30 s at 11,000 *xg*. The flow-through was discarded and the column placed back into the collection tube and the silica membrane dried by centrifuging for 1 min at 11,000 *xg*. The spin column was then placed in a fresh 1.5 ml tube and 15 - 30 µl of

Elution Buffer C was added directly onto the silica membrane, incubated at room temperature for 1 min, and centrifuged again for 1 min at 11,000 *xg*. The DNA was collected and the concentration measured using a Nano Drop 2000 instrument (Thermo Scientific) and stored at -20°C.

### *3.1.9* **Transformation of *E. coli* NovaBlue competent cells**

NovaBlue and Rosetta 2 competent *E. coli* cells (Novagen) were kept at -80°C in tubes of 20 μl. The cells were thawed on ice for two to five minutes and then were finger-flicked to ensure complete re-suspension. One microlitre of a ligation reaction was added to each tube separately, mixed gently, put on ice and incubated for 5 min. Then the cells were heat-shocked at 42°C for exactly 30 s and returned to ice for 2 min. Eighty microlitres of SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose) was added to each tube and either incubated at 37°C in a shaking incubator for 30 minutes, if the selection antibiotic was kanamycin, or spread directly on LB-agar (+antibiotic) plates. The plates were then incubated in a 37°C incubator overnight.

## 3.2 Protein expression and purification methods

### 3.2.1 Protein expression (test cultures)

A single bacterial colony was transferred to 3 ml of LB with ampicillin (50 μg/ml) and chloramphenicol (34 μg/ml) and incubated at 37°C with shaking until an $OD_{600}$nm of ~ 0.6-0.8 was reached and then stored at 4°C. (A glycerol stock was also made by removing 150 µl of culture, and adding 50 µl 50% (v/v) glycerol then stored at -80°C to create a frozen stock.) Three hundred microlitres of this culture was used to start two 30 ml cultures (1:100) containing antibiotics (30 µl ampicillin (50 mg/ml) and 30 µl chloramphenicol (34 mg/ml)) plus 60 µl 20% (w/v) glucose. A 1 ml aliquot was removed, centrifuged for 1 min at 13,000 *xg,* the supernatant discarded and the pellet stored at -80°C (un-induced control). Then the 30 ml cultures were incubated with shaking at 37°C until an $OD_{600}$nm ~ 0.8-1.0 was obtained.

The cultures were induced with 30 µl 1 M IPTG (final concentration 1 mM) and grown overnight at 37°C (or 16°C) with shaking. Aliquots of 1 ml were removed at 1 h, 3 h, and overnight post-induction, centrifuged for 1 min at 13,000 *xg*, the supernatant removed and pellets stored at -80°C. The remaining ~30 ml culture was centrifuged at 8,000 *xg* for 10 min, the supernatant was removed and the pellet stored at -80°C.

The cell pellets obtained from the 1 ml aliquots (un-induced control, plus 1 h, 3 h and overnight pellets at 37°C and 16°C) were re-suspended in 200 µl 1x SDS loading buffer) then carefully sonicated using 3 x 15 s bursts with 30 s intervals on ice. Then samples were heated to 95°C for 5 minutes to be analysed by SDS-PAGE and western blotting.

### 3.2.2  Protein expression for subsequence purification

Three millilitre LB cultures were prepared from the bacterial glycerol stock with the corresponding antibiotics (3 µl ampicillin (50 mg/ml)) and 3 µl chloramphenicol (34 mg/ml)) and incubated at 37°C with shaking. The cultures were grown until an $OD_{600}$nm of ~ 0.6 - 0.8 was obtained and then stored at 4°C. The following day, 300 µl of culture was used to seed two 30 ml cultures (1:100) containing antibiotics (30 µl ampicillin (50 mg/ml) and 30 µl chloramphenicol (34 mg/ml)) plus 60 µl 20% (w/v) glucose. Then the 30 ml cultures were grown at 37°C with shaking until an $OD_{600}$nm of ~ 0.8 - 1.0 was reached. This was used to seed a 300 ml culture (1:10) containing 300 µl ampicillin (50 mg/ml) and 600 µl 20% (w/v) glucose and incubation was continued until an $OD_{600}$nm of ~ 0.8 - 1.0 was obtained. The cultures were then induced with 1 mM IPTG (330 µl of 1 M stock) and incubated overnight at 20°C with shaking. The culture was centrifuged at 8,000 *xg* for 10 min, the supernatant discarded and the pellets stored at -80°C for subsequent purification.

### 3.2.3  Protein purification (His-Trap FPLC)

The bacterial pellets were re-suspended in 30 ml of lysis buffer supplemented with 30 µl of protease inhibitor cocktail (all at a concentration of 1 mg/ml) and 100 µl of 100 mM PMSF. Then, 0.1 mg/ml lysozyme was added and the mix was incubated on ice for 15 min. Cells were lysed by sonication on ice using 3 x 15 s bursts at 30 s intervals and centrifuged in Oakridge tubes at 25,000 xg for 20 min at 4°C. The supernatant was collected and filtered through 1 µm syringe pre-filters, and then through 0.45 µm syringe filters. A 1 ml His-Trap column (GE Healthcare) was washed with 3 volumes of water, followed by 3 volumes of lysis buffer containing 0.1 mM PMSF using an FPLC (AKTA Purifier UPC10; GE Healthcare) in a cold cabinet at 4°C. Filtered supernatant was added to the washed column using a 10 ml super-loop.

The column was washed with lysis buffer containing 0.1 mM PMSF until no more protein eluted.

Gradient elution with imidazole (Buffers A and B, Table 2.2) was carried out using 20 column volumes of elution buffer containing 0.1 mM PMSF and 0.5 ml fractions collected. Ten microlitres of each fraction were removed and added to 10 µl water and 10 µl 3x SDS-PAGE loading buffer and analysed by SDS-PAGE. The remaining fractions were stored at -80°C.

### 3.2.4   Amicon buffer exchange

After confirmation by SDS-PAGE and western blot analysis, the pooled eluate was loaded in a 15 ml Amicon filter tube 30 K (Merck Millipore). The tube was centrifuged at 4000 $xg$ for 30 min to 1 h at 4°C and centrifugation continued until approximately 250-500 µl of product remained in the filter tube.

The product was then transferred to a 1.5 ml microcentrifuge tube and diluted 1:10 with Buffer A for Mono S purification.

### 3.2.5   Protein purification (Mono S FPLC)

The FPLC system was prepared in the same way as for His-Trap purification preparation, however using buffers used for Mono S purification. The column was loaded at the same flow rate and the flow through collected. The column was washed with buffer A and a 20 ml linear gradient of KCl (50 mM – 1 M) was performed. Fractions were collected based upon the 280 nm peak observed upon the real time chromatogram, Fractions were analysed by SDS-PAGE and western blot and stored in elution buffer (buffer B).

### 3.2.6   SDS-PAGE

Two 10% SDS-PAGE gels were prepared in Bio-Rad mini-protean apparatus. Ten millilitres of separating gel were prepared by combining 3.4 ml 30% acrylamide, 2.7 ml of separating

buffer (Table 2.2), 100 µl of 10% (w/v) APS and 10 µl TEMED. The separating gel mixture was poured between the plates and overlaid with 100% ethanol and left to set for >30 min. The ethanol was washed off and the plates dried using 3MM filter paper. The 5% stacking gel was prepared by combining 1.34 ml 30% acrylamide, 2.6 ml stacking gel buffer, 10 µl of 10% (w/v) APS and 10 µl TEMED and poured onto the separating gel. A well-forming comb was put in place (10 or 15 wells) and the gel left to set for >30 min.

Samples were re-suspended in 3x SDS-PAGE loading buffer and heated at 95°C for 5 min. The gel was unwrapped and the white strip and comb were removed and the wells washed with water. The SDS-PAGE apparatus was assembled and 1x TGS added to the reservoir and the outer tank filled with ~2 cm of buffer. One microlitre of protein marker, 10 µl un-induced and 5 µl induced of protein samples were loaded, and electrophoresis carried out at 125V for 2 h. One SDS-PAGE gel was stained with InstantBlue (Expedeon) for about 15 min to examine overall protein staining, while the other gel was transferred to Immobilon FL membrane (EMD Millipore) at 25 V for 1.5 h in WTB. Following transfer, the membrane was analysed by immunoblotting using anti-His-tag antibodies.

### 3.2.7 Western Blotting Protocol

The Immobilon FL PVDF membrane was prepared by placing it in methanol for 15 s then washing in $dH_2O$ for 1 min and then placed in transfer buffer for 1 to 2 min. A plate snapper was used to remove the gel and then the gel rinsed briefly in transfer buffer.

The transfer was set up the as following order: sponge, filter paper, gel, membrane, filter paper, and sponge and run for 1.5 h at 25 V.

The membrane was rinsed in PBS for a few minutes then blocked for 1 h in Odyssey blocker (diluted 1:1 in PBS).

The membrane was then incubated with the primary antibody, a mouse anti-His monoclonal antibody (Novagen) at a dilution of 1:1000 overnight at 4°C. The membrane was then washed three times in PBS-Tween-20 for 5 min before incubation with a 1:10,000 dilution of Alexa Fluor 680 goat anti-mouse secondary antibody (Thermo Fisher) for 1 h at room temperature and protected from light. The membrane was washed three times for 5 min in PBS-Tween-20. Finally, the membrane was washed in PBS for 5 min to remove excess Tween-20 and scanned using an Odyssey imaging system (LI-COR).

### 3.3 Enzyme Activity Methods

### 3.3.1 NaBH$_4$-trapping assay

The trapping assay was carried out using 300 nM of purified proteins and 5 nM oligonucleotide substrate in total volume of 10 µL of reaction buffer and 1 µL of 1 M NaBH$_4$ solution added prior to incubating the reaction at 30°C for 20 min. The reaction was stopped by the addition of 10 µL of 3x SDS-PAGE loading buffer and heated to 95°C for 5 min. Samples were loaded onto 10% SDS-PAGE gels and electrophoresis carried out at 125V for 2 h, then the gel was visualised using an Odyssey imaging system.

### 3.3.2 Denaturing Polyacrylamide/Urea Gel Preparation

The gel was prepared by adding 16.8 g urea into a 50 ml Falcon tube followed by 8 ml of 5x TBE and 10 ml of 30% (w/v) polyacrylamide gel solution. The mixture was made up to 40 ml with dH$_2$O and mixed thoroughly. Then, 250 µL of 10% (w/v) APS and 25 µL TEMED were added just before pouring the gel mix into the vertical gel cassette and a 10 well comb fitted on top. The gel was left for approximately 30 min to set and samples were loaded and electrophoresis carried out at 300 V for 90 min.

### 3.3.3 DNA Glycosylase/lyase Activity Assays

The assays were carried out using 300 nM of purified proteins and 5 nM of oligonucleotide substrate in a total volume of 10 µL of reaction buffer (Table 2.3). The reaction mixtures were incubated at 30°C for 20 min and the substrate and product bands separated by electrophoresis through 20% denaturing polyacrylamide/urea gels and visualised using the Odyssey imaging system.

Increasing protein concentrations (150 nM, 300 nM and 600 nM) and incubation times 20 min and 30 min) and different temperatures (30°C and 37°C) were used in order to optimise the

reaction conditions. The reactions were stopped using 10 µl of loading buffer (Table 2.3). then heated for 5 min at 95°C.

# 4 Results

## 4.1 Overview

In order to express hNEIL3-FL and the four truncated versions of hNEIL3 in *E. coli,* the

pETDuet-2 expression vector described by Lui *et al.* (2012) was used. To achieve this,

pETDuet-2 containing the murine NEIL3 (MmNEIL3) cDNA was used as the starting material.

Briefly, the ORF6-MmNEIL3 DNA (see Figure 4.14) was removed from the vector using NdeI

and XhoI restriction enzymes and then replaced with the ORF6 sequence fused to one of the

hNEIL3 cDNAs that had been obtained by PCR. This was the first stage of the project.

However, direct cloning of the ORF6-hNEIL3 DNAs into pETDuet-2 using restriction digests

with NdeI and XhoI was not possible, since there is an NdeI restriction site at position 864 in

the hNEIL3 cDNA (Figure 4.1). In addition, pETDuet-2 lacked any other restriction sites that

could have been used in place of NdeI. The only exception to this was hNEIL3-843, because

the NdeI restriction site at position 864 is beyond the end of the hNEIL3-843 sequence.

Therefore, to overcome this problem and achieve successful cloning of the majority of the

hNEIL3 cDNAs into pETDuet-2, a multistep cloning strategy had to be designed.

First, the cDNAs of hNEIL3-(843, 1044, 1236, 1506 and full length) were amplified (Figure

4.5) from the original commercial clone, pCMV6-AC-hNEIL3 and then transferred to the

pET30b-ORF6 vector, which had previously been prepared in our lab and was obtained from

lab stocks (Balis, 2013). After successful cloning of the various hNEIL3 cDNA inserts into

pET30b-ORF6 to generate an ORF6-hNEIL3 overlapping fusion sequence (

Figure 4.2), two different strategies were required to prepare the final pETDuet-2 constructs.

For ORF6-hNEIL3-843, the combined sequence was simply released by NdeI and XhoI double-

digest (Figure 4.15) and subsequently cloned into the pETDuet-2 expression vector. However, for the construction of pETDuet-2-hNEIL3-(1044, 1236, 1506 and full length) clones, a modified cloning strategy had to be used, because direct cloning of the hNEIL3 inserts was not possible due to the NdeI site in the hNEIL3 cDNA sequence at position 864 (Figure 4.1) Thus, with the exception of hNEIL3-843, all other ORF6-hNEIL3 sequences were first PCR amplified from pET30b-ORF6-hNEIL3 (Figure 4.12 and Figure 4.13) and the amplicons digested with XhoI. The amplicons thus had a blunt 5′-end and a XhoI site at the 3′-end to enable cloning into a prepared pETDuet-2 vector (Figure 4.15). For this, pETDuet-2 was first cut with NdeI and the product incubated with mung bean endonuclease to create a blunt end. Following DNA purification, the treated linear plasmid was incubated with XhoI to create the 3′-end. Thereafter the ORF6-hNEIL3 inserts were cloned into pETDuet-2 using standard procedures.
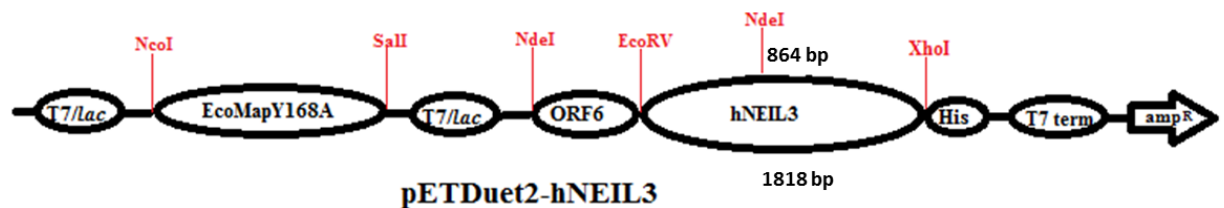


**Figure 4.1 Map of the pET30a and pETDuet-2 expression vectors containing human NEIL3. Adapted from (Liu *et al.,* 2012)**
 ORF6-hNEIL3 insert restriction map

Figure 4.1 shows ORF6 at the start of the insert with NdeI site used in the cloning, and an XhoI site at end of the insert followed by His-tag i also shows the multi cloning site of pETDuet vector accommodating EcoMapy168A, ORF6-hNEIL3

**CATATG**AAAATCGAAGCAGGTAAACTGGTACAG**AAGGAG**ATTAACTG**AT**G

NdeI

RBS

hNEIL3 Start codon
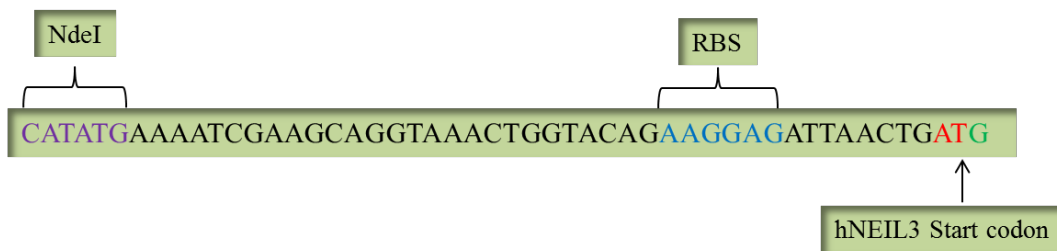
**Figure 4.2 Sequence of 0RF6-hNEIL3**

Figure 4.2 represents the junctions between NEIL3 cDNA sequences and the ORF6 sequence at the 5′ end, shows the NdeI site, the ORF6 and the red highlight (AT) shows the overlapping nucleotides at the end of ORF6 sequence and the start of hNEIL3.



**Figure 4.3 Illustrations of the hNEIL3 inserts after cloning into pETDuet-2.**

Vector map of pETDuet-2-hNEIL3 shows EcoMapY168A, the leading ORF6 sequence followed by hNEIL3 and the His tag downstream the C-terminal of hNEIL3.

Using this method, four out of the five hNEIL3 cDNAs were successfully cloned into pETDuet-2 vector indicating the positions of EcoMap, ORF6 and hNEIL3-(843, 1044, 1506 and full length) (Figure 4.3)., Cloning of pETDuet-2-hNEIL3-1236 was not successful and it was decided to work with the other four clones only. The successful clones were checked by restriction enzyme digests and confirmed by DNA sequencing. Of these, three out of the four constructs [hNEIL3-(843, 1044, and full length)] were successfully expressed in *E. coli* and the proteins purified by FPLC using His-Trap column for [hNEIL3-(843 and 1044)], utilizing the 6 histidine codon sequence (His-tag) following the C-terminal of hNEIL3 inserts in the pETDuet-2 expression vector, and His-Trap followed by Mono S columns for hNEIL3-FL (Figure 4.17-Figure 4.27). Expression and purification of the recombinant proteins was analysed by SDS-PAGE and western blotting and the purified proteins tested for activity using enzyme activity assays (Figure 4.28-Figure 4.50). The recombinant hNEIL3 proteins proved to be active on a variety of oxidised DNA bases in different contexts, including single-stranded DNA and at different positions in a model DNA replication fork.

## 4.2 Cloning

### 4.2.1 Cloning of hNEIL3 cDNA inserts into the pET30b-ORF6 vector

The pCMV6-AC-hNEIL3 plasmid DNA was first checked for the presence of hNEIL3 by restriction digestion with EcoRI and XhoI, before being subjected to PCR to amplify the target hNEIL3 cDNAs. The two enzymes cut in the multiple cloning sites (MCS) either side of the hNEIL3 cDNA insert in the vector.
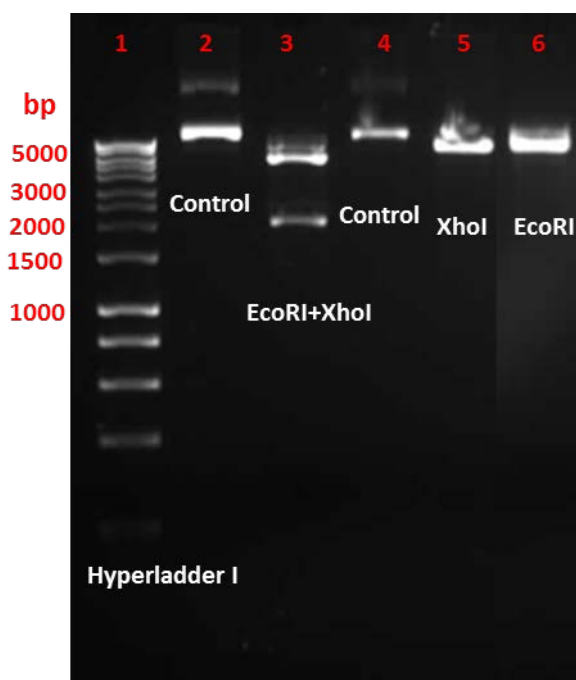


**Figure 4.4 Restriction digests of pCMV6-AC-hNEIL3 with EcoRI and XhoI.**
Lane 1, Hyperladder I; lane 2, uncut plasmid; lane 3, double-digest (EcoRI and XhoI); lane 4, uncut plasmid; lane 5, single digest with XhoI; lane 6, single digest with EcoRI.

In Figure 4.4, the single digest with either EcoRI or XhoI shows one band of about 5000 bp confirming that both enzymes cut the plasmid once and linearize it, while a double-digest reveals two bands, one around 2000 bp in length (hNEIL3 cDNA is 1818 bp) confirming the presence of hNEIL3 cDNA. The additional base pairs (approximately 182 bp) are due to the

positions where the two enzymes cut the vector in the MCS and the distance to the hNEIL3 sequence.

## 4.2.2. Amplification of hNEIL3 cDNAs

After confirming the presence of the hNEIL3 in pCMV6-AC, PCR reactions were carried out using the plasmid DNA as template to amplify the different hNEIL3 cDNAs. While a common forward primer (lacking the first AT nucleotides of the hNEIL3 ATG) was used for all PCR reactions, reverse primers containing a XhoI restriction site were designed according to the length of each hNEIL3 insert (Table 3.1).



**Figure 4.5 Amplification of full-length and truncated hNEIL3 cDNAs from pCMV6-AC-hNEIL3.**
Lane 1, Hyperladder I; lane 2, hNEIL3-843; lane 3, hNEIL3-1044; lane 4, hNEIL3-1506; lane 5, hNEIL3-1236; lane 6, hNEIL3-full-length.

The bands observed in Figure 4.5 indicate that all the PCR reactions were successful and gave rise to bands of the expected size. However, the full-length hNEIL3 cDNA is relatively faint compared to the other amplicons. Following purification, the PCR products were then ligated into the pET30b-ORF6 vector.

## 4.2.3. Preparation of hNEIL3 cDNAs and pET30b-ORF6 vector for ligation.

The hNEIL3 cDNAs amplified in the previous step were prepared for ligation to the ORF6 sequence in the previously constructed pET30b-ORF6 vector to create ORF6-hNEIL3 inserts that will then be cloned into the pETDuet-2 expression vector. To begin, pET30b-ORF6 was digested with EcoRV and XhoI restriction enzymes to yield compatible ends with the hNEIL3 cDNA inserts.
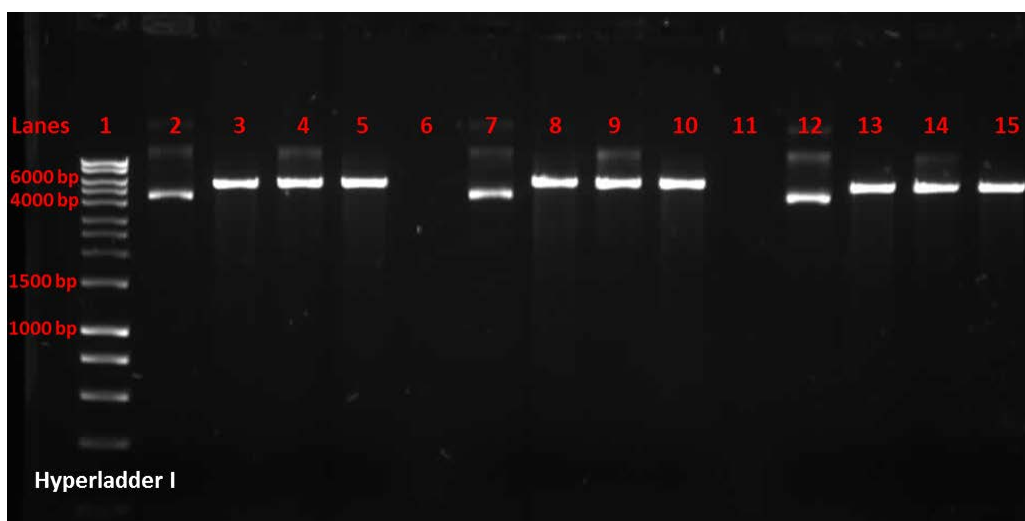


**Figure 4.6 Restriction digests of pET30b-ORF6 with XhoI and EcoRV.**
Lane 1, Hyperladder I; lane 2, uncut plasmid; lane 3, XhoI; lane 4, XhoI and EcoRV; lane 5, EcoRV; lane 7, uncut plasmid; lane 8, XhoI; lane 9, XhoI and EcoRV; lane 10, EcoRV; lane 12, uncut plasmid; lane 13, XhoI; lane 14, XhoI and EcoRV; lane 15, EcoRV.

As shown in Figure 4.6, pET30b-ORF6 was subject to single- and double-digests with XhoI and EcoRV. The vector was linearized creating a XhoI site overhang at the 3′ end and a blunt end at the 5′ end resulting from digestion with EcoRV. These sites are compatible with the 5′ blunt end in the hNEIL3 PCR products and a 3′ XhoI site overhang resulting from digestion of

the inserts with XhoI. The result shows the expected band size of linearized vector between 5000 and 6000 bp in lanes 3, 4, 5, 8, 9, 10, 13, 14, 15 (vector size 5421 bp).

Reverse primers for all hNEIL3 inserts were designed with a XhoI site to create an overhang compatible with the XhoI site of the ORF6 sequence in the vector (Table 3.1). All hNEIL3 PCR products were subject to restriction digestion with XhoI, however, the agarose gel picture is not shown as this digest will only reduce the length of the DNA by a few base pairs and therefore the difference cannot be visualized in an agarose gel.

Following preparation of the hNEIL3 inserts and the pET30b-ORF6 vector, the inserts were ligated with the vector using T4 DNA ligase at a 3:1 insert/vector ratio. The ligation mixtures were transformed into *E. coli* NovaBlue competent cells using the transformation method described in Section 3.1.63.1.9.

### 4.2.4. Testing *E. coli* colonies for successful ligation and transformation of hNEIL3 inserts into pET30b-ORF6

Following overnight growth of the transformed *E. coli* cells on LB-agar plates, a number of colonies were picked and grown in 5 ml LB (30 mg/ml kanamycin) culture overnight at 37°C. The DNA was extracted as described in Section 3.1.2 and the plasmid DNA was subjected to restriction digestion with NdeI to confirm the successful ligation of the inserts into the pET30b-ORF6 vector. NdeI cuts the pET30b-ORF6-hNEIL3 vector at the start of the ORF6 sequence and at position 864 bp in the hNEIL3 cDNA, thus in the presence of hNEIL3 inserts (except for hNEIL3-843), NdeI cuts both the vector and hNEIL3 cDNA resulting in a band of approximately 909 bp (864 bp of hNEIL3 + 45 of ORF6) (Figure 4.7). This reaction was used

as a confirmatory test for the ligation of the hNEIL3 cDNAs into pET30b-ORF6 except of course, for hNEIL3-843.
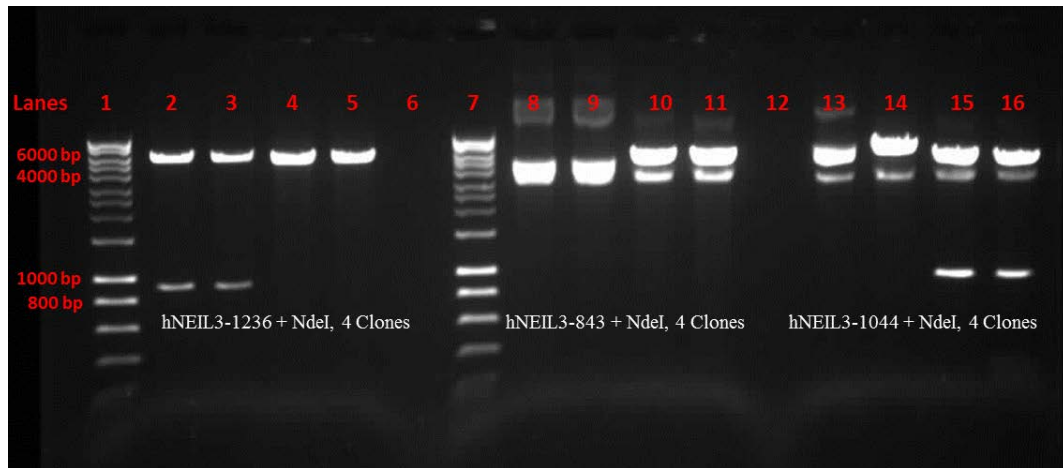


**Figure 4.7 Restriction digests of pET30b-ORF6-hNEIL3 (1236, 843, and 1044) with NdeI.**
Lane 1, Hyperladder I; lanes 2 - 5, hNEIL3-1236; lane 7, Hyperladder I; lanes 8 – 11, hNEIL3-843; lanes 13 - 16 hNEIL3-1044.

In Figure 4.7, lanes 2 and 3 contain plasmids with the hNEIL3-1236 insert as shown by the presence of the linearized vector between 5000 bp and 6000 bp (5421 bp) and a second band between 800 bp and 1000 bp. As NdeI cuts the hNEIL3 cDNA at 909 bp, this confirms the presence of hNEIL3-1236 in the vector. Conversely, lanes 4 and 5 show only one band at the size of linearized vector indicating the absence of the hNEIL3-1236 insert in the vector (negative colonies). Figure 4.7 also shows that in lanes 15 and 16 positive clones of hNEIL3-1044 were obtained, while those in lanes 13 and 14 lacked the hNEIL3 insert.

For pET30b-ORF6-hNEIL3-843, only one band, representing the linearized vector can be seen in lanes 8 and 9 in Figure 4.7, as NdeI cuts only at the start of the ORF6 sequence and not in the hNEIL3-843 insert. The double bands observed in lanes 10 and 11 in Figure 4.7 most likely represent incomplete digestion of the plasmid, something also observed in lanes 13 – 16.

Confirmation that the clones in lanes 8 – 11 contained hNEIL3-843 using a separate restriction enzyme digest is shown in Figure 4.8. Here, hNEIL3-843 was subject to a double-digest with NdeI and XhoI to release the ORF-hNEIL3-843 insert from the vector confirming the presence of the insert in both clones tested as shown in Figure 4.8 (lanes 3 and 5).
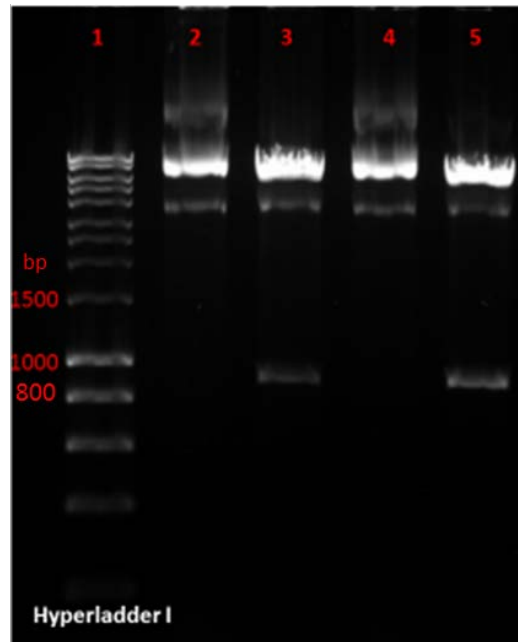


**Figure 4.8 Restriction digests of pET30b-ORF6-hNEIL3-843 with NdeI and XhoI.**
Lane 1, Hyperladder I; lanes 2 and 4, uncut plasmid; lanes 3 and 5, double-digest with XhoI and NdeI.

In Figure 4.8, the ORF6-hNEIL3-843 insert was released from both clones tested (lanes 3 and 5), and bands of the predicted size, between 800 bp and 1000 bp obtained (843 bp hNEIL3+47 ORF6), as NdeI cuts at the start of ORF6 and XhoI cuts the end of hNEIL3-843. The bands were cut from the agarose gel and purified as described in Section 3.1.83.1.8 for cloning into pETDuet-2.

Similarly, pET30b-ORF6-hNEIL3 (1506 and full length) was subjected to an NdeI restriction digest to confirm the presence of inserts in the picked clones.

**Figure 4.9 Restriction digests of pET30b-ORF6-hNEIL3 cDNAs (1506, full length) with NdeI: (C for Control, D for digest).**

Lane 1, Hyperladder I; lane 3, hNEIL3-FL clone 1 NdeI digest; lane 4, hNEIL3-FL clone 1 uncut; lane 6, hNEIL3-FL clone 2 uncut; lanes 5 and 7, hNEIL3-FL clone 2 NdeI digest; lane 9, hNEIL3-1506 clone 1 uncut; lane 10, hNEIL3-1506 clone 1 NdeI digest; lane 11, hNEIL3-1506 clone 2 uncut; lane 12, hNEIL3-1506 clone 2 NdeI digest; lane 13, hNEIL3-1506 clone 3 uncut; lane 14, hNEIL3-1506 clone 3 NdeI digest.

The results shown in Figure 4.9 indicate that hNEIL3-1506 was successfully ligated into the pET30b-ORF6 vector. Following restriction enzyme digestion with NdeI, the three clones of hNEIL3-1506 show an upper band at the size of the linearized vector and a lower band at position of the hNEIL3 cut by NdeI (909 bp), confirming the successful cloning of hNEIL3-1506 into the pET30b-ORF6 vector. The loss of the top band in lane 10 and 12 is due to efficient plasmid digestion in comparison to controls in lanes 9, 11 and 13 where the plasmid DNA samples were not fully digested by NdeI restriction enzyme.

For the full length hNEIL3 the two clones were negative for the presence of the insert as no band shows at the predicted position of the hNEIL3 cut by NdeI (Figure 4.9), indicating that

the ligation and transformation of full-length hNEIL3 into pET30b-ORF6 vector had not been successful. Therefore, the ligation and transformation experiment was repeated for pET30b-ORF6-hNEIL3-FL and the restriction digest analysis repeated. The result of the NdeI digest, shown in Figure 4.9, confirmed the successful ligation of hNEIL3-full length into pET30b-ORF6.
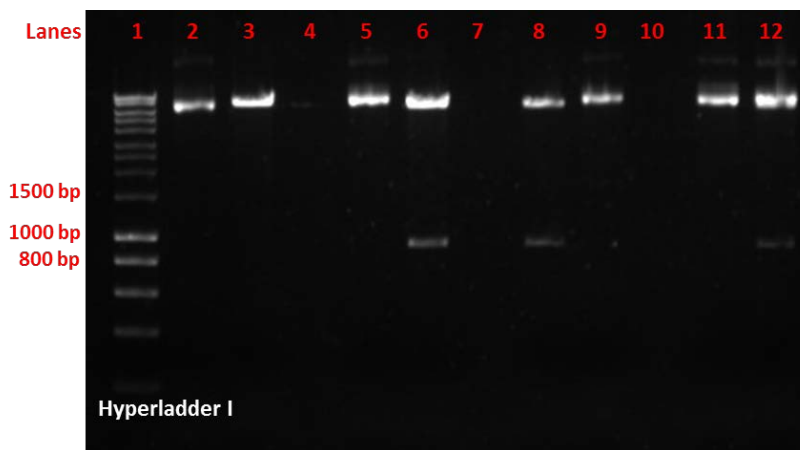


**Figure 4.10 Restriction digests of pET30b-ORF6-hNEIL3-FL with NdeI.**
Lane 1, Hyperladder I; lanes 2, 3, 5, 6, 8, 9, 11 and 12, NdeI digests.
Lanes 2, 3, 5, 9 and 11 show only one band of the linearized vector indicating that the insert was not successfully ligated with the vector. However, lanes 6, 8 and 12 show a linearized vector band size between 5000 bp and 6000 bp (5421 bp) and a lower band between 800 bp and 1000 bp where NdeI cuts hNEIL3 confirming the successful ligation of the insert with vector.

**4.2.5. Cloning of ORF6-hNEIL3 inserts into pETDuet-2.**

**4.2.6. Cloning of ORF6-hNEIL3-843 inserts into the pETDuet-2 expression vector.**

In order to prepare pETDuet-2 for ligation with ORF6-hNEIL3-843, it was subjected to restriction digestion with NdeI and XhoI. This created compatible ends for the ORF6-hNEIL3-843 insert to allow ligation into pETDuet-2. Ligation of the insert and vector and transformation of *E. coli* competent cells was carried out as described in Sections 3.1.6 and 3.1.9.

After transformation of *E. coli,* six colonies were picked and tested for the presence of the hNEIL3-843 insert. DNA extracted from the colonies was subject to double-digestion with NdeI and XhoI to confirm the successful cloning of hNEIL3-843 into pETDuet-2.
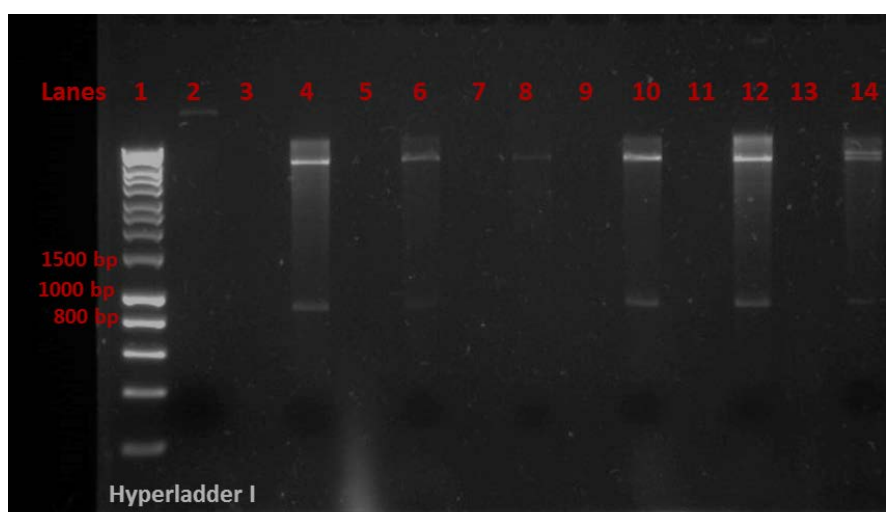


**Figure 4.11 Double digest of pETDuet-2-ORF6-hNEIL3-843 with NdeI and XhoI.**
Lane 1, Hyperladder I; lane 2, uncut plasmid; lanes 4, 6, 8, 10, 12 and 14, hNEIL3-843, no samples were loaded in lanes 5, 7,9,11 and 13.

In Figure 4.11, only the clone in lane 8 did not contain the cDNA insert, while the clones in lanes 4, 6, 10 and 12 were positive showing an upper band equivalent to the size of the linearized plasmid and lower band at the expected size of hNEIL3-843.

**4.1.7. Cloning of ORF6-hNEIL3-(1044, 1236, 1506 and full length) inserts into the pETDuet-2 expression vector.**

For the ORF6-hNEIL3-(1044, 1236, 1506 and FL) inserts, a different cloning strategy was designed. This consisted of amplifying each of the hNEIL3 cDNAs from pET30b-ORF6-hNEIL3 (Figure 4.12 andFigure 4.13) and then digesting the PCR product with XhoI to create a XhoI 3′-end overhang and a blunt end at the 5′-end of the cDNA to be compatible with the XhoI and blunt ends of pETDuet-2.

pETDuet-2 was first prepared using NdeI to linearize the vector and then blunting the 5′-end NdeI overhang with mung bean nuclease. Following clean-up, the linearized vector was then cut with XhoI. This resulted in linearized plasmid DNA with a blunt end at the 5′-end and a XhoI overhang at the 3′-end, Thereafter, ligation of the cDNA inserts to linearized pETDuet-2 was carried out to prepare pETDuet-2-ORF6-hNEIL3-(1044, 1236, 1506 and FL) constructs before transformation into *E. coli* competent cells.
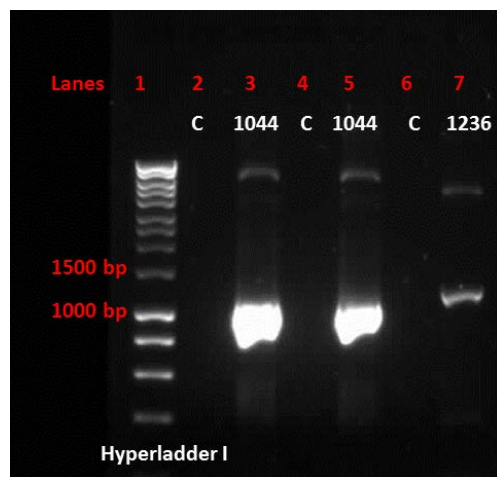


**Figure 4.12 Amplification of ORF6-hNEIL3 cDNAs 1044 and 1236.**
Lane 1, Hyperladder I; lanes 2, 4 and 6, control samples (no primers were added); lanes 3 and 5, ORF6-hNEIL3-1044; lane 7, ORF6-hNEIL3-1236.

Figure 4.12 shows that the ORF6-hNEIL3-1044 and ORF6-hNEIL3-1236 inserts were successfully amplified from pET30b-ORF6-hNEIL3, with bands of approximately the expected

71

size being obtained. However, the PCR product for ORF6-hNEIL3-1236 was much fainter than for the 1044 equivalent.

In order to construct the pETDuet-2-ORF6-hNEIL3-FL expression vector, PCR of ORF6-NEIL3-FL was carried out using different annealing temperatures (74ºC and 76ºC) in order to determine the optimal PCR conditions (Figure 4.13).
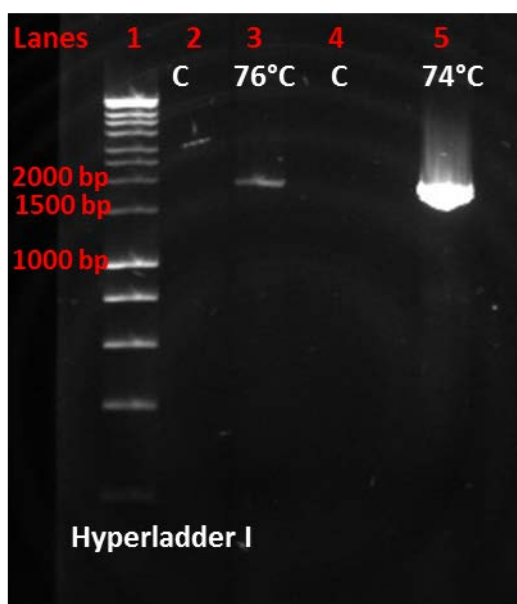


**Figure 4.13Amplification of ORF6-hNEIL3-FL by PCR at two annealing temperatures.**
Lane 1, Hyperladder; lanes 2 and 4, control samples (no primers added); lane 3, annealing temperature 76°C; lane 5, annealing temperature 74°C.

In Figure 4.13 it can be seen that ORF6-hNEIL3-FL was successfully amplified from the pET30b-ORF6-hNEIL3-FL plasmid DNA in both samples, although using an annealing temperature of 74°C resulted in more PCR product and this temperature was used for the subsequent experiments.

The ORF6-hNEIL3 PCR products (inserts) were subject to restriction digestion with XhoI to create compatible ends with pETDuet-2 as previously described. Then the inserts were purified using the PCR clean up method described Section 3.1.8 to be cloned into pETDuet-2.

### 4.1.8. Cloning ORF6-hNEIL3 inserts into pETDuet-2.

The original pETDuet-2-ORF6-MmNEIL3 vector was subjected to a double-digest with NdeI and XhoI to release the ORF6-MmNEIL3 DNA sequence from the expression vector and confirm that the correct plasmid had been obtained.
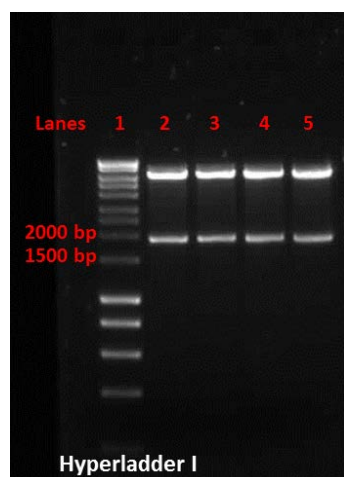


**Figure 4.14 Double-digest of pETDuet-2-ORF6-MmNEIL3 with NdeI and XhoI.**
Lane 1, Hyperladder I; lanes 2 - 5, pETDuet-2-ORF6-MmNEIL3 double-digest with NdeI and XhoI.

In Figure 4.14, ORF6-MmNEIL3 was released from all four clones of pETDuet-2-ORF6-MmNEIL3 revealing an upper band corresponding to the size of the vector and a lower band at the size of ORF6 and the NEIL3 coding sequence from *M. musculus* (1821 bp). Thus the ORF6-MmNEIL3 DNA was released from pETDuet-2 in all four clones.

In order to prepare pETDuet-2 for ligation with the ORF6-hNEIL3 sequences, the vector was first digested with NdeI. The restriction enzyme reaction mixture was incubated at 37°C for 90 min, and then the NdeI enzyme was inactivated by placing the reaction tube at 65°C for 20 min. The reaction mix was then incubated with mung bean nuclease to blunt the NdeI overhang and the DNA was purified using a PCR-Clean Up kit (Section 3.1.7 and 3.1.8 ).

After purification, pETDuet-2 was digested with XhoI and all of the reaction mix was loaded onto a 0.8% agarose gel and subjected to electrophoresis to separate the ORF6-MmNEIL3 DNA from the pETDuet-2 vector (Figure 4.15). The bands were visualized on a UV-transilluminator and the upper band corresponding to the pETDuet-2 plasmid cut from the gel, purified and subsequently used for ligation with hNEIL3 inserts.
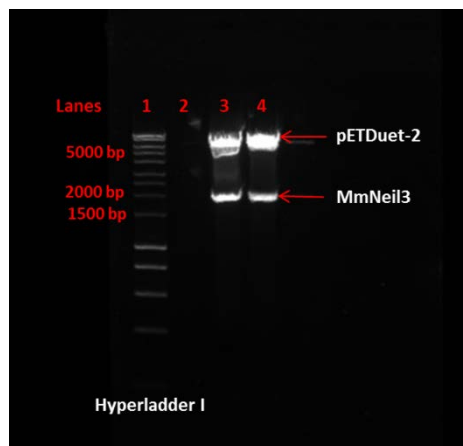


**Figure 4.15 pETDuet-2-ORF6-MmNEIL3 cut with (NdeI + Mung Bean Nuclease + XhoI)**
Lane 1, Hyperladder I; lanes 3 and 4, pETDuet-2-ORF6-MmNEIL3 cut with (NdeI + mung bean nuclease + XhoI).

In lanes 3 and 4 of Figure 4.15, an upper band corresponding to pETDuet-2 and a lower band at the size of the ORF6-MmNEIL3 are evident, confirming the release of ORF6-MmNEIL3 from the plasmid.

T4 DNA ligase protocol was used as described in Section 3.1.6 to achieve a 3:1 insert/vector ratio to ligate the ORF6-hNEIL3 inserts into the pETDuet-2 expression vector. The ligation mixtures were transformed into *E. coli* NovaBlue competent cells using the transformation method described in Section 3.1.9, Colonies were picked for overnight cultures then tested for the presence of the inserts in the expression vector by restriction enzyme digest with NdeI (because NdeI cuts the ORF6-hNEIL3 insert at position of 909 bp (Figure 4.16 A, B and C) revealing a band at 909 bp on the agarose gel).
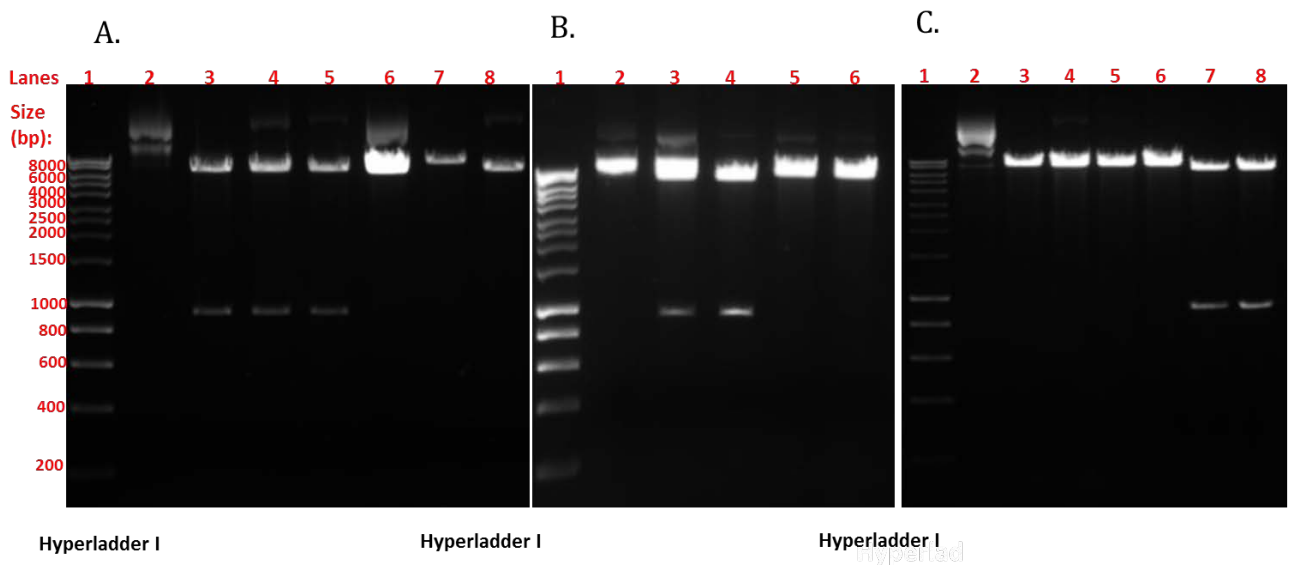


**Figure 4.16 Restriction Digests of A. pETDuet-2-ORF6-hNEIL3-1044, B. pETDuet-2-ORF6-hNEIL3-1506 and C. pETDuet-2-ORF6-hNEIL3-FL with NdeI.**
A. Lane1, Hyperladder I; lane 2, uncut plasmid cDNA; lanes 3 - 8, pETDuet-2-ORF6-hNEIL3-1044 with NdeI. B. Lane1, Hyperladder I; lane 2, uncut plasmid; lanes 3 - 6 pETDuet-2-ORF6-hNEIL3-1506 clones. C. Lane 1, Hyperladder I; lane 2, uncut plasmid cDNA; lanes 3 - 8 pETDuet-2-ORF6-hNEIL3-FL digested with NdeI.

Figure 4.16 A shows positive clones in lanes 3 - 5 indicated by the presence of a band around 900 bp where NdeI cuts the ORF6-hNEIL3 insert, while the clones in lanes 6 - 8 were negative. Figure 4.16 B shows that the clones in lanes 3 and 4 are positive for the presence of the insert with an upper band corresponding to the linearized vector and a lower band between 800 bp and 1000 bp indicating digestion of hNEIL3 by NdeI, while the clones in lanes 5 and 6 do not

show any insert present.**Error! Reference source not found.** Figure 4.16 C shows that clones in lanes 7 and 8 were positive for the insert DNA, indicated by the presence of a band about 900 bp, while the clones in lanes 3 - 6 did not contain the correct insert.

Following these experiments, lab stock were made of all the positive clones of the pETDuet-2-ORF6-hNEIL3 cDNAs (843, 1044, 1506 and FL) and stored in -80ºC for future use. Unfortunately, despite several attempts, the ORF-hNEIL3-1236 could not be cloned into pETDuet-2. The reasons for this are not clear.

**4.1.9. DNA Sequencing of the ORF6-hNEIL3 inserts in pETDuet-2.**

Plasmid DNA of the pETDuet-2-ORF6-hNEIL3-(843, 1044, 1506 and FL) clones were sent to

Eurofins for sequencing using forward and reverse primers for each clone. The results obtained

confirmed the successful cloning of all constructs, (Sequences data presented in Appendix).

## 4.3   Protein Expression

### 4.3.1   Protein expression test results

Protein expression was first tested for each hNEIL3 truncation by growing 30 ml cultures under different conditions and proceeding with 1 ml aliquots for analysis by SDS-PAGE and western blotting. As the results show (Figure 4.17, Figure 4.19, Figure 4.21 andFigure 4.23) all hNEIL3 protein expression tests were successful and also confirmed the co-expression of the EcoMap gene (32 kDa), while western blotting with an His-Tag antibody confirmed the expression of hNEIL3 his-tagged proteins revealing bands at the predicted protein sizes of: hNEIL3-(843) 31.2 kDa, hNEIL3-(1044) 38.6 kDa, hNEIL3-(1506) 55.7 kDa and hNEIL3-full length 67.3 kDa.

To determine the optimum conditions for protein over-expression, the small bacterial cultures were induced with IPTG and incubated at different temperatures for increasing lengths of time. Figure 4.17 shows the SDS-PAGE and western blot analyses of such an experiment with pETDuet-2-hNEIL3-843, where bacterial cultures were induced and incubated at either 16°C or 37°C for 3 h and 18 h. At 37°C no appreciable difference in expression was observed following incubation for 3 h or 18 h (Figure 4.17B, lanes 3 - 4), however, at 16°C, significant recombinant protein expression was only observed after incubation for 18 h (Figure 4.17B, lanes 6 - 7).
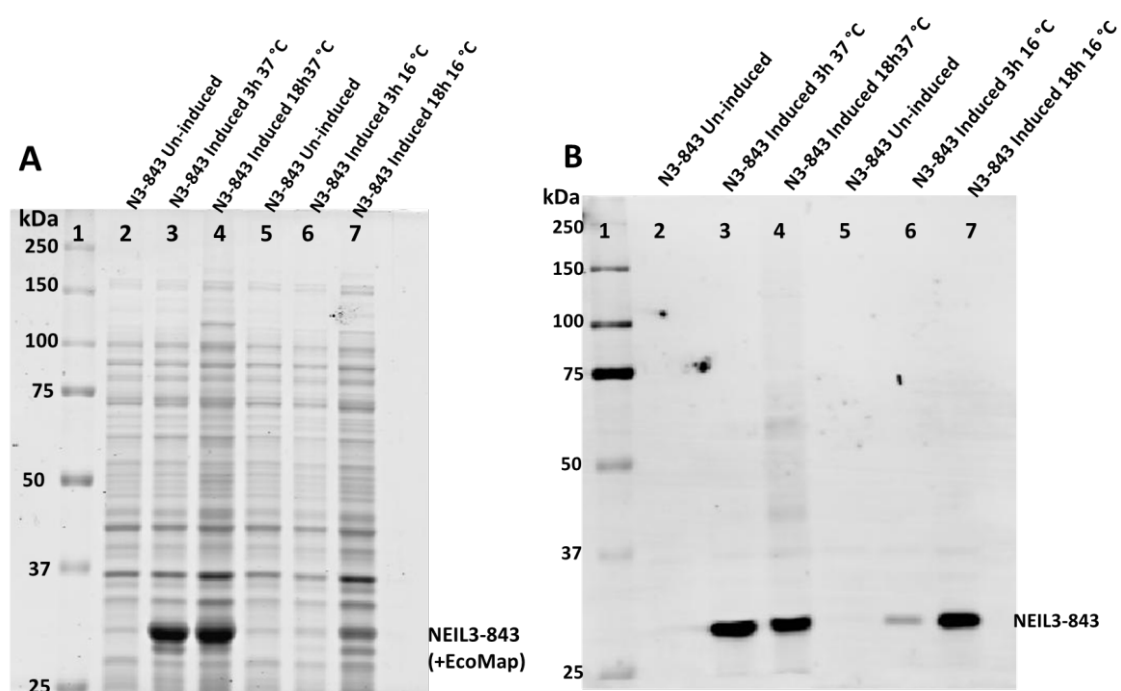
**Figure 4.17 SDS-PAGE and western blot analyses of pETDuet-2-hNEIL3-843 test protein expression.**

(**A**) SDS-PAGE gel, and (**B**) western blot of hNEIL3-843. Lane 1, protein size marker; lane 2, hNEIL3-843 uninduced; lane 3, hNEIL3-843 induced for 3 h at 37°C; lane 4, hNEIL3-843 induced for 18 h at 37°C; lane 5, hNEIL3-843 uninduced; lane 6, hNEIL3-843 induced for 3 h at 16°C; lane 7, hNEIL3-843 induced for 18 h at 16°C.

In summary, analysis of Figure 4.17A shows that a band of approximately size of hNEIL3-843 (31.2 kDa) is present in most of the induced samples (lanes 3, 4 and 7). However, as the molecular weight of the co-expressed EcoMap is 32 kDa, it is not possible to tell by SDS-PAGE alone, if one or both proteins have been expressed. However, in Figure 4.17B, the western blot clearly identifies the His-tagged NEIL3-843 with His-tag antibodies, thus confirming the expression of hNEIL3-843 in *E. coli* cells. The expression of both proteins is more clearly observed in later Figures (*e.g.*Figure 4.19), where the molecular weight of the co-expressed proteins is significantly different.

Following these test results, a 400 ml culture was induced with 1 mM IPTG and incubated at 16°C for 18 h and the protein extracted and subjected to His-Trap column purification. The results are given in Figure 4.18, which shows both the SDS-PAGE and western blotting of protein fractions following His-Trap purification of hNEIL3-843.
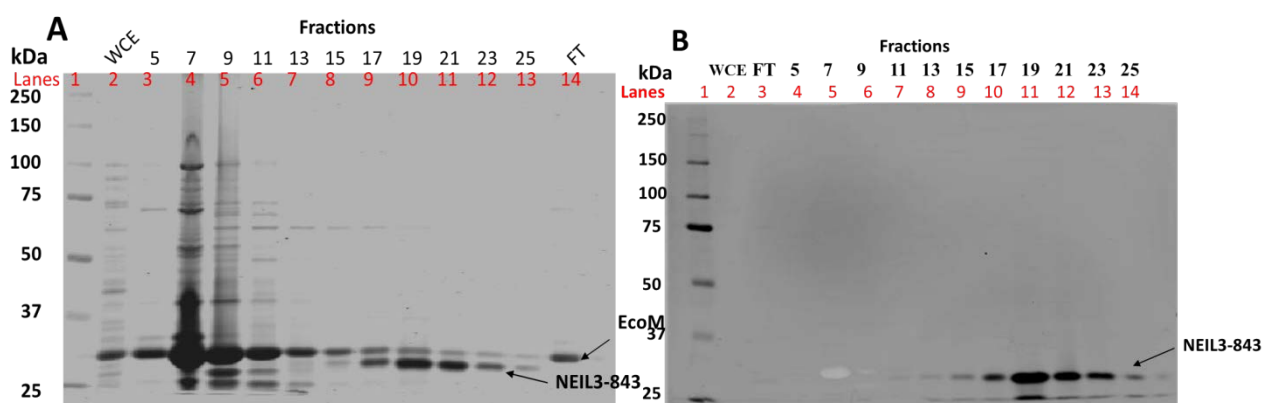


**Figure 4.18 SDS-PAGE (A) and western blotting (B), of column fractions following His-Trap FPLC purification of hNEIL3-843.**
(**A**) SDS-PAGE of column fractions, lane 1, protein size marker; lane 2, whole cell extract (WCE), lanes 3 – 13, column fractions 11 - 25; lane 14, flow-through (FT). (**B**) Western blot of column fractions. Lane 1, protein size marker; lane 2, whole cell extract; lane 3, flow-through; lanes 4 – 14, column fractions 5 - 25.

Figure 4.18 confirms the successful expression and purification of hNEIL3-843 protein. Most of the *E. coli* proteins and EcoMap were eluted between fractions 7 and 9 while the target protein, hNEIL3-843) was eluted between fractions 17 and 23. Bands of the size of the target protein can be clearly seen after SDS-PAGE in Figure 4.18A and were confirmed to be the His-tagged hNEIL3-843 by western blotting as shown in Figure 4.18B. also show smaller size bands, suggested being degradation product of the recombinant protein.

Following the successful expression and purification of hNEIL3-843, similar experiments were undertaken with hNEIL3-1044 and the results of the test expression are given in Figure 4.19 The results are similar to those obtained for hNEIL3-843 and again an incubation temperature of 16°C and incubation period of 18 h was chosen.
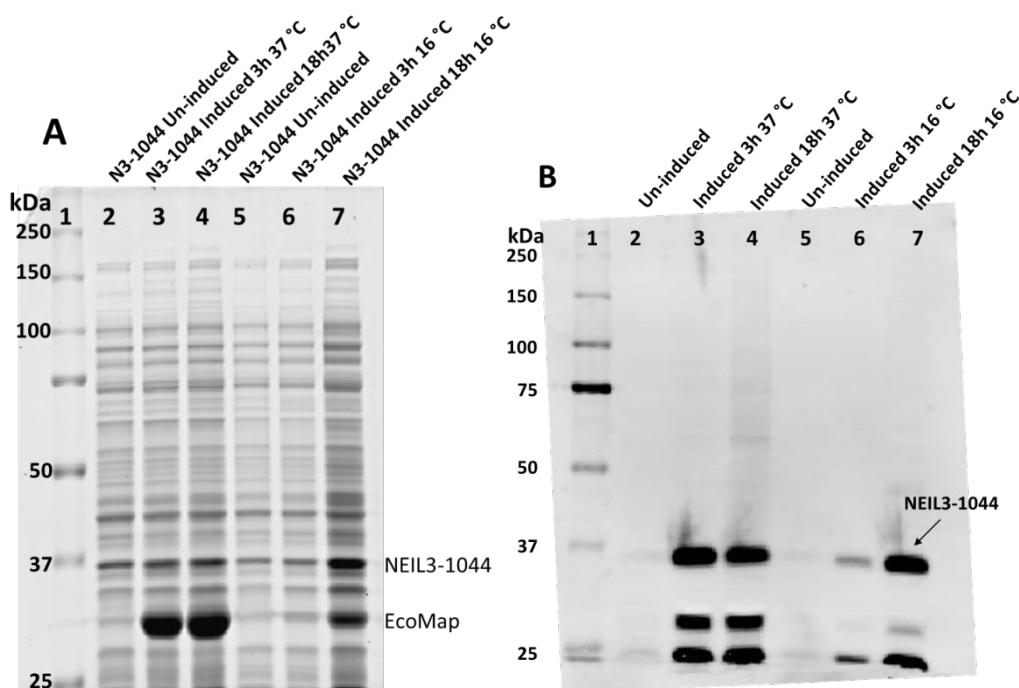


**Figure 4.19 SDS-PAGE (A) and western blot (B) analyses of pETDuet2-hNEIL3-1044 test protein expression.**
Lane 1, protein size marker; lane 2, hNEIL3-1044 uninduced; lane 3, hNEIL3-1044 induced for 3 h at 37°C; lane 4, NEIL3-1044 induced for 18 h at 37°C; lane 5, hNEIL3-1044 uninduced; lane 6, hNEIL3-1044 induced for 3 h at 16°C; lane 7, hNEIL3-1044 induced for 18 h at 16°C.

Following SDS-PAGE, Figure 4.19A shows two strong staining polypeptide bands in the induced samples; the first around 37 kDa, equivalent to the size of hNEIL3-1044, and the second approximately 32 kDa, equivalent to the size of EcoMap. The western blot with a His-Tag antibody in Figure 4.19B confirms the identity of the larger band as hNEIL3-1044, although other smaller bands are observed, especially in the samples incubated at 37°C. These bands are much reduced when expression was carried out at 16°C (lanes 6 – 7) and are not

visible in the equivalent western blot for hNEIL3-843 (Figure 4.17B). They are also not present following purification of hNEIL3-1044 (Figure 4.20). Therefore, it may be concluded that protein degradation had occurred in these particular samples and that this is not a general problem encountered with expression of this recombinant hNEIL3 truncated version.
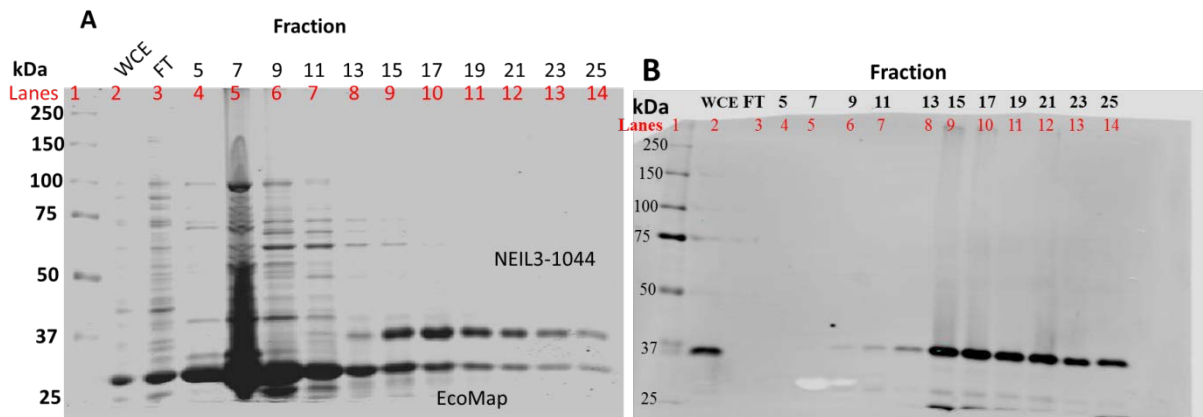


**Figure 4.20 SDS-PAGE (A) and western blot (B) of column fractions following His-Trap FPLC purification of hNEIL3-1044.**
(**A**) SDS-PAGE gel, (**B**) western blot of purified hNEIL3-1044 protein. Lane 1, protein size marker; lane 2, hNEIL3-1044 whole cell extract; lane 3, flow through; lanes 4 - 14, hNEIL3-1044 fractions 5 - 25.

SDS-PAGE and western blotting of the His-Trap protein purification of hNEIL3-1044 are shown in Figure 4.20A and Figure 4.20B. Both confirm the expression and purification of the target hNEIL3-1044 protein, with most of the recombinant protein eluting in fractions 15 - 21. Similar to Figure 4.18 show smaller size bands, suggested being degradation products of the recombinant protein.

Following on from the previous experiments, Figure 4.21 shows the SDS-PAGE and western blot of pETDuet-2-hNEIL3-1506 test protein expression. Although there is no clear abundant band at the predicted size of hNEIL3-1506 after SDS-PAGE at either incubation temperature

(Figure 4.21A), the western blot of the same samples shows clear bands at the predicted size of 55 kDa for hNEIL3-1506, indicating a low level of recombinant protein expression. However, good induction of EcoMap expression is observed in the samples incubated at both 37°C (lanes 3 and 4), and after 18 h at 16°C (lane 7), indicating that recombinant protein expression was achieved as expected. The lower level of expression, compared with the smaller hNEIL3 proteins, is also indicated by the relatively faint bands obtained by western blotting when compared with the results of the hNEIL3-(843 and 1044) expression. The western blot also indicates the presence of an extra low molecular weight band that is smaller than the hNEIL3-1506 protein. As this band is observed both in the test expression and after protein purification (**Error! Reference source not found.**), it must be concluded that this represents a protein degradation product that contains the C-terminus plus His-Tag part of the recombinant protein.

**Figure 4.21 SDS-PAGE (A) and western blot (B) analyses of pETDuet-2-hNEIL3-1506 test protein expression.**

**(A)** SDS-PAGE and **(B)** western blot. Lane 1, protein size marker; lane 2, hNEIL3-1506 uninduced; lane 3, hEIL3-1506 induced for 3 h at 37°C; lane 4, hNEIL3-1506 induced for 18 h at 37°C; lane 5, hNEIL3-1506 uninduced; lane 6, hNEIL3-1506 induced for 3 h at 16°C; lane 7, hNEIL3-1506 induced for 18 h at 16°C.
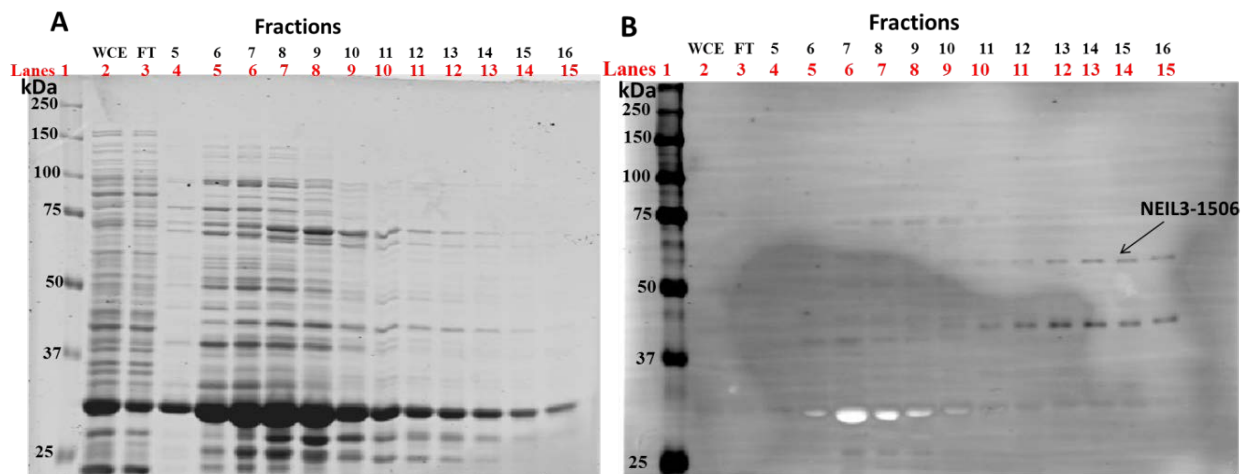


**Figure 4.22 SDS-PAGE (A) and western blot (B) of column fractions following His-Trap FPLC purification of hNEIL3-1506.**

Lane 1, protein size marker; lane 2, hNEIL3-1506 whole cell extract; lane 3, flow through; lanes 4 - 15 hNEIL3-1506 fractions 5 - 16.

Analysis of the SDS-PAGE in Figure 4.22A shows no specific band at the size of the target protein (55 kDa) while in Figure 4.22B, the western blot of the purified protein fractions, bands at the size of the target protein can be observed in fractions 12 – 16. Smaller bands are also visible in Figure 4.22B, which most likely indicate a degradation product of the target protein (hNEIL3-1506). The bands at the predicted protein size are relatively faint due to the low expression levels and therefore, the amount of the purified protein obtained.
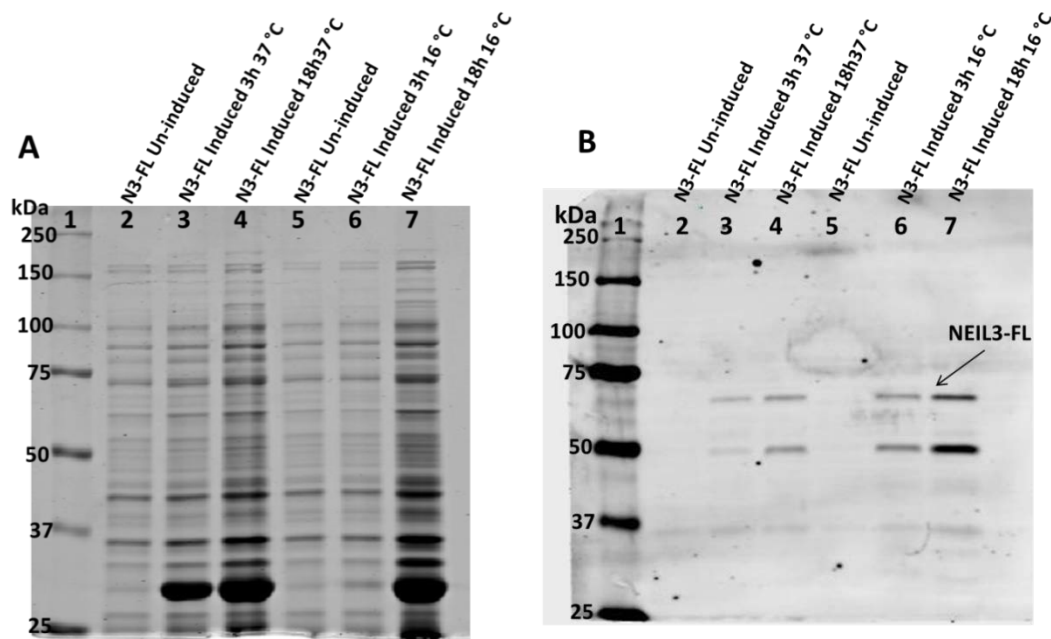


**Figure 4.23 SDS-PAGE (A) and western blot (B) analyses of pETDuet-2-hNEIL3-FL test protein expression.**

SDS-PAGE **(A)** and western blot **(B)**. Lane 1, protein size marker; lane 2 hNEIL3-FL un-induced; lane 3, hNEIL3-FL induced for 3 h at 37°C; lane 4, hNEIL3-FL induced for 18 h at 37°C; lane 5, hNEIL3-FL un-induced; lane 6, hNEIL3-FL induced for 3 h at 16°C; lane 7, hNEIL3-FL induced for 18 h at 16°C.

Figure 4.23A again shows no specific bands at the size of the hNEIL3-FL protein in induced samples but again shows high expression of EcoMap (lanes 3, 4 and 7). In Figure 4.23B, the western blot of the hNEIL3-FLwith His-tag antibodies shows two bands in the induced samples, one the approximate size of the target protein (67.3 kDa) and one at about 50 kDa. As both of

these proteins are detected by the anti:His-tag antibody, it can be assumed that the 50 kDa band represents a proteolytic degradation product of hNEIL3-FL that contains the C-terminal domains, but lacks the N-terminal DNA glycosylase active site.



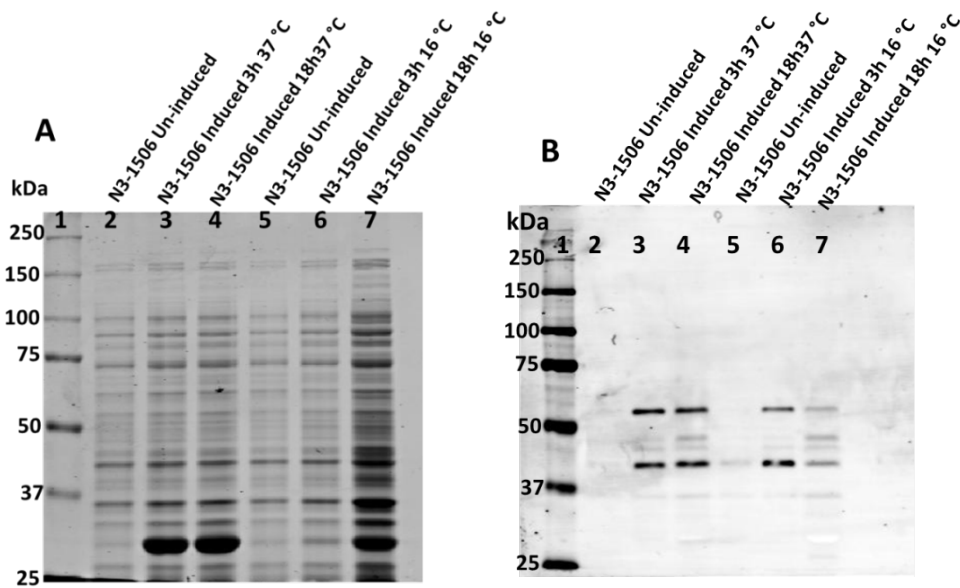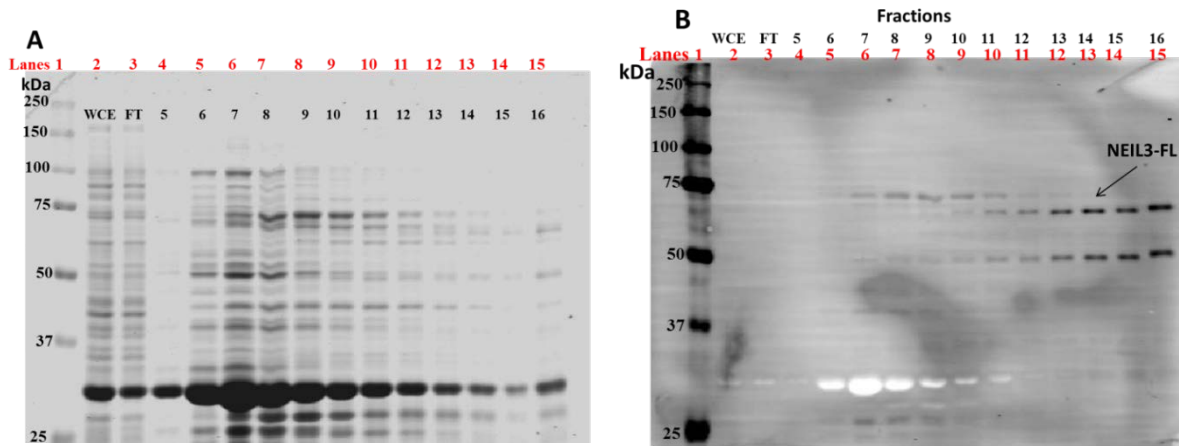**Figure 4.24 SDS-PAGE (A) and western blot (B) of fractions following His-Trap FPLC purification of hNEIL3-FL.**
Lane 1, protein size marker; lane 2, hNEIL3-FL whole cell extract; lane 3, flow through; lanes 4 – 15, hNEIL3-FL fractions 5 - 16.

**Error! Reference source not found.** and Figure 4.24 show the result of hNEIL3-FL test expression and His-Trap purification. Similar to the results observed for hNEIL3-1506, the absence of a significant band at the size of the target protein in SDS-PAGE is due to the low amount of hNEIL3-FL protein obtained after induction, while the western blot pictures show bands at the size of the target protein confirming the expression and purification of hNEIL3-FL. The western blot also shows that the suspected C-terminal degradation product co-elutes with hNEIL3-FL and thus is not separated by the His-Trap purification method, further confirming its identity, at least as a His-tagged protein.

## 4.3.2 Protein purification of hNEIL3-FL for activity assays

While 400 ml cultures were enough to yield useable quantities of hNEIL3-843 and hNEIL3-1044, to obtain sufficient amounts of the hNEIL3-FL protein, a much larger culture of pETDuet-2-ORF6-hNEIL3-FL (1.2 L) was induced with 1 mM IPTG and incubated for 18 h at 16°C. The bacterial cells were lysed and the soluble supernatant prepared as described in Section 3.2.2. Following His-Trap purification, the fractions were analysed by SDS-PAGE and western blotting and the results are shown in Figure 4.25A and B. This shows bands of the expected size for hNEIL3-FL and the 50 kDa putative degradation product of the hNEIL3-FL protein co-eluting principally in fractions 11 – 23.
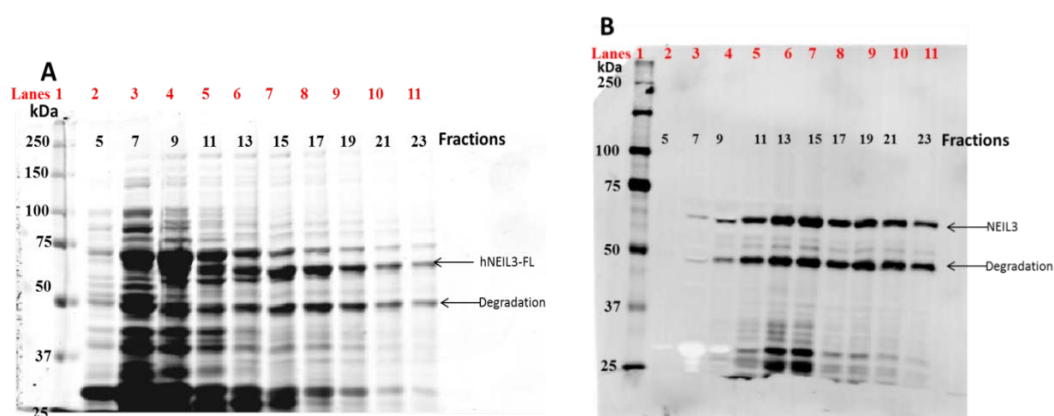


**Figure 4.25 SDS-PAGE (A) and western blot (B) of fractions following His-Trap FPLC purification of hNEIL3-FL.**
Lane 1, protein size marker; lanes 2 – 11, hNEIL3-FL fractions 5 - 23.

The partially purified protein was then subjected to Mono-S ion exchange chromatography for further purification and to reduce any bacterial protein contamination. The resulting fractions analysed by SDS-PAGE and western blotting using an anti-His antibody (Figure 4.26A and B). Figure 4.26A and B show clear bands at the expected size of hNEIL3, while the bands at the size of the degradation protein products appeared to be weaker in comparison with hNEIL3, indicating a preferential purification of intact hNEIL3-FL.
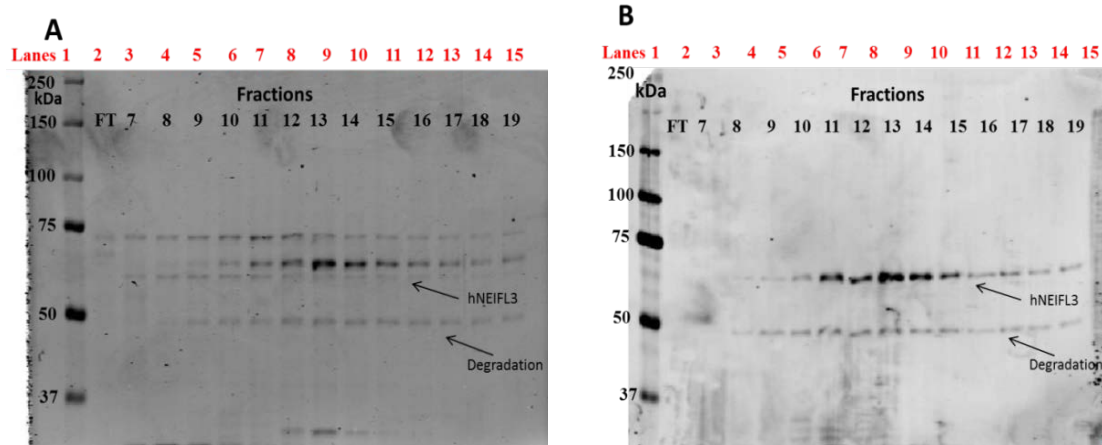
**Figure 4.26 SDS-PAGE (A) and western blot (B) of fractions following Mono S FPLC purification of hNEIL3-FL.**

**A** and **B**: Lane 1, protein size marker; lane 2, flow through; lanes 2 to 15, fractions 7 - 19.

The western blot membrane in Figure 4.26B was stored at -20°C and later incubated with an anti-NEIL3 antibody. The result, given in Figure 4.27, confirmed that the bands observed using the anti:His tag antibody were due to hNEIL3 protein. However, the western blot does not show any preferential purification of the degradation product from hNEIL3-FL and instead shows a very consistent pattern across all the fractions. This must be due to the specificity of the polyclonal antibody (Section 3.2.7) to the hNEIL3 polypeptide.
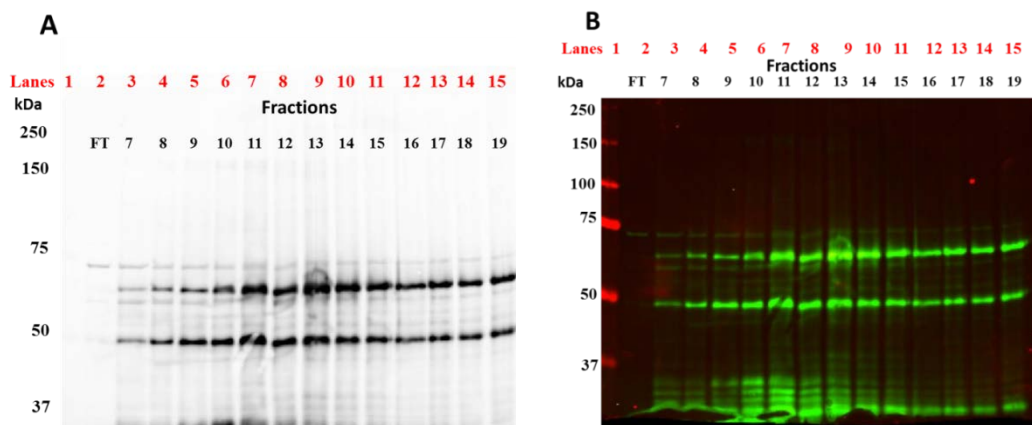
**Figure 4.27 Western blot (using anti:NEIL3 antibody) of fractions following Mono S FPLC purification of hNEIL3-FL.**

**A** and **B**: one membrane scanned with two different contrast/channels. Lane 1, protein size marker; lane 2 flow through; lanes 2 - 15, hNEIL3-FL fractions 7 - 19.

Fractions containing hNEIL3-FL were pooled and concentrated using an Amicon centrifugation filter with a 30,000 MW cut off (Section 3.2.4). These were then used for DNA glycosylase activity assays.

## 4.4 DNA Glycosylase Activity Assays

### 4.4.1 hNEIL3 purity and stability determination

The DNA glycosylase activity of the purified recombinant hNEIL3 proteins was tested by incubating the proteins with a variety of 5′-fluorescently tagged oligonucleotide substrates containing different site-specific oxidised bases for 20 min at 30ºC. The reaction mixtures were then loaded onto 10% denaturing polyacrylamide/urea gels for electrophoresis and visualised with a fluorescent scanner. The DNA glycosylase activity of the hNEIL3 proteins were tested single- and double-stranded oligonucleotides containing a single oxidised base (5-hydroxyuracil (5-OHU), thymine glycol (TG) or 7,8-dihydro-8-oxoguanine (8-oxoG)). These oxidised bases were also placed at one of three different positions in a model DNA replication fork and the activity of the different hNEIL3 proteins analysed.

To confirm that the observed activity was from the recombinant hNEIL3 proteins and not contaminating bacterial DNA glycosylases that had co-purified, a sodium borohydride trapping experiment was undertaken. This technique forms a covalent linkage between the DNA glycosylase and the fluorescently labelled substrate that can be separated by SDS-PAGE. This achieved by the generation of a Schiff base intermediate product, generated through the associated AP- lyase of a bifunctional DNA glycosylase, which is reduced by $NaBH_4$ to form a covalently bound DNA-protein complex. Thus, the molecular weight of the active proteins can be determined. Figure 4.28 shows the result of one such experiment where hNEIL3-843 and hNEIL3-FL were incubated with a single-stranded oligonucleotide containing 5-OHU. One band, corresponding to the size of hNEIL3-843 protein-DNA complex (31.2 kDa + the molecular weight of the cleaved DNA substrate and 67.3 kDa + molecular weight of the cleaved substrate) can be seen in lanes 3 and 5 respectively. This confirms that the DNA glycosylase

activity observed in the activity assays is due to the hNEIL3 proteins and not to contamination by the Fpg or Nth proteins of *E. coli*.
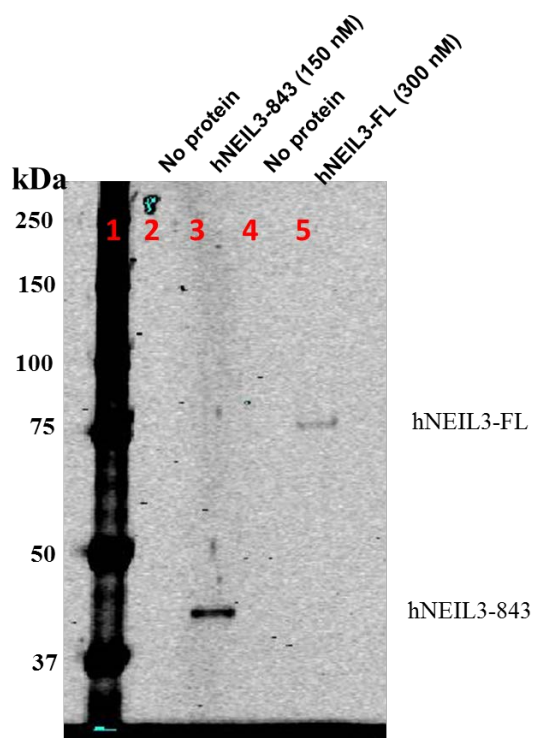


**Figure 4.28    NaBH₄-trapping of hNEIL3-843 and hNEIL3-FL on a single-stranded oligonucleotide containing a single 5-OHU lesion.**
Lane 1, Molecular weight marker; lane 2, blank; lane 3, hNEIL3-843; lane 4, blank; lane 5, hNEIL3-FL.

As a predominant degradation product had been observed during protein purification of hNEIL3-FL, it was decided to test the stability of hNEIL3-FL under the reaction conditions used. Therefore, the protein was incubated in reaction buffer at 30°C for increasing lengths of time, up to 30 min. Figure 4.29 shows that, surprisingly, only one band corresponding to the size of hNEIL3-FL is present in each lane and that there is no discernible degradation in any of the samples at any of the time points, indicating that the hNEIL3 protein is stable under the

reaction conditions used. The result also indicates that the degradation product observed throughout the protein purification procedure was removed by the final Amicon concentration step while preparing the Mono S fractions for enzyme assays.
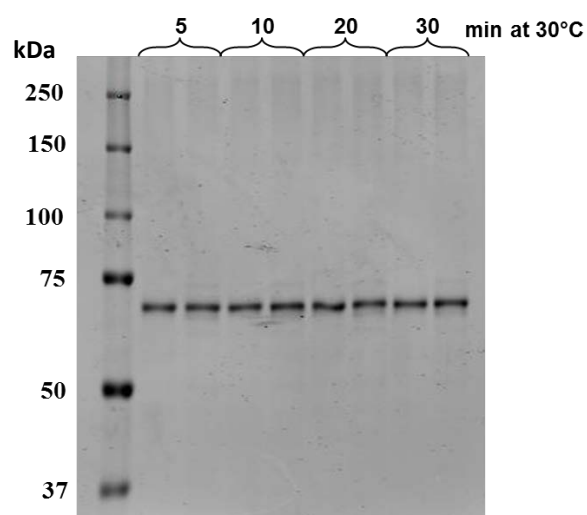


**Figure 4.29  SDS-PAGE of hNEIL3-FL following incubation at 30°C for increasing times in DNA glycosylase reaction buffer.**
Duplicate samples of hNEIL3-FL (150 nM) were mixed with the activity assay reaction buffer with no oligonucleotides and the samples incubated for 5, 10, 20, and 30 min at 30°C.

### 4.4.2 Oligonucleotide incision assays.

To start the biochemical analysis of the recombinant hNEIL3 proteins, it was decided to test their DNA glycosylase/lyase activity on ssDNA substrates containing an oxidised pyrimidine base at a known position within the sequence, as the murine NEIL3 homologue is known to be active on such substrates (Liu et al., 2012). In Figure 4.30, increasing concentrations of hNEIL3-843 were incubated at 30ºC or 37ºC for 20 min with a 39 mer ssDNA oligonucleotide substrate containing 5-OHU at position 20 to determine optimal reaction conditions for subsequent experiments. The results in Figure 4.30 show that incision of the oligonucleotide depended on the presence of hNEIL3-843, there being no product band in the control lanes. At 37ºC cleavage of the substrate appears to be hNEIL3-843 concentration dependent, with the highest level of product observed with 600 nM of hNEIL3-843 (lane 5). In contrast, at 30ºC there was maximal incision of the substrate oligonucleotide when 300 nM hNEIL3-843 was used (lane 3). Based on these results it was decided to use 300 nM hNEIL3-843 and 30ºC for subsequent activity assays.
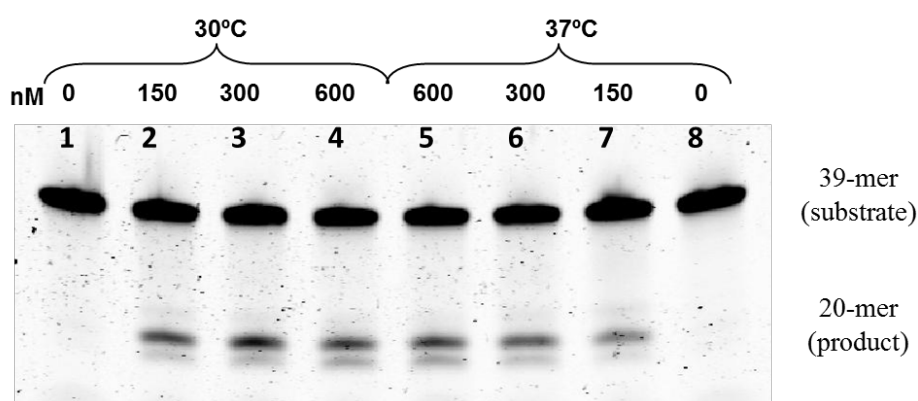


**Figure 4.30  DNA glycosylase/lyase activity of hNEIL3-843 on 5-OHU in ssDNA at 30ºC and 37ºC.**
Samples were incubated with increasing concentrations of hNEIL3-843 (0 - 600 nM) for 20 min at 30ºC (lanes 1 - 4) or 37ºC (lanes 5 – 8).

As the incised product in Figure 4.30 is only a small proportion of the amount of substrate on the gel and a previous report has found MmNEIL3 to be a monofunctional DNA glycosylase (Krokeide et al., 2013) the experiment in Figure 4.30 was repeated at 30°C with the addition of NaOH to half the reaction mixtures following the enzyme incubation. Alkali are known to cause hydrolysis of abasic sites and therefore the experiment was designed to test if hNEIL3-843 was acting primarily as a monofunctional DNA glycosylase, *i.e.* without the additional AP-lyase activity of a bifunctional DNA glycosylase. The result show in Figure 4.31 shows that there was little difference in the amount of product obtained after the addition of NaOH to the samples, indicating that hNEIL3-843 acts as a bifunctional DNA glycosylase under these reaction conditions. However, Figure 4.31 does show that the addition of NaOH increases the migration of the product band. This indicates that hNEIL3-843 acts primarily through β-elimination (leaving a 3′ α,β-unsaturated aldehyde) on these substrates while the samples where NaOH was added produced a product equivalent to β-δ elimination (a phosphate group at the 3′-end).
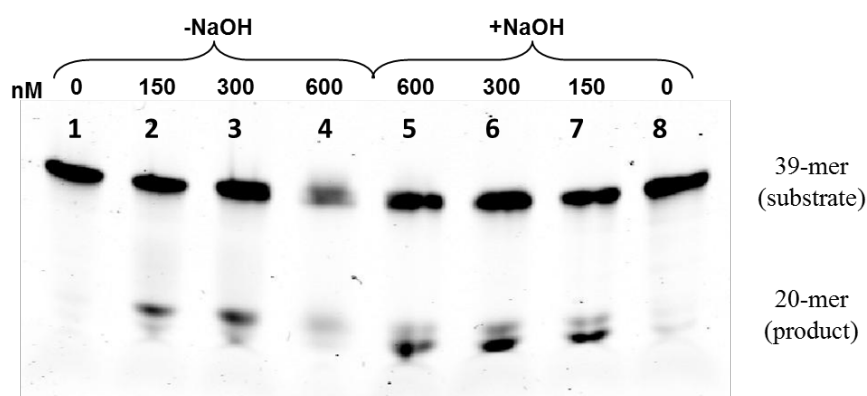


**Figure 4.31 DNA glycosylase / lyase activity of hNEIL3-843 with on 5-OHU in ssDNA +/-NaOH.** The reaction mixtures were incubated for 20 min at 30ºC. Following incubation, 1 µl of 1 M NaOH was added to half the samples (lanes 5 - 8) and heated to 95°C for 3 min and then the reaction stopped by the addition of formamide / bromophenol loading buffer.

In order to determine the influence of the RanBP zinc finger domain on the activity of hNEIL3 (Figure 4.33) an oligonucleotide containing 5-OHU was incubated at 30°C with increasing concentrations of either hNEIL3-843 or hNEIL3-1044. The result shown in Figure 4.32 indicates that both proteins are active on the substrate and that the migration of the product is the same, suggesting a β-elimination product. Thus, the addition of the RanBP zinc finger domain to the DNA glycosylase domain does not affect the enzyme activity in this assay.
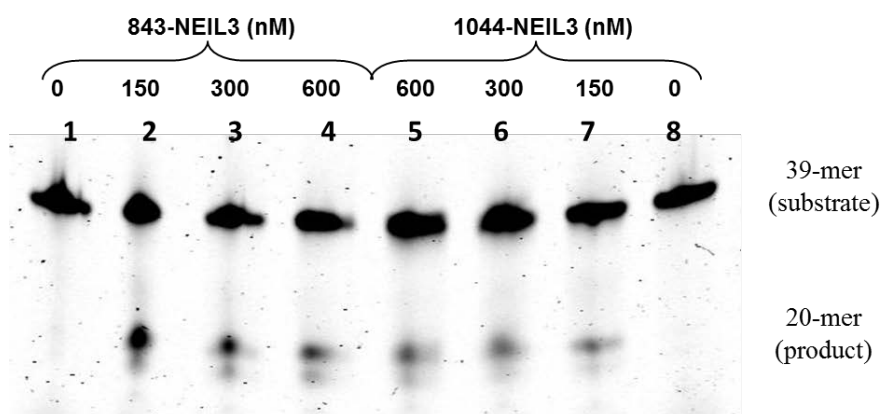


**Figure 4.32 DNA glycosylase / lyase activity of hNEIL3-843 and hNEIL3-1044 on 5-OHU in ssDNA.**
Reactions were incubated for 20 min at 30°C with increasing amounts of hNEIL3-843 (lanes 1 – 4) or hNEIL3-1044 (lanes 5 – 8) and the reaction was stopped by the addition of formamide / bromophenol blue loading buffer.

In a similar manner, the DNA glycosylase / lyase activity of hNEIL3-FL was determined (Figure 4.33). Increasing concentrations of the purified enzyme were incubated with the 5-OHU containing oligonucleotide, with and without the subsequent addition of NaOH Figure 4.33. The results indicate that hNEIL3-FL is a bifunctional DNA glycosylase that acts through β-elimination. Comparison of this result with that for hNEIL3-843 (Figure 4.31) shows an almost identical result and indicates that the additional C-terminal domains on the full-length hNEIL3 appear to have no effect on the DNA glycosylase activity of hNEIL3 on this substrate.
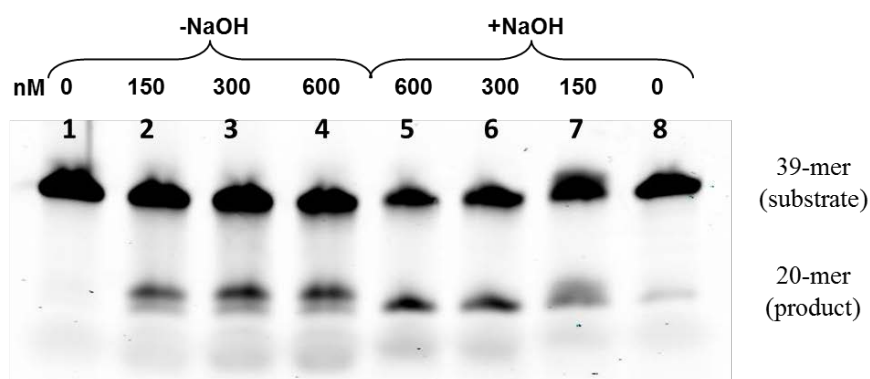
**Figure 4.33 DNA glycosylase / lyase activity of hNEIL3-FL on 5-OHU in ssDNA +/- NaOH.**
Reactions were incubated for 20 min at 30°C with increasing concentrations of hNEIL3- FL.
One microlitre of 1 M NaOH was added to samples 5 - 8 at the end of the reaction and incubated
at 95°C for 3 min before the addition of formamide / bromophenol blue loading buffer.

Next, the DNA glycosylase / lyase activity of hNEIL3-843 on ssDNA substrates containing
different oxidised bases was analysed (Figure 4.34). The results are shown at two different
settings of the scanner, as the 5-OHU containing oligonucleotide was labelled with Alexa Fluor
680 which fluoresces more intensely under the conditions used for the other substrate
oligonucleotides that were labelled with IRDye700. The results show that hNEIL3-843 is active
on ssDNA substrates containing 5-OHU or TG, producing a 20 mer product and has a weak
activity on 8-oxoG, but no activity on the uracil containing substrate (Figure 4.34). Similar
results were obtained with hNEIL3-FL (Figure 4.35), where again only weak activity was
observed on 8-oxoG and no activity was evident when uracil was present in the substrate
oligonucleotide. This again suggests the C-terminal domains had no effect on the DNA
glycosylase / lyase activity of hNEIL3 on these substrates.
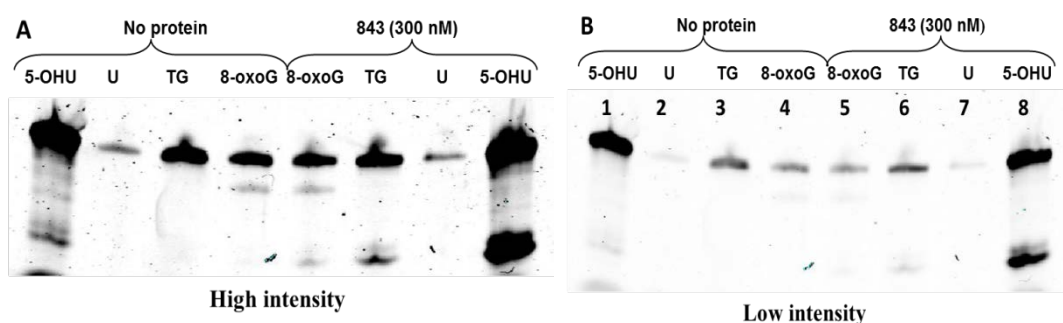
**Figure 4.34 DNA glycosylase / lyase activity of hNEIL3-843 on single-stranded oligonucleotide substrates containing different oxidised bases.**

hNEIL3-843 was incubated for 20 min at 30ºC with ssDNA substrates containing either 5-OHU, U, TG or 8-oxoG.
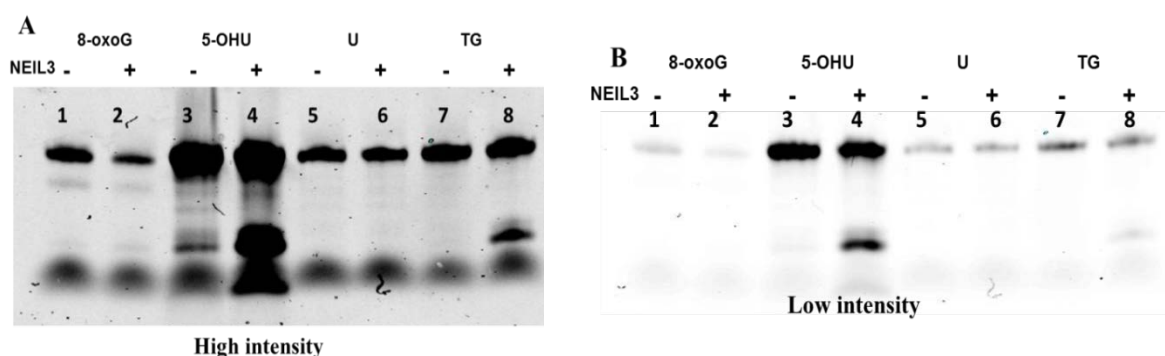


**Figure 4.35  DNA glycosylase / lyase activity of hNEIL3-FL on single-stranded oligonucleotide substrates containing different oxidised bases.**

hNEIL3-FL was incubated for 20 min at 30ºC with ssDNA substrates containing either 5-OHU, U, TG or 8-oxoG.

To confirm the results in Figure 4.35, the experiments were repeated with the addition of the bacterial DNA glycosylases Nei and Nth as controls and the results quantified (Figure 4.36 and Figure 4.37). Figure 4.36 shows that hNEIL3-FL is active on a ssDNA substrate containing 5-OHU processing 8.4% of the substrate (Figure 4.36A, B) and TG processing 14.3% of the substrate (Figure 4.36C and D), but shows only a weak activity on 8-oxoG processing 3.6% of the substrate (Figure 4.37A, B) and no activity on uracil with only background levels of

fluorescence obtained (Figure 4.37C, D). While detectable levels of incision were obtained for

Nei on the 5-OHU and TG substrates, no activity was observed with either Nei or Nth on uracil

or 8-oxoG (Figure 4.37). This agrees with results reported in the literature (Hazra *et al.,* 2000)

that these bacterial proteins have no, or only minimal activity on ssDNA. Of course, this result

gives further proof that the observed activity is a result of hNEIL3 action and not from

contaminating *E. coli* proteins.
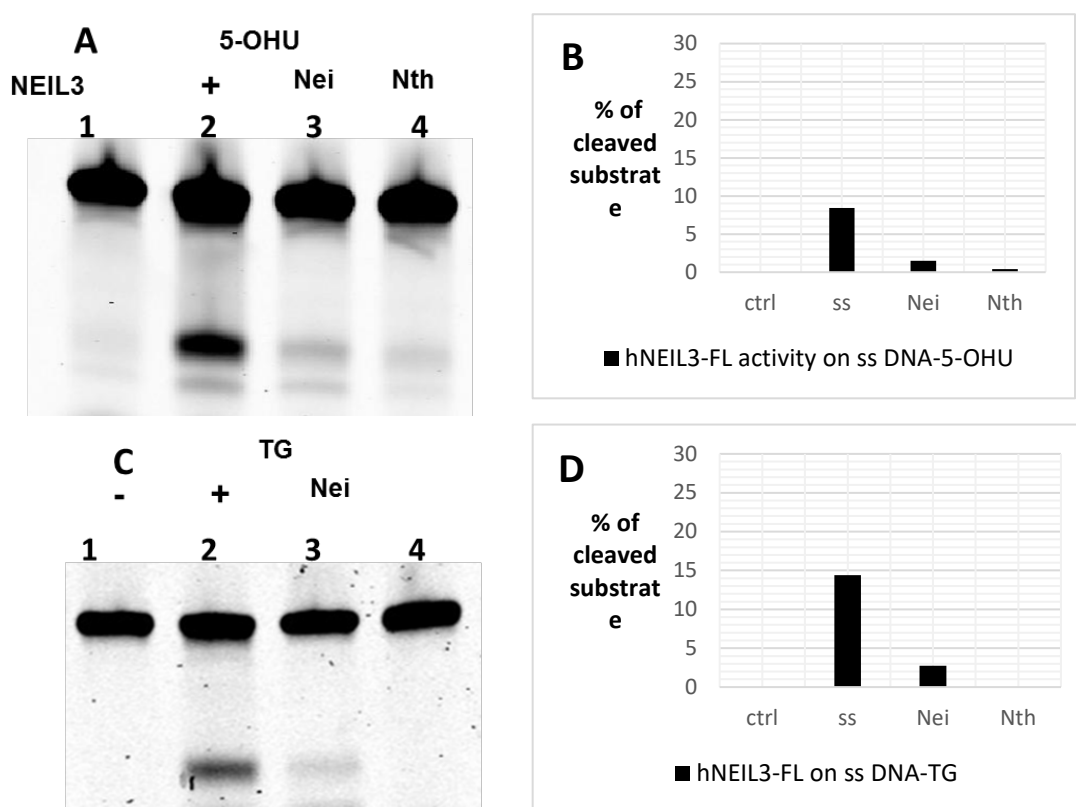


**Figure 4.36 DNA glycosylase / lyase activity of hNEIL3-FL, Nei and Nth on ssDNA substrates containing 5-OHU (A, B) and TG (C, D).**

Reactions were incubated for 20 min at 30ºC with hNEIL3-FL (150 nM), Nei (15 nM) or Nth (15 nM).

**Figure 4.37 DNA glycosylase / lyase activity of hNEIL3-FL, Nei and Nth on ssDNA substrates containing 8-oxoG (A, B) or uracil (C, D).**

Reactions were incubated for 20 min at 30°C with hNEIL3-FL (150 nM), Nei (15 nM) or Nth

(15 nM).

To compare the activity of hNEIL3 on single-stranded and double-stranded DNA, hNEIL3-843

and hNEIL3-FL were incubated for increasing times with either 5-OHU in ssDNA or 5-OHU

base paired with G in a double-stranded oligonucleotide (Figure 4.38). The result confirms that

both proteins are active on ssDNA substrates containing 5-OHU but show little or no activity

on 5-OHU in dsDNA. Furthermore, increasing the incubation time beyond 5 min did not result

in any increase in product, suggesting that either the substrate was limiting, that the enzyme

was irreversibly bound to the substrate or product, or that the enzyme had lost activity. In Figure

4.38B, no activity was observed for either protein on dsDNA substrates containing a 5-OHU:G base pair.
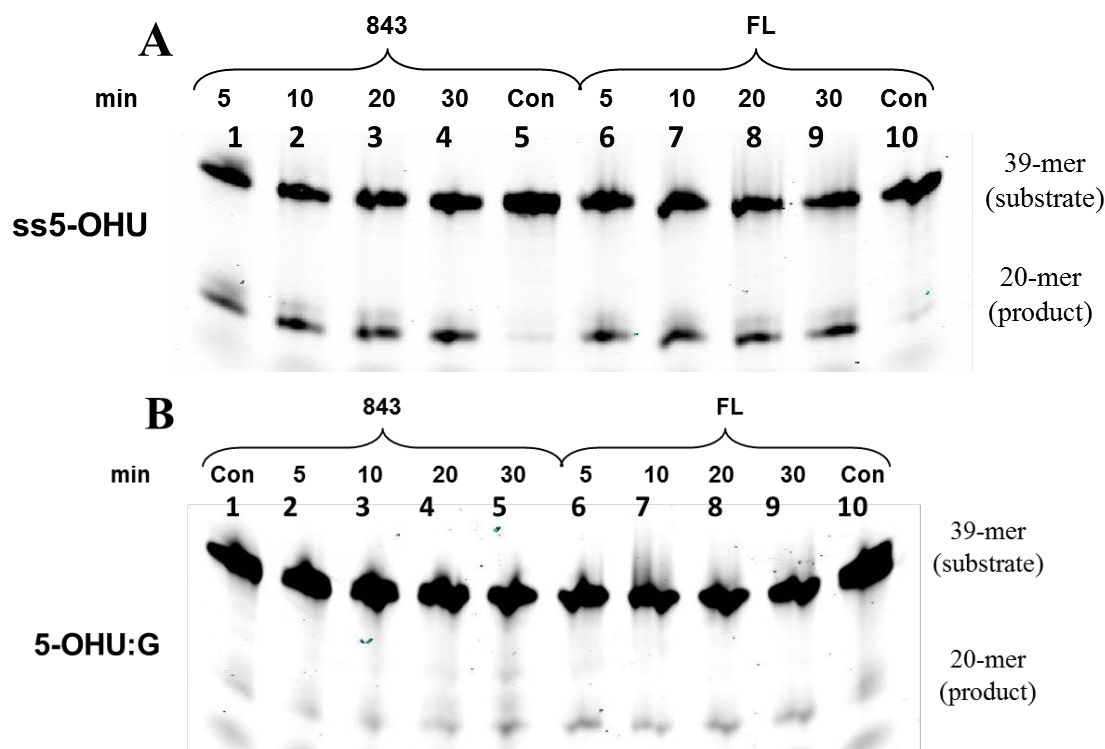


**Figure 4.38 Time course activity assay of hNEIL3-843 and hNEIL3-FL on ssDNA containing 5-OHU (A) and dsDNA containing 5-OHU:G (B).**
Reactions were incubated for 5, 10, 20 and 30 min at 30°C with hNEIL3-843 (lanes 1 – 5) or hNEIL3-FL (lanes 6 – 10).

Because the DNA glycosylase/lyase activity of hNEIL3 did not increase after 5 min incubation, a time-course experiment was undertaken with a much shorter time-frame. Figure 4.40 shows a time-course experiment with a shorter time scale (0, 5, 30, 60, 300 s) using hNEIL3-FL with ssDNA and dsDNA substrates containing 5-OHU and 5-OHU:G respectively. The result shows increasing activity on the ssDNA substrate up to 300 s (5% of the substrate has been incised, (Figure 4.39A, B), and no or very little activity on the dsDNA substrate (2.2% of the substrate

incised, (Figure 4.39C, D), A similar result was obtained when TG was substituted for 5-OHU (Figure 4.40), with significant incision only observed after 300 s for the ssDNA substrate (5% of the substrate incised, (Figure 4.40A, B) and only a minor activity on TG in a dsDNA context (1.4% of the substrate incised, (Figure 4.40B),
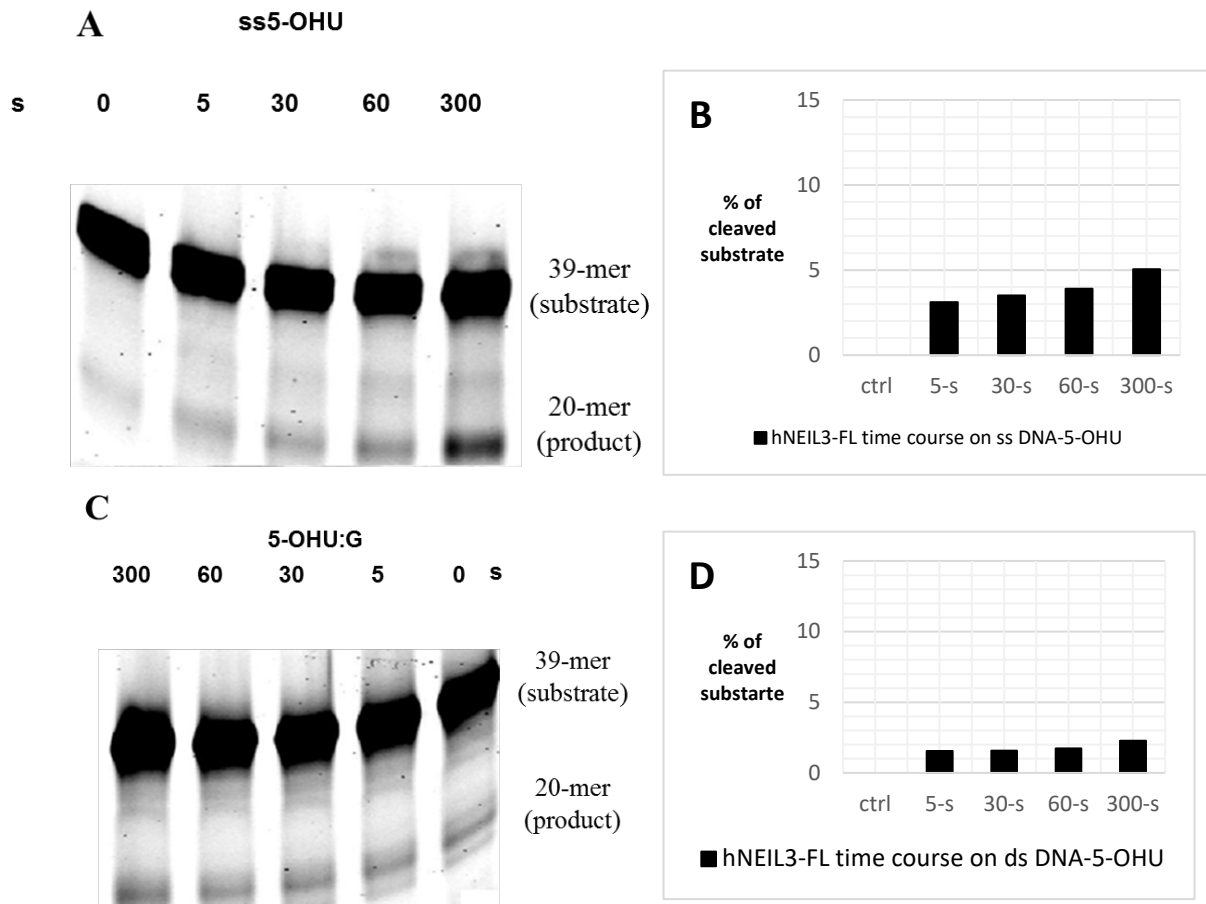


**Figure 4.39 Time course activity assay and quantitative analysis of hNEIL3-FL on (A) ssDNA substrate containing 5-OHU and (B) dsDNA substrate containing 5-OHU:G.**
Single and double-stranded oligonucleotides containing 5-OHU and 5-OHU:G respectively were incubated for 0, 5, 30, 60 and 300 seconds at 30ºC with hNEIL3-FL.
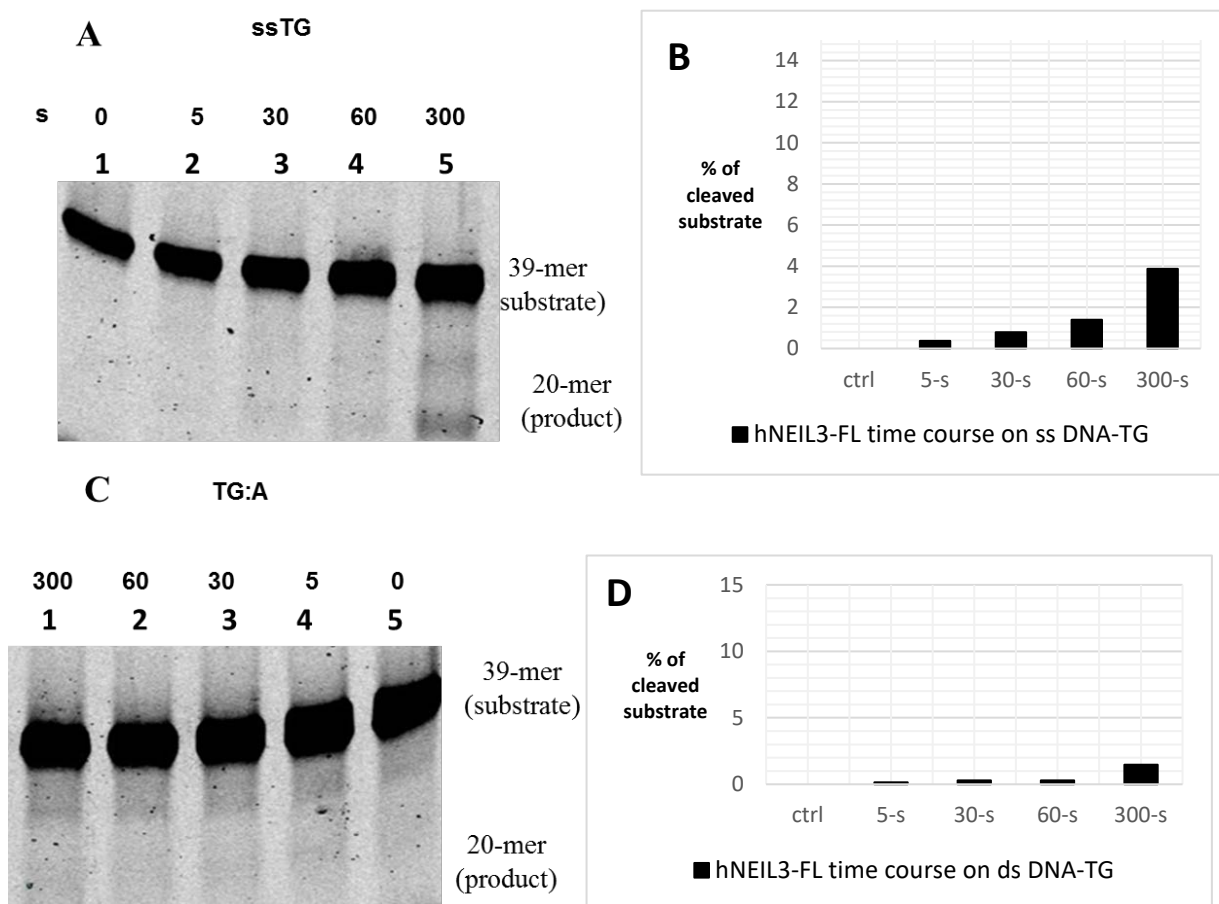
**Figure 4.40 Time course activity assay and quantitative analysis of hNEIL3-FL on (A) ssDNA substrate containing TG and (B) dsDNA substrate containing TG:A.**

Single and double-stranded oligonucleotides containing TG and TG:A respectively were incubated for 0, 5, 30, 60 and 300 seconds at 30ºC with hNEIL3-FL.

In order to confirm the activity assays on dsDNA substrates and the mode of action of the bifunctional DNA glycosylase activity of hNEIL3, the experiments were repeated with the addition of the bacterial Nei (which acts through β/δ-elimination) and Nth (which acts through β-elimination) and the results are shown in Figure 4.41. Figure 4.41A confirms a weak activity of hNEIL3-FL on 5-OHU when base-paired with G (Figure 4.41B) shows that 2.9% of the substrate has been incised). Interestingly, the mode of action is shown to be β/δ-elimination, as the product band migrates to the same position as that of Nei (β/δ-elimination) and further than that of the product of the Nth (β-elimination) reaction. β/δ-elimination generates a 5′ phosphate which migrates further in the gel when compared to the 5′ PUA generated through β-elimination of the phosphodiester backbone.

Weak activity was also observed for 8-oxoG when base-paired with C (1.6% of the substrate has been incised, (Figure 4.41D). Again the product migrates in line with the Nei product, but as Nth is not active on the substrate it is not possible to definitively assign a β/δ-elimination mechanism to this reaction, although it does seem likely. Finally, hNEIL3-FL showed no activity on TG when base-paired with A (0.7% of the substrate has been incised, (Figure 4.41F), while the bacterial control proteins both showed good activity on this substrate
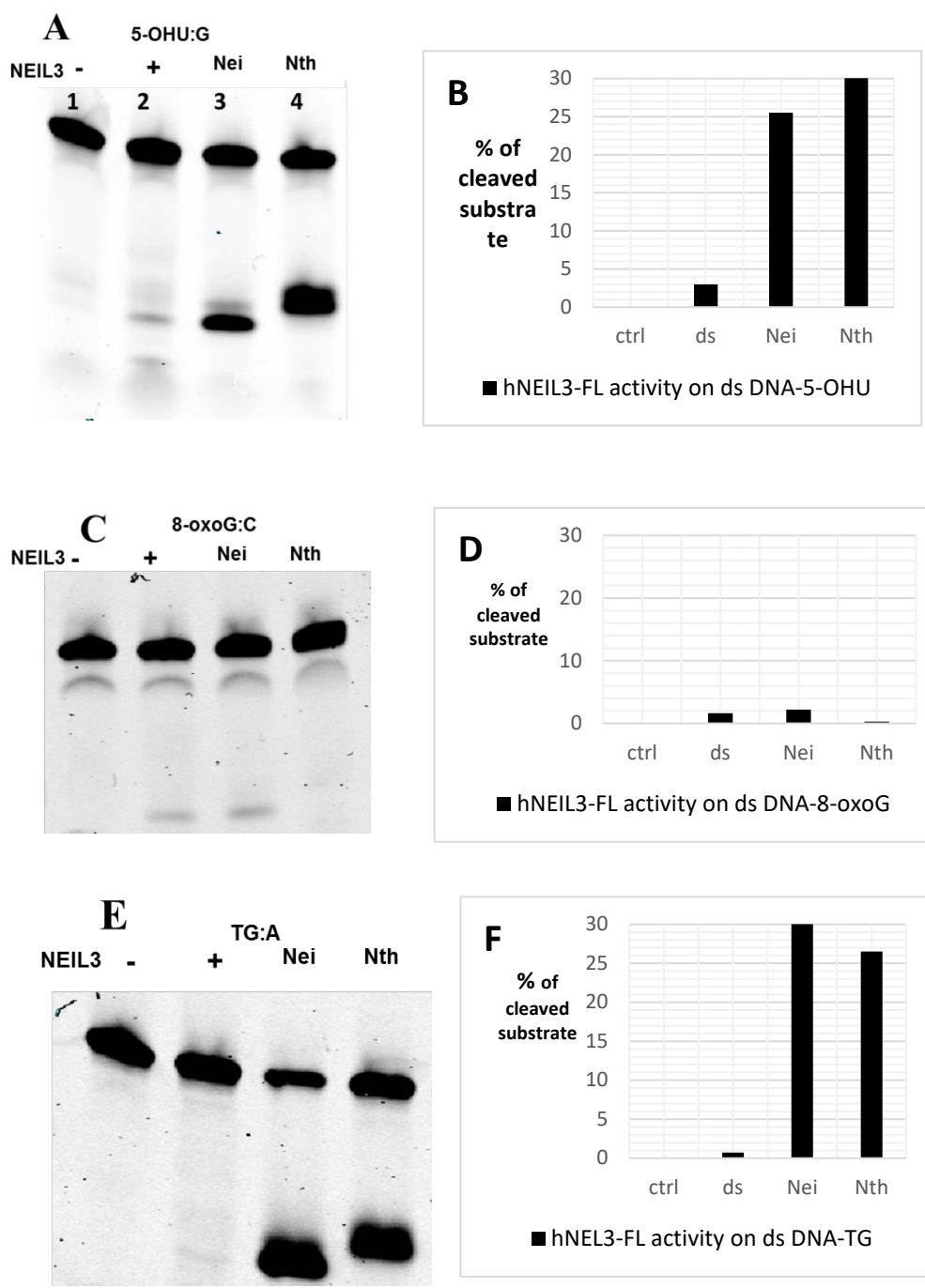
**Figure 4.41 DNA glycosylase / lyase activity and quantitative analysis of hNEIL3-FL, Nei and Nth on dsDNA substrates containing either 5-OHU:G (A, B), 8-oxoG:C (C, D), or TG:A (E, F).** Reactions were incubated at 30ºC for 20 min with either hNEIL3-FL (150 nM), Nei (15 nM) or Nth (15 nM).

To further examine the activity of hNEIL3-FL on oxidised bases in different contexts in dsDNA, the protein was incubated with dsDNA substrates containing 5-OHU, TG, or 8-oxoG paired with each of the four possible bases Figure 4.42. For 5-OHU, weak activity was observed when T, C, or G was opposite the oxidised cytosine (Figure 4.42A) This confirms the previous finding of weak activity of hNEIL3-FL on 5-OHU:G in Figure 4.41A For TG, the strongest activity was observed when paired with C (Figure 4.42B). A more intense band than in the control was also observed when TG was paired with A (Figure 4.42B), however, this was not routinely observed in subsequent experiments. The presence of a significant product band in the no-protein control suggests that this base pair was susceptible to degradation by the NaOH treatment alone. For 8-oxoG, weak activity was observed for every base pair, although again there was significant product in one of the no-protein controls (8-oxoG:T; (Figure 4.42C). These experiments indicate that hNEIL3 could act on oxidised bases in a dsDNA context *in vivo* and further experiments were designed to test this using a model DNA replication fork structured DNA oligonucleotide substrates.
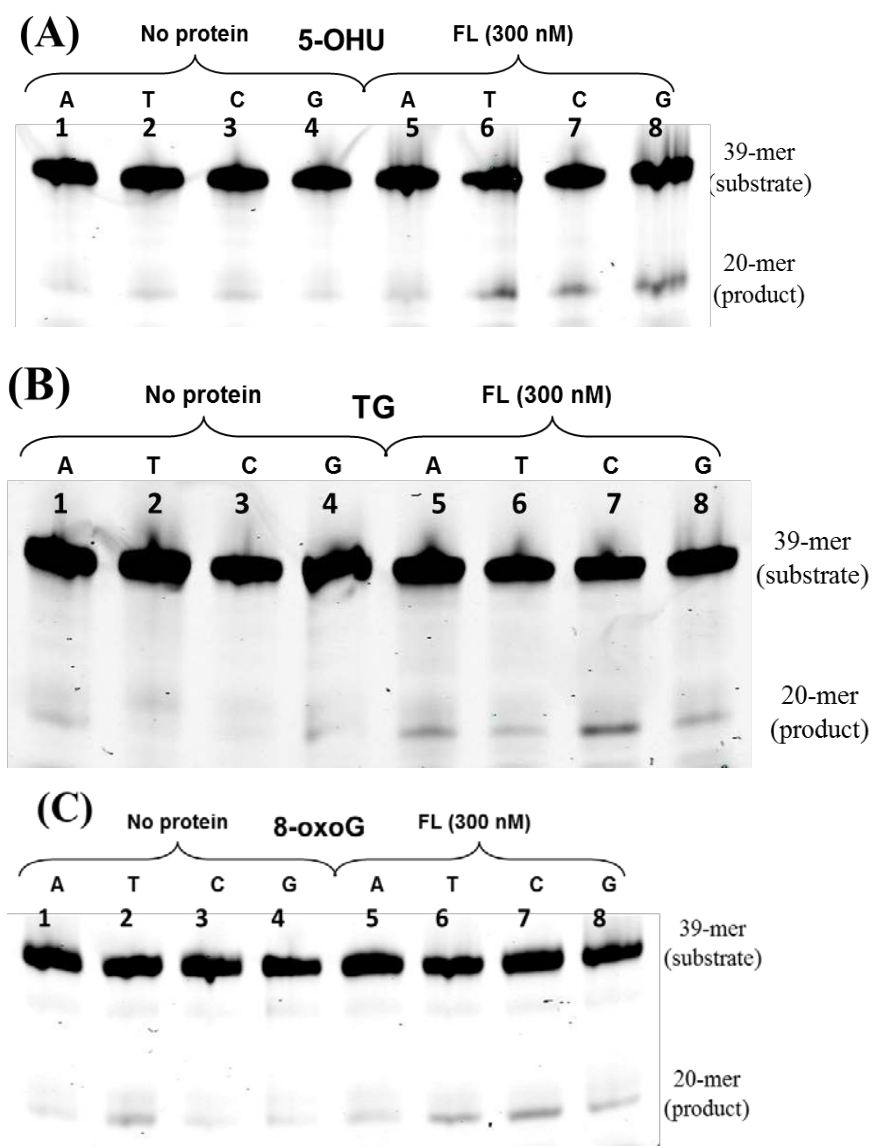
**Figure 4.42 DNA glycosylase / lyase activity of hNEIL3-FL on dsDNA substrates containing 5-OHU (A), TG (B) and 8-oxoG (C) paired with all four possible bases.**
Reactions were incubated at 30ºC for 20 min with hNEIL3-FL. Following incubation, 1 µl of 1 M NaOH was added and the reaction stopped by the addition of formamide/bromophenol blue loading buffer.

### 4.4.3 Incision of model replication fork substrates containing a single oxidised base.

For the last set of experiments, a model replication fork was designed to test the DNA glycosylase / lyase activity of hNEIL3-843 and hNEIL3-FL on oxidised bases in three different contexts in the fork. Details of the construct design are given in Table 2.4 but briefly, 'Fork-4' refers to an oxidised base after the fork in the ssDNA region. 'Fork' refers to an oxidised base in the last nucleotide in the dsDNA region of the fork, while 'Fork+4' refers to an oxidised base four nucleotides in front of the fork in the dsDNA. Each Figure shows the activity of hNEIL3 on these substrates as well as control reactions with the oxidised base in ssDNA, dsDNA and a negative control where no protein was added to the reaction.
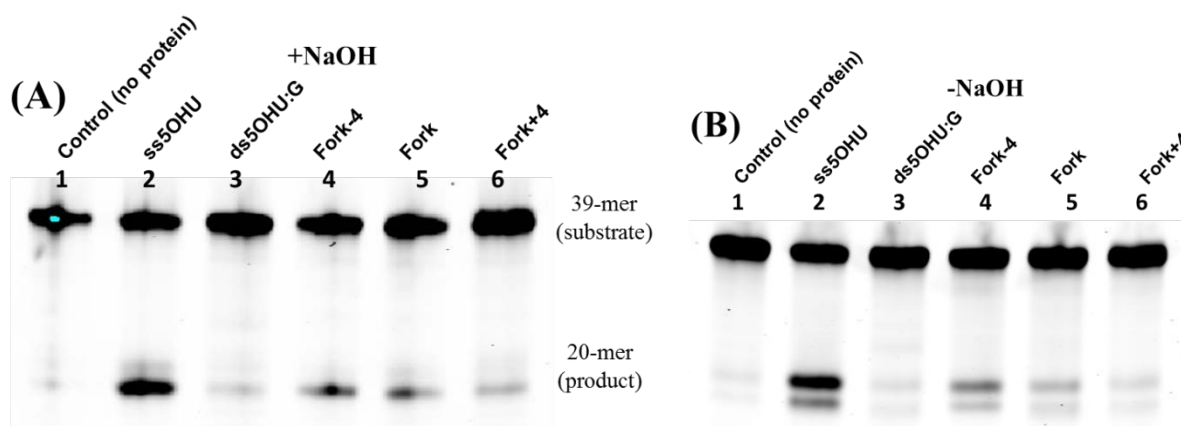


**Figure 4.43 DNA glycosylase / lyase activity of hNEIL3-843 on 5-OHU in model replication fork substrates.**
hNEIL3-843 was incubated at 30ºC for 20 min with the different substrates. In (A) reactions were stopped by the addition of 1 µl of 1 M NaOH, while in (B), the reactions were stopped by the addition of formamide / bromophenol blue loading buffer only.

In Figure 4.43 hNEIL3-843 was incubated with the five DNA substrates containing 5-OHU as the oxidised base. In Figure 4.43A, NaOH was added after the enzyme reaction to analyse the

3′-end of the labelled fragment. In agreement with the results of the initial experiments, hNEIL3-843 was most active on 5-OHU in ssDNA (lane 2). Comparison of the two gels also confirms that hNEIL3-843 is acting through β-elimination, the product band migrating further following NaOH treatment (Figure 4.43A). In this experiment, activity on the model replication fork substrates was weaker than on the ssDNA oligonucleotide, with activity highest when 5-OHU is in the -4 position (ssDNA) and least in the +4 region (dsDNA). All product bands migrate the same distance indicating a similar mode of incision (β-elimination).

When the experiment was repeated with hNEIL3-FL (Figure 4.44), a similar pattern of incision was observed, *i.e.* the oxidised base in the ssDNA context was the preferred substrate. However, strikingly, the incision product of hNEIL3-FL at the +4 position migrates further in the gel than the other product bands (Figure 4.44B, lane 6), indicating a β/δ-lyase mode of action. This is in contrast to what is observed with hNEIL3-843 (Figure 4.43, lane 6) where the band position is in line with β-elimination cleaving product, and is highly reproducible (*e.g.* Figure 4.45). Figure 4.45 also confirms the β/δ mode of action as the lower band migrates in line with the product of Nei, which has β/δ lyase activity in contrast to the β-elimination mechanism of Nth (Koketsu et al., 2004) The same, faster migrating band is also observed when hNEIL3-FL is incubated with 5-OHU in a dsDNA oligonucleotide (Figure 4.44, lane 3), however, in Figure 4.44, and to a lesser extent in Figure 4.45, the activity of hNEIL3-FL is greater in the context of the model replication fork. This is, to my knowledge, the first time the C-terminal domains of hNEIL3 have been shown to influence the bifunctional DNA glycosylase activity of the protein.
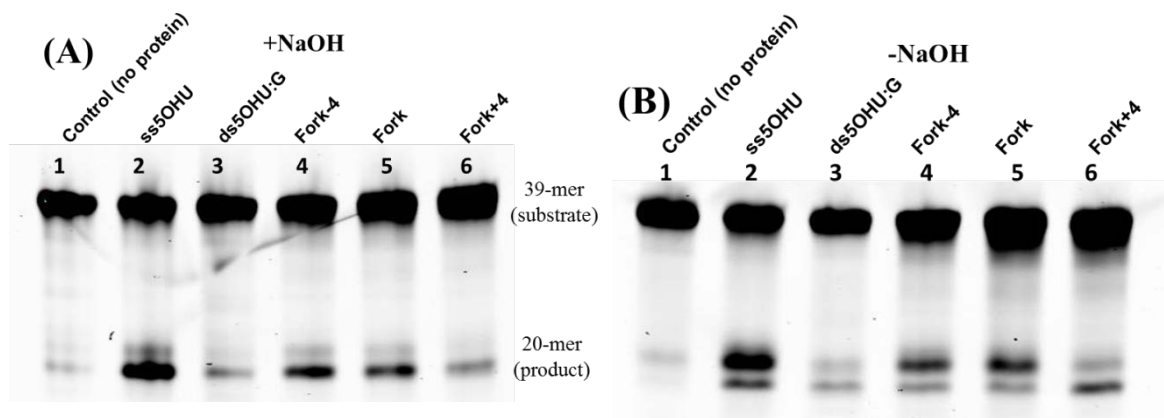
**Figure 4.44  DNA glycosylase / lyase activity of hNEIL3-FL on 5-OHU in model replication fork substrates.**

hNEIL3-FL was incubated at 30ºC for 20 min with the different substrates. The reactions were stopped by the addition of formamide/bromophenol blue loading buffer.
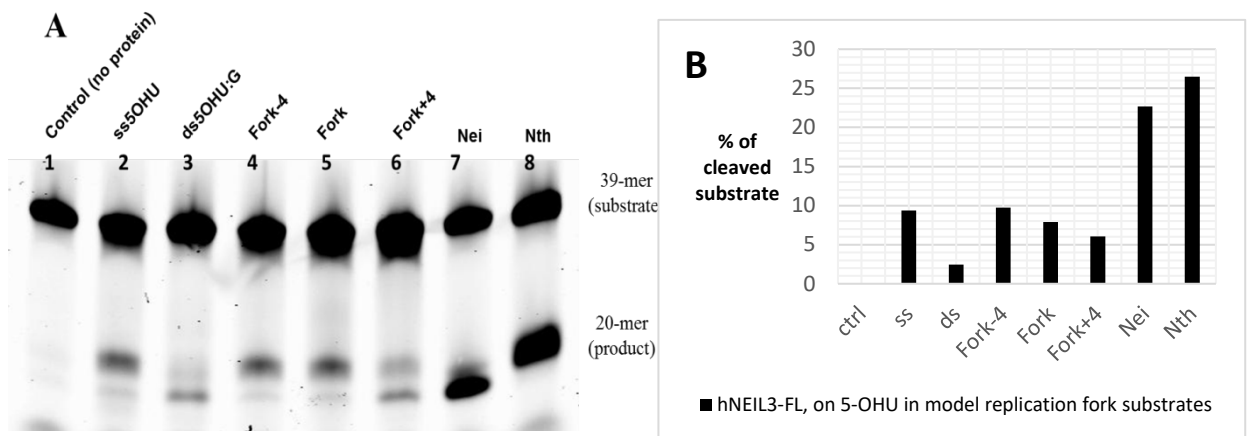


**Figure 4.45 DNA glycosylase / lyase activity (A) and quantitative analysis (B) of hNEIL3-FL on 5-OHU in model replication fork substrates.**

hNEIL3-FL was incubated at 30ºC for 20 min with the different substrates. The reactions were stopped by the addition of formamide/bromophenol blue loading buffer.

To investigate the mechanism of strand incision further, an experiment was performed with APE-1 added to the reaction mixture. This would test if the hNEIL3 was acting primarily as a monofunctional DNA glycosylase and also help to confirm the β/δ lyase activity observed at

the +4 position in the fork in Figure 4.44and Figure 4.45 (6.036% of 5-OHU substrate in +4 fork position has been incised (Figure 4.45B). The results presented in Figure 4.46 indicate that the addition of APE-1 did not affect the method of incision when ssDNA or dsDNA was the substrate (lanes 2 & 3), nor did it have any effect when 5-OHU was in the -4 position of the model replication fork (lane 3), the banding pattern being almost identical to that in Figure 4.44. However, the addition of APE1 did alter the migration of the product band when 5-OHU was placed at the fork or at the +4 position, in the dsDNA region. In both, the product band migrates between the bands obtained through $\beta-$ and $\beta/\delta-$ lyase activity and thus indicating the activity of APE-1 on the abasic site caused by the monofunctional DNA glycosylase activity of hNEIL3-FL.



**Figure 4.46 DNA glycosylase / lyase activity of hNEIL3-FL + APE1 on 5-OHU in model replication fork substrates.**
Oligonucleotide substrates were incubated at 30ºC for 20 min with NEIL3-FL (300 nM) and APE1 (15 nM).
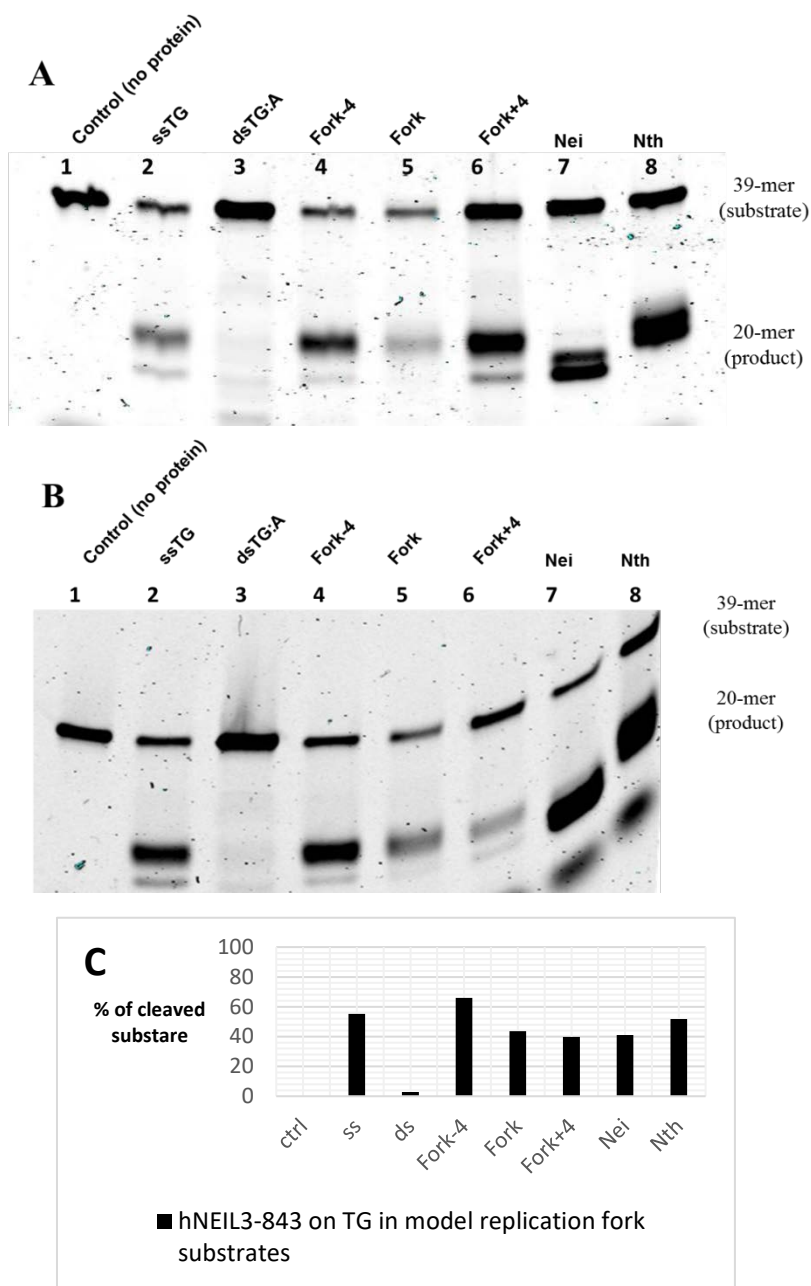
**Figure 4.47 DNA glycosylase / lyase activity A, B (repeat) and quantitative analysis (C) of hNEIL3-843 on TG in model replication fork substrates.**

hNEIL3-843 was incubated at 30ºC for 20 min with the different substrates. Experiment (A) was repeated to produce B. The quantification data was sourced from gel Figure 4.47B.

Next, hNEIL3-843 was incubated with the same substrates but containing TG (Figure 4.47). In agreement with previous experiments, no activity of hNEIL3-843 was observed on the dsDNA substrate (Figure 4.47A, B, lane 3 and Figure 4.41E), but good activity was observed on TG in ssDNA (lane 2). Activity at position -4 (ssDNA) on the model replication fork is similar to that of the ssDNA oligonucleotide (lanes 2 and 4). The activity at the Fork and position +4 is less clear, with one gel showing strong incision activity at position +4 and the other not (39.69% of TG substrate in fork +4  position has been incised (Figure 4.47C)). However, comparison with Figure 4.49, where the same experiment is shown but with hNEIL3-FL, suggests that hNEIL3 can act as a bifunctional DNA glycosylase in the double-stranded region of the replication fork (compare lane 6 in Figure 4.47and Figure 4.48).

What is striking about the results in Figure 4.48 is that the mechanism of incision by hNEIL3-FL at TG is by β-elimination and not β/δ-lyase as observed for 5-OHU. Again this result is highly reproducible, only the relative activities at the different positions altering between experimental repeats (6.87% of TG substrate in fork +4 position has been incised Figure 4.48C).

**Figure 4.48 DNA glycosylase / lyase activity A, B (repeat) and quantitative analysis (C) of hNEIL3-FL on TG in model replication fork substrates.**

A and B are repeat experiments. hNEIL3-FL was incubated at 30ºC for 20 min with the different substrates. The quantification data was sourced from Figure 4.48A.

Finally, the experiments were repeated with 8-oxoG as the oxidised base in the model DNA replication fork structure. The results, shown in Figure 4.49 and Figure 4.50 are less clear than for the oxidised pyrimidines. hNEIL3-843 shows only weak activity against 8-oxoG, principally when it is in a ssDNA context (Figure 4.49, lane 2 with 3.28 % of 8-oxoG substrate in ssDNA context has been incised Figure 4.49C), However, weak activity can also

be observed at both positions -4 and +4 in the fork structure (lanes 4 and 6). This is the first evidence that hNEIL3 can release 8-oxoG from dsDNA, albeit a weak activity. Comparison with the faint band seen in the Nei lane (lane 7), incision of the abasic site by hNEIL3-843 appears to be by β-elimination.
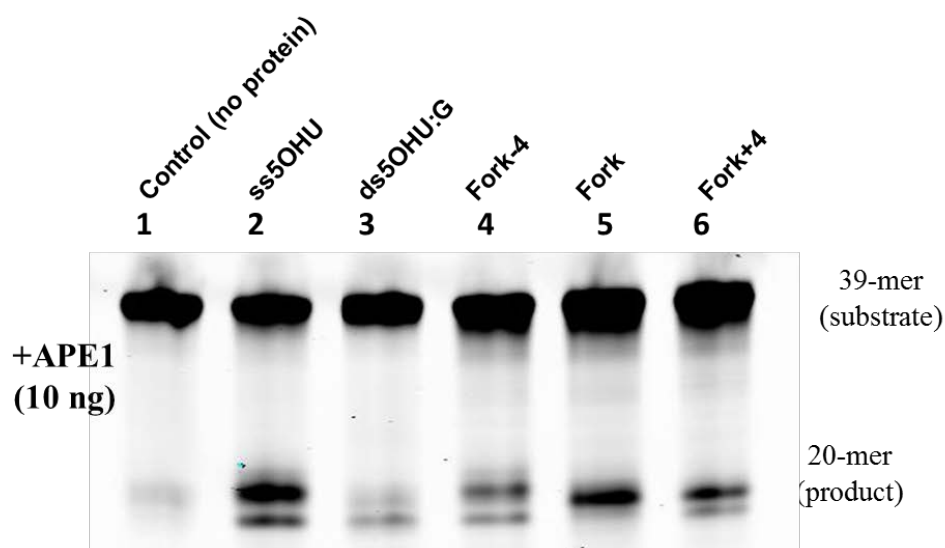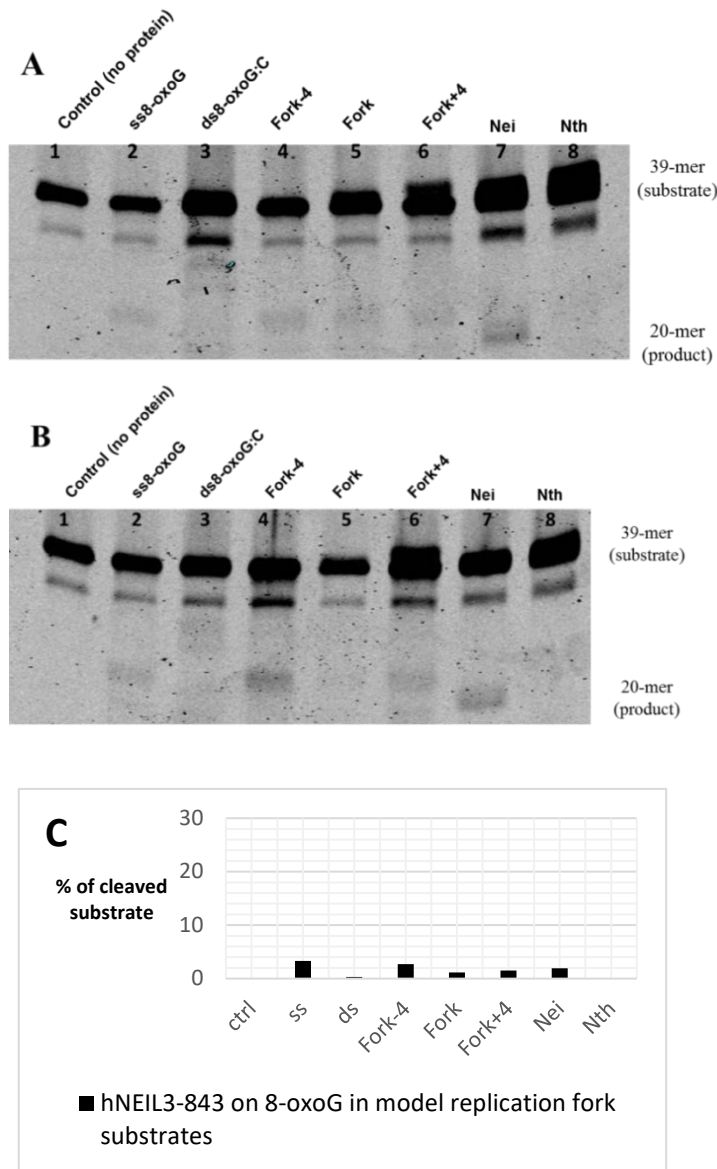


**Figure 4.49 DNA glycosylase / lyase activity A, B (repeat) and quantitative analysis (C) of hNEIL3-843 on 8-oxoG in model replication fork substrates.**
A and B are repeat experiments. hNEIL3-843 was incubated at 30ºC for 20 min with the different substrates. The quantification data was sourced from Figure 4.49A.

Activity of hNEIL3-FL on 8-oxoG is also weak, however, significant incision of the dsDNA substrate is observed, which is completely lacking in hNEIL3-843 (compare lane 3 in Figure 4.49 and Figure 4.50, 2.83 % of  dsDNA 8-oxoG:C substrate has been incised Figure 4.50C while only 0.29% dsDNA 8-oxoG:C substrate has been incised Figure 4.49C). Interestingly, the product band appears to migrate in line with the product of Nei, again indicating that the C-terminal domains of hNEIL3 can influence the DNA glycosylase / lyase activity of the protein. Also visible in Figure 4.37A,  it is evident that the small amount of incision carried out by hNEIL3-FL at the abasic site left after removal of 8-oxoG, is by $\beta$-elimination, similar to what was observed for 5-OHU (Figure 4.44 lane 2). A summary of all these findings is given in Table 4.1.
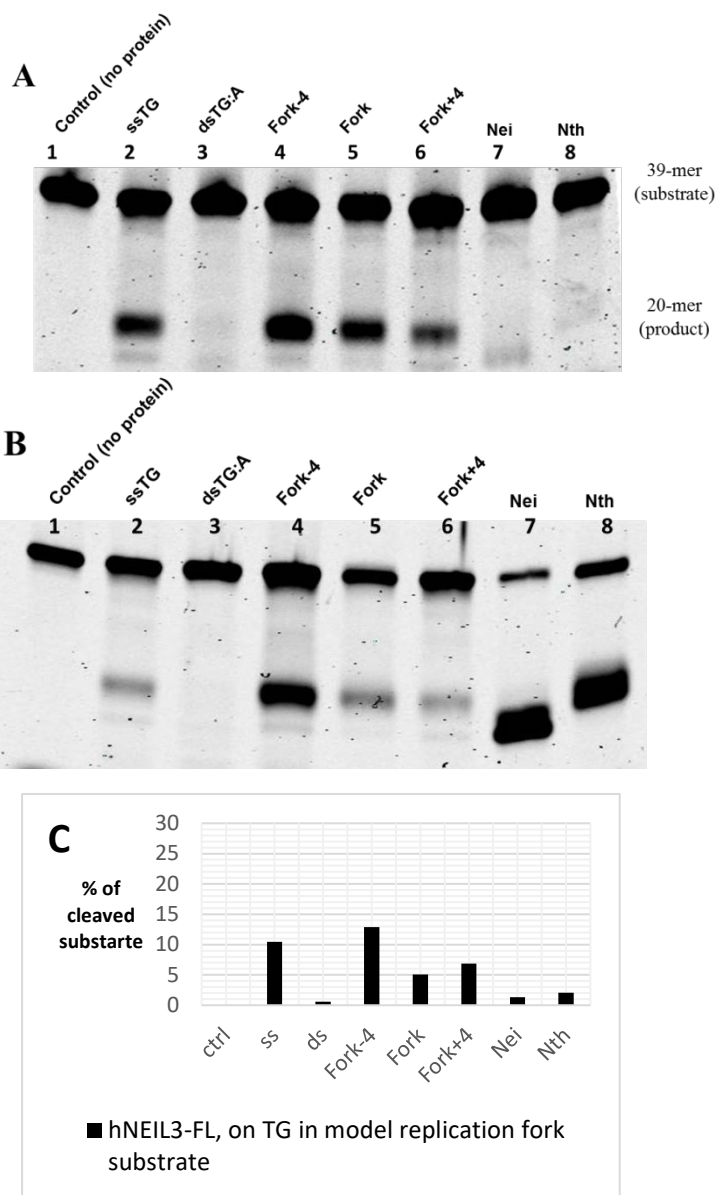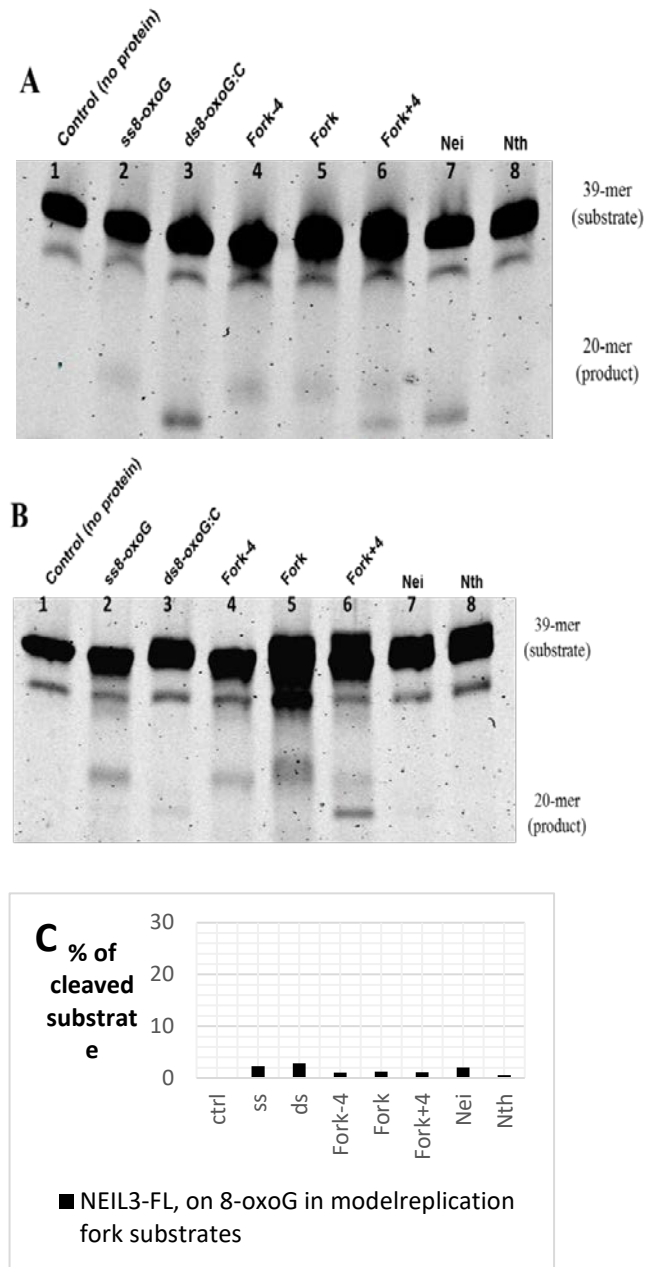
**Figure 4.50 DNA glycosylase / lyase activity A, B (repeat) and quantitative analysis (C) of hNEIL3-FL on 8-oxoG in model replication fork substrates.**

A and B are repeat experiments. hNEIL3-FL was incubated at 30ºC for 20 min with the different substrates. The quantification data was sourced from Figure 4.50A.

**Table 4.1  Summary of hNEIL3 Enzyme Activity.**

| Oligonucleotide substrate | hNEIL3-843 | hNEIL3-FL |
|---|---|---|
| ssDNA(5-OHU) | β-elimination | β-elimination |
| ssDNA(TG) | β-elimination | β-elimination |
| ssDNA(8-oxoG) | β-elimination | β-elimination |
| dsDNA(5-OHU:G) | Weak β-elimination | Weak β/δ-elimination |
| dsDNA(TG:A) | No activity | No activity |
| dsDNA(8-oxoG:C) | No activity or weak β-elimination | Weak β/δ-elimination |
| Fork-4 DNA(5-OHU) | β-elimination | β-elimination |
| Fork DNA(5-OHU) | β-elimination | β-elimination |
| Fork+4 DNA(5-OHU) | β-elimination | β/δ-elimination |
| Fork-4 DNA(TG) | β-elimination | β-elimination |
| Fork DNA(TG) | β-elimination | β-elimination |
| Fork+4 DNA(TG) | β-elimination | β-elimination |
| Fork-4 DNA(8-oxoG) | β-elimination | β-elimination |
| Fork DNA(8-oxoG) | β-elimination | β-elimination |
| Fork+4 DNA(8-oxoG) | β-elimination | β/δ-elimination |

# 5 Discussion

Data presented in this thesis show successful expression and purification of full-length and truncated versions of the human NEIL3 protein in an active form. The recombinant proteins have been characterised using enzyme activity assays, using different oxidised bases in ssDNA, dsDNA and in different positions in a replication fork context. The results obtained give an insight into the role of hNEIL3 in the repair of oxidised DNA base damage during the replication process. Furthermore, when comparing the activity of the truncated DNA glycosylase only domain (hNEIL3-843) and hNEIL3-FL, data shows that the C-terminal tail of NEIL3 has an influence on the AP lyase activity of the hNEIL3 protein and that hNEIL3-FL uniquely acts through both β-elimination and/or β,δ-elimination AP lyase activity, and that this is dependent on the substrate lesion and the position of the oxidised base at the replication fork.

Human NEIL3 has been reported to be difficult to express in *E. coli* and this has limited the biochemical characterization and functional studies of the protein in order to understand its role in cell survival (Liu et al, 2012). The expression vector pETDuet-2 was first described by Liu *et al.* (2012) and used to express active murine NEIL3 protein. This vector possesses two significant parts, one is a short sequence, ORF6, in front of the coding sequence of NEIL3, and the other is an engineered version of the *E. coli* amino peptidase, EcoMapY168A that has greater activity on the N-terminal methionine in hNEIL3 than the endogenous enzyme. The expression vector also possesses a six-histidine fusion tag downstream of the target gene that allows protein purification using affinity chromatography (Liu et al, 2012).

The first aim of the project was to express both full-length and truncated versions of hNEIL3 using the pETDuet-2 expression system to allow their biochemical characterisation in a number

of different assays. In this way it was hoped that the role of the extended C-terminal domains can be determined through their interaction with different DNA substrates.

Cloning of mouse NEIL3 into the pETDuet-2 expression vector was achieved by amplifying mouse NEIL3 cDNA and adding NdeI and XhoI restriction sites to the forward and reverse primers respectively, then digesting the PCR product with NdeI and XhoI and cloning the digested cDNA into the pETDuet-2 expression vector by linearizing the vector with the same restriction enzymes to create compatible ends (Liu *et al*., 2012). For hNEIL3, this cloning strategy could be used for the hNEIL3-843 truncation only. For the ORF6-hNEIL3-1044, 1236, 1506 and full-length cDNAs, direct cloning into the pETDuet2 expression vector using an NdeI and XhoI double-digest was not possible because hNEIL3 possesses an NdeI restriction site at position 864 bp. This means that the NdeI enzyme cuts the hNEIL3 cDNA preventing cloning of any longer hNEIL3 fragment into the pETDuet2 expression vector. To overcome this obstacle, a different cloning strategy had to be designed. Due to destruction of the EcoRV site in pETDuet2 during the construction of the vector, and the presence of multiple EcoRV restriction sites in the EcoMapY168A DNA sequence, the double-stranded ORF6 sequence was synthesised (Eurofins) and ligated into pET30-b (Balis, 2013). Therefore, the ORF6-hNEIL3 inserts were first constructed in pET30-b-ORF6 and then transferred to the pETDuet-2 expression vector.

After cloning the hNEIL3 inserts into the pET30b-ORF6 vector was achieved and confirmed, a novel strategy was used to clone the ORF6-hNEIL3 inserts into pETDuet-2 expression vector. An ORF6-hNEIL3 forward primer was designed to amplify ORF6-hNEIL3 inserts from pET30b-ORF6-hNEIL3 then the PCR products were digested with XhoI to create a XhoI site compatible with the XhoI site in the expression vector. The pETDuet-2 expression vector was

prepared for cloning in three steps, first the vector was linearized by NdeI restriction enzyme digest, second the NdeI site overhang resulting from the NdeI restriction digest was removed using mung bean nuclease and finally the vector was digested with XhoI. By the end of this process compatible ends for cloning ORF6-hNEIL3 inserts in the correct reading frame into the pETDuet-2 expression vector were created.

Human NEIL3 protein expression in *E. coli* was carried out starting with test protein expression of pETDuet-2-ORF6-hNEIL3 (843, 1044, 1506 and full length) using a modified protein expression protocol of Liu *et al.* (2012). Different incubation times (3 h and 18 h) and temperatures (16°C and 37°C) were tried after inducing the *E. coli* cultures with 1 mM IPTG. Protein was extracted and recombinant hNEIL3 expression, confirmed by SDS-PAGE and western blotting, before hNEIL3 was expressed using larger *E. coli* cultures (300 ml) induced with IPTG and incubated at 16°C for 18 h. The His-tagged proteins were then purified and the expression and purification of hNEIL3-(843, 1044, 1506 and full length) was confirmed by SDS-PAGE and western blotting.

The SDS-PAGE and western blotting results, following the His-Trap FPLC purification of hNEIL3-(843 and 1044) purified proteins confirmed that the proteins were pure enough to carry out the enzyme activity assays. However, for hNEIL3-FL, expression and purification was repeated using larger cultures (1200 ml) and the same expression protocol was followed. The protein was purified using His-Trap chromatography (Section 3.2.3) and fractions were tested by SDS-PAGE and western blotting and subsequently underwent further purification using Mono S FPLC (Section 3.2.5). At the end of the expression and purification process the successful expression and purification of recombinant hNEIL3-(843, 1044 and FL) was confirmed by SDS-PAGE and western blotting.

In order to characterize the purified recombinant hNEIL3-(843, 1044 and FL) proteins, hNEIL3 proteins were tested by enzyme activity assays to determine their DNA glycosylase / AP lyase activity. Both hNEIL3-843 and hNEIL3-FL were also subjected to a $NaBH_4$-trapping assay and the result proved that the observed activity was due to the recombinant proteins (Figure 4.28). The stability of hNEIL3-FL during the activity assays was also tested and the protein proved to be stable under the reaction conditions (Figure 4.29).

The biochemical analysis of the recombinant hNEIL3 proteins was assessed by incubating the purified proteins with oligonucleotides containing a single oxidised base. In order to test their DNA glycosylase / AP lyase activity, a 39 mer ssDNA oligonucleotide substrate containing 5-OHU at position 20 was initially used. The reaction was carried out at 30ºC or 37ºC for 20 min to determine the optimal reaction conditions (Figure 4.30). All three hNEIL3 proteins proved to be active on the substrate using 300 nM of the proteins. The observed activity of the recombinant hNEIL3 proteins on the ssDNA oligonucleotide substrate was through β-elimination, as when the proteins were tested in the presence of NaOH, the addition of NaOH produced product bands equal to β/δ elimination. The results were similar for all hNEIL3 proteins (843, 1044 and FL), indicating that the hNEIL3-FL is a bifunctional DNA glycosylase that acts predominantly through β-elimination and the C-terminal domain in hNEIL3-FL has no effect on those substrates under these reaction conditions.

At this point it was decided to proceed with enzyme activity assays for hNEIL3-843, which contains the N-glycosylase domain and hNEIL3-FL which contains all the protein domains. The DNA glycosylase / AP lyase activity of hNEIL3-843 and hNEIL3-FL on ssDNA substrates containing different oxidised bases (5-OHU, TG, 8-oxoG and U) was tested and both proteins

proved to be active on ssDNA substrates containing 5-OHU or TG acting through β-elimination, while only weak activity was observed on 8-oxoG and no activity on the uracil containing substrate. The experiments were repeated with the addition of the bacterial DNA glycosylases Nei and Nth as positive controls to confirm β,δ (Nei) and β-elimination (Nth) reaction products.

Time course experiments were carried out to determine the time needed to complete the reaction, the results show that the maximum percentage of 20 mer cleaved product is reached in 5 min. The DNA glycosylase / AP lyase activity of the recombinant hNEIL3 proteins were then tested on double-stranded oligonucleotides containing a single oxidised base (5-OHU:G, TG:A and 8-oxoG:C), the result show weak activity on double-stranded oligonucleotides containing (5-OHU:G and TG:A) and no activity on the substrate containing (8-oxoG:C).

The hNEIL3 recombinant proteins were then tested on a model replication fork that was designed with the oxidised base in three different positions (Table 2.4): (i) at position -4, the oxidised base resides after the fork, in a ssDNA context, (ii) in 'Fork' the oxidised base is the last nucleotide in the dsDNA region of the fork, and (iii) at +4 the oxidised base position is four nucleotides in front of the fork in the dsDNA.

In Figure 4.43, the results show that hNEIL3-843 is most active on the ssDNA substrate containing 5-OHU confirming the previous results with high activity on 5-OHU in the Fork-4 position (ssDNA) and least activity in the +4 region (dsDNA) and acting through β-elimination. However, while hNEIL3-FL shows a similar pattern of incision for the 5-OHU containing ssDNA substrate and Fork-4 position (ssDNA), with 5-OHU at the Fork+4 position (dsDNA), the product band migrates further in the gel and in line with the product of Nei incision (Figure

4.45), indicating that hNEIL3-FL is acting through β,δ-lyase mode of action on those substrates. A similar faster migrating, but weaker, band is also observed when hNEIL3-FL is incubated with 5-OHU in a dsDNA oligonucleotide, but the activity is greater when the oxidised base is at the Fork+4 position (dsDNA) in the model replication fork. This indicates, for the first time, that the C-terminal domains of hNEIL3 have an effect on the mechanism of the bifunctional DNA glycosylase activity of the protein.

These experiments were repeated on the same model replication fork substrates but containing TG; for hNEIL3-843 the results were in agreement with the 5-OHU substrates while, in contrast, hNEIL3-FL acts through β-elimination in all of the TG substrates and not β/δ-lyase as observed for 5-OHU (Figure 4.48). The experiments were also repeated using substrates containing 8-oxoG, but hNEIL3-843 showed only weak activity on ssDNA and at both Fork-4 and Fork+4 positions in the replication fork structure (Figure 4.49). Comparison with the product of Nei incision, indicates that hNEIL3-843 incised the abasic site through β-elimination (Figure 4.49).

For hNEIL3-FL activity on 8-oxoG, the results again show only weak activity in a ssDNA context and incision by β-elimination (Figure 4.50). However, in both repeats there is significant incision of the 8-oxoG containing oligonucleotide at the Fork+4 position and the incision is through β,δ lyase activity (Figure 4.50). This incision activity was essentially absent when hNEIL3-843 was used (lane 3 in Figure 4.49 and Figure 4.50). Thus these results are in agreement with the hNEIL3-FL activity observed for 5-OHU and confirm that the C-terminal domains of hNEIL3 can influence the DNA glycosylase / AP lyase activity of the protein.

NEIL3 is normally expressed in specific highly proliferating cells, however, cancer cells commonly display overexpression of NEIL3. As NEIL3 expression is highest in the S and G2 phases of the cell cycle, this suggested a role in the maintenance of successful DNA replication. Data presented here shows that hNEIL3 undertakes DNA glycosylase and AP-lyase activity dependent on the lesion position or type of the substrate (TG always β and 5-OHU β or β,δ-elimination), and that the C-terminal domain of NEIL3 has influence on the type of the bifunctional lyase activity of the NEIL3 protein.

Here it is shown that, of the three substrate lesions tested, NEIL3 has a preference for DNA containing 5-OHU and TG, oxidised bases that are also processed by NEIL1, NEIL2 and NTH1 (Wallace, 2014). However, the finding that NEIL3 is expressed during the S/G2-phases of the cell cycle suggests that NEIL3 has a role of recognizing and removing oxidative lesions during DNA replication particularly at the replication fork. NEIL1 is also reported to be cell cycle regulated with the highest activity found during S-phase (Hazra and Mitra, 2006). Complementary data obtained in our lab (Albelazi *et al.* in preparation) indicates efficient activity of hNEIL1 on 5-OHU and TG in dsDNA and dsDNA close to the replication fork, while NEIL3 shows preference for ssDNA and ssDNA in the model replication fork, with no activity in dsDNA containing TG, suggesting that NEIL1 and NEIL3 have overlapping and complementary functions in the repair of oxidized bases at the replication fork.

While NEIL3 shows only weak or no activity on DNA substrates containing 8-oxoG, NEIL3 is active on the further oxidation products of 8-oxoG, Sp and Gh (Krokeide *et al.,* 2013), although NEIL1 has been reported to show greater activity on those substrates in enzyme activity assays (Krokeide *et al.,* 2013; Martin *et al.,* 2017). On the other hand, Krokeide *et al.* (2013) showed that NEIL3 incised oxidised DNA lesions by β-elimination unlike NEIL1 and NEIL2, which

act through β,δ-elimination. Data presented here show that NEIL3 also possesses a β,δ-elimination mode of action depending on the damaged base type and position in the model replication fork.

In recently published work, the lack of NEIL3 led to telomere dysfunction causing anaphase bridges (Zhou *et al*., 2017). Also, it has been shown that NEIL3 has a role in the repair of ICLs (Semlow *et al*., 2016; Klattenhoff *et al.,* 2017; Martin *et al*., 2017). Along with the bacterial homolog Nei, NEIL1 and NEIL3 can unhook three- and four-stranded psoralen-induced ICL structures by hydrolysis of the glycosylic bond between the base and deoxyribose sugar. Uniquely, hNEIL3 achieved this through a mechanism that did not generate toxic double-strand breaks and therefore subsequent activation of the DNA damage response pathway and potential deleterious consequences to the cell (Martin *et al.,* 2017).

Thus, the overexpression of hNEIL3 in cancer cells may give them a growth advantage and could be vital during S phase in relation to the repair of the DNA replication thus safe guarding the genomic integrity of the newly synthesised DNA in daughter cells. Taken together, all this data suggests that the inhibition of hNEIL3 activity in cancer cells may be a good target to overcome the resistance of some cancer cells to specific genotoxic agents that are commonly used in cancer therapy to induce ICLs in DNA.

Thus, data presented here supports a role for hNEIL3 as a bifunctional DNA glycosylase, particularly in DNA replication, with AP lyase activity on oxidized bases within ssDNA and dsDNA at the replication fork either through β or β,δ-elimination depending on the context of the oxidised base lesion. This activity is, to my knowledge, unique to NEIL3 as all other bifunctional DNA glycosylases act through either β-elimination or β,δ-elimination alone. Furthermore, work presented here, and that performed subsequently using the recombinant

proteins discussed in this thesis, have shown that hNEIL1 and hNEIL3 have complementary roles at the DNA replication fork, both in the release of oxidised pyrimidines and in the resolution of ICLs. Further work is now required to examine the control of hNEIL3 levels in cells, the proteins it interacts with and the effect of reducing or deleting hNEIL3 in specific cells to determine any increased efficacy of clinically relevant genotoxic cancer agents.

# 6  References

Arai, T. Kelly, V.P.,  Minowa, O. T., Noda, Nishimura, S. (2006)  The study using wild-type and Ogg1 knockout mice exposed to potassium bromate shows no tumor induction despite an extensive accumulation of 8-hydroxyguanine in kidney DNA, *Toxicology* **221,** 179–186.

Aller, P., Rould, M. A., Hogg, M., Wallace, S. S. & Doublié. (2007) A structural rationale for stralling of a replicative DNA polymerase at the most common oxidative thymine lesion thymine glycol. *Proc. Natl. Acad. Sci. USA,* **104**, 814-818.

Balis, L. (2013). Molecular cloning of the hNEIL3 cDNA into the bicistronic pETDuet-2 Expression vector. MSc University of Salford, Salford.

Bandaru, V., Sunkara, S., Wallace, S. & Bond, J.P. (2002) A novel human DNA glycosylase that removes oxidative damage and is homologous to *Escherichia coli* endonuclease VIII. *DNA Repair,* **1**, 517-529.

Banerjee, D.,  Mandal, S.M.,  Das,  A., Hegde, M. L., Das, S.,  Bhakat, K. K., Boldogh, I., Sarkar, P. S., Mitra, S., Hazra, T. K. (2011) Preferential repair of oxidized base damage in the transcribed genes of mammalian cells. *J. Biol. Chem*, **286**, 6006-6016.

Bjelland S, Seeberg E. (1996) Different efficiencies of the Tag and AlkA DNA glycosylases from Escherichia coli in the removal of 3-methyladenine from single-stranded DNA. *FEBS Lett,* **397,** 127–129.

Boiteux, S., Gajewski, E., Laval, J., Dizdaroglu, M. (1992) Substrate specificity of the *Escherichia coli* Fpg protein (formamidopyrimidine-DNA glycosylase): excision of purine lesions in DNA produced by ionizing radiation or photosensitization. *Biochemistry,* **31,** 106-110.

Broderick, P., Bagratuni, T., Vijayakrishnan, J.,  Lubbe, S., Chandler, I., Houlston, R. S. (2006) Evaluation of NTHL1, NEIL1, NEIL2, MPG, TDG, UNG and SMUG1 genes in familial colorectal cancer predisposition. *BMC Cancer* **6,** 243

Cadet, J., Douki, T., Frelon, S., Sauvaigo, S., Pouget, JP., Ravanat, JL. (2002) Assessment of oxidative base damage to isolated and cellular DNA by HPLC-MS/MS measurement. *Free Radical Biology and Medicine*, **33**, 441–449.

Cadet, J., Douki, T., Ravanat, J. L. (2010) Oxidatively generated base damage to cellular DNA, *Free Radic. Biol. Med*, **49,** 9–21.

Cadet, J. & Wagner, J. R. (2013) DNA Base Damage by Reactive Oxygen Species, Oxidizing Agents, and UV Radiation. *Cold Spring Harbor Perspectives in Biology,* **5**, a012559.

Cappelli, E. Taylor, R. Cevasco, M. *et al.*, (1997) Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. *J. Biol. Chem.* **272(38)**, 23970–23975.

Chakraborty, A., Wakamiya, M., Venkoca-Canova, T. *et al.* (2015) *NEIL2*-null Mice Accumulate Oxidized DNA Bases in the Transcriptionally Active Sequences of the Genome and are Susceptible to Innate Inflammation. *J. Biol. Chem.,* **250**, 24636-24648

Chan, M.K., Ocampo-Hafalla, M.T., Vartanian, V., Jaruga, P., Kirkali, G. *et al.,* (2009) Targeted deletion of the genes encoding NTH1 and NEIL1 DNA N-glycosylases reveals the existence of novel carcinogenic oxidative damage to DNA, *DNA Repair,* **8**, 786-794.

Chetsanga, C.J. & Lindahl, T. (1979) Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *Escherichia coli. Nucleic Acids Res*, **6**, 3673–3684.

Chin, K., de Solorzano, C.O., Knowles, D., Jones, A., Chou, W., Rodriguez, E.G., Kuo, W.L., Ljung, B.M., Chew, K., Myambo, K., *et al*. (2004). *In situ* analyses of genome instability in breast cancer. *Nat. Genet,* **36,** 984-988.

Couvé-Privat, S., Gaëtane, M., Rosselli, F., Saparbaev, M. (2007) Psoralen-induced DNA adducts are substrates for the base excision repair pathway in human cells. *Nucleic Acids Research,* **17**, 5672-5682.

Dalhus B, Laerdahl JK, Backe PH, Bjoras M. (2009) DNA base repair–recognition and initiation of catalysis. *FEMS Microbiol Rev*, **33,** 1044–1078.

Dallosso, A. R., Dolwani, S., Jones, N., Jones, S., Colley, J., *et al.,* (2008) Inherited predisposition to colorectal adenomas caused by multiple rare alleles of MUTYH but not OGG1, NUDT1, NTH1 or NEIL 1, 2 or 3, *Gut*, **57,** 1252-1255.

Davis, A. & Chen, D. (2013). DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res*, **2(3),** 130–143.

David, S., O'Shea, V., Kundu, S. (2007) Base-excision repair of oxidative DNA damage. *Nature*, 447, 941-950.

Deans, J. & West, S. (2011) DNA Interstrand crosslink repair and cancer. *Nature Reviews Cancer,* **11**, 467-480.

Demple B, DeMott MS. (2002) Dynamics and diversions in base excision DNA repair of oxidized abasiclesions. *Oncogene*, **21**, 8926–8934.

Dey, S., Maiti, A.K., Hegde, M.L., Hegde, P.M., Boldogh, I., Sarkar, P.S., Abdel Rahman, S.Z., Sarker, A.H., Hang, B., Xie, J., Tomkinson, A.E., Zhou, M., Shen, B., Wang, G., Wu, C., Yu, D., Lin, D., Cardenas, V., Hazra, T.K., (2012) Increased risk of lung cancer associated with a functionally impaired polymorphic variant of the human DNA glycosylase NEIL2. *DNA Repair (Amst.)* **11,** 570-578.

Dou, H., Mitra, S., Hazra, T.K. (2003) Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *J. Biol. Chem.* **278,** 49679-49684.

Doublie, 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. *Proc. Natl. Acad. Sci. USA,* **10,** 10284-10289.

Drohat, A.C., & Maiti, A. (2014) Mechanisms for enzymatic cleavage of the N-glycosidic bond in DNA, *Org. Biomol. Chem*, **12**, 8367–8378.

Fromme JC, Verdine GL. (2002) Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM. *Nat Struct Biol*, **9 (7),** 544–552.

Fromme JC, Verdine GL. (2003) DNA lesion recognition by the bacterial repair enzyme MutM. *J Biol Chem,* **278 (51),** 51543–51548.

Forsbring, M., Vik, E.S., Dalhus, B., Karlsen, T.H., Bergquist, A., Schrumpf, E., Bjoras, M., Boberg, K.M., Alseth, I. (2009) Catalytically impaired hMUTYH and NEIL1 mutant proteins identified in patients with primary sclerosing cholangitis and cholangiocarcinoma, *Carcinogenesis,* **30,** 1147–1154.

Fortini, P. & Dogliotti, E. (2007) Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. *DNA Repair*, **4**, 398-409.

Friedman, J.I., Stivers, J.T. (2010) Detection of damaged DNA bases by DNA glycosylase enzymes. *Biochemistry,* **49**, 4957-67.

Greenberg, M.M., (2012) The formamidopyrimidines: purine lesions formed in competition with 8-oxopurines from oxidative stress, *Acc. Chem. Res.* **45**, 588–597.

Guo, Y., Wallace, S.S., Bandaru, V. (2009). A novel bicistronic vector for overexpressing Mycobacterium tuberculosis proteins in *Escherichia coli*. *Protein Expr Purif*, **65**, 230-237.

Hakem R. (2008) DNA-damage repair; the good, the bad, and the ugly. *EMBO J.* **27**, 589-605.

Hailer, M.K., Slade, P.G., Martin, B.D., Rosenquist, T.A., 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 (Amst.)* **4,** 41-50.

Halliwell B. (2000). Why and how should we measure oxidative DNA damage? *Am J Clin Nutr.* **72**, 1082–10877.

Hazra, T.K., Izumi, R., Venkataraman, R., Kow, Y.W., Dizdaroglu, M. & Mitra, S. (2000) Characterization of a novel 8-oxoguanine-DNA glycosylase activity in *Escherichia coli* and identification of the enzyme as endonuclease VIII. *J. Biol. Chem.*, **275**, 27762-27767.

Hazra, T.K., Izumi, T., Boldogh, I. *et al.* (2002) Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc. Natl. Acad. Sci. USA,* **99**, 3523-3528.

Hazra, T. & Mitra, S. (2006) Purification and characterization of NEIL1 and NEIL2, members of a distinct family of mammalian DNA glycosylases for repair of oxidized bases. *Methods In Enzymology,* **408**, 33-48.

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 Res,* **18**, 27-47.

Hegde, P. M., Dutta, A., Sengupta, S. *et al.* (2015) The C-terminal Domain (CTD) of human DNA glycosylase NEIL1 is required for forming BERosome repair complex with DNA replication proteins at the replicating genome: Dominant negative function of the CTD. *J. Biol.Chem.*, **290**, 20919-20933.

Hegde, M. L., Izumi, T. & Mitra, S. *et al.* (2012) Oxidized Base Damage and Single-Strand Break Repair in Mammalian Genomes: Role of Disordered Regions and Post translational Modifications in Early Enzymes. *Progress in Molecular Biology and Translational Science*, **110**, 1236-153.

Hewitt, G., Jurk, D,. Marques, F. D. M.,Correia-Melo, C., Hardy, T., Gackowska, A., Anderson, R., Taschuk, M., Mann, J., Passosa, J. F. (2012) Telomeres are favoured targets of a persistent DNA damage response in ageing and stress-induced senescence. *Nat. Commun.,* **3**, 708.

Henderson, T., Delaney, C., Gu, F., Tannenbaum, S., Essigmann, J. (2002) Oxidation of 7,8-dihydro-8-oxoguanine affords lesions that are potent sources of replication errors *in vivo*. *Biochemistry*, **41**, 914–921.

Hildrestrand1, G., Neurauter1, C., Diep, D. *et al.* (2009) Expression patterns of Neil3 during embryonic brain development. *BMC Neuroscience*, **10**, 2202-2210.

Hailer, MK., Slade, PG., Martin, BD.,. Rosenquist, TA.,. Sugden, KD. (2005) Recognition of the oxidized lesions spiroiminodihydantoin and guanidinohydantoin in DNA by the mammalian base excision repair glycosylases NEIL1 and NEIL2. *DNA Repair (Amst.)*, 41–50.

Hitomi K, Iwai S, Tainer JA. (2007) The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair. *DNA Repair*, **6,** 410–428.

Hoeijmakers, J. (2009) DNA Damage, aging, and cancer. *New England Journal of Medicine*, 361, 1475-85.

Houtgraaf, J. H., Versmissen, J., Giessen W.J.V. (2006) A concise review of DNA damage checkpoints and repair in mammalian cells. *Cardiovasc. Revasc. Med.*, **7**, 165-172.

Hu, J., de Souza-Pinto, N.C., Haraguchi, K., Hogue, B.A., Jaruga, P., Greenberg, M.M., Dizdaroglu, M., Bohr, V.A. (2005) Repair of formamidopyrimidines in DNA involves different glycosylases: role of the OGG1, NTH1, and NEIL1 enzymes, *J. Biol. Chem.* **280**, 40544–40551.

Hsu, G.W., Ober, M., Carell, T., Beese, L.S. (2004). Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature*, 431, 217–221.

Hung, R. J., Hall, J., Brennan, P., Boffetta, P. (2005) Genetic polymorphisms in the base excision repair pathway and cancer risk: a HuGE review, *Am. J. Epidemiol*, **162,** 925–942.

Izumi, T., Wiederhold, L.R., Roy, G., Roy, R., Jaiswal, A., Bhakat, K.K., Mitra, S., Hazra, T.K. (2003) Mammalian DNA base excision repair proteins: their interactions and role in repair of oxidative DNA damage. *Toxicology,* **193,** 43–65.

Jaruga, P., Birincioglu, M., Rosenquist, T.A., Dizdaroglu, M. (2004) Mouse NEIL1 protein is specific for excision of 2,6-diamino-4-hydroxy-5-formamidopyrimidine and 4,6-diamino 5 formamidopyrimidine from oxidatively damaged DNA. *Biochemistry,* **43,** 15909–15914.

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,** 1053–1055.

Jena, N. & Mishra, P.C. (2013) Is FapyG mutagenic?: Evidence from the FT study. *Chem. Phys. Chem.*, **14**, 3236-3270.

Koketsu, S., Watanabe, T., Nagawa, H. (2004) Expression of DNA repair protein: MUTYH, NTH1 and MTH1 in colorectal cancer, *Hepato-Gastroenterology,* **51,** 638–642

Klattenhoff, A.W., Thakur, M., Chu, C. S., Ray, D., Habib, S.L., Kidane, D. (2017) Loss of NEIL3 DNA glycosylase markedly increases replication associated double strand breaks and enhances sensitivity to ATR inhibitor in glioblastoma cells. *Oncotarget,* **8,** 112942-112958.

Krokan, H.E. & Bjørås, M. (2013) Base excision repair. *Cold Spring Harbor Perspectives in Biology,* **5**, a012583.

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,** 160-164.

Krokeide, S.Z., Laerdahl, J.K., Salah, M. *et al.* (2013) Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroiminodihydantoin and guanidinohydantoin. *DNA Repair*, **12**, 1160- 1164

Lee, C.Y., Delaney, J.C.,  Kartalou, M., *et al.* (2009) Recognition and processing of a new repertoire of DNA substrates by human 3-methyladenine DNA glycosylase (AAG). *Biochemistry* **48** 1850–1861.

Lindahl, T. (1974) An *N-Glycosidase* from *Escherichia coli* That Releases Free Uracil from DNA Containing Deaminated Cytosine Residues. *Proc. Natl. Acad. Sci. USA,* **9**, 3649-3653.

Liu, M., Bandaru, V., Holmes, A., Averill, A.M., Cannan, W., & Wallace, S.S. (2012). Expression and purification of active mouse and human NEIL3 proteins. *Protein Expr Purif.,* **84**, 130-139.

Liu, M., Bandaru, V., Bond, JP. *et al*. (2010) The mouse ortholog of NEIL3 is a functional DNA glycosylase *in vitro* and *in vivo*. *Proc. Natl Acad. Sci. USA*, **107**, 4925–4930.

Liu, M., Doublié, S. & Wallace, S. (2013) NEIL3 the final frontier for the DNA glycosylases that recognize oxidative damage. *Mutation Research*, **743,** 4-11.

Liu, W F., Yu, S S., Chen, G J., & Li, Y Z. (2006). DNA damage checkpoint, damage repair, and genome stability. *Acta Genetica Sinica*, **33**, 381-390.

Luna, L. (2013). Loss of Neil3, the major DNA glycosylase activity for removal of hydantoins in single stranded DNA, reduces cellular proliferation and sensitizes cells to genotoxic stress. *Biochim. Biophys. Acta-Molecular Cell Research*, **1833**, 1157-1164.

Markkanen, E., Dorn, J. & Hübscher, U. (2013) MUTYH DNA glycosylase: the rationale for removing undamaged bases from the DNA. *Frontiers in Genetics*, **4**, 1-20.

Martin, P.R., Couvé, S.,  Zutterling, C.,  Albelazi, M.S.,  Groisman, R., Matkarimov, B.T., Parsons, J.L., Elder, R.H., Saparbaev, M.K. (2017) The human DNA glycosylases NEIL1 and NEIL3 excise psoralen-induced DNA-DNA cross-links in a four-stranded DNA structure. *Scientific Reports,* **7,** 17438.

Maynard, S., Schurman, S.H., Harboe, C., de Souza-Pinto, N.C., & Bohr, V A. (2009). Base excision repair of oxidative DNA damage and association with cancer and aging. *Carcinogenesis*, **30,** 2-10.

Massaad, M., Zhou, J., Tsuchimoto, D. *et al.* (2016) Deficiency of base excision repair enzyme NEIL3 drives increased predisposition to autoimmunity. *Journal of clinical Investigation*, **126**, 4219-4236.

McAuley-Hecht KE, Leonard GA, Gibson NJ, Thomson JB, Watson WP, Hunter WN, Brown T. (1994).  Crystal structure of a DNA duplex containing 8-hydroxydeoxyguanine-adenine base pairs. *Biochemistry*, **33**:10266–10270.

McCullough, A. K., Dodson, M. L,. Lloyd, R. S. (1999) Initiation of base excision repair: glycosylase mechanisms and structures, *Annu. Rev. Biochem.* **68** 255–285

Meeker, A.K., Hicks, J.L., Platz, E.A., March, G.E., Bennett, C.J., Delannoy, M.J., De Marzo, A.M. (2002). Telomere shortening is an early somatic DNA alteration in human prostate tumorigenesis. *Cancer Res,* **62,** 6405-6409.

Morgan, M.T., Bennett, M.T., Drohat, A.C. (2007) Excision of 5-halogenated uracils by human thymine DNA glycosylase. Robust activity for DNA contexts other than CpG,  *J. Biol. Chem.* **282,** 27578–27586.

Melamede, R.J., Hatachat, Z., Kow, Y.W., Ide, H. & Wallace, S.S. (1994)  Isolation and characterization of endonuclease VIII from *Esherichia coli*. *Biochemistry,* **33**, 1255-1264.

Morland, I., Rolseth, V., Luna, L., Rognes, T., Bjoras, 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. *Nucleic Acids Res*, **30**, 4926– 4936.

Naegeli, H. (1997). *Mechanisms of DNA damage recognition in mammalian cells*. Landes Bioscience, Austin USA

Nickoloff, J. A. & Hoekstra, M. F. (1998). *DNA damage and repair*. Humana Press, Totowa New Jersey USA

Niles, J.C.,  Wishnok, J.S.,  Tannenbaum, S.R. *et al.* (2001)  Spiroiminodihydantoin is the major product of the 8-oxo-7,8-dihydroguanosine reaction with peroxynitrite in the presence of thiols and guanosine photooxidation by methylene blue, *Org.Lett*, **3**, 963–966

Palm,W., and de Lange, T. (2008) How Shelterin Protects Mammalian Telomeres. *Annual Review of Genetics,* **42,** 301-334.

Palles, C., Cazier, J., Howarth, K. *et al.* (2013) Germline mutations in the proof-reading domains of *POLE* and *POLD1* predispose to colorectal adenomas and carcinomas. *Nature Genetics*, **45**, 136-144.

Regnell, C., Hildrestrand, G., Sejersted, Y. *et al.* (2012) Hippocampal adult neurogenesis is maintained by Neil3-dependent repair of oxidative DNA lesions in neural progenitor cells. *Cell Rep*, **2**, 503-510.

Rudolph, P., Alm, P., Olsson, H., Heidebrecht, H.J., Ferno, M., Baldetorp, B., Parwaresch, R. (2001) Concurrent overexpression of p53 and c-erbB-2 correlates with accelerated cycling and concomitant poor prognosis in node-negative breast cancer. *Hum. Pathol,* **32,** 311–319.

Prakash, A., & Borgstahl, G.E. (2012) The Structure and Function of Replication Protein A in DNA Replication. The Eukaryotic Replisome: a Guide to Protein Structure and Function. *Subcell. Biochem.* **62,** 171-196

Prasad R, Beard WA, Strauss PR, Wilson SH. Human DNA polymerase beta deoxyribose phosphate lyase. Substrate specificity and catalytic mechanism. *J Biol Chem*, **273,** 15263–15270.

Poulsen HE. (2005). Oxidative DNA modifications. *Exp Toxicol Pathol*, **57,** 161–169.

Rolseth, V., Krokeide, S. Z., Kunke, D., Neurauter, C. G., Suganthan, R., Sejersted, Y., Hildrestrand, G. A., Bjøras, M., Luna, L. (2013). Loss of Neil3, the major DNA glycosylase activity for removal of hydantoins in single stranded DNA, reduces cellular proliferation and sensitizes cells to genotoxic stress. *Biochim. Biophys. Acta,* **1833,** 1157–1164.

Roy, LM., Jaruga, P., Wood, T.G., McCullough, AK., Dizdaroglu, M., Lloyd, R.S. (2007) Human polymorphic variants of the NEIL1 DNA glycosylase, *J. Biol. Chem*, **282**, 15790–15798.

Russo, MT., De Luca, G., Degan, P. & Parlanti, E. (2004) Accumulation of the oxidative base lesion 8-hydroxyguanine in DNA of tumor-prone mice defective in both. *Cancer Res*. **64,** 4411–4.

Sampath, H., McCullough, A. K., Lloyd, R. S. (2012) Regulation of DNA glycosylases and their role in limiting disease, *Free Radic. Res.* **46,** 460–478.

Serre L, Pereira de Jesus K, Boiteux S, Zelwer C, Castaing B. (2002) Crystal structure of the lactococcus lactis formamidopyrimidine-DNA glycosylase bound to an abasic site analogue-containing DNA. *EMBO J*. **21,** 2854–2865.

Sejersted, Y., Hildrestrand, GA., Kunke, D. *et al.* (2011) Endonuclease VIII-like 3 (Neil3) DNA glycosylase promotes neurogenesis induced by hypoxia-ischemia. *Proc Natl Acad Sci U S A*, **108**, 18802–18807.

Sjolund, A. A., Senejani,A. G., Sweasy, J. B., (2013)  MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles, *Mutat. Res.* **743-744**. 12-25

Skarpengland, T., Laugsand L E. & Janszky, I. (2015) Genetic variants in the DNA repair gene NEIL3 and the risk of myocardial infarction in a nested case–control study. The HUNT Study. *DNA Repair (Amst).* **28**, 21–27.

Sleeth, K. M., Robson, R. L., Dianov, G. L., *et al.* (2004) Exchangeability of Mammalian DNA Ligases between Base Excision Repair Pathways. *Biochemistry*, **43 (40),** 12924–12930.

Slupphaug, G., Mol, C.D., Kavli, B., Arvai, A.S., Krokan, H.E., Tainer, J.A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*, **92**, 384-87

Takao, M., Oohata, Y., Kitadokoro, K. et al. (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,* **14,** 261-270.

Thiviyanathan, V., Somasunderam, A., Volk, D.E. *et al.* (2008) Base-pairing Properties of the Oxidized Cytosine Derivative, 5-Hydroxy Uracil. *Biochem. Biophys. Res. Commun.,* **366**, 752-757.

Wang, J. C. & Gautier, J. (2010) The Fanconi anemia pathway and ICL repair: implications for cancer therapy. *Crit. Rev. Biochem. Mol. Biol.,* **45**, 424-439.

Wallace, S.S. (2013). DNA Glycosylases Search for and Remove Oxidized DNA Bases. *Environ Mol Mutagen*, **54**, 691-704.

Wallace, S., Murphy, D. & Sweasy, J. (2012). Base excision repair and cancer. *Cancer Letters,* **327**, 73-89.

Wallace, S. (2014). Base excision repair: A critical player in many games. *DNA repair*, **19,** 14-26.

Wiederhold, L., Leppard, J B., Kedar, P., *et al*. (2004) AP endonuclease independent DNA base excision repair in human cells, *Mol. Cell*. **15,** 209–220.

Wiederholt, C. J., Patro, J.N., Jiang, Y. L., Haraguchi, K. Greenberg, M. M., (2005) Excision of formamidopyrimidine lesions by endonucleases III and VIII is not a major DNA repair pathway in *Escherichia coli*, *Nucleic Acids Res*. **33** 3331–3338.

Wilson, D.M. & Bohr, V.A. (2007) The mechanics of base excision repair and its relationship to aging and disease. *DNA Repair (Amst)*, **6**, 544-559.

Wyatt, M.D. (2013) Advances in understanding the coupling of DNA base modifying enzymes to processes involving base excision repair, *Adv. Cancer Res.* **119**, 63–106.

Xie, Y., Yang, H., Cunanan, C., Okamoto, K. *et al*. (2004) Deficiencies in mouse MUTYH and Ogg1 result in tumor predisposition and G to T mutations in codon 12 of the K-ras oncogene in lung tumors. *Cancer Res*. **64,** 3096–102.

Yoon, J., Bhatia, G., Prakash, S. & Prakash, L. (2010) Eror-free replicative bypass of thymine glycol by the combined action of DNA polymerases κ and ζ in human cells. *Proc Natl Acad. Sci. USA,* **107**, 14116-14121.

Zhang, Q.-M., Yonekura, S.-I., Takao, M., Yasui, A., Sugiyama, H., & Yonei, S. (2005). DNA glycosylase activities for thymine residues oxidized in the methyl group are functions of the hNEIL1 and hNTH1 enzymes in human cells. *DNA Repair,* **4,** 71-79.

Zhou, J., Liu, M., Fleming, AM., Burrows, CJ., Wallace, SS. (2013). Neil3 and NEIL1 DNA glycosylases remove oxidative damages from quadruplex DNA and exhibit preferences for lesions in the telomeric sequence context. *J. Biol. Chem.* **288**, 27263-27272.

Zhou, P.T., Li, B., Ji, J., Wang, M.M., Gao, C.F. (2015) A systematic review and metaanalysis of the association between OGG1 Ser326Cys polymorphism and cancers, *Med. Oncol,* **32,** 472.

Zhou, J., Chan, J., Lambelé, M. *et al*. (2017) Neil3 repairs telomere damage during S phase to secure chromosome segregation at mitosis. *Cell Reports.* **20,** 2044-2056.
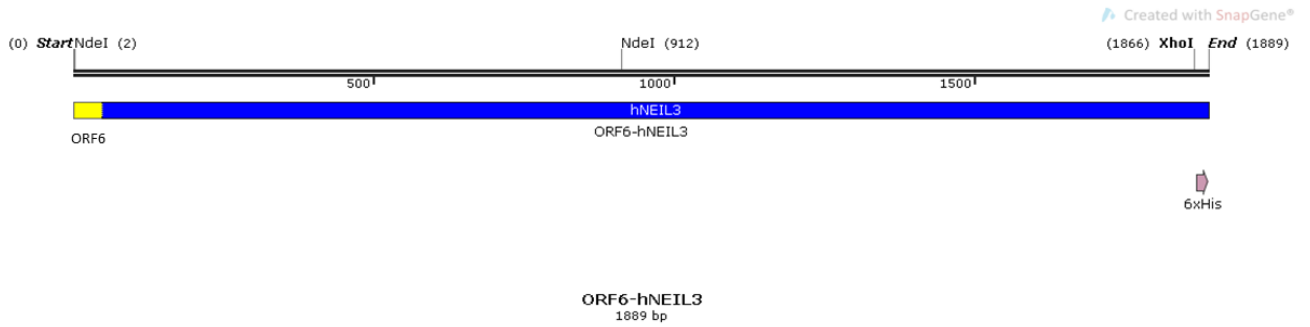
# 7   Appendix

## 7.1   Human Neil3-FL Sequence results

CATATGAAAATCGAAGCAGGTAAACTGGTACAGAAGGAGATTAACTGATGGTGGAAGGAC

CAGGCTGTACTCTGAATGGAGAGAAGATTCGAGCGCGGGTGCTCCCGGGCCAGGCGGTGACCG

GCGTGCGGGGAAGCGCTCTGCGGAGTCTGCAGGGCCGCGCCTTGCGGCTCGCAGCCTCCACG

GTTGTGGTCTCCCCGCAGGCTGCTGCACTGAATAATGATTCCAGCCAGAATGTCTTGAGCCTGTT

TAATGGATATGTTTACAGTGGCGTGGAAACTTTGGGGAAGGAGCTCTTTATGTACTTTGGACCAA

AAGCTTTACGGATTCATTTCGGAATGAAAGGCTTCATCATGATTAATCCACTTGAGTATAAATATAA

AAATGGAGCTTCTCGTGTTTTGGAAGTGCAGCTCACCAAAGATTTGATTTGTTTCTTTGACTCATC

AGTAGAACTCAGAAACTCAATGGAAAGCCAACAGAGAATAAGAATGATGAAAGAATTAGATGTAT

GTTCACCTGAATTTAGTTTCTTGAGAGCAGAAAGTGAAGTTAAAAAACAGAAAGGCCGGATGCTA

GGTGATGTGCTAATGGATCAGAACGTATTGCCTGGAGTAGGGAACATCATCAAAAATGAAGCTCT

CTTTGACAGTGGTCTCCACCCAGCTGTTAAAGTTTGTCAATTAACAGATGAACAGATCCATCACCT

CATGAAAATGATACGTGATTTCAGCATTCTCTTTTACAGGTGCCGTAAAGCAGGACTTGCTCTCTC

TAAACACTATAAGGTTTACAAGCGTCCCAATTGTGGTCAGTGCCACTGCAGAATAACTGTGTGCC

GCTTTGGGGACAATAACAGAATGACATATTTCTGTCCTCACTGTCAAAAGAAAATCCTCAACATG

TTGACATATGCAAGCTACCGACTAGAAATACTATAATCAGTTGGACATCTAGCAGGGTGGATCAT

GTTATGGACTCCGTGGCTCGGAAGTCGGAAGAGCACTGGACCTGTGTGGTGTGTACTTTAATCAA

TAAGCCCTCTTCTAAGGCATGTGATGCTTGCTTGACCTCAAGGCCTATTGATTCAGTGCTCAAGA

GTGAAGAAATTCTACTGTCTTTAGCCACTTAATGAAGTACCCGTGTAATACTTTTGGAAAACCTC

ATACAGAAGTCAAGATCAACAGGAAAACTGCATTTGGAACTACAACTCTTGTCTTGACTGATTTTA

GCAATAAATCCAGTACTTTGGAAAGAAAAACAAAGCAAAACCAGATACTAGATGAGGAGTTTCAAA

ACTCTCCTCCTGCTAGTGTGTGTTTGAATGATATACAGCACCCCTCCAAGAAGACAACAAACGATA

TAACTCAACTATCCAGCAAAGTAAACATATCACCTACAATCAGTTCAGAATCTAAATTATTTAGTCC

AGCACATAAAAAACCGAAAACAGCCCACTACTCATCACCAGAGCTTAAAAGCTGCAACCCTGGAT

ATTCTAACAGTGAACTTCAAATTAATATGACAGATGGCCCTCGTACCTTAAATCCTGACAGCCCTC

GCTGCAGTAAACACAACCGCCTCTGCATTCTCCGAGTTGTGAGGAAGGATGGGGAAAACAAGGG

CAGGCAGTTTTATGCCTGTCCTCTACCTAGAGAAGCACAATGTGGATTTTTTTGAATGGGCAGATTT

GTCCTTCCCATTCTGCAACCATGGCAAGCGTTCCACCATGAAAACAGTATTGAAGATTGGACCTA

ACAATGGAAAGAATTTTTTTGTGTGTCCTCTTGGGAAGGAAAAACAATGCAATTTTTTCCAGTGGG

CAGAAAATGGGCCAGGAATAAAAATTATTCCTGGATGCTAACTCGAGCACCACCACCACCA

CCAC



ORF6-hNEIL3
1889 bp

hNEIL3 FL forward

>FR07649289

CTATAGGGGATTGTGAGCGGATAACAATTCCCCATCTTAGTATATTAGTTAAGTATAAGA

AGGAGATATACATATGAAAATCGAAGCAGGTAAACTGGTACAGAAGGAGATTAACTGAT

GGTGGAAGGACCAGGCTGTACTCTGAATGGAGAGAAGATTCGAGCGCGGGTGCTCCCGG

GCCAGGCGGTGACCGGCGTGCGGGGAAGCGCTCTGCGGAGTCTGCAGGGCCGCGCCTTG

CGGCTCGCAGCCTCCACGGTTGTGGTCTCCCCGCAGGCTGCTGCACTGAATAATGATTCC

AGCCAGAATGTCTTGAGCCTGTTTAATGGATATGTTTACAGTGGCGTGGAAACTTTGGGG

AAGGAGCTCTTTATGTACTTTGGACCAAAAGCTTTACGGATTCATTTCGGAATGAAAGGC

TTCATCATGATTAATCCACTTGAGTATAAATATAAAAATGGAGCTTCTCGTGTTTTGGAAG

TGCAGCTCACCAAAGATTTGATTTGTTTCTTTGACTCATCAGTAGAACTCAGAAACTCAAT

GGAAAGCCAACAGAGAATAAGAATGATGAAAGAATTAGATGTATGTTCACCTGAATTTA

GTTTCTTGAGAGCAGAAAGTGAAGTTAAAAAACAGAAAGGCCGGATGCTAGGTGATGTG

CTAATGGATCAGAACGTATTGCCTGGAGTAGGGAACATCATCAAAAATGAAGCTCTCTTT

GACAGTGGTCTCCACCCAGCTGTTAAAGTTTGTCAATTAACAGATGAACAGATCCATCAC

CTCATGAAAATGATACGTGATTTCAGCATTCTCTTTTACAGGTGCCGTAAAGCAGGACTT

GCTCTCTCTAAACACTATAAGGTTTACAAGCGTCCCAATTGTGGTCAGTGCCACTGCAGA

ATAACTGTGTGCCGCTTTGGGGACAATAACAGAATGACATATTTCTGTCCT

hNEIL3-FLength reverse

>FR07649198

CTACACAGTGAGATCTCGCAATTCCGAAAAATAAACAATATGACTCCGTCCTCGGTGGGA

TACTTAAAGCATTACGGCCCATCTTACTATAGTACATGTGACTGTTAGGCTAGCACTCAA

GGCTTATAGGATTCAATGATCAAAGACAGAAGAAAGATTTTATCGTACATTGCCCACGT

CAGTGAAGGGCTCGGGTGATTATGTTTGAAAACCTCGATGCAGATGTCGAGATCAACAAT

GACAGTACAATTGGAATCTACAGCTCTGGTATTTGATGATTTTAGGCAATAAATGCGATA

TTTAGTAAGGCAAGACACAGCAAAGCCCGTTATATGATGAGGAGTTCCAAGACTCTCCTC

GAGCTAGGGTGTGGTTGGATGAGATCCAGCACCCATCCAAGAAGACAACAAACGATATA

ACTCACCTGTCCAGCAAAGTAAACATATCACGGACAATCAGTTCAGAATATAAATCATTT

AGTCCAGCACAAAAAAAACCGAAAACAGCCCAGTACTCATCACCAGAGATTAAAAGCTG

CATCCACGGATATTCTAACAGTGAACTTCAAATTAATATGACGGATGGCCCTCGTACCTT

AAATCCTGACAGCCCTCGCTGCAGTAAACACAACCGCCTCTGCATTCTAGGAGTTGTGAG

GAAGGATGGGGAAAACAAGGGCAGGCAGTTTTATGCCTGTCCTCTGCCTAGAGAAGCAC

AACGTGGATTTTTTGAATGGGCAGATTTGTCCTTCCCTTTCTGCAACCAAGGCAAGAGTTC

CACCATGATAACAGTATTGAAGATTGGCCGTAACAATGGAAAGAATTTTTTTGTGTGTCC

TCTTGGGAGGGAAAAACAATGCAATTTTTTCCAGTGGGCAGAAAATGGGCCAGGAATAA

AAAGTATTCCTGGACGCCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACA

AAGCCTTTAAAGTATTATTGGGTTTTC


hNEIL3-1506 forward

>FR07649287

TAGGGGATTGTGAGCGGATAACAATTCCCCATCTTAGTATATTAGTTAAGTATAAGAAGG

AGATATA<mark>CATATG</mark>AAAATCGAAGCAGGTAAACTGGTACAGAAGGAGATTAACTGATGGT

GGAAGGACCAGGCTGTACTCTGAATGGAGAGAAGATTCGAGCGCGGGTGCTCCCGGGCC

AGGCGGTGACCGGCGTGCGGGGAAGCGCTCTGCGGAGTCTGCAGGGCCGCGCCTTGCGG

CTCGCAGCCTCCACGGTTGTGGTCTCCCCGCAGGCTGCTGCACTGAATAATGATTCCAGC

CAGAATGTCTTGAGCCTGTTTAATGGATATGTTTACAGTGGCGTGGAAACTTTGGGGAAG

GAGCTCTTTATGTACTTTGGACCAAAAGCTTTACGGATTCATTTCGGAATGAAAGGCTTCA

TCATGATTAATCCACTTGAGTATAAATATAAAAATGGAGCTTCTCGTGTTTTGGAAGTGC

AGCTCACCAAAGATTTGATTTGTTTCTTTGACTCATCAGTAGAACTCAGAAACTCAATGG

AAAGCCAACAGAGAATAAGAATGATGAAAGAATTAGATGTATGTTCACCTGAATTTAGTT

TCTTGAGAGCAGAAAGTGAAGTTAAAAAACAGAAAGGCCGGATGCTAGGTGATGTGCTA

ATGGATCAGAACGTATTGCCTGGAGTAGGGAACATCATCAAAAATGAAGCTCTCTTTGAC

AGTGGTCTCCACCCAGCTGTTAAAGTTTGTCAATTAACAGATGAACAGATCCATCACCTC

ATGAAAATGATACGTGATTTCAGCATTCTCTTTTACAGGTGCCGTAAAGCAGGACTTGCT

CTCTCTAAACACTATAAGGTTTACAAGCGTCCCAATTGTGGTCAGTGCCACTGCAGAATA

ACTGTGTG

hNEIL3 (1506) reverse

>FR07649197

TCAGTTTCTTTTGACGAACACGTGTTGACTTATCGTACGGATCGCACGGGCTGTTTGTACA

GGCTTCAGCTCCAATTGCATGTCAGCATGCGTCATGGATGCCAGTCGTGTTTACTGCATAT

CTGCTTCGTTCATTTCAGCATGCTCTCTTACCACGTGCTGTAAAAGCAGGTCTTGCTCTAT

CTTAACCATATAATGATTACTAAGTGTCTCACTTGTGATCAGTACATACCGGATTATCTTG

TGAGTCTTCTTTGAGAACAATTAACAGAATGACATAGTTCTGTGCTCACTGTCACAAAGA

AATATCATAAACATGTTA<mark style="background-color: yellow; color: red;">CATATG</mark>CAAGCTACACGAGTAGAAATACCTATAATCAGTTGG

ACATCTAACCAGGGTTGGATCATGTTATGGAATCCGTGGCTCGGAAGTCGGAAGAGCTCT

GGACTCTGTGTGGTGTGTACTTTAATCAATAAGCCCTCTTATAAGGCCATGTGATGCTTGC

TTGACCTCAAGGCCTATTGATTCAGTGCTCAAGAGTGAAGAAAATTCTACTGTCTTTAGCC

ACTTAATGAAGTACCCGTGTAATACTTTTGGAAAACCTCATACAGAAGTCAAGATCAACA

GGAAAACTGCATTTGGAACTACAACTCTTGTCTTGACTGATTTTAGCAATAAATCCAGTA

CTTTGGAAAGAAAAACAAAGCAAAACCAGATACTAGATGAGGAGTTTCAAAACTCTCCT

CCTGCTAGTGTGTGTTTGAATGATATACAGCACCCCTCCAAGAAGACAACAAACGATATA

ACTCAACTATCCAGCAAAGTAAACATATCACCTACAATCAGTTCAGAATCTAAATTATTT

AGTCCAGCACATAAAAAACCGAAAACAGCCCACTACTCATCACCAGAGCTTAAAAGCTG

CAACCCTGGATATTCTAACAGTGAACTTCAAATTAATATGACAGATGGCCCTCGTACCTT

AAATCCTCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAA

GAGTATTGCCCCCGAACCCCAAAAAACAACCTT

hNEIL3-1044 forward

>FR07649286

ATAGGGGATTGTGAGCGGATAACAATTCCCCATCTTAGTATATTAGTTAAGTATAAGAAG

GAGATATA<mark>CATATG</mark>AAAATCGAAGCAGGTAAACTGGTACAGAAGGAGATTAACTGATGG

TGGAAGGACCAGGCTGTACTCTGAATGGAGAGAAGATTCGAGCGCGGGTGCTCCCGGGC

CAGGCGGTGACCGGCGTGCGGGGAAGCGCTCTGCGGAGTCTGCAGGGCCGCGCCTTGCG

GCTCGCAGCCTCCACGGTTGTGGTCTCCCCGCAGGCTGCTGCACTGAATAATGATTCCAG

CCAGAATGTCTTGAGCCTGTTTAATGGATATGTTTACAGTGGCGTGGAAACTTTGGGGAA

GGAGCTCTTTATGTACTTTGGACCAAAAGCTTTACGGATTCATTTCGGAATGAAAGGCTTC

ATCATGATTAATCCACTTGAGTATAAATATAAAAATGGAGCTTCTCGTGTTTTGGAAGTG

CAGCTCACCAAAGATTTGATTTGTTTCTTTGACTCATCAGTAGAACTCAGAAACTCAATGG

AAAGCCAACAGAGAATAAGAATGATGAAAGAATTAGATGTATGTTCACCTGAATTTAGTT

TCTTGAGAGCAGAAAGTGAAGTTAAAAAACAGAAAGGCCGGATGCTAGGTGATGTGCTA

ATGGATCAGAACGTATTGCCTGGAGTAGGGAACATCATCAAAAATGAAGCTCTCTTTGAC

AGTGGTCTCCACCCAGCTGTTAAAGTTTGTCAATTAACAGATGAACAGATCCATCACCTC

ATGAAAATGATACGTGATTTCAGCATTCTCTTTTACAGGTGCCGTAAAGCAGGACTTGCT

CTCTCTAAACACTATAAGGTTTACAAGCGTCCCAATTGTGGTCAGTGCCACTGCAGAATA

ACTGTGTGCCGCTTTGGGGACAATAACAGAATGACATATTTCTGTCCTCACTGTCAAAAA

AAAATCCTCA

hNEIL3-1044 reverse

>FR07649196

GGAGTGCAGGCTCACCTAAGATTTTATTGATTTCTTTGCAGTCAGCAGCAGAATTCAGAA

ATGTCAATGATAAGCCTCCAGAGACTAGAGAGTGATGACAGATTCAGATGTATGTTCACA

TGAATTTAGTTTCAAGAGAGCAGAAAGTGAAGTTAAAAAACAGAAAAGGCCAGATGAAA

GGTGATGTGCTAATGGATCGGAACGTATTGCCCGGAGTAGGGAACATCATCAAAAATGA

AGCTCTCTTTGACAGTGGTCTCCACCTAGCTGTTAAAGTTTGTCAATTAACAGATGAACAG

ATCCATCACCTCATGAAAATGATACGTGATTTCAGCATTCTCTTTTACGGGTGCCGTAAAG

CAGGACTTGCTCTCTCTAAACACTATAAGGTTTACAAGCGTCCCAATTGTGGTCAGTGCC

ACTGCAGAATAACTGTGTGCCGCTTTGGGGACAATAACAGAATGACATATTTCTGTCCTC

ACTGTCAAAAAGAAAATCCTCAACATGTTGA<mark style="background-color:yellow">CATATG</mark>CAAGCTACCGACTAGAAATACTA

TAATCAGTTGGACATCTAGCAGGGTGGATCATGTTATGGACTCCGTGGCTCGGAAGTCGG

AAGAGCACTGGACCTGTGTGGTGTGTACTTTAATCAATAAGCCCTCTTGTAAGGCATGTG

ATGCTTGCTTGACCTCAAGGCCTATTGATTCACTCGAGCACCACCACCACCAACCAGTGA

GATCCGGCTGCTAACAAAGCCCCGGAACGCTAAAAGGGGGTCCCCCC


hNEIL3-843 forward

>FR07649194

GATTGTGAGCGGATAACAATTCCCCATCTTAGTATATTAGTTAAGTATAAGAAGGAGATA

TA<mark style="background-color:green">CATATG</mark><span style="color:purple">AAAATCGAAGCAGGTAAACTGGTACAGAAGGAGATTAACTGAT</span><span style="color:green">GGTGGAAG</span>

<span style="color:green">GACCAGGC</span>TGTACTCTGAATGGAGAGAAGATTCGAGCGCGGGTGCTCCCGGGCCAGGCG

GTGACCGGCGTGCGGGGAAGCGCTCTGCGGAGTCTGCAGGGCCGCGCCTTGCGGCTCGC

AGCCTCCACGGTTGTGGTCTCCCCGCAGGCTGCTGCACTGAATAATGATTCCAGCCAGAA

TGTCTTGAGCCTGTTTAATGGATATGTTTACAGTGGCGTGGAAACTTTGGGGAAGGAGCT

CTTTATGTACTTTGGACCAAAAGCTTTACGGATTCATTTCGGAATGAAAGGCTTCATCATG

ATTAATCCACTTGAGTATAAATATAAAAATGGAGCTTCTCGTGTTTTGGAAGTGCAGCTC

ACCAAAGATTTGATTTGTTTCTTTGACTCATCAGTAGAACTCAGAAACTCAATGGAAAGC

CAACAGAGAATAAGAATGATGAAAGAATTAGATGTATGTTCACCTGAATTTAGTTTCTTG

AGAGCAGAAAGTGAAGTTAAAAAACAGAAAGGCCGGATGCTAGGTGATGTGCTAATGGA

TCAGAACGTATTGCCTGGAGTAGGGAACATCATCAAAAATGAAGCTCTCTTTGACAGTGG

TCTCCACCCAGCTGTTAAAGTTTGTCAATTAACAGATGAACAGATCCATCACCTCATGAA

AATGATACGTGATTTCAGCATTCTCTTTTACAGGTGCCGTAAAGCAGGACTTGCTCTCTCT

AAACACTATAAGGTTTACAAGCGTCCCAATTGTGGTCAGTGCCACTGCAGAATAACTGTG

TGCCGCTTTGGGGACCATAACAGAATGACATATTTCTGTCCTCACTGTCAAAAACTCGAG

CACCACCACCACCACCACTGAAATCCGGCTGCTAACAAAGCCCGAAAGGAAAC


hNEIL3-843 reverse

>FR07649195

GAGAAGATTCGAGCGCGGGTGCTCCCGGGCCAGGCGGTGACCGGCGTGCGGGGAAGCGC

TCTGCGGAGTCTGCAGGGCCGCGCCTTGCGGCTCGCAGCCTCCACGGTTGTGGTCTCCCC

GCAGGCTGCTGCACTGAATAATGATTCCAGCCAGAATGTCTTGAGCCTGTTAATGGATA

TGTTTACAGTGGCGTGGAAACTTTGGGGAAGGAGCTCTTTATGTACTTTGGACCAAAAGC

TTTACGGATTCATTTCGGAATGAAAGGCTTCATCATGATTAATCCACTTGAGTATAAATAT

AAAAATGGAGCTTCTCGTGTTTTGGAAGTGCAGCTCACCAAAGATTTGATTTGTTTCTTTG

ACTCATCAGTAGAACTCAGAAACTCAATGGAAAGCCAACAGAGAATAAGAATGATGAAA

GAATTAGATGTATGTTCACCTGAATTTAGTTTCTTGAGAGCAGAAAGTGAAGTTAAAAAA

CAGAAAGGCCGGATGCTAGGTGATGTGCTAATGGATCAGAACGTATTGCCTGGAGTAGG

GAACATCATCAAAAATGAAGCTCTCTTTGACAGTGGTCTCCACCCAGCTGTTAAAGTTTG

TCAATTAACAGATGAACAGATCCATCACCTCATGAAAATGATACGTGATTTCAGCATTCT

CTTTTACAGGTGCCGTAAAGCAGGACTTGCTCTCTCTAAACACTATAAGGTTTACAAGCG

TCCCAATTGTGGTCAGTGCCACTGCAGAATAACTGTGTGCCGCTTTGGGGACAATAACAG

AATGACATATTTCTGTCCTCACTGTCAAAAACTCGAGCACCACCACCACCACCACTGAGA

TCCGGCTGCTAACAAAGCCCGA