University of Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

MASTER'S THESIS

Study program/specialization: Biological chemistry	Autumn/Spring semester, 2020/2021 Open access
Author: Celine Lorentsen	
Supervisor(s): Svein Bjelland	
Title of master's thesis: Replacing Pro240 with Gly in hSMUG1 increases uracil excision in R-loop but abolish 5-hydroxymethyluracil excision in ssDNA	
Credits: 60 stp.	
Keywords: DNA damage DNA glycosylases hSMUG1 P240G hSMUG1 S26R/E35D Excision BER	Number of pages: 64 +supplemental material/other: 104 Stavanger, 23.07.2021

Acknowledgments

I would like to express eternal gratitude to my supervisor, Professor Svein Bjelland, for accepting me as his student for the second time, for always believing in me, and for being available for supervising around the clock. I highly appreciate his extraordinary knowledge and his detailed-oriented mindset. I would like to express my deepest gratitude to my laboratory supervisor, Dr. Marina Alexeeva, for her invaluable patience, kindness, and not least her ability to contribute in several areas. I will thank her for challenging me in the lab and making me more independent, I will always remember the support she provided me during this year. I would like to offer a special thank you to Professor Peter Ruoff for being understanding during a difficult year and guiding me through complicated kinetics. I would like to extend my sincere thanks to Dr. Dugassa Nemie-Feyissa for his assistance and helpful advice in the laboratory. I wish to acknowledge Professor Kåre Bredeli Jørgensen for plenty of valuable suggestions and patience when supervising me in constructing organic structures and engineer Eli Drange Vee for helping me with mathematical challenges. A special thank you to the group of Professor Hilde Nilsen (UiO) for the gift of enzymes that were crucial in this project. I would like to thank my laboratory partners, Shimaa Mehanna, Cyrell Ann Ruales, and Trond Bærheim for their social and mental support, countless coffee breaks, and for making this year memorable. Finally, I would like to acknowledge the support provided by my family, I want to thank my mother and sister for always having faith in me, and my nephews for reminding me of the importance of laughter and play. A warm thank you to my dear father, his past advice always guides me in the present. And deep gratitude is expressed to my loved one for serving me love, mental support, and nutritious dinner 365 days this year.

Abstract

Deamination and oxidation of pyrimidines are two major spontaneously damaging events in DNA. The DNA glycosylase SMUG1 initiates repair of deaminated or oxidized bases by breaking the glycosidic bond between the damaged base and deoxyribose, however, has been less studied than most similar enzymes.

In the present study, we investigated the glycosylase activity of hSMUG1 P240G and S26R/E35D mutant proteins, in addition to commercial and in-house purified wild type enzyme, using DNA oligonucleotide substrates of different structure [single-stranded (ss), bubble, and R-loop] with uracil (U) inserted at a specific site, and ssDNA with 5-hydroxymethyluracil (hmU) inserted at a defined site. We demonstrate, for the first time, that hSMUG1 P240G and S26R/E35D retained the glycosylase activity for all substrates except hmU in ssDNA, and showed higher relative activity for uracil in R-loop than bubble DNA. In contrast, wild type hSMUG1 was highly active for hmU in ssDNA and demonstrated higher activity for uracil in bubble than R-loop DNA.

Table of contents

List of figures

List of tables

Abbreviations

UDG: uracil DNA-glycosylase

hSMUG: human single-strand selective monofunctional uracil-DNA glycosylase

hUNG: human uracil-*N*-glycosylase

ssDNA: single-stranded DNA

dsDNA: double-stranded DNA

AP site: apurinic/apyrimidinic site

APE: apurinic/apyrimidinic endonuclease

ROS: reactive oxygen species

BER: base excision repair

C, G, A, T, U: cytosine, guanine, adenine, thymine, uracil

hpmU: 5-(hydroperoxymethyl)uracil

hU: 5-hydroxyuracil

hmU: 5-hydroxymethyluracil

fU: 5-formyluracil

FU: 5-fluorouracil

cU: 5-carboxyuracil

ClU: 5-chlorouracil

εC: 3,N^4-ethenocytosine

alx: alloxan

ida: isodialuric acid

O: oxanine

NHEJ: non-homologous end joining

CSR: class switch recombination

SHM: somatic hypermutation

AD: antibody diversification

AID: activation-induced cytosine deaminase

APOBEC: apolipoprotein B mRNA-editing enzyme, catalytic polypeptide

Pol: polymerase

Ig: immunoglobulin

C_H: heavy chain

BCR: B cell receptor

1 Introduction

1.1 Endogenous base damages in DNA and their repair

Endogenous mutagenic agents, *i.e.* water, alkylating agents, and reactive oxygen species (ROS) generated from normal metabolic processes, constantly attack and damage cellular DNA including its bases. An insufficient response to DNA damage can have devastating consequences and are considered a major factor in mutagenesis, carcinogenesis, aging, and neurological diseases (1-6). The base excision repair (BER) pathway corrects the majority of endogenous base damages in DNA, thus being responsible for maintaining genome stability for long-term cell survival (7-10).

1.1.1 Deamination of cytosine to uracil

Adenine (A), guanine (G), Cytosine (C), and 5-methylcytosine (mC) contain reactive amine sites that may be lost in a hydrolytic deamination reaction. Cytidines spontaneously deaminate in water at a slightly higher rate compared to both adenosine and guanosine (2) A water molecule can perform a nucleophilic attack at C4 within the DNA base cytosine, resulting in spontaneous deamination and generation of the abnormal DNA base uracil (U) (Figure 1). The deamination mechanism includes the cleavage of the C4-N4 bond, protonation of N3, and

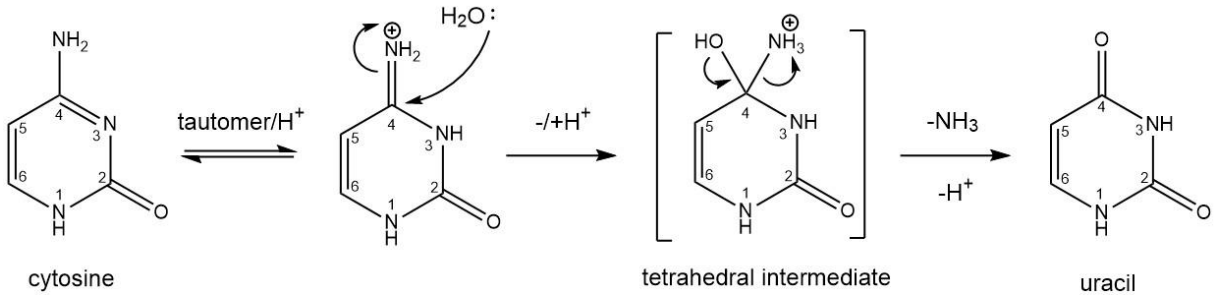


Figure 1. The mechanism of hydrolytic deamination of cytosine into the abnormal uracil base in DNA. The bases are illustrated in anti-conformation, which is typical for standard Watson-Crick base pairing At low pH, NH_2^+ protonation occurs following deprotonation of water. The structure is constructed using ChemDraw Professional 20.1 (PerkinElmer).

elimination of ammonia (10). Deamination of cytosine to uracil generates a premutagenic U:G mispair and the uracil lesion (which in RNA pairs with A) is read as thymine (T) if not repaired prior to replication, causing a G:C → A:T transition mutation (Figure 2). Indeed, this mutation is frequently found in human cancers (11,12). Next to hydrolytic depurination, deamination of

cytosine is considered the major spontaneous damaging event in DNA (13), estimated to occur 60–500 times a day in human cells (3). Uracil is the most prevalent non-canonical base in DNA, it is also a natural intermediate in the synthesis of thymidine, thus uracil can also be generated by misincorporation of dUMP instead of dTMP opposite A during DNA replication (Figure 2), creating a non-mutagenic and less harmful U:A pair (1,10,11,14). Besides, mutations can be caused by translesion DNA polymerases after uracil removal, by insertion of an incorrect base opposite the abasic or apurinic/apyrimidinic (AP) site (4,14).

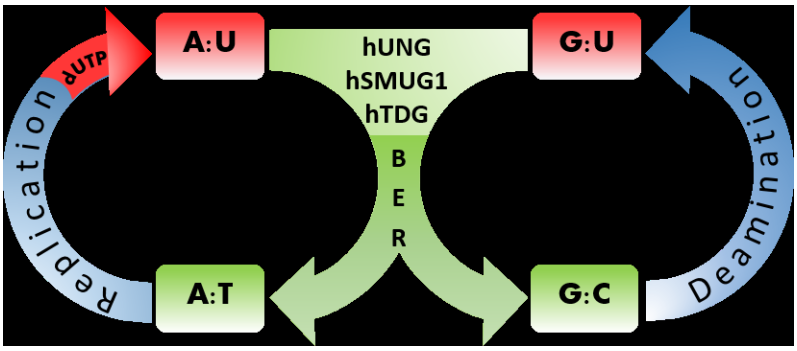


Figure 2. Introduction of uracil in DNA. Deamination of cytosine to uracil in DNA causes the G:C to A:T transition (ringht), while introduction through dUTP is non-mutagenic (left).

As opposed to spontaneous deamination of cytosine where uracil must be faithfully repaired prior to replication, enzymatic deamination of cytosine at the immunoglobulin (Ig) loci initiates the physiological process of antibody diversification (AD) and depends on the uracil lesion not being faithfully repaired at certain sites. AD is initiated by antigen stimulation which activates B cells to perform antibody class switching by intrachromosomal recombination (CSR) within special G-rich tandem repeats named switch regions (S regions). The recombination is initiated by the B-cell-specific protein activation-induced cytidine deaminase (AID), a single-strand DNA-specific enzyme that converts C to U in the donor and acceptor regions of the subsequent end-joining DNA fragments (15). Uracil is primarily excised by uracil-DNA *N*-glycosylase (UNG) (16-22) and the resulting AP site is incised by an AP endonuclease (APE) (23). However, a recent unpublished study (24) has suggested that UNG itself might be able to incise the DNA backbone (Figure 3) without any involvement of an APE. This may explain the significant CSR activity found in cells deficient in both APE1 and APE2, which are the enzymes likely to perform this activity according to the classical model (23,25). Another uracil-DNA glycosylase (UDG), single-strand selective monofunctional UDG 1 (SMUG1), partially substitutes for UNG and assist in antibody diversification by excising uracil

in somatic hypermutation patterns (26). After the AP sites have been incised, MMR proteins are engaged in transforming the single-stranded (ss) breaks into double-stranded (ds) breaks. Eventually, the ends of the two S regions are suitable for the non-homologous end joining (NHEJ) system, resulting in the production of IgG, IgA, or IgE without changing antigen specificity (15). Another important process caused by B cell activation initiated by AID is somatic hypermutation (SHM). The two processes are both dependent on T cell/antigen stimulation and occur in the germinal center of secondary lymphoid organs (spleen, lymph nodes, and tonsils). While CSR produces antibody isotypes (*e.g.*, switches IgM to IgG), SHM produces B cell receptors and antibodies with high antigen affinity by introducing point/missense mutations (and occasionally deletions or insertions) into immunoglobulin variable (V) regions (27,28).

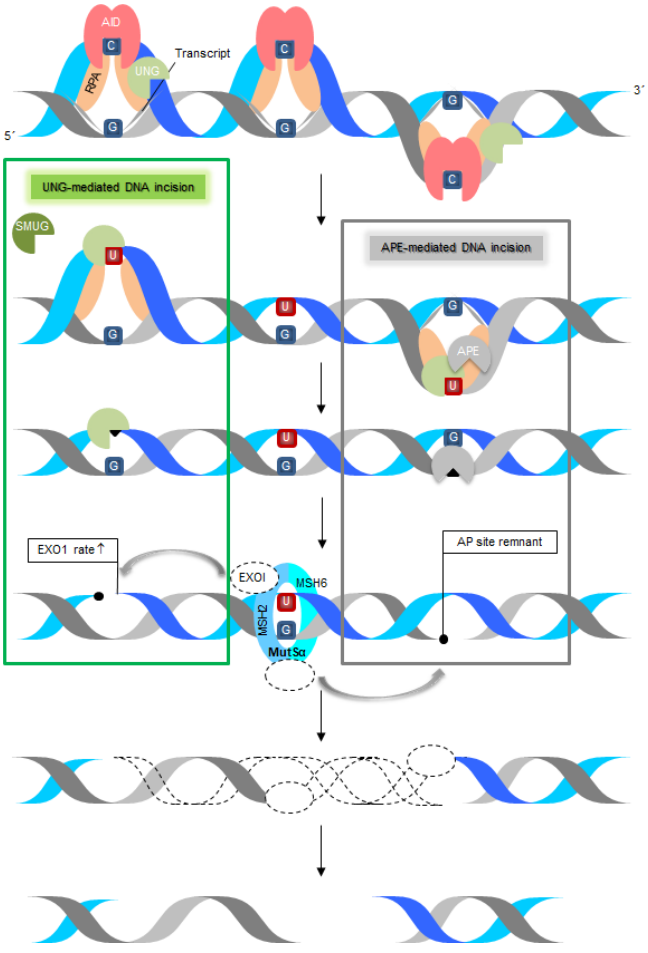


Figure 3. UNG-mediated DNA incision in CSR initiated by AID. (Courtesy of S. Bjelland).

The origin of the most common mutation in cancer (G:C → A:T transition) (12) has remained elusive. However, characteristic mutational signatures of cytosine deaminases in the APOBEC-family (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide family) have been identified in many cancers and are the second most common mutational signature after aging (29), suggesting deamination of cytosine to uracil as a common initiating event for this prevalent mutation. It has been shown higher genomic uracil levels in B-cell lymphoma cell lines in comparison to non-lymphoma cancer cell lines, and the levels were highly correlated with AID mRNA. Also, AID knockdown significantly reduced genomic uracil content, indicating that AID-induced U:G mismatches may be a primary cause of mutations in B-cell malignancies. The same study also reported an inverse correlation of genomic uracil level with excision activity and expression of the two UDGs SMUG1 and UNG (12).

1.1.2 Oxidation of pyrimidines

Reactive oxygen species (ROS) are formed in all aerobic organisms as a by-product of cellular metabolism. Also, exogenous sources like ionizing- and ultraviolet radiation as well as pollutants create ROS in cells. ROS oxidize several cell components including DNA, resulting in base lesions and strand breaks. Although base lesions generally induce minor distortions in the DNA helix, they may arrest DNA synthesis or transcription causing cell death. In aerobic organisms, oxidative DNA damage is a major cause of mutagenesis and cell death. Base oxidations occur in both DNA and the nucleotide pool. Several oxidized bases have been identified, and studies of the repair of specific lesions have been performed both *in vitro* and *in vivo* (5,8,9,30).

Oxidized DNA base products were first described for thymine, where the 5,6-double bond and the 5-methyl group are the major sites for ROS attack. Attack on its 5,6-double bond turns the planar aromatic ring into a non-planar structure, which may ring-open or fragment into different products. Both Endonuclease III-like protein 1 (hNth1) and hNeil1 exhibit activity against non-aromatic structures in DNA, while hSMUG1 is targeting structures with retained aromaticity. Free radical attack on the methyl group forms the unstable 5-hydroperoxymethyluracil (hpmU), which subsequently converts to 5-hydroxymethyluracil (hmU) and/or 5-formyluracil (fU) (Figure 8). Certain *E. coli* bacteriophages have replaced thymine with hmU in DNA, therefore, normal coding properties of this altered base were first suggested. This complied with studies (of both bacteria and mammalian cells) failing to

demonstrate mutagenicity or DNA synthesis arrest caused by hmU (8,30), and hmU glycosylase-deficient cells show normal and fertile phenotypes (31). Later studies have reported both toxicity and mutagenicity of hmdU to exponentially growing cells (32,33). For instance, the incorporation of both adenine and guanine opposite hmU in DNA gives rise to AT → GC transitions.

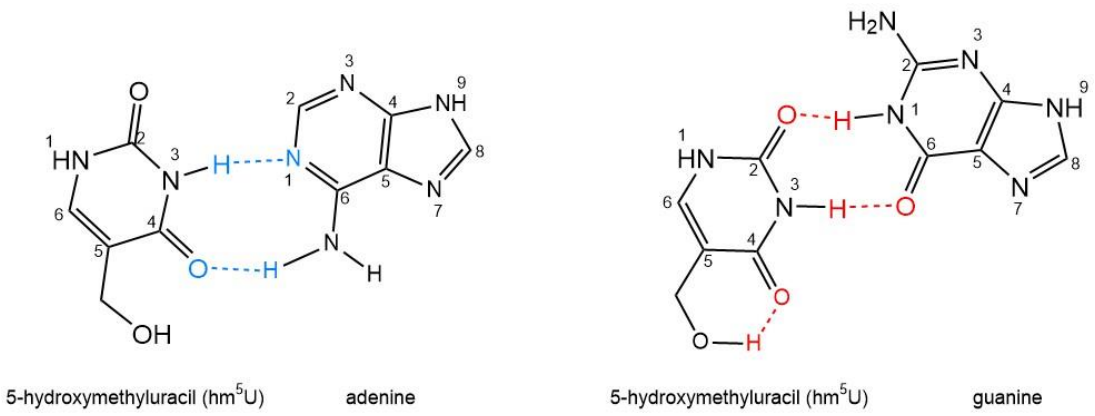


Figure 4. Illustration of hmU in base pair with adenine and guanine. Since hmU is a thymine analog, base pair with adenine will maintain the Watson-Crick geometry. hmU:G mismatch is a wobble base pair, stabilized by an intra-residue hydrogen bond (34,35). Blue dotted line, cognate base pairs; red dotted line, non-cognate base pairs. Constructed using ChemDraw Professional 20.1 (PerkinElmer).

The hmU:A base pair maintains the Watson-Crick geometry by a stabilizing inter-residue hydrogen bond between the 5-hydroxymethyl group and a nearby 5' guanine base, while the hmU:G mispair is a wobble base pair stabilized by an intra-residue hydrogen bond between the 5-hydroxymethyl group and its own O4 carbonyl group (Figure 4). hmU can also be formed by deamination of mC and subsequent oxidation of the methyl group (or in reverse order) in CpG sites generating a premutagenic hmU:G mispair, potentially resulting in G:C → A:T transitions by incorporation of dAMP opposite hmU (8,36). Although the activity for excision of hmU from DNA in mammalian cells was demonstrated 37 years ago (37), it was attributed to SMUG1 20 years ago (38). An *in vivo* proof that SMUG1 is the major DNA glycosylase for hmU is the almost complete loss of such activity in SMUG1-deficient mice (31). According to a study, the substrate preference for hmU removal is ssHmU>HmU:G>HmU:A (18). SMUG1 is the main glycosylase for removal of the oxidized thymine lesion, fU (39,40), excising the lesion from ssDNA and opposite all normal bases in dsDNA with the highest activity when opposite non-cognate C or T (8,41). Oxidation of the formyl group of fU produces cU which is also removed by SMUG1 (42), the damage has been synthetically produced and stably incorporated into DNA (8). In addition, cdU might arise in the cellular nucleotide precursor pool through oxidative

reactions (8). 5-hydroxyuracil (hU) derived from cytosine is another important oxidation product excised from DNA by SMUG1 (40).

In human cells, oxidized pyrimidines are mainly excised by hNTH1, hNEIL1, or hNEIL2, and oxidized purines by hOGG1. In contrast, SMUG1 seems to excise oxidized bases that are inefficiently removed by hNTH1, hNEIL1, and hOGG1 (9). The importance of hSMUG1 in removing hmU, fU, and hU from DNA has been demonstrated using HeLa cell extract, where almost all activity for these base lesions was neutralized by SMUG1 antibodies. In conclusion, SMUG1 is the primary repair enzyme for pyrimidines bearing an oxidized methyl group at ring C5 in both ss- and dsDNA (31,40).

1.1.3 Base excision repair

Living organisms have evolved various DNA repair mechanisms to protect the cellular genome from endogenous and exogenous attacks (5). The BER pathway (Figure 5), which is conserved from bacteria to humans, is a fundamental mechanism. The repair process is initiated by specific DNA glycosylases that together recognize many different base damages (Table 1). BER takes place by different mechanisms depending on the glycosylase used and the cells physiological state. In the major short-patch pathway, a single nucleotide gap is generated, whereas in the long-patch pathway two to ten nucleotides are removed. Usually, five steps are used to describe the BER pathway: base excision, DNA incision, end processing, gap-filling, and ligation (Figure 5) (3,18). The active site of the DNA glycosylases can only interact with an extrahelical base, therefore, in the first step of base excision, DNA glycosylases bind to the minor groove of DNA and the damaged base is flipped out of the major groove and into the active site of the glycosylase. In the DNA molecule, the DNA bases are covalently bonded to the C1 position of the deoxyribose sugar molecule. Monofunctional DNA glycosylases mainly activate a water molecule by particular amino acid residues to perform a nucleophilic attack of the C1 position of the damaged nucleotide, thus the damaged base is excised by hydrolysis in which the N-glycosidic bond between the base and the deoxyribose sugar molecule is cleaved, leaving an AP site (3).

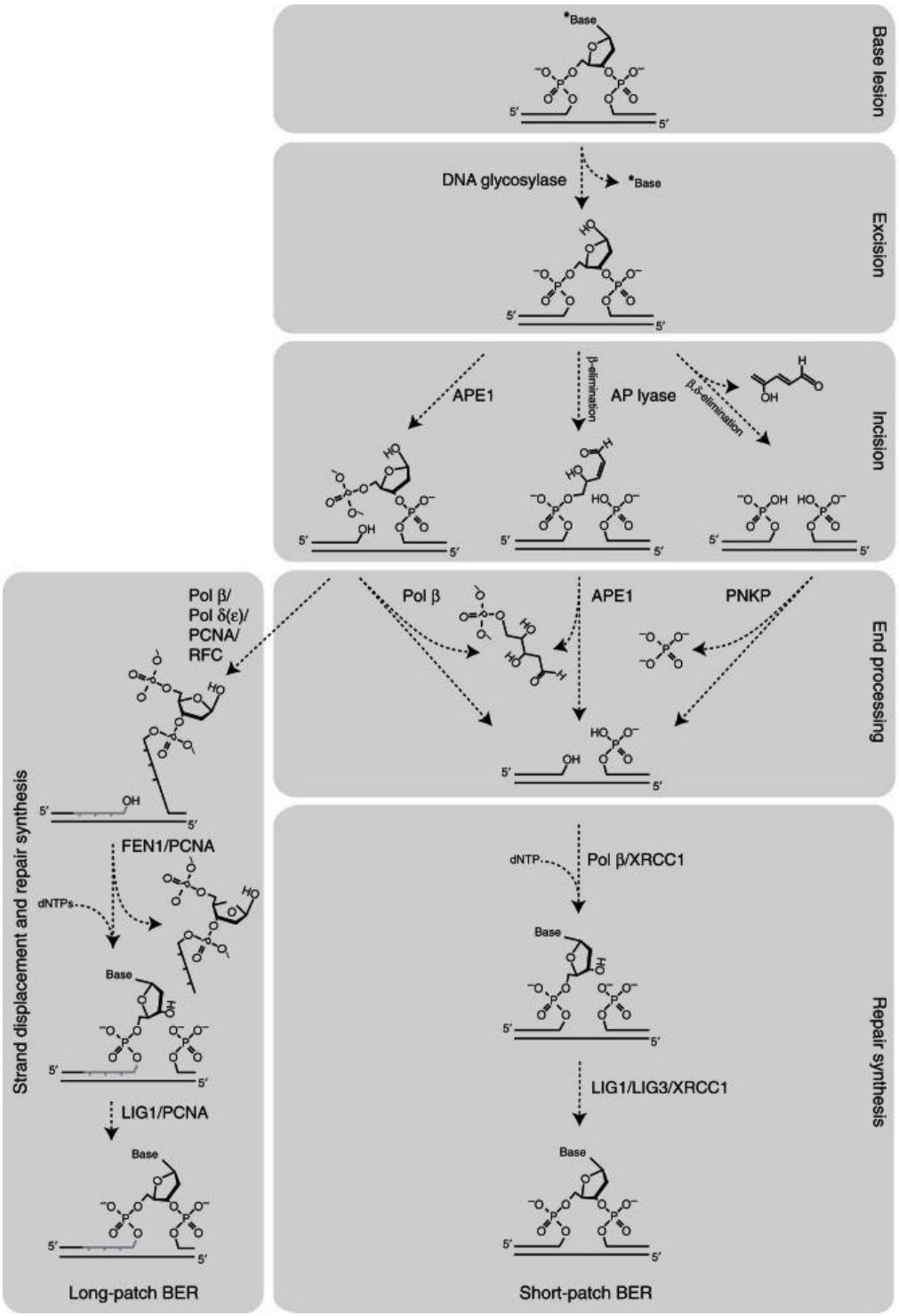


Figure 5. Pathways of base excision repair (BER). The repair process takes place in five core steps including excision of the damaged base, strand incision, end processing, and repair synthesis comprising gap filling and ligation. BER takes place by short-patch repair or long-patch repair (3).

The two next steps, incision and end processing, depending on whether the glycosylase is monofunctional or bifunctional. In the case of a monofunctional glycosylase, another enzyme, AP-endonuclease 1 (APE1), cleaves the DNA strand 5’ of the abasic site, generating 3’-OH and 5’-deoxyribose phosphate (5’-dRP) ends. The latter is then removed to make the desirable

substrate with a 5’ phosphate end. Bifunctional glycosylases have lyase activity in addition and two possible short-patch pathways are described. In one case, the glycosylase performs β-elimination generating an unsaturated hydroxyaldehyde linked to the 3’ end (3’-dRP) and the desired phosphate at the 5’ end. The 3’dRP is efficiently removed by APE1 that generates a 3’-OH end, preparing the intermediate for the polymerase step. In the other case, the glycosylase performs β,δ-elimination cleaving at both sides of the AP site, releasing the unsaturated deoxyribose. This generates a one nucleotide gap flanked by 3’-phosphate and 5’-phosphate ends. The 3’-phosphate is then removed by PNKP (3). In the fourth step, Pol β inserts the correct nucleotide (dCMP or dTMP) and generates the 5’-phosphate end required for the fifth termination step of BER; ligation by DNA ligase 3 aided by the scaffold protein XRCC1, resulting in a correct restored DNA sequence. The alternative long-patch pathway is generally active in proliferating cells, using mainly replication proteins. It requires pol ε or/and δ, proliferating cell nuclear antigen (PCNA) used as a sliding clamp and replication factor C (RFC). An endonuclease (FEN1) removes the displaced nucleotides, Pol β stimulates the repair synthesis and the long-patch pathway is terminated by nick sealing by ligase 1 (3).

1.1.3.1 Uracil-DNA glycosylases

DNA glycosylases are DNA repair enzymes capable of excising modified or mismatched bases, derived from alkylation, oxidation, or deamination, thereby initiating the BER pathway. DNA glycosylases are all substrate-selective with some overlapping specificities (Table 1). In mammals, 11 distinct DNA repair enzymes are known and SMUG1, UNG, thymine-DNA glycosylase (TDG), and methyl-CpG-binding domain protein 4 (MBD4) remove uracil from DNA (3,14). DNA glycosylases are found in all three domains of life and are generally well conserved from bacteria to humans (30), with some exceptions; the conservation is limited to the enzymatic core domain and the mammalian glycosylases have amino- and/or carboxy-terminal extensions compared to prokaryotic glycosylases (3). Some DNA glycosylases recognize base lesions both in base pairs (dsDNA) and in ssDNA (*e.g.* UNG, SMUG1, and Nei-like DNA Glycosylase 1 (NEIL1)). Importantly, the DNA glycosylases are able to recognize bases carrying minor modifications, hidden in undistorted DNA in the presence of a large excess of normal bases (3).

Table 1. Mammalian DNA glycosylases and their respective substrates and mouse knockout.

Enzyme	**Known substrates**	**Mouse knockout (3)**
UNG2[B]	ssU > U:G > U:A, fU in ssDNA and dsDNA, alx, hU, ida (3)	Partial defect in CSR, skewed SHM, B-cell lymphomas
UNG1[M]	Like UNG2 (3)	Unknown
SMUG1[B]	U:G > U:A > ssU (3,11,31,40,42,43) hmU (3,18,31,38,40,42,44), fU (39,40,42,45), hU (40,42,46) , εC (3,18,47), cU (42), ida (40,46), alx (40,46), X (48), O (49), ClU (42) , FU (18,42,50,51)	Viable and fertile, SMUG1/UNG/MSH triple knockout reduced longevity
TDG[M]	U:G > T:G, hmU in dsDNA, FU (3)	Embryonic lethal, epigenetic role in development
MBD4[M]	U:G, T:G, hmU in CpG context, εC, FU in dsDNA (3)	Viable and fertile, C to T transitions, intestinal neoplasia
MPG[M]	m^3A, m^7G, m^3G, Hx, εA (3)	Viable and fertile, triple knockouts in MPG/ AlkBH2/AlkBH3 hypersensitive to inflammatory bowel disease
OGG1[B]	oxoG:C, Fapy:C (3)	Viable and fertile, OGG1/ MUTYH double knockouts cancer prone
MUTYH[M]	A:oxoG/C/G (3)	OGG1/MUTYH double knockouts cancer prone
NTHL1[M]	Tg, FapyG, hC, hU in dsDNA (3)	Viable and fertile, NTHL1/NEIL1 double knockouts cancer prone
NEIL1[M]	Tg, FapyG, FapyA, oxoG, hU, DHU, Sp and Gh in ss and dsDNA (3)	Viable and normal at birth, obese after 7 months, NTHL1/NEIL1 double knockouts cancer prone
NEIL2[M]	Like NEIL1 (3)	Unknown
NEIL3[B]	FapyG, FapyA, Sp, and Gh in ssDNA (3)	Viable and fertile, memory and learning deficit

[M], monofunctional; [B], bifunctional; ss, single-stranded; ds, double-stranded; hmU, 5-hydroxymethyluracil; fU, 5-formyluracil; hU, 5-hydroxyuracil; εC, 3,N^4-ethenocytosine; cU, 5-carboxyuracil; ida, isodialuric acid; alx, alloxan; X, xanthine; O, oxanine; ClU, chlorouracil; Hx, hypoxanthine; εA, 1,N^6-ethenoadenine; oxoG, 8-oxoguanine; FapyG, 2,6-diamino-4-hydroxy-5-formamidopyrimidine; FapyA, 4,6-diamino-5-formamidopyrimidine; Tg, thymine glycol; hC, 5-hydroxycytosine; DHU, dihydrouracil; Sp, spiroiminodihydantoin; Gh, guanidinohydantoin.

1.1.3.2 SMUG1 incision

In hSMUG1 excision, the N-glycosidic bond cleavage is suggested to be an S_N1-like reaction generating a uracil anion and AP site C1' cation (Figure 7). The S_N1 reaction utilizes unfavorable atomic clashes in U-DNA as reaction energy, and because hSMUG1 carries the Asn85 unable to activate water (as opposed to the polar Asp145 seen in UNG), this mechanism might be suitable for hSMUG1. Also, observations indicate that hUNG uracil excision activity is more affected by the replacement of Asp145, compared to hSMUG1 activity when replacing Asn85. However, Asn85 in hSMUG1 is suggested to coordinate the reactive water molecule to attach the deoxyribose oxocarbenium ion via the activation of the uracil anion. The backbone of the carbonyl group of Asn96 seen in xSMUG when bound to uracil (Asn85 in hSMUG1) coordinates water by a hydrogen bond (52).

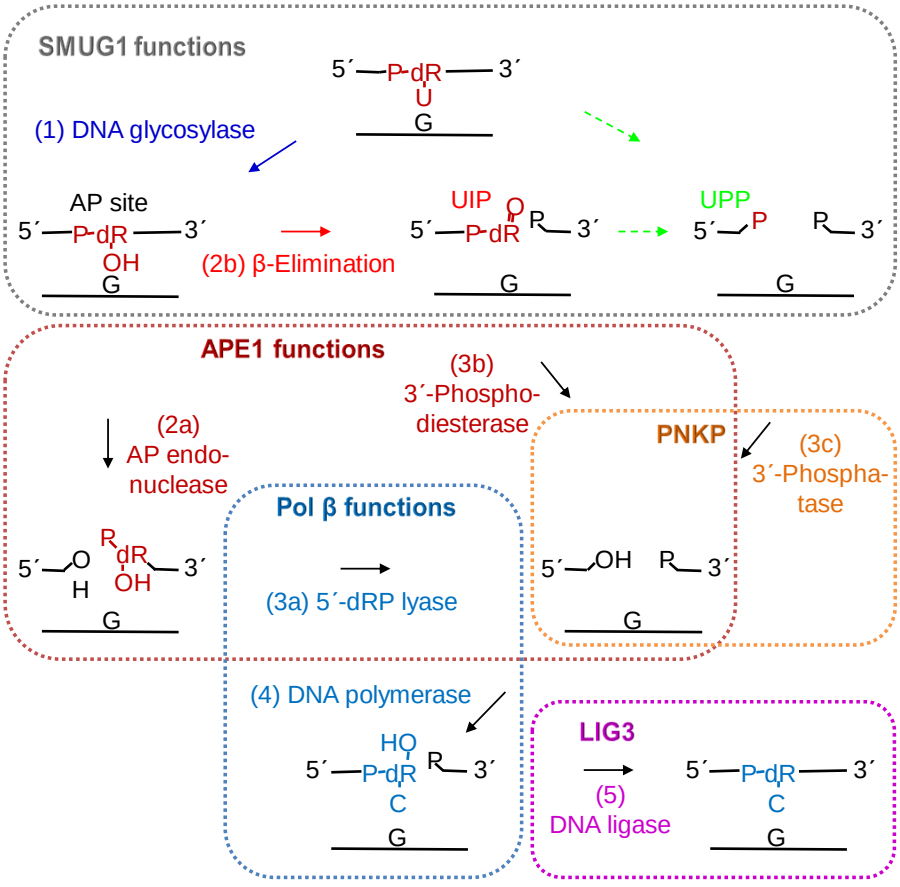


Figure 6. Suggested steps in the human BER pathway after SMUG1 recognition of uracil-DNA. First, SMUG1 excise uracil (step 1; blue), then the enzyme is eighter replaced by APE1 (dark red) for AP site incision (step 2a), or the enzyme itself incise the AP site (step 2b; red), leaving behind a 3'-α,β-unsaturated aldehyde (UIP). UIP can then be processed by APE1 (step 3b) producing a 3'-OH end, or alternatively processed (green broken arrows) producing a 3'-phosphate end (UPP). UPP is a substrate for the bifunctional PNKP (orange), generating the designated 3'OH end. The same 3'OH end is also generated by dRP-lyase activity (step 3a) by the repair DNA polymerase (Pol β; dark blue) after APE1 incision. The one nucleotide gap can then be filled with the correct dCMP (step 4) by the repair DNA polymerase β (Pol β; dark blue). The termination of BER is achieved by nick-sealing (step 5) by DNA ligase III (LIG3; purple). Removed residues, dark red; incorporated residues, dark blue; dR, deoxyribose (Alexeeva et al. 2019).

DNA glycosylases bind to labile AP sites with different strengths after base removal to prevent possible DNA strand breaks and collapse by premature hydrolytic cleavage. hSMUG1 competes with hAPE1 in AP site binding (53) and incise the phosphodiester backbone of uracil-DNA (Figure 6) (52). Because hSMUG1 exhibits a tight interaction (as opposed to hUNG) to the AP site and interact with both DNA strands (53,54), the induced strain might provide sufficient reaction energy. Incision by hSMUG1 is suggested to follow the bifunctional pathway (Figure 7), β-elimination occurs by deprotonation of C2' and formation of C1' enolate intermediate at the formyl group. The latter generates a 3'-α,β-unsaturated aldehyde designated uracil-DNA incision product (UIP), and a 5'-phosphate. The reaction mechanism of UIP is unclear and might be processed by hAPE1 resulting in a 3'OH end.

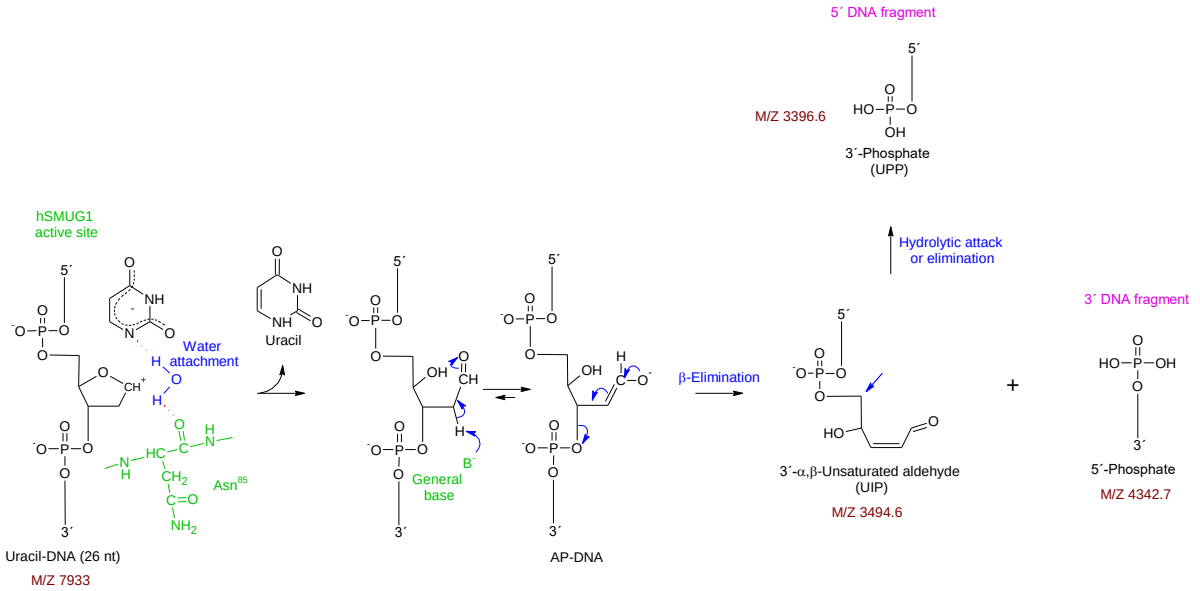


Figure 7. Proposed reaction mechanism for UIP and UPP generation in hSMUG1. Courtesy of Svein Bjelland.

Alternatively, hSMUG1 process UIP by generating a 3'-phosphate designated uracil-DNA processing product (UPP). UPP is then a substrate for the bifunctional polynucleotide kinase/phosphatase (PNKP), resulting in the desired 3'OH end that facilitates gap-filling (Figure 6). The same prime end also results from the traditionally AP site incision by APE1 after DNA polymerase β (Pol β) completes the dRP-lyase activity. The U-DNA incision activity by hSMUG1 is also suggested to be a part of CSR and SHM mechanisms (52). Both UIP and UPP are well-known products of bifunctional glycosylases and are shown *in vitro* to be substrates for hAPE1 and hPNKP respectively (55).

1.2 SMUG1

1.2.1 Structure and enzymology

SMUG1 repair base damages and initiates the fundamental BER pathway, thus protecting DNA against cytotoxicity and mutagenesis (18,51,53,56-60). Twelve different substrates have been allocated to SMUG1 (Figure 8). Current evidence suggests that SMUG1 is a major backup glycosylase for UNG in the removal of uracil in both U:G mismatches (11,31,40,42,43) and antibody diversification (26,31,60) and that both SMUG1 and UNG2 together coordinate the initial steps of BER (18,53,59). The uracil activity of UNG is significantly higher than the activity of SMUG1 (18). However, SMUG1 has a broader substrate specificity compared to UNG (Figure 8), being the primary repair enzyme for oxidized pyrimidines (31,40); it removes 5-methyl-oxidized thymine lesions like 5-hydroxymethyluracil (hmU) (18,31,38,40,42,44), 5-formyluracil (fU) (39,40,42,45) and 5-carboxyuracil (cU) (42). In addition, SMUG1 removes oxidized cytosine products, including 5-hydroxyuracil (hU) (40,42,46), isodialuric acid (ida) (40,46), and alloxan (alx) (40,46). Besides, the guanine deamination products xanthine (X) (48) and oxanine (O) (49), the highly mutagenic alkylated cytosine 3,N^4-ethenocytosine (εC) (18,47), the chemotherapeutic agent 5-fluorouracil (FU) (18,42,50,51), and the halogenated uracil 5-chlorouracil (ClU) (42) are all excised and removed by SMUG1. The substrate preference of SMUG1 is ssDNA > dsDNA (G and A pair) at low salt concentration, and dsDNA (G pair) > dsDNA (A pair) > ssDNA at a high salt concentration (58) and U > HmU >> εC > FU (18).

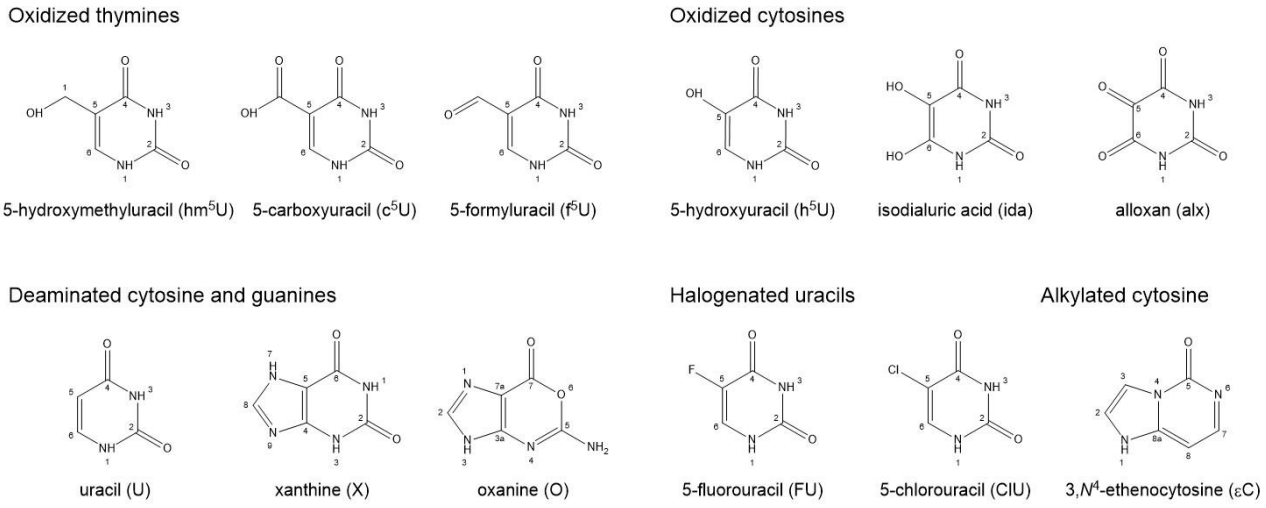


Figure 8. Substrates allocated SMUG1. Constructed using ChemDraw Professional 20.1 (PerkinElmer).

The crystal structure of hSMUG1 has not yet been achieved, but a detailed structure was predicted by homology modeling (61) using structure-based alignment of *Xenopus leavis* (African clawed frog) SMUG1 (xSMUG1) (47), the gram-negative metal-reducing proteobacterium *Geobacter metallireducens* SMUG1 (GmeSMUG1) (62), and alignment of binding and catalytic residues (Figure 9) (61). The vertebrate SMUG1 has the highest sequence similarity to bacteria and not insects, including 69.9 % similarity to GmeSMUG1 (53). SMUG1 (family III UDG), UNG (family I UDG), and TDG (family II UDG) belong to the same large (structural) UDG superfamily. The superfamily comprising six UDG families in total, in which five of them, included all mentioned, show <10 % overall sequence homology, but still share many of the same functions. They share a common fold of the core domains and an overall similar architecture, thus delicate structure differences make these enzymes exhibit different substrate preferences (63).

Figure 9. Sequence alignment of hSMUG1, xSMUG1, and GmeSMUG1. Motifs accountable for recognition using the intercalating loop (yellow), lesion binding (red), and catalysis (green) are depicted. Blue letters, Trp residues of hSMUG1; asterisks, identical residues; colons, conserved residues; dots, residues with some conserved physicochemical properties (61).

The active site of hSMUG1 includes the Arg243 residue in the N-terminal conserved in TDG/family II (rather than Asp found in family I) and the His239 residue in the C-terminal motif conserved in UNG/family I (rather than Asn found in family II). The base of the SMUG1 pocket has, like UNG, Asn as a major discriminatory factor in substrate specificity (rather than the extended hydrophobic pocket of TDG). In UNG, an aromatic residue appears as a steric

barrier to binding of bases modified at C5, SMUG1 (and TDG) possesses the small Gly instead. Thus, SMUG1 is a hybrid of two uracil glycosylase families and displays a more relaxed active site compared to the strict active site of UNG, which is nearly specific for uracil (47,64,65).

Mutational, structural, and kinetic studies have revealed an extensive network of contacts in the SMUG1-DNA catalytically active complex, and the functions of some amino acids have been suggested (Figure 10). The sequences His239–Lys249 are called the intercalating loop (66). The residues in the region penetrate DNA at the lesion site and are responsible for base damage recognition, base flipping, and stabilization of the base in an extrahelical state (64).

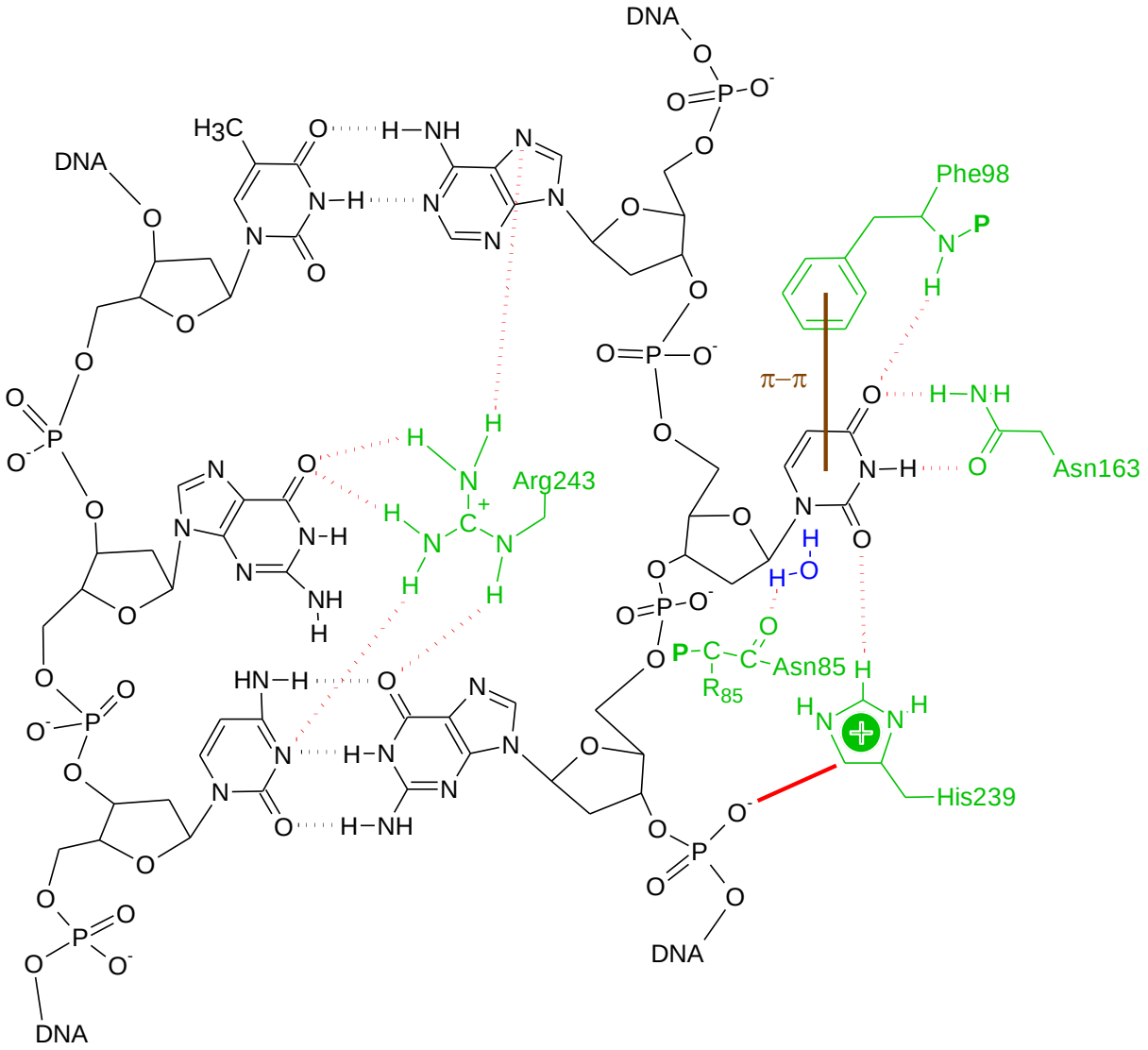


Figure 10. Active site interactions of hSMUG1. DNA structure in black; protein amino acid residues in green; catalytic water in blue; P, protein; R, amino acid residue. Red unbroken line, ionic interaction; red dotted line, hydrogen bond; black dotted line, H bond between cognate bases; brown unbroken line, van der Waals interaction. Courtesy of Svein Bjelland.

A pre-steady state kinetic analysis has outlined four steps involved in the interaction of hSMUG1 with DNA (Figure 11). Step 1 includes nonspecific initial DNA binding involving intercalating loop movement of the enzyme, DNA melting at the lesion site, and possibly damaged base flipping into an extrahelical state (61). By π–π stacking of Phe98 and hydrogen bonding of Asn163 to the flipped-out base, the enzyme discriminates between the pyrimidines and stabilizing them in the base-binding pocked. The region Gly87–Met91 recognizes the C5 substituent through water-bridged hydrogen bonds with uracil or direct hydrogen bonds with oxidated pyrimidines (hU, hmU, and fU). The movement of the intercalating loop is most likely associated with the "wedging" strategy for lesion search. It has been reported that unrelated *Escherichia coli* DNA glycosylases with different structures and specificity share this function in substrate discrimination, including insertion of a 3–4 amino acid "wedge" into DNA and forcing the lesion into a shallow binding pocket (67).

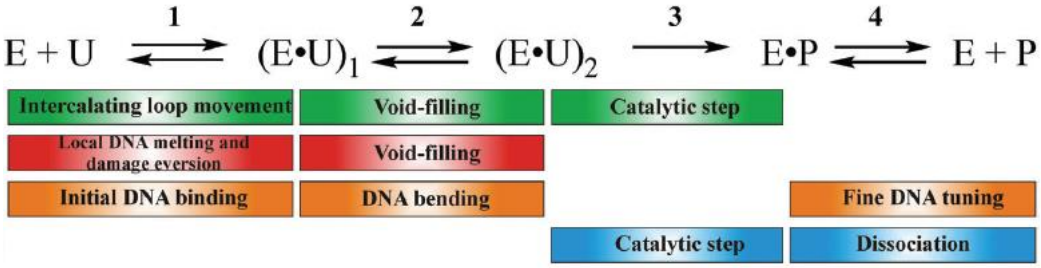


Figure 11. A four-step pre-steady-state kinetic scheme illustrating the interaction of hSMUG1 with DNA. Step 1 includes DNA binding, intercalating loop movement, DNA melting, and possibly base flipping. Step 2 comprises void filling. Step 3 includes hydrolysis of the N-glycosidic bond. Step 4 represents a conformational change of DNA and product release (61).

The crystal structure of the xSMUG1-DNA complex also reveals a wedge formed by a loop followed by a short α-helix, indicating that the C-terminal part of the wedge interacts with the distal strand of DNA suggesting an invasive interaction and an enhanced disruption and distortion of the DNA duplex compared to the complex formed by UNG (47,53). Step 2 of the model (Figure 11) comprises void filling by amino acids of the intercalating loop (61), and the responsible amino acid has been dedicated to Arg243 which forms a network of hydrogen bonds to adjacent bases (Figure 10) (64). In step 3, Asn85 and His239 catalyze the hydrolysis of the *N*-glycosidic bond connected to the damaged base in an S_N1-like reaction, and His239 processes the DNA backbone via electrostatic interaction with a phosphate group. Step 4 represents DNA conformational change and a slow product release (61,66). It has been established that the rate-limiting stage of the enzymatic process of SMUG1 is the release of the enzyme-product complex (66). hSMUG1 substrate selectivity has been suggested to partly depend on the size

of the 5-substituent. The apparent rate constant was demonstrated to decrease with increasing size for U, fU, and ClU when opposite G, but for the oxidation damaged pyrimidines, fU, cU, and hmU, the rate constant increased with increasing size. Also, The 5-substituents that easily interact with water are more likely to displace water molecules and further form hydrogen bonds with both amino acids and water molecules, hydrophilic compounds are therefore favorable. The results of the study also indicate that reduced duplex stability affects hSMUG1 discrimination by favoring unstable mispairs over paired bases (42).

1.2.2 Cellular functions

The protein SMUG1 is coded by the gene *SMUG1* and is located within a small region of chromosome 12 together with two other human UDGs (UNG and TDG). SMUG1 was first suggested to take place relatively late in the evolutionary perspective, and was only found in higher eukaryotes (vertebrates and insects); however, BLAST search reveals orthologs in a wide variety of species in both prokaryotes and eukaryotes, including proteobacteria, planctomycetes, and non-vertebrates. In some non-vertebrates and insects, SMUG1 is the only known UDG, indicating that it might be the only uracil-glycosylase in these species (53). SMUG1 shows low tissue specificity, with ubiquitous expression in the brain, skin, placenta, and 24 other tissues (NCBI). Many alternatively spliced transcript variants exist of this gene, including a human 0.7 kb transcript (in comparison to the 1.6 kb transcript coding the functional hSMUG1) with an open reading frame lacking a carboxy-terminal domain necessary for catalytic activity. Regulations of the *SMUG1* gene expression are linked to a 2000 base pair region, located upstream of the first transcribed (not translated) exon. It has been demonstrated that nucleotide excision repair (NER) proteins are involved in SMUG1 stimulation, including the global genome NER-initiating factor XPC and Cockayne syndrome group B (CSB) (68). SMUG1 is also able to remove uracil from nucleosomes, although less efficient when compared to naked DNA (69).

Unlike hUNG2, hSMUG1 is not cell cycle dependent, and is nearly constitutively expressed through the cell cycle, including in quiescent cells. hSMUG1 is predominantly localized in the nucleoli, whereas hUNG is excluded from this part of the cell and mainly accumulates in the S-phase at the replication foci (18,43). *Ung*-deficient mice lack the mutator phenotype (with increased spontaneous mutation frequency) seen in *ung*-deficient bacterial and yeast, and neutralizing antibodies have been used to identify SMUG1 as the major UDG in *Ung*-deficient mice (43). Several studies suggest that UNG2 has a major role in both pre-

replicative excisions of uracil (deaminated cytosine) and post-replicative repair of misincorporated uracil at the replication fork (18,70), while human SMUG1 inhibits cell proliferation (53). *In vitro*, hSMUG1 binds tightly and competes with hAPE1 for AP site-binding, while hUNG rapidly leaves the site and thereby stimulating AP site binding and cleavage by hAPE1. An important SMUG1 motif specific for the binding of the AP site has been identified in a previous study and mutations in this motif increased the catalytic turnover, while cell proliferation was no longer haltered. Their findings suggested that SMUG1 may be more important in the repair of deaminated cytosine in non-replicating than replicating cells (53). *Smug*$^{-/-}$ mice revealed accumulation of genomic hmU with the highest level in the brain (26-fold compared to wild type). The same mice also showed reduced uracil excision activity in brain extracts. *Ung*$^{-/-}$ mice revealed only a slight increase in uracil levels, but when *Smug*$^{-/-}$ and *Ung*$^{-/-}$ were combined, the uracil levels were increased up to 25-fold compared to wild type, indicating that SMUG1 prevents uracil accumulation and is an important uracil glycosylase also when functional UNG is present in cells. Whole genome sequencing (WGS) of tumors from *Ung/Smug1*-deficient mice showed accumulation of mutations, mainly GC → AT transitions at CpG dinucleotides (59).

Because radiation is cytotoxic to cells, it has been suggested that SMUG1 and UNG suppression might be of relevance to the efficacy of cancer radiation therapy. Another important feature by SMUG1 in cellular protection is the excision activity of FU to reduce drug cytotoxicity (50). FU has been extensively used as a treatment in many common cancers, due to its cytotoxic effect on cells, and it is suggested that FU is incorporated into DNA, resulting in FU accumulation and cytotoxicity. siRNA-mediated knockdown of Smug1 in *Ung*$^{-/-}$ mouse embryo fibroblasts reveal a hypersensitivity to FU treatment. Furthermore, SMUG1, but not UNG, remove FU from DNA to protect against cell killing. The mechanism is suggested to give drug resistance in tumors and is a predictive biomarker of drug response (71).

Mice studies indicate a role for SMUG1 in cancer (31). Correlative studies also suggest a role for SMUG1 in cancer by showing associations between SMUG1 expression and prognosis. Low SMUG1 expression is associated with breast cancer, including high-grade histological characterization of tumors, increased cell proliferation, and increased mutations in tumor suppressor genes. The study suggests that low SMUG1 levels predict response to adjuvant breast cancer therapy (72). A transcriptome-wide association study (TWAS) revealed an association between decreased genetically predicted expression of *SMUG1* and increased risk of pancreatic cancer (73). Inactivation mutations in the *SMUG1* gene are associated with a

poor prognosis in colon cancer (33 % survival compared to 90 % in patients without the mutations) and were found together with mutations in genes (APC, KRAS, PIK3CA, and p53) usually linked to colon cancer (74). High SMUG1 expression is associated with a good prognosis in ovarian cancer and *SMUG1* was one of the top three most significant genes in the prognostic signature used. However, increased *SMUG1* expression is frequently found in many cancers (bladder, gastric, breast, esophageal, and cervical) with poorly differentiated and chemoradiotherapy-resistant tumors (75).

1.2.2.1 RNA processing

There is growing evidence that DNA repair proteins, including SMUG1 (68), have moonlighting functions due to their additional involvement in RNA metabolism (56). Like DNA molecules, RNA molecules constantly undergo chemical modifications. While only a few functional modifications are identified in DNA, more than 140 types of RNA modifications are known to have regulatory functions. Other modifications are unclear in whether they are lesions or have functional roles. Some modifications become stable components of long-lived RNA, *e.g.* post-transcriptional pseudouridine modifications in human telomeric RNA component (hTERC) and ribosomal RNA (rRNA) (76). SMUG1 might be a moonlighting protein, by contributing to both DNA repair and RNA processing (Figure 12) (56).

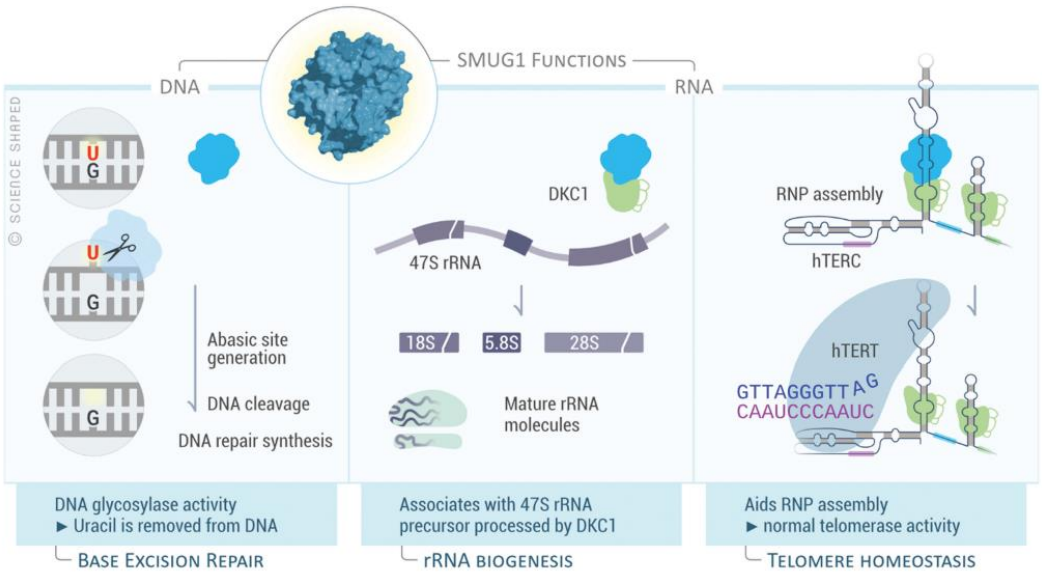


Figure 12. SMUG1 as a DNA repair glycosylase in BER and as an RNA processing enzyme in rRNA biogenesis and telomere homeostasis. Figure by Ellen Tenstad/Science Shaped (76).

SMUG1 interacts with the DKC1-containing H/ACA ribonucleoprotein. Site-directed mutagenesis showed that Glu29, Gly33, and Glu231 of hSMUG1 are necessary for DKC1 binding (44). DKC1 is the main dyskerin pseudouridine (Ψ) synthase in mammalian cells and is involved in the assembly of small nucleolar ribonucleoproteins (snoRNPs) and the biogenesis and posttranscriptional processing of many non-coding RNA molecules (*e.g.* rRNA and hTERC) (44,76,77). Due to the tremendous number of RNA molecules in human cells, enzymes potential involvement in genomic maintenance may have a large impact in a clinical context (56,78).

1.2.2.1.1 rRNA biogenesis

Ribosomal RNA (rRNA) is the most abundant cellular RNA molecule and defects in the assembly of the molecule have been linked to diseases (79). The modification 5-hydroxymethyluridine (hmrU) (the ribose variant of the SMUG1 substrate hmdU) is found in significant amounts in human rRNA and accumulation of this lesion has been reported to correlate with an increase of misprocessed and degradation marked (*i.e.* polyadenylated) rRNA and consequently reduced level of mature rRNA. Excision of 5-hydroxymethyldeoxyuridine (hmdU) by hSMUG1 in single-stranded RNA has been demonstrated *in vitro*. In SMUG1 depleted cells (in particular SMUG1 and DKC1 depleted), both hmrU accumulation in mature rRNA (*i.e.* 28S and 18S) and downregulated mature rRNA (*i.e.* 47S) has been shown, suggesting excision of hmrU from RNA by SMUG1 *in vivo* (Figure 12). Therefore, SMUG1 has been concluded as an hmrU regulator in the quality control of rRNA (76). SMUG1-DKC1 interaction has been demonstrated in both nucleoli and Cajal bodies; the former is the organelle where rRNA synthesis and processing take place, and the latter a region in the nucleus enriched with proteins and RNA molecules usually found in metabolically active cells (*e.g.* neurons) and proliferative cells (*e.g.* embryonic and tumor cells) (44).

1.2.2.1.2 Telomere maintenance

A recent study revealed that SMUG1 (or SMUG1-DKC1 interaction) also is required for regulating the amount of mature hTERC levels by regulating the levels of modified bases located between the CR4–CR5 domain and the H-box, a region enriched by SMUG1 substrates and important for binding to DKC1. hTERC serves as a template for the human telomerase reverse transcriptase (hTERT) in telomere lengthening at the ends of the chromosomes

(76,80,81). As shown in SMUG1-knockout cells, increased levels of base modifications in hTERC are related to reduced binding to DKC1, leading to a downregulated level of mature hTERC unable to provide sufficient telomerase activity. This results in dramatic telomere shortening, suggesting SMUG1 as a promoter for telomere homeostasis. The same study also observed no differences in transcription initiation or kinetics, suggesting SMUG1 activity in downstream steps (*e.g.* RNA degradation). Mature hTERC levels could not be obtained after silencing the RNA-degradation machinery, indicating SMUG1-activity in targeting hTERC to the RNA-degradation system (*e.g.* exosome or another degradation pathway) through increased base modification. The modification(s) is(are) not identified, although hmrU is most likely a candidate, due to the accumulation of this modification in SMUG1 depleted cells. Also, as mentioned, hmdU is recognized and processed by SMUG1 in human rRNA, implying that this modification might exist in other RNA molecules (*e.g.* hTERC) (76).

Studies support a dual role for SMUG1 in telomere maintenance, either as a DNA repair enzyme excising modified bases at telomers and initiating BER (81) or as an RNA processing enzyme regulating levels of hTERC. Reduced levels of hTERC were the limiting factor for telomerase activity, indicating RNA processing as the main function for SMUG1 in telomere homeostasis (76). Another study showed that smug1$^{-/-}$ mice exhibit telomerase deficiency and increased telomere defects, leading to inhibited bone marrow proliferation (56). Due to the dramatic telomere shortening in SMUG1-knockout cells, it would be highly interesting to understand its effect on cancer cells and whether SMUG1 depletion might limit cancer cell growth in certain circumstances. More research is needed to elucidate whether the cells rely on DNA repair or RNA processing in different conditions or species (76).

2 Aim of the project

Deamination and oxidation of bases are both major spontaneous damaging events frequently occurring in DNA, and many of the damages result in mutagenesis and cytotoxicity. Understanding the roles of DNA repair enzymes is crucial, nevertheless, many DNA repair enzymes, including human SMUG1, are not well understood.

This project aimed to purify recombinant *E. coli*-produced mutant proteins of the human SMUG1 enzyme, perform enzyme activity assays for different DNA substrates, and further analyze and possibly elucidate amino acid residues involved in catalysis. Many DNA repair proteins with multifunctional properties, like hSMUG1, are involved in diseases and might have undiscovered functions of high importance for cell homeostasis.

3 Materials and Methods

3.1 Enzyme

A modified pETM-11 plasmid encoding for the full length of mutated hSMUG1 (P240G or S241A), N-terminal His6-tag, T7/lac promoter, and kanamycin resistance was a kind gift from professor H. Nilsen (UiO). The plasmid was transformed into competent *Escherichia coli (E. coli)* BL21(DE3) host cells or Rosetta (a BL21(DE3) derivative), and the protein was expressed and purified (Appendix A). The His6-tag was cleaved using in-house produced and purified Tobacco Etch Virus (TEV) cysteine protease (Appendix B). To calculate the molecular weight of the mutants the Swiss Bioinformatics Resource Portal Expacy was used and one amino acid (*in casu*, P or A) from the FASTA sequence found in the UniProt Knowledgebase (Swiss-Prot) was substituted by another amino acid (*in casu*, G or A).

3.2 Oligonucleotide substrates

Single-stranded DNA containing uracil at a specific site were supplied with synthetically incorporated Cy3 fluorophore (Integrated DNA technologies): /5Cy3/TAGACATTGCCCTCGAGGTA/dU/CATGGATCCGATTTCGACCTCAAACCTAG ACGAATTCCG [DNA60-dU, protected by phosphorothioate at each end, used as control substrate]; /5Cy3/CCACACAAAGGG/dU/AAAGCCGGGGCA [DNA25-dU]. Equimolar amounts (100 pmol/µl) of the labeled DNA60-dU strand (1 µl) and unlabeled complementary strand (1 µl) were annealed in sterile filtrated STE buffer (8 µl; 10 mM Tris pH 8.0, 50 mM NaCl, and 1 mM EDTA) by heating at 95 °C for 4 min in a thermocycler, followed by cooling to room temperature for 2 h. The substrate was then diluted in sterile filtrated TE buffer (90 µl) to make 1 pmol/µl substrate. The ssDNA substrates were prepared by diluting 100 pmol/µl ssDNA stock (1 µl) in 1× TE buffer (99 µl) to make 1 pmol/µl substrate. The substrates were stored at –4 °C in darkness.

3.3 Production of recombinant hSMUG1 mutant proteins

3.3.1 Preparation of chemically competent cells and transformation

To transform cells with the vector carrying the gene of interest more efficiently, the cell walls must be chemically altered so that foreign DNA can pass through more easily. The procedure was performed in sterile conditions. The *E. coli* BL21(DE3) (strain lacking Ion and ompT proteases for protein expression) (Merck) and its Rosetta derivative (designed to enhance expression of eukaryotic proteins containing codons rarely used by *E. coli*) (Kind present from Dr. Dugassa (UiS), originally from Novagen) were recovered from –80 °C stock by streaking the frozen bacteria using a sterile loop on LB-agar plates. The LB agar plates with BL21(DE3) and Rosetta were incubated at 37 °C overnight (o/n). A colony was inoculated in sterile LB-broth (3 ml) and incubated at 37 °C with orbital shaking at 220 rpm o/n. Next, the o/n culture (200 µl) was scaled up with sterile LB-broth (25 ml) and grew until OD_{600} reached 0.3–0.5 at 37 °C and orbital shaking at 220 rpm. The sample was transferred on ice and centrifuged for 10 min at 4 °C and 2400 g. To make the cells chemically competent, the pellet was resuspended in ice-cold and sterile 100 mM $CaCl_2$ (3 ml) and incubated on ice for 30 min before centrifuged for 10 min at 4 °C and 1960 g. The pellet was again resuspended in ice-cold and sterile 100 mM $CaCl_2$ (400 µl). The competent cells were divided into aliquots in sterile tubes and transformed within 24 h.

In order to transform, competent cells (50 µl/100 µl) and the plasmid (50–100 ng) containing the gene of interest (*SMUG1* substituted by eighter P240G or S241A) were incubated for 30 min on ice. The sample was heat shocked for 45 sec (BL21(DE3)) and 30 sec (Rosetta) at 42 °C in the water bath and then transferred on ice. Super Optimal broth with Catabolite repression (SOC) media was added (500 µl) to the bacteria and incubated for 1 h at 37 °C with orbital shaking at 220 rpm. The bacterial sample was centrifuged for 5 min at room temperature and 5800 g, and the supernatant was discarded. The pellet was resuspended in the remaining media (approximately 50 µl). The sample was plated on LB-agar with corresponding antibiotics *i.e.* kanamycin (50 µg/ml) and chloramphenicol (34 µg/ml) for Rosetta; kanamycin (50 µg/ml) for BL21(DE3). The plates were incubated at 37 °C o/n.

3.3.2 Expression test

In order to test for efficient expression of the target gene, and determine the optimal time of expression, *SMUG1* expression tests of both BL21(DE3) and Rosetta were performed. An o/n culture was made by inoculating one colony of each transformed strain (BL21(DE3) and Rosetta) in LB-broth (3 ml) and the above-mentioned corresponding antibiotics and incubated at 37 °C and 220 rpm orbital shaking. The o/n culture (1000 µl) was scaled up in LB-broth (25 ml) and corresponding antibiotics. The culture was grown at 37 °C with vigorous shaking at 220 rpm until OD_{600} reached 0.4–0.6. The OD was measured using a cuvette (Brand® semi-micro cuvette) and scanned by Smart Spec Plus spectrophotometer (Bio-Rad). The sample was split into aliquots (3 ml) and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37 °C in 1 h, 2 h, 3 h, and 4 h, and at 23 °C o/n. A non-induced sample was made prior to the induction. For all samples, the equivalent of 1 ml of cells at OD_{600} = 0.8 (*i.e.* $0.8/OD_{600}$ of sample = volume in ml) was aliquoted for all samples, centrifuged for 1 min at room temperature and 16200 g. The pellets were stored at –20 °C for later SDS-PAGE analysis.

For analysis, the frozen pellets were resuspended in 1× Laemmli sample buffer (100 µl), boiled for 10 min at 95 °C, cooled down to room temperature, and centrifuged for 5 min at room temperature and 16200 g. The samples (10 µl) were analyzed using SDS-PAGE to determine the most effective strain for *SMUG1* expression.

3.3.4 Expression using autoinduction

Expression of the T7 promoter and the following downstream DNA coding for the protein of interest are induced after the growing cells have consumed the glucose content of the medium and subsequently forced to use lactose (82). A transformed colony with the gene of interest was inoculated in MDG (non-inducing minimal medium) (3 ml) with corresponding antibiotics (see section 3.3.1) and incubated o/n at 37 °C and 220 rpm orbital shaking. The o/n culture (1000 µl) was scaled up and inoculated in ZYM-5052 (1000 ml), a complex medium for autoinducing the T7 promoter of the plasmid. The culture together with a negative control was incubated o/n at 28 °C and 220 rpm orbital shaking. The cells were harvested by centrifugation for 20 min at 4 °C and 5400 g. The supernatant was discarded, and the pellets were frozen in liquid nitrogen and stored at –80 °C.

3.3.5 Expression using IPTG induction

An o/n culture was made using a colony of the transformed cells, inoculating it in LB-broth (25 ml) with corresponding antibiotics (see section 3.3.1) and incubated o/n at 37 °C and 220 rpm orbital shaking. A small-scale of the same content was used as a negative control. The culture was put on ice and scaled up using LB-broth (1000 ml) and corresponding antibiotics (see section 3.3.1). The scaled-up culture was incubated at 37 °C and 220 rpm orbital shaking until OD_{600}≈0.6. A non-induced sample was made before induction using a calculated volume (ml) of 0.8/OD_{600}. The non-induced sample was centrifuged for 5 min at room temperature and 11000 g, and the pellet was frozen for later SDS-PAGE analysis. The rest of the culture was induced using IPTG (1 mM) and incubated for 2 h at 37 °C and 220 rpm orbital shaking. The cells were harvested by centrifuge for 20 min at room temperature and 5400 g. The pellets were frozen at –20 °C for later purification.

3.4 Purification of recombinant hSMUG1 mutant protein

To purify the protein, different strategies were attempted due to the instability of the protein. TALON Metal Affinity Resins charged with cobalt were used for three methods: Batch Purification, Äkta Start Protein Purification System with HiTrap TALON crude column (1 ml, prepacked with cobalt-based chromatography media) (GE Healthcare), and gravity purification. When a peristaltic pump was used to load the protein, HisTrapHP column (5 ml) (GE Healthcare) prepacked with Ni Sepharose High Performance (HP) affinity resin was applied. Both resins are compatible with immobilized metal affinity chromatography (IMAC), containing a tetradentate chelator having a prominent affinity and specificity for His-tagged proteins (83).

3.4.1 First affinity purification

The pellets (section, 3.3.4 Expression using autoinduction) were thawed and lysis buffer (7 ml) containing freshly added β-Mercaptoethanol (2 mM) was added for each gram pellet and resuspended. Lysozyme (1 mg/ml) and Complete EDTA-free protease inhibitor (1 tablet, Thermo Scientific, lot VC2936735) were added and incubated at 4 °C for 30 min with gentle shaking. Tergitol/NP-40 (0.5 %), $MgCl_2$ (5 mM), DNase (40 µg/ml) and RNase (5 µg/ml) were added and incubated at 4 °C for 20 min with gentle shaking. The lysate was sonicated

(Sonifier® ultrasonic) with high-frequency ultrasonic energy (on ice and with stirrer) for 10 min at 30 % amplitude to disrupt the cells. The sonication was performed manually in 5-sec intervals. Insoluble debris was removed by centrifugation for 30 min at 4 °C and 48380 g, and the crude extract (the supernatant) was collected for purification. From this step, the procedure was continued by either the Batch Method or Äkta Start Protein Purification System.

3.4.1.1 The Batch Method

Resuspended Talon beads (2 ml) were equilibrated according to the manufactures instructions. Briefly, the storage buffer was removed by centrifugation for 5 min at 4 °C and 500 g. The beads were washed with MilliQ water (10× beads volume) and centrifuged for 5 min at 4 °C and 500 g. Buffer A (binding buffer, 50 nM Tris pH 7.5, 300 nM NaCl, and 2 mM fresh β-Merchaptoethanol) was added (10 ml), incubated for 10 min at 4 °C, centrifuged for 5 min at 4 °C and 500 g, and repeated twice. The final step in equilibrating the beads included adding buffer A containing 10 mM imidazole (10 ml), incubating for 5 min at 4 °C, centrifuged for 5 min at 4 °C, 500 g, and discarded the supernatant. The crude extract (from the above section) was incubated with the equilibrated beads for 30 min at 4 °C with gentle shaking. Some extract was also saved for SDS-PAGE analysis. The extract with Tallon beads was spined down by centrifugation for 5 min at 4 °C and 500 g, and the flow-through (supernatant) was collected for SDS-PAGE analysis. Buffer A with 10 mM imidazole (10 ml) was added, incubated for 5 min at 4 °C with gentle shaking, and removed by centrifugation for 5 min at 4 °C and 500 g. The step was repeated twice. Buffer A with 100 mM imidazole (elution buffer) (2 ml) was added, incubated for 10 min at 4 °C with gentle shaking, and centrifuged for 5 min at 4 °C and 500 g. The protein fractions (elutions) were collected in pre-cooled microtubes and the step was repeated twice. The elutions were saved for SDS-PAGE analysis and further purification steps.

3.4.1.2 Äkta Start Protein Purification System

The Akta start purification system was prepared by washing the pumps and the fractionation tube with water and Buffer A (50 mM TRIS pH 7.5 and 300 nM NaCl). Pump B was washed with Buffer B (50 mM TRIS pH 7.5, 300 nM NaCl, and 500 nM Imidazole). The HisTrapHP column (5 ml) was washed with water and then Buffer A in 3 column volumes (CV). The crude extract (from section 3.4.1) was loaded (saved a small sample for SDS-PAGE analysis) to the column using a peristaltic pump with a flow rate of 2–3 ml/min. Then the column was connected to the Äkta Start, and all operations were performed using a 1 ml/min flow rate. The column

was washed by 3 CV of buffer A, then washed by 2 CV of 5 % Buffer B (Buffer A containing 25 mM imidazole). The flow-through was collected for SDS-PAGE analysis. The protein was eluted using 30 ml gradient elution from 5 % to 100 % of buffer B (corresponding to buffer A with 25–500 mM of imidazole). The fraction of eluted protein was 1 ml. Samples corresponding to the peaks on the chromatogram were collected for SDS-PAGE analysis and further purification steps.

3.4.2 Second purification step using affinity chromatography

For both above-mentioned methods, the eluted fractions of the target protein were mixed with TEV protease (0.5 mg/ml final, in-house stock) and pooled in dialysis buffer (containing freshly added 2 mM β-Mercaptoethanol) using Pre-wetted RC Tubing MWCO 15 kDa (Spectra/Por). The protein was dialyzed o/n at 4 °C for His-tag cleavage and imidazole removal. The dialysis buffer was changed once, and the protein samples were pooled again in 2 h.

Second affinity chromatography was performed using Äkta Start Protein Purification System with HiTrap Talon crude column (1 ml) in order to separate the cleaved His-tag, TEV protease (containing His6xtag), and cleaved and uncleaved hSMUG1 P240G-His-tag. The Äkta Start and the column were equilibrated using buffer A (50 mM TRIS pH 7.5 and 300 nM NaCl). The dialyzed protein sample was loaded into the column using a sample valve with a rate of 1 ml/min. The flow-through containing cleaved hSMUG1 P240G was collected in 1 ml fractions. The column was connected to the Äkta Start System and washed with Buffer A (3 CV) with a flow rate of 1 ml/min. The gradient was run with a target concentration of 100 % Buffer B (50 mM TRIS pH 7.5, 300 nM NaCl, and 500 nM Imidazole) with a length of 30 min and 1 ml fractionation size.

The crude extract, flow-through, and elutions were analyzed by SDS-PAGE. The fractions containing the protein of interest were further up-concentrated by ultrafiltration using VIVASPIN 15R, 10000 MWCO (Sartorius). The tubes were placed at the correct angle (according to the manual) in a swing rotor and centrifuged at 3000 g for 5 min and repeated four times. The protein concentration was estimated by Nanodrop using buffer A as blanc. The protein was stored at –80 °C in 50 % glycerol.

3.4.3 Affinity purification using gravity column

Gravity purification was performed twice, to purify the protein and to remove the cleaved His-tag respectively. The purification was also completed without His-tag cleavage and removal in some assays. The pellets from IPTG induction (See Human *SMUG1* expression using IPTG induction) were thawed and resuspended in lysis buffer (10 ml/l media used, comprising 20 mM Tris pH 7.5, 10 mM Imidazole, 150 mM NaCl, 0.2 % NP-40, 1 µl PMSF, and freshly added 2 mM β-Mercaptoethanol). The lysate was added lysozyme (1 mg/ml) and incubated with gentle shaking for 30 min at 4 °C. DNase was added (40 µg/ml) and the lysate was again incubated with gentle shaking at 4 °C for 15–30 min. The lysate was sonicated at 10 % altitude for 1 min (10 sec on/off × 6) and centrifuged for 30 min at 4 °C and 5400 g. Resuspended Talon beads (1 ml) were equilibrated using water (4 ml) and wash buffer 2 (4 ml, comprising lysis buffer w/o NP-40). The crude extract was added to the beads (saved 20 µl for SDS-PAGE analysis), and the flow-through was collected for later SDS-PAGE analysis. The beads were washed with wash buffer 2 (5 ml, comprising lysis buffer w/o NP-40), wash buffer 3 (5 ml, comprising wash 2 buffer with 1 M NaCl), and wash buffer 4 (5 ml, comprising wash 2 buffer with 20-50 mM imidazole w/o PMSF). The protein was eluted using elution buffer (2 ml, comprising wash 2 buffer with 330 mM imidazole w/o PMSF). The beads were regenerated using MES (20 Mm) and dH_2O (5× beads volume each), stored at 4 °C in 20 % ethanol (1× beads volume) to be reused three times. The samples were mixed with 2x Laemili sample buffer in a 1:1 ratio. The non-induced pellet was resuspended in 1× Laemili sample buffer (100 µl). All samples were spun down using a mini-centrifuge, boiled for 5 min at 95 °C, and spun down again. Stain-free Western C protein standard (BioRad) (3 µl), the induced samples (20 µl), and the non-induced sample (10 µl) were loaded to the gel wells and ran for 25 min at 300 V. The gel was visualized using ChemiDoc™ Touch Imaging System. The protein concentrations were measured in three parallels using a NanoDrop instrument and the elution buffer as blanc. Eluted protein (20 µl) was mixed with 2× Leamili sample buffer in a 1:1 ratio and saved at 4 °C for "before TEV cleavage" SDS-PAGE analysis. For His-tag cleavage and imidazole removal, the protein was pooled with TEV (0.5 mg/ml, in-house stock) in dialysis buffer (2 l) containing freshly added β-Mercaptoethanol (2 mM) using pre-wetted RC tubing MWCO 16 kDa. The protein was dialyzed o/n at 4 °C.

To remove the cleaved His-tag, gravity purification was performed again. The Talon beads then bind the His-tag only, while the protein of interest appears in the flow-through. The protein fractions from the dialysis pool were added to equilibrated Talon beads and incubated

for 30 min at 4 °C. The flow-through containing the protein of interest was collected, and the beads were washed using wash 2 and wash 4 (5× beads volume each). The His-tag protein was eluted using elution buffer (4× beads volume). The beads were regenerated and reused three times. A Nanodrop instrument was used to measure three parallels of protein concentration using dialysis buffer (without imidazole) as a blanc. The protein (now without His-tag) was stored in 50 % glycerol at –80 °C.

3.4.4 SDS-PAGE analysis

A 12 % polyacrylamide gel (for TRIS-Glycine buffer system) was made (according to the BioRad kit manual) using a spacer plate with a 1.0 mm integrated spacer. All samples to be analyzed were mixed with 2× Laemmli buffer (1× if resuspending a pellet) in a 1:2 ratio and heated for 5–10 min at 95 ºC. The samples (10–20 µl) and stain-free Precision Plus Protein™ WesternC protein standard (BioRad) (3 µl) were loaded to the gel wells. Bio-Rad Mini-Protean Tetra Cell system was used to run the gel at 220 V for 43 min. The gel was visualized using ChemiDoc^{TM} Touch Imaging System (BioRad), using the stain-free application.

3.4.5 Western blotting of hSMUG1 P240G

In order to verify the protein, Western blotting was performed for the purified fractions. The SDS-PAGE of the protein of interests (both with/without His-tag) were run under the same conditions described above. The Trans-Blot Turbo transfer pack (BioRad) was assembled with the SDS PAGE gel according to the manufacture manual. The blotting was performed using the Trans-Blot Turbo Transfer System (BioRad) for 3 min. The membrane was blocked using 5 % BSA in PBST (0.1 % Tween 20) for 1 h at room temperature with gentle shaking. The membrane was incubated with either 1:2000 primary rabbit monoclonal anti-SMUG1 antibody (Abcam, cat.# ab192240) in 5 % BSA in PBST or 1:2000 His-tag anti-rabbit antibody (Cell signaling, cat.# 2365) in 5 % BSA in PBST and incubated at 4 °C o/n with gentle shaking. The membrane was washed using PBST (0.1 % Tween 20) for 5 min with shaking 3 times. A secondary antibody solution was made: antibody for molecular weight standard, Precision ProteinTM Strep Tactin-HRP Conjugate (Bio-Rad) was mixed with a secondary antibody for SMUG1, 1:2000 Goat anti-rabbit-HRP IgG (Therma scientific, cat.# 31460) or a secondary antibody for His-tag, mouse-IgGk BP-HRP conjugated (Santa Cruz, cat.# sc-516102) into 5 % BSA in PBST. The solution was added to the membranes and incubated with

shaking for 1.5 h at room temperature. Then, the membrane was washed using PBST (0.1 % Tween 20) for 5 min with shaking, 3 times. To develop the membrane a substrate solution Clarity Western ECL Substrate (BioRad) was prepared in a 1:1 ratio, added to the semi-dried membrane by a pipette, and incubated at room temperature for 5 min. The excess substrate was removed by gently pressing the membrane between a double-layer plastic cover. The membrane was imaged using ChemiDoc™ Touch Imaging System (BioRad), with a chemiluminescence setup.

3.5 hSMUG1 activity assay for excision of uracil-DNA

Assays using different enzyme concentrations were performed to indirectly determine the enzymatic activity for different DNA substrates (52). The procedure was performed in darkness due to the light-sensitive fluorescent label at the 5'-end of the oligos. A reaction mixture (20 µl total) containing 5× HEPES buffer (4 µl, 225 mM HEPES, pH 7, 10 % glycerol, and 2 mM EDTA) with freshly made DTT (5 mM final), 1 M KCl (1.4 µl, 70 mM final), 10 mg/ml BSA (1 µl, 0.5 mg/ml final), and dH_2O (11.6 µl) was made, and then 1 pmol/µl substrate (1 µl) (see Appendix A, Hybridization of Cy3-U:G substrate) and varying concentrations (diluted in 1× HEPES buffer comprising 1 mM DTT) of the purified enzyme (1 µl) was added. The mixture was centrifuged for 1 min at room temperature and 1500 g. The samples were incubated at 37 °C for 20 min and the reaction was terminated by the addition of a stop-solution (45 µl) containing 20 mM EDTA and 0.5 % (w/v) sodium dodecyl sulfate (SDS), supplemented with 10 mg/ml proteinase K (1 µl) (the samples were kept at room temperature during these steps, due to precipitation of the stop-solution when kept on ice). The samples were incubated for 10 min at 37 °C before precipitation of DNA with ice-cold 96 % ethanol containing 0.1 M sodium acetate (150 µl) supplemented with 16 µg tRNA (1.6 µl). The samples were inverted multiple times and incubated in darkness at –20 °C o/n. The samples were then centrifuged directly from the freezer for 15 min twice (in two different tube directions) at 4 °C and 13000 g. The supernatant was discarded and the pellets were added ice-cold 70 % ethanol (300 µl). The samples were centrifuged for 5 min at the same conditions, and the supernatant was discarded by tapping the tubes on a tissue. The samples were centrifuged for 1 min at the same conditions, and the residual supernatant was removed by a pipette. The pellets were dried for 10–40 min (in darkness and on ice) to completely evaporate all ethanol. To be able to analyze the enzymatic excision of uracil which results in an alkali-labile AP site the pellets were resuspended in NaOH (10 µl) and heated for 10 min at 90 °C.

Denaturing PAGE was performed to indirectly identify and quantify the glycosylase activity. A 20 % PAGE gel containing 3 % formamide or 8 M urea (for Taurine and TBE buffer system respectively) was made using a spacer plate with a 0.75 mm integrated spacer. The samples were mixed with denaturing loading buffer (10 µl, 80 % formamide, 1 mM EDTA pH 8.0, 1 % (w/v) blue dextran) without boiling to avoid thermolysis (chemical decomposition by heating). The samples (5 µl) were loaded into thoroughly washed gel wells. Bio-Rad Mini-Protean Tetra Cell system was used to run the gel at 100 V or 200 V (for Taurine and TBE buffer system respectively) for 1.5 h in darkness. The gel was visualized without dissembling the glass plates using a Typhoon^{TM} Variable Mode PhosphorImager with fluorescence setup (532 nm excitation wavelength in the green region and 580BP emission filter). The results were quantified using the ImageQuant 5.1 software (Molecular Dynamics, Sunnyvale, CA). For detailed buffer descriptions and all protocols see Appendix A.

3.5.1 Kinetic model

To describe the hSMUG1 excision activity for uracil in single-stranded DNA the experimental data were adapted into two kinetic curve fit models: Saturation adaption includes adsorption and desorption due to unspecific binding of the enzyme to its substrate; linear fit includes a proportional relation between the rate of the catalysis (velocity) and the concentration of the enzyme at a defined substrate concentration. A Three-Staged Kinetic Model has been described for hSMUG1 excision and incision (52). The model describes the rapid initial uracil (U) excision that produces an AP site in DNA (DNA_{ap}) by the following reaction

$$E + DNA \xrightarrow{k_1} DNA_{ap}{\bullet}U{\bullet}E \xrightarrow{k_p^{ex}} DNA_{ap} + U + E \quad (1)$$

and the reaction velocity (v_{ex}) is calculated as uracil (U) generation divided by 20 min

$$v_{ex} = \frac{U\,(nM)}{20\,(\mathrm{min})} \quad (2)$$

In order to generate the graphs, two different functions were used. When the data was adapted to linear fit kinetics the following function was used

$$g(x) = a \times x \quad (3)$$

where a = 0.43 and x = $[E]_0$

and when the data was assumed to follow saturation kinetics the following function was used

$$f(x) = \frac{v_{max} \times x}{(K_m + x)} \quad (4)$$

where km = 1×10^{-5} and v_{max} = 0.01.

```
1
2  set terminal pdfcairo dashed enhanced size 18 cm, 14 cm font "Arial, 20"
3
4
5  set title "hSMUG1 P240G"
6  set output 'P240G-ssU-DNA-all-CL.pdf'
7  set key off
8  set grid
9  set xlabel "[E]_0 (nM)"
10 set ylabel "v (nM/min)"
11
12 km=1.e-5
13 vmax=0.01
14 f(x)=vmax*x/(km+x) # fit using saturtion kinetics
15
16 g(x)=a*x #linear fit
17 a=0.34
18
19 fit g(x) 'P240G-ssU-DNA-all-CL.txt' via a
20
21 fit f(x) 'P240G-ssU-DNA-all-CL.txt' via km, vmax
22
23 plot "P240G-ssU-DNA-all-CL.txt" title "v" pt 22 lt rgb "dark-green" lw 3, g(x)
   title "g(x)=a*x" lw 3 lt rgb "dark-red", f(x) title "f(x)=Vmax*x/(KM+x)"  lw 3 lt
   rgb "dark-blue"
```

Figure 13. The script used to generate plots and kinetic parameters using all independently measured data values. Both the linear fit ($g(x) = a \times x$) and the assumed saturation adapted ($f(x) = v_{max} \times x)/(K_m + x)$) functions are shown. Courtesy of Peter Ruoff.

```
1
2  set terminal pdfcairo dashed enhanced size 18 cm, 14 cm font "Arial, 20"
3
4
5  set title "hSMUG1 P240G"
6  set output 'P240G-ssU-DNA-ave-CL.pdf'
7  set key off
8  set grid
9  set xlabel "[E]_0 (nM)"
10 set ylabel "v (nM/min)"
11
12 km=1.e-5
13 vmax=0.01
14 f(x)=vmax*x/(km+x) # fit using saturtion kinetics
15
16 g(x)=a*x #linear fit
17 a=0.34
18
19 fit g(x) 'P240G-ssU-DNA-ave-CL.txt' via a
20
21 fit f(x) 'P240G-ssU-DNA-ave-CL.txt' via km, vmax
22
23 plot "P240G-ssU-DNA-ave-CL.txt" title "v" with errorbars pt 22 lt rgb "dark-green"
   lw 3, g(x) title "g(x)=a*x" lw 3 lt rgb "dark-red", f(x) title
   "f(x)=Vmax*x/(KM+x)"  lw 3 lt rgb "dark-blue"
```

Figure 14. The script used to generate plots and kinetic parameters using the average (± SD) of the independently measured data values. Both the linear fit ($g(x) = a \times x$) and the assumed saturation adapted ($f(x) = v_{max} \times x)/(K_m + x)$) functions are shown. Courtesy of Peter Ruoff.

Two different scripts were used to develop the plot and the kinetic parameters from the experimental data. One where all independently measured data values were used (Figure 13) and one where the average (± SD) of the independently measured data values was used (Figure 14). The plots and the kinetic parameters were generated using command prompt (Windows) via Gnuplot, which is freely available at http://www.gnuplot.info/.

4 Results

4.1 Recombinant hSMUG1 mutant protein production and purification

4.1.1 hSMUG1 P240G

hSMUG1 P240G mutant protein was interesting based on the suggestion that the P240 residue most likely contributes to catalysis by being present in the intercalating loop region of the active site of the hSMUG1 DNA glycosylase (64,66). Proline (P) is a rigid residue restricting rotations relative to adjacent residues (*i.e.* His and Ser), while glycine (G) provides more flexibility for conformational changes in the active sites region having little influence on the preceding and following residue (84). The mutated protein was produced in *E. coli* BL21(DE3) cells or its Rosetta derivative as a six-His-tag-fusion protein (referred to as His-tag) using the pET System (Figure 15A).

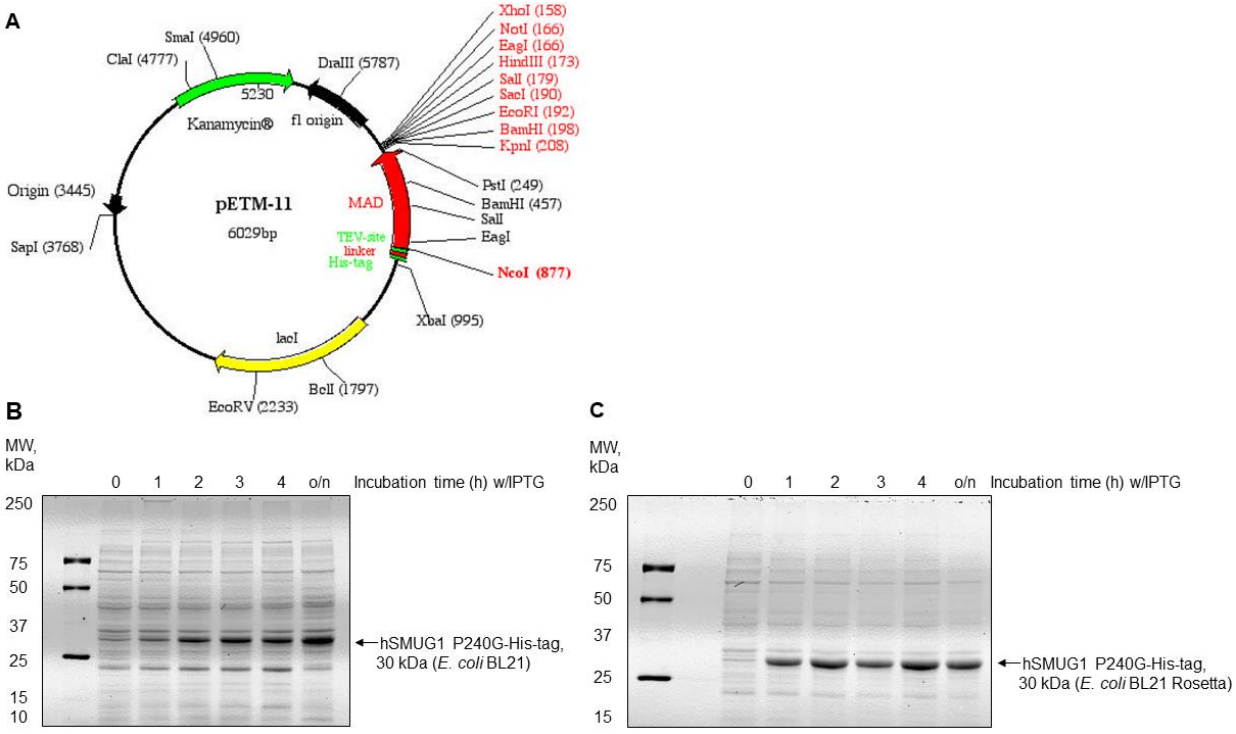


Figure 15. Production of mutated hSMUG1 P240G-His-tag following gene expression induced by isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM) in both E. coli BL21(DE3) strain and its Rosetta derivative. ***(A)*** *The constructed pETM-11 vector (6029 bp) (EMBL Protein Expression and Purification Facility) used for expression comprises a T7 lac promoter upstream of the foreign gene producing the hSMUG1 P240G mutant protein.* ***(B)*** *The gene corresponding to hSMUG1 P240G-His-tag was expressed in BL21(DE3) cells in 0–4 h (37 ºC) and o/n (23 ºC) respectively.* ***(C)*** *The same gene expressed in the BL21(DE3) Rosetta derivative in 0–4 h (37 ºC) and o/n (23 ºC) respectively. For corresponding OD values and Western blot membrane representing the BL21(DE3) Rosetta gel, see Appendix B, Production and purification of hSMUG1 P240G. Abbreviations: o/n, overnight; MW, molecular weight; kDa, kilodalton; bp, base pair.*

A test of *in vivo* protein overproduction was performed, to define the optimal strain and induction time for the foreign *SMUG1* gene (See Appendix B, Production and purification of hSMUG1 P240G for more information). The results indicated that growth overnight gives the highest yield in BL21(DE3) (Figure 15B) while 2 h is sufficient for its Rosetta derivative (Figure 15C) also exhibiting a lower yield of other (contaminating) proteins. However, the differences were not significant, so both strains were used in further attempts of protein production. Following affinity purification using TALON Metal Affinity Resins charged with Co^{2+} for high affinity and specificity for His-tag by the Batch Method (83), purified hSMUG1 P240G-His-tag (Table 2, Lot No. 3) showed a band at the expected size (30 kDa) in SDS-PAGE analysis (Figure 16A). Finally, hSMUG1 P240G was dialyzed with in-house produced TEV protease (0.5 mg/ml, Appendix B, Purification of TEV protease) to cleave the His-tag,

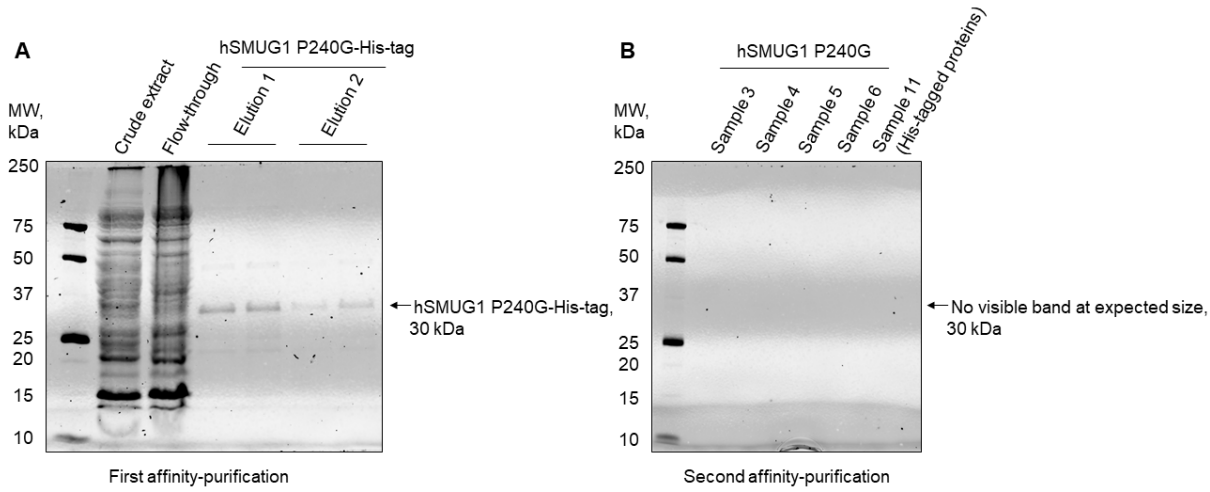


Figure 16. hSMUG1 P240G-His-tag was produced by autoinduction of the corresponding gene using the pET System (Figure 15A) in E. coli BL21(DE3) and further purified and analyzed by SDS-PAGE before and after His-tag cleavage by TEV. ***(A)*** *SDS-PAGE of Batch purified (Materials and Methods 3.4.1.1) hSMUG1 P240G-His-tag.* ***(B)*** *SDS-PAGE of hSMUG1 P240G after dialysis with TEV for His-tag cleavage and separation of the components by affinity chromatography using Äkta Start Protein Purification System (Materials and Methods 3.4.2). Sample 3–6, flow-through of hSMUG1 P240G; sample 11, elution of His-tagged proteins (including TEV, remnants of uncleaved hSMUG1 P240G, and His-tag itself). Samples 4 and 5 were later used in all activity assays using the P240G variant, revealing the most active UDG purified. For all production and purification steps of hSMUG1 P240G, see Appendix B, Lot No. 3.*

and second affinity-purified using HiTrap Talon crude column (prepacked with cobalt-based resins) (Figure 17) in order to separate the His-tag, the TEV protein, and the uncleaved protein from the cleaved hSMUG1 P240G. SDS-PAGE analysis showed no visible band at the expected size after the second affinity purification (Figure 16B), although this particular enzyme represents the most active UDG (Figure 21C, Lot No. 3). Gravity purified hSMUG1 P240G-His-tag (Table 2, Lot No. 4) was analyzed by SDS-PAGE (Figure 18A) and identified by Western blot (Figure 18B) using a monoclonal antibody specific for SMUG1.

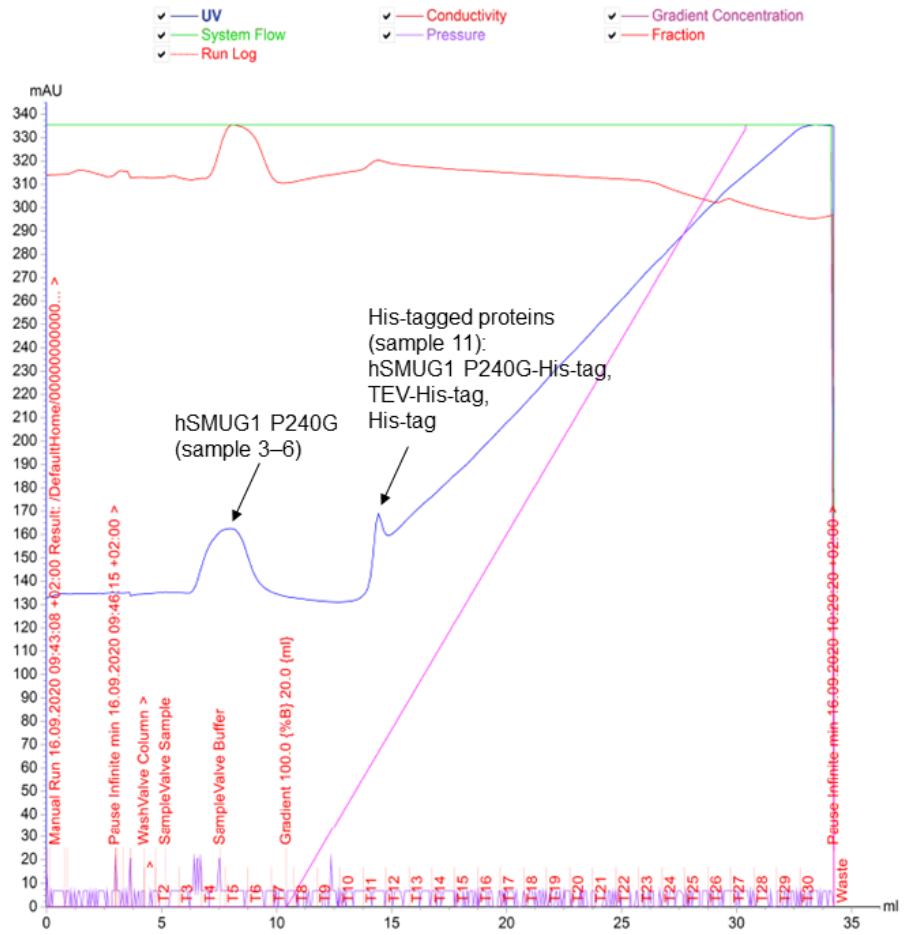


Figure 17. Second affinity purification using Äkta Start Protein Purification System was performed in order to separate cleaved hSMUG1 P240G from the His-tagged proteins (i.e. TEV-His-tag, remnants of hSMUG1 P240G-His-tag, and the His-tag itself). The chromatography was performed using a HiTrap column (1 ml) prepacked with cobalt-based chromatography resins designed for purification of His-tagged recombinant proteins. Equilibriating buffer (50 mM TRIS pH 7.5 and 300 nM NaCl) was loaded into the system and the dialyzed protein was loaded into the column using a sample valve with a rate of 1 ml/min. The gradient was run with a target concentration of 100 % elution buffer (50 mM TRIS pH 7.5, 300 nM NaCl, and 500 nM Imidazole) and the samples were collected in 1 ml fractions (Materials and Methods 3.4.2). The chromatogram shows the flow-through peaks representing hSMUG1 P240G (sample fraction 3–6) and the elution peak representing His-tagged proteins (sample fraction 11). Sample fractions 4 (0.109 mg/ml) and 5 (0.099 mg/ml) were mixed and further up-concentrated by centrifuge ultrafiltration to 0.16 mg/ml and used for all hSMUG1 P240G activity assays.

His-tag cleavage of hSMUG1 P240G-His-tag (Table 2, LOT no. 6) was demonstrated by first running SDS-PAGE (Figure 19A and B left panels) of eluted hSMUG1 P240G-His-tag (the fused protein) collected after the first affinity purification and hSMUG1 P240G flow-through (cleaved protein) collected after the second affinity purification. The same gel was then blotted into a Western blot membrane, and thereby, His-tag (Figure 19A, right panel) and SMUG1 (Figure 19B, right panel) were detected using Goat anti-rabbit-HRP IgG and Rabbit monoclonal anti-SMUG1 antibody, respectively.

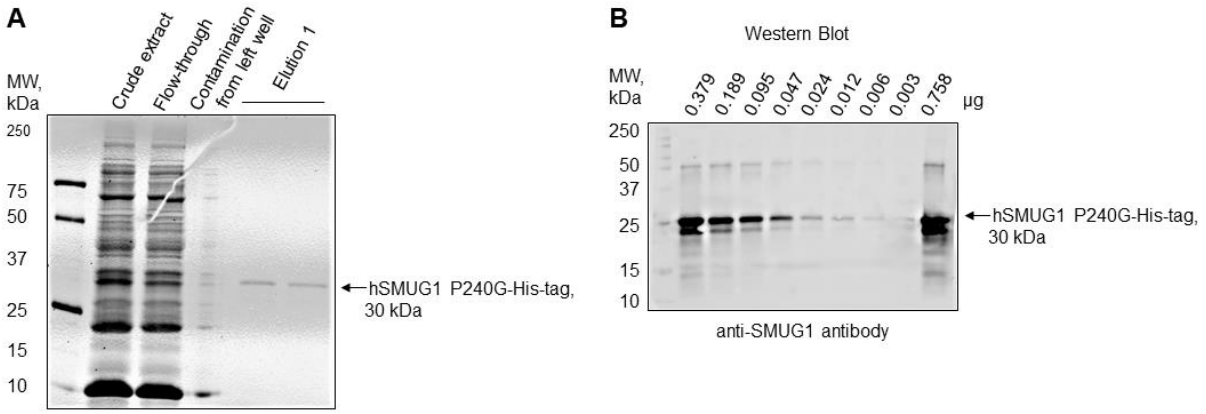


Figure 18. The plasmid (Figure 15A) containing hSMUG1 P240G-His-tag was expressed in E. coli BL21(DE3) Rosetta and further gravity purified (Materials and Methods 3.4.3) and analyzed before His-tag-cleavage by TEV. ***(A)*** *SDS-PAGE analysis of hSMUG1 P240G-His-tag (representing Lot No. 4).* ***(B)*** *Western blot of hSMUG1 P240G-His-tag in 1:2 serial dilutions, using Rabbit monoclonal anti-SMUG1 primary antibody and Goat anti rabbit-HPR IgG secondary antibody in 1:2000 dilutions and 5 % BSA. Precision ProteinTM Strep Tractin-HPR Conjugate was used for detection of the stain-free protein standard WesternC. Abbreviations: MW, molecular weight; kDa, kilodalton.*

The figure representing the Western blot membrane with His-tag specific antibody (Figure 19A, right panel) clearly shows that hSMUG1 P240G-His-tag is present at the expected size before His-tag cleavage, while no hSMUG1 P240G appearance is present after the His-tag cleavage, although it is present as a faint band when analyzed by SDS-PAGE (Figure 19A, left panel).

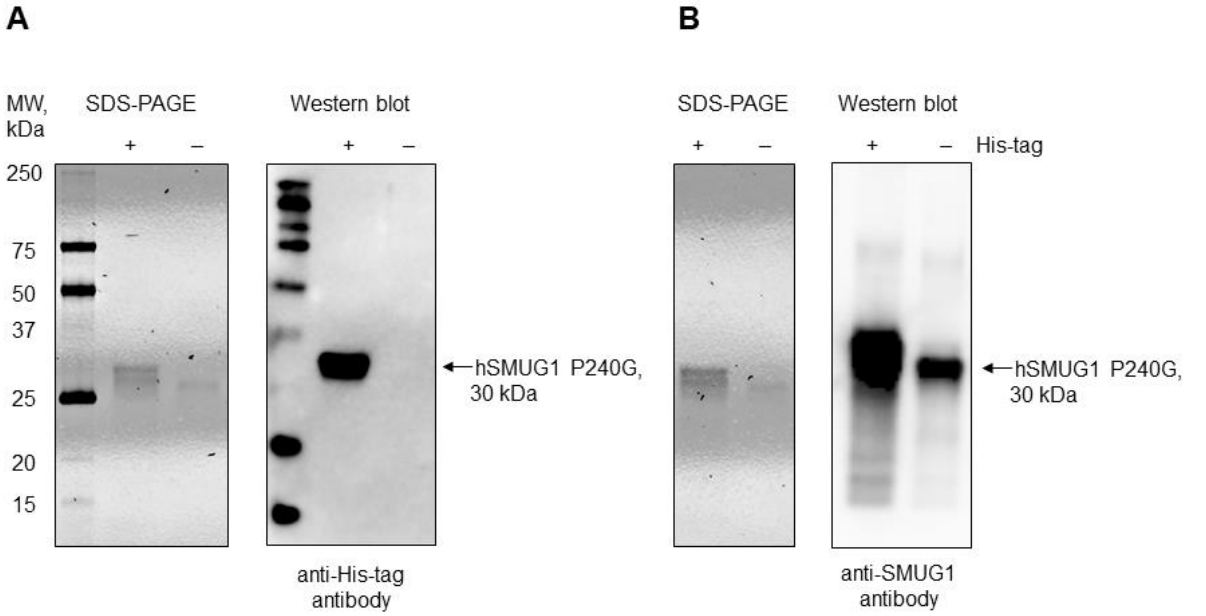


Figure 19. Identification of hSMUG1 P240G produced in E. coli BL21(DE3) Rosetta before and after TEV-cleavage and second affinity purification for His-tag removal. One SDS-PAGE gel (A and B left panels) was blotted into one Western blot membrane (A and B right panels), the membrane was then cut into two pieces, and thereby detected by two different antibodies. ***(A)*** *SDS-PAGE gel (left panel) and the corresponding Western blot membrane (right panel) using Rabbit Anti-His-tag primary antibody and Mouse-IgGk BP-HRP conjugated secondary antibody in 1:2000 dilutions and 5 % BSA for detection of possible protein with His-tag. No hSMUG1 P240G was detected by the anti-His-tag antibody after cleavage and removal of the His-tag. Precision Protein™ Strep Tractin-HPR Conjugate was used for detection of the stain-free protein standard WesternC.* ***(B)*** *SDS-PAGE gel (left panel) and the corresponding Western blot membrane (right panel) using Rabbit monoclonal anti-SMUG1 primary antibody and Goat anti rabbit-HPR IgG secondary antibody in 1:2000 dilutions and 5 % BSA for detection, showing that hSMUG1 is present both before and after His-tag cleavage. Abbreviations: MW, molecular weight; kDa, kilodalton.*

At the figure representing the membrane with SMUG1 specific antibody (Figure 19B, right panel), the widest band shows the presence of hSMUG1 P240G before His-tag cleavage, while a more narrow band illustrates hSMUG1 P240G after His-tag cleavage, indicating that some protein is lost during the second purification. In conclusion, hSMUG1 P240G is properly cleaved due to the fact that hSMUG1 P240G is not present at the membrane after the His-tag has been cleaved, illustrated by using His-tag specific antibody, simultaneously, analyzing the same sample, hSMUG1 P240G is identified when using SMUG1 specific antibody. For all hSMUG1 P240G purifications attempts and their corresponding SDS-PAGE gels, see Appendix B, Production and purification of hSMUG1 P240G.

4.1.2 hSMUG1 S241A

A second hSMUG1 mutant protein with Ser241 replaced by Ala was tested for *in vivo* overproduction under the same conditions as described for P240G above. The results indicated that 1 h incubation time gave a high yield in BL21(DE3) strain (Figure 20A, left panel) while 3 h was better for the BL21(DE3) Rosetta derivative (Figure 20A, right panel).

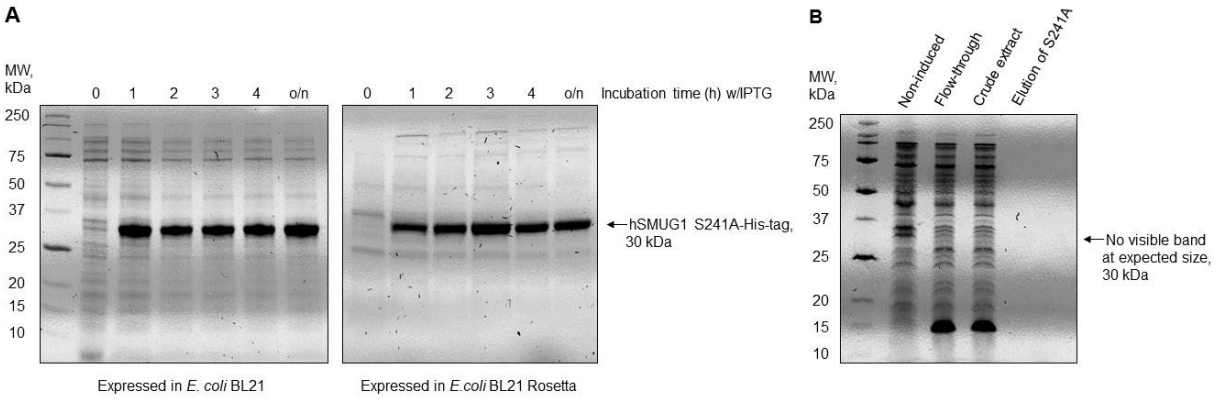


Figure 20. Production and purification of hSMUG1 S241A-His-tag (representing Lot No. 8). ***(A)*** *Expression test of 1 mM IPTG induced SMUG1 S241A using E. coli BL21(DE3) strain and its Rosetta derivative respectively. Expression is illustrated after 0–4 h ((37 ºC)) and o/n (23 ºC). See details in Appendix B, Production and purification of hSMUG1 S241A.* ***(B)*** *SDS-PAGE analysis of 1 h and 0.8 mM IPTG induced hSMUG1 S241A after gravity purification (Materials and Methods 3.4.3). No visible band at the expected size. Abbreviations: o/n, overnight; MW, molecular weight; kDa, kilodalton; IPTG, Isopropyl β-D-1-thiogalactopyranoside.*

In order to purify hSMUG1 S241A mutant protein, the corresponding foreign gene was expressed in *E. coli* BL21(DE3) in a larger scale under similar conditions with one exception; 0.8 mM IPTG (instead of 1 mM) was used for induction. The protein was then gravity purified, dialyzed without the TEV protease, and analyzed by SDS-PAGE (Figure 20B). No visible band was shown at the expected size of the eluted protein. This might be the result of different

explanations; a lower IPTG concentration and a short incubation time were used, the protein was produced at a larger scale which means somewhat different conditions, and the mutant substitution was different compared to P240G which might result in a more unstable protein.

4.2 hSMUG1 P240G and S26R/E35D mutant proteins were active on uracil in single-stranded DNA

To determine DNA glycosylase activity, an oligodeoxyribonucleotide (1 pmol) containing a damaged base residue (*in casu*, uracil) inserted at a specific location (Figure 21A) was incubated together with both wild type and mutated hSMUG1 (variable concentrations). Since uracil excision produces an alkali-labile AP site, the glycosylase activity can be indirectly determined by separating cleaved and uncleaved DNA by denaturing polyacrylamide gel electrophoresis (PAGE) (Figure 21B) following *e.g.* NaOH-treatment in the presence of heat, which cleaves AP sites in a β/δ-elimination reaction (85).

Table 2. Overview of the eight purified Lots of mutated hSMUG1 including the purification date, type of protein substitution, type of E.coli BL21(DE3) strain used, the method used for purification, information about tag-removal, the achieved concentration after purification and dilution in 50 % storage buffer (stock concentration), the number of pmol loaded into the reaction for activity assay, and the measured UDG activity in percent. The amount loaded (pmol) is calculated by using the corresponding molecular weight of the mutant proteins (NEB/TB, 29861.73 Da; P240G, 29821.67 Da; S241A, 28945.73 Da; S26R/E35D, 29930.84) which is found using the Swiss Bioinformatics Resource Portal (Expasy) and the amino acid sequence (UniProt) by manually substituting the corresponding amino acid. Abbreviations: Nd, not detected; UDG, uracil-DNA glycosylase.

Lot No.	**Date**	**Mutated hSMUG1**	***E. coli* strain**	**Method**	**W/His-tag[2]**	**Stock[3] concentration (mg/ml)**	**Loaded (pmol)**	**UDG activity (%)**
1	19.08.21	P240G	BL21	Batch	Yes	0.039	1.31	68.99 59.81[1]
2	24.08.21	P240G	BL21	Batch + Äkta	No	0.5325	–	Nd
3	16.09.21	P240G	BL21	Batch + Äkta	No	0.16	5.37	99.48
4	02.10.21	P240G	BL21 Rosetta	Gravity	Yes	0.101	3.39	3.44 0.00[1]
5	16.10.21	P240G	BL21 Rosetta	Gravity	Yes	0.01885	0.63	0.66 0.92[1]
6	03.11.21	P240G	BL21 Rosetta	Gravity	No	Nd	Nd	6.21
7	11.11.21	P240G	BL21	Gravity	No	0.2583	8.66	10.18
8	09.12.21	S241A	BL21	Gravity	Yes	0.038	1.31	1.09 0.00[1]

[1]dsU-DNA substrate
[2]Presumably removed by a second chromatography purification (Äkta)
[3]After diluted in storage buffer (50 % glycerol)

In order to determine possible DNA glycosylase activity, seven of the purified protein batches (Table 2) were tested by individually adding the enzyme stock concentrations (1 µl) directly into reaction mixtures before analyzing the exposed substrate (Figure 21C). Two of the batches were determined significantly active (Figure 21C, Lot No. 1 and 3), exhibiting high excision activity for ssU-DNA. hSMUG1 P240G-His-tag was also tested for double-stranded DNA (Figure 21C, Lot No. 1) exhibiting a reduction in activity for this substrate, as compared to the negligible reduction seen in wild type (Figure 21C, wt). Since the enzyme from Lot No. 1 contains His-tag, Lot No. 3 (hSMUG1 P240G without His-tag) was further used to characterize its activity for uracil in single-stranded DNA.

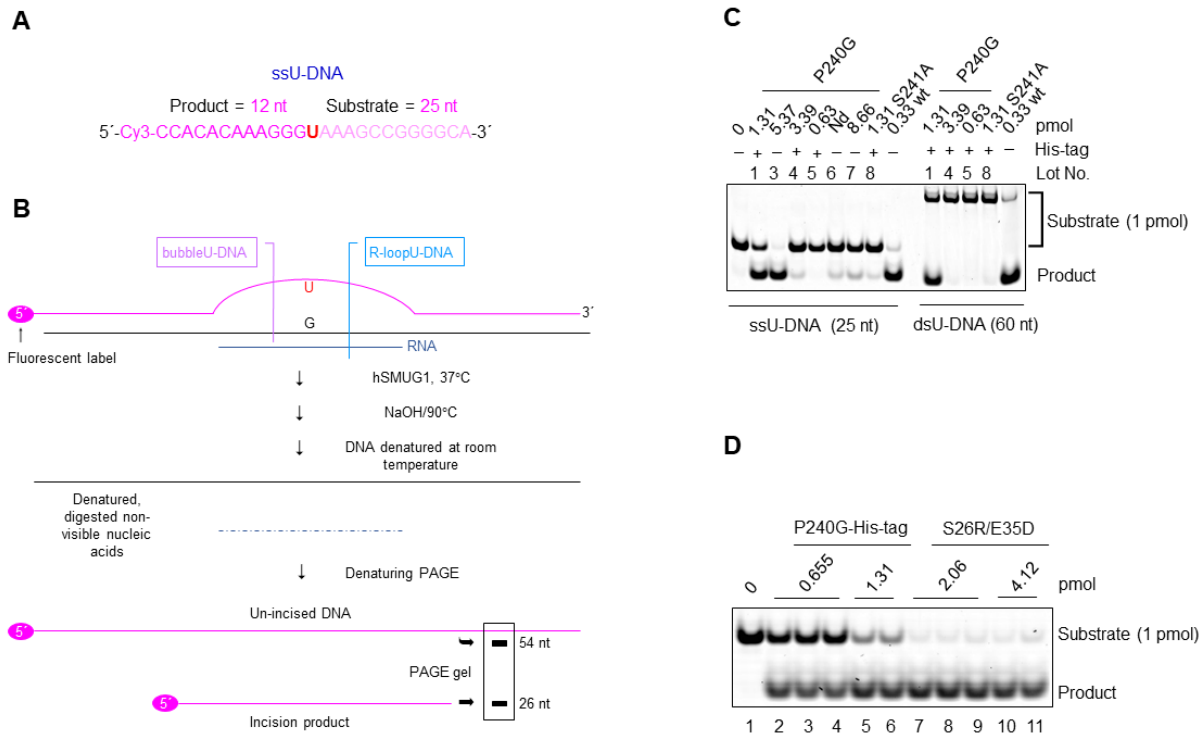


Figure 21. Excision activity of uracil in single-stranded DNA by hSMUG1 mutants. ***(A)*** *The illustrated 5'fluorescently labeled substrate (ssU-DNA) was used for all assays.* ***(B)*** *A simplified scheme of the general base excision assay used (Courtesy of Svein Bjelland). hSMUG1 mutants (variable concentrations) was incubated with the 5'fluorescently labeled ssU-DNA (1 pmol) in 45 mM HEPES, pH 7.5, 5 mM DTT, 0.4 mM EDTA, 70 mM KCl, and 0.5 mg/ml bovine serum albumin (BSA) at 37 ºC for 20 min (see Materials and methods 3.5).* ***(C)*** *Activity assay of seven purified hSMUG1 mutant proteins (P240G and S241A) using the stock concentrations directly into the reaction mix to elucidate possible uracil-DNA glycosylase activity before the exposed substrate was analyzed. Four batches were tested for dsU-DNA (Material and Methods, oligonucleotide substrates) in addition to ssU-DNA. No enzyme was used as a negative control and commercially purified wild type (wt) hSMUG1(NEB) was used for positive control.* ***(D)*** *Activity assay of purified hSMUG1 mutant proteins (P240G-His-tag and S26R/E35D) by adding both the stock concentrations and the stock concentrations in 50 % glycerol directly into the reaction mix to determine possible glycosylase activity. No enzyme was used as a negative control. The gels were quantified using ImageQuant 5.1 software (Molecular Dynamics, Sunnyvale CA). Abbreviations: Nd, not detected; nt, nucleotides; U, uracil, ssU-DNA, single-stranded uracil-DNA; dsU-DNA, double-stranded uracil-DNA; wt, wild type(NEB); NEB, New England BioLabs; P240G, substitution of P to G at site 240; S241A, substitution of S to A at site 241; S26R/E35D, substitution of S and E for R and D at sites 26 and 35 respectively.*

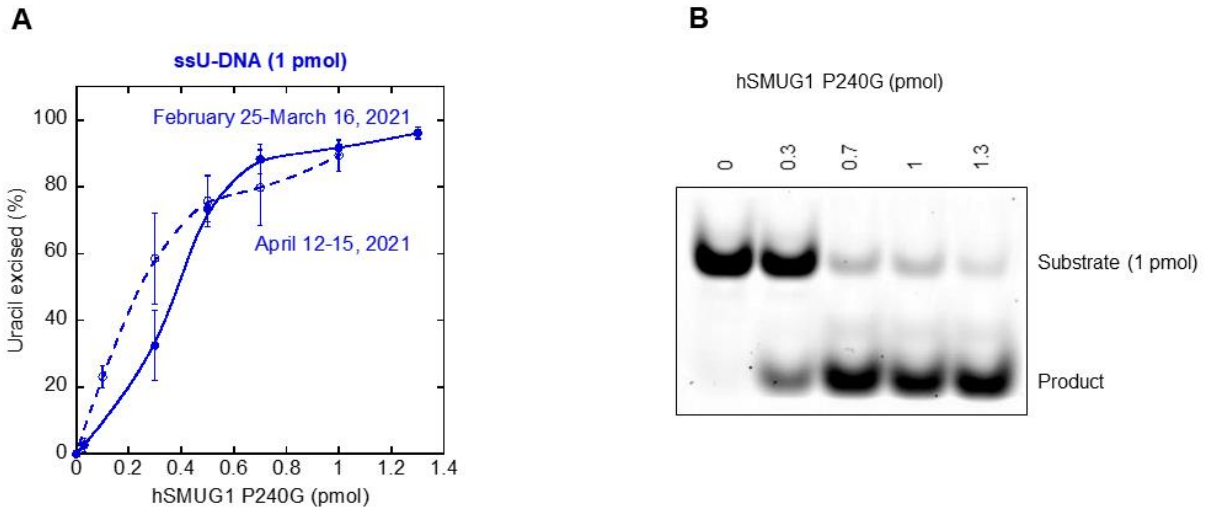


Figure 22. hSMUG1 P240G excision activity at ssU-DNA. ***(A)*** *Activity analysis of purified hSMUG1 P240G, showing excision (%) as a function of increasing enzyme amounts (0–1.3 pmol). February 25-March 16, 2021; The first trial, representing the main collection of data. April 12-15, 2021; The second trial, representing a smaller data collection, performed at a later period for comparison with a different purified mutant protein (S26R/E35D) and other substrates (experiments performed by other members of the research group). Both experiments are exerted under the same conditions, except for the periods of the implementation. Each value represents the average (±SD) of 4–12 independent measurements.* ***(B)*** *hSMUG1 P240G incubated with ssU-DNA with increasing enzyme amounts (0–1.3 pmol) under the conditions explained in Figure 21B. Abbreviation: ssU-DNA, single-stranded uracil-DNA. Graph generated using KaleidaGraph (Courtesy of Svein Bjelland).*

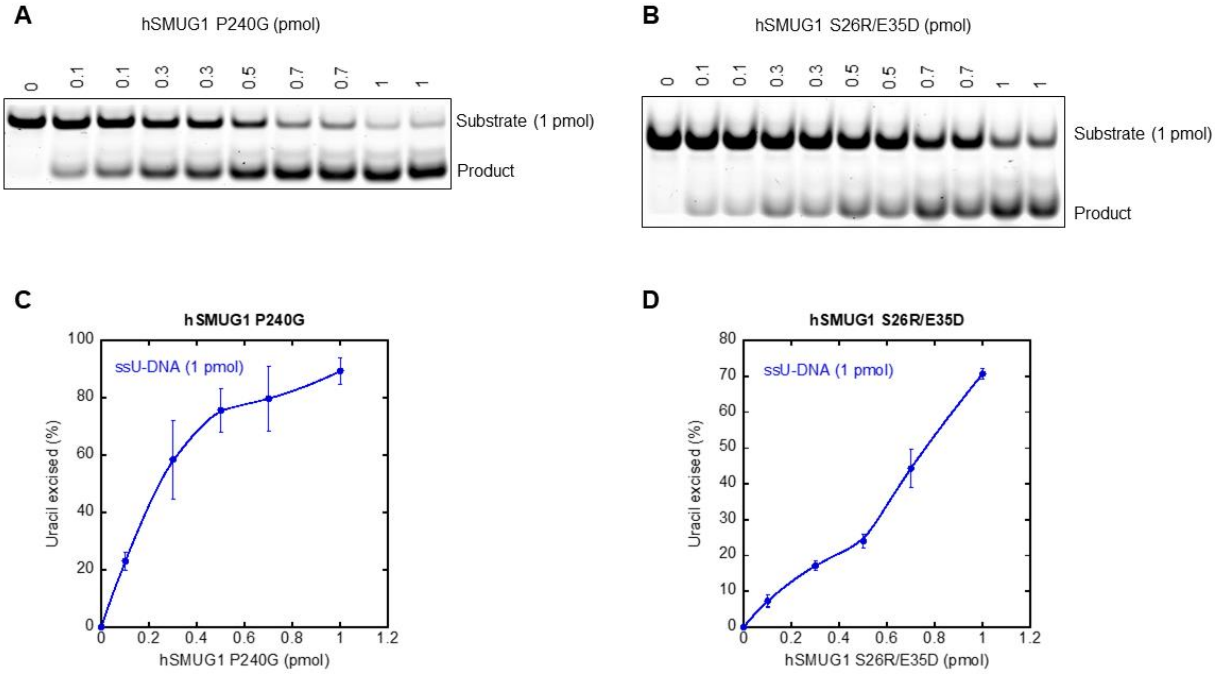


Figure 23. Excision of ssU-DNA by two different hSMUG1 mutant proteins (P240G and S26R/E35D) under the conditions as described in Figure 21B. ***(A)*** *Excision of uracil in single-stranded DNA by hSMUG1 P240G with increasing enzyme amount (0–1 pmol).* ***(B)*** *Excision of uracil in single-stranded DNA by hSMUG1 S26R/E35D with increasing enzyme amount (0-1 pmol).* ***(C)*** *Uracil excision (%) by hSMUG1 P240G is illustrated as a function of the increasing enzyme amount.* ***(D)*** *Uracil excision (%) by hSMUG1 S26R/E35D is illustrated as a function of increasing enzyme amount. Each value represents the average (±SD) of 4–12 independent measurements. Abbreviation: ssU-DNA, single-stranded uracil-DNA. Graphs generated using KaleidaGraph (Courtesy of Svein Bjelland).*

Another purified mutant, hSMUG1 S241A did not show significant activity for ssU-DNA (Figure 21C, Lot No. 8). Using the same method, another in-house purified (by another member of the research group) (86) mutant, hSMUG1 S26R/E35D was confirmed active against uracil in ssDNA by adding the stock concentration (1 µl) in addition to the stock concentration in 50 % glycerol (1 µl) directly into the reaction mix and later analyzing the 20 minutes exposed substrate (Figure 21D, lanes 7–11). The positive control (Figure 21C, wt) showed expected excision activity and the negative control (Figure 21C/D) showed no activity.

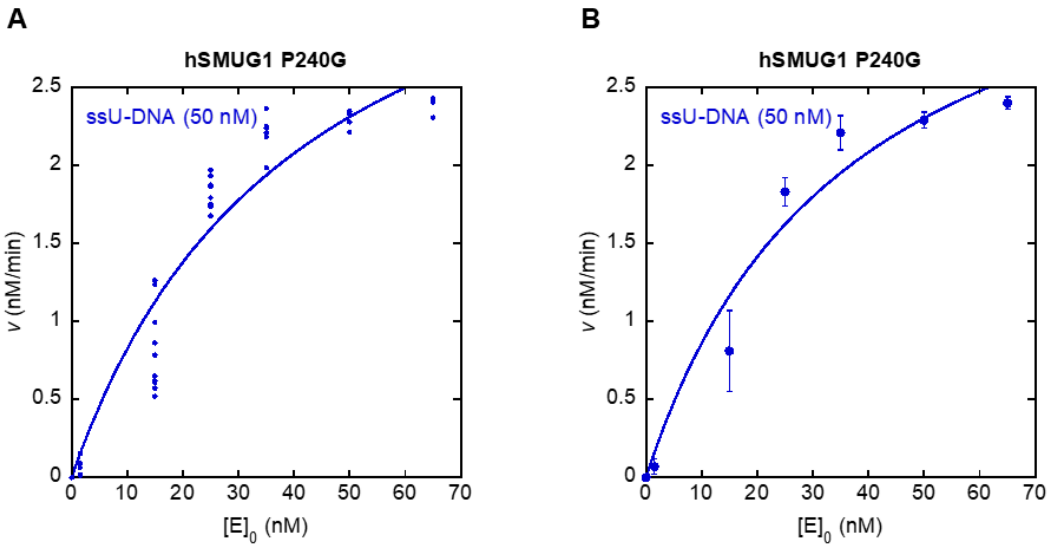


Figure 24. Saturation adapted kinetics of hSMUG1 P240G showing the excision velocity (v) in nM/min at ssU-DNA as a function of increasing (0–70 nM) initial enzyme concentrations, $[E]_0$ (nM). Both plots represent values from the first attempt comprising the main data collection. ***(A)*** *Enzyme kinetics of hSMUG1 P240G showing all independently measured data points.* ***(B)*** *Enzyme kinetics of hSMUG1 P240G where each value represents the average (±SD) of 4–10 independent measurements. Abbreviations: ssU-DNA, single-stranded uracil-DNA; nM, nanomolar. Graphs generated using KaleidaGraph (Courtesy of Svein Bjelland).*

Further, the 5' end-labeled ssU-DNA (Figure 21A) was incubated with increasing amounts of hSMUG1 P240G in individually repeated experiments (Figure 22B). The P240G mutant variant removed virtually all uracil residues present in ssDNA at the highest enzyme concentration examined at the first trial (Figure 22A, solid line), revealing that Pro240 substituted for Gly in hSMUG1 does not obstruct the glycosylase activity exhibited by the native SMUG1 enzyme for uracil in ssDNA. All the individually repeated experiments (Figure 24A) and the average (±SD) (Figure 24B) were adapted into saturation kinetics due to a curved behavior of the data points. From the second hSMUG1 P240G trial (Figure 23A), the curve reveals the abrupt increase in excision (%) at initial enzyme concentrations (Figure 23C) normally demonstrated by glycosylases, and when comparing results from different experiment periods (Figure 22A, solid and dotted lines), hSMUG1 P240G shows no obvious activity variation with time. From the second trial, experimental data of all (Figure 24A, upper panels)

and the average (±SD) (Figure 24A, lower panels) were adapted into saturation kinetics (3.5.1 Kinetic model, equation 4), and their respective kinetic parameters are summarized in Table 3.

Repeated experiments of in-house purified S26R/E35D (Figure 23B) demonstrating that neither of the substitutions in the double-mutant obstructs the UDG activity completely. The double mutant, S26R/E35D exhibited a slight reduction in activity compared to the P240G variant (purified by the same method) and were able to remove over 70 % of all uracil residues at the highest enzyme concentration (Figure 23D), however, it is important to keep in mind that inactivated enzymes may affect the results and different mutant proteins can therefore not be directly compared. All the experimental data (Figure 25B, upper panel) and the average (±SD) (Figure 25B, lower panel) were adapted into linear kinetics (3.5.1 Kinetic model, equation 3), therefore, no kinetic parameters were established. No uracil-excision occurred without any enzyme (Figure 23A/B and Figure 22B). Results from all repeated experiments are presented in Appendix B, Activity assays for excision of uracil-DNA by hSMUG1 P240G or S26R/E35D.

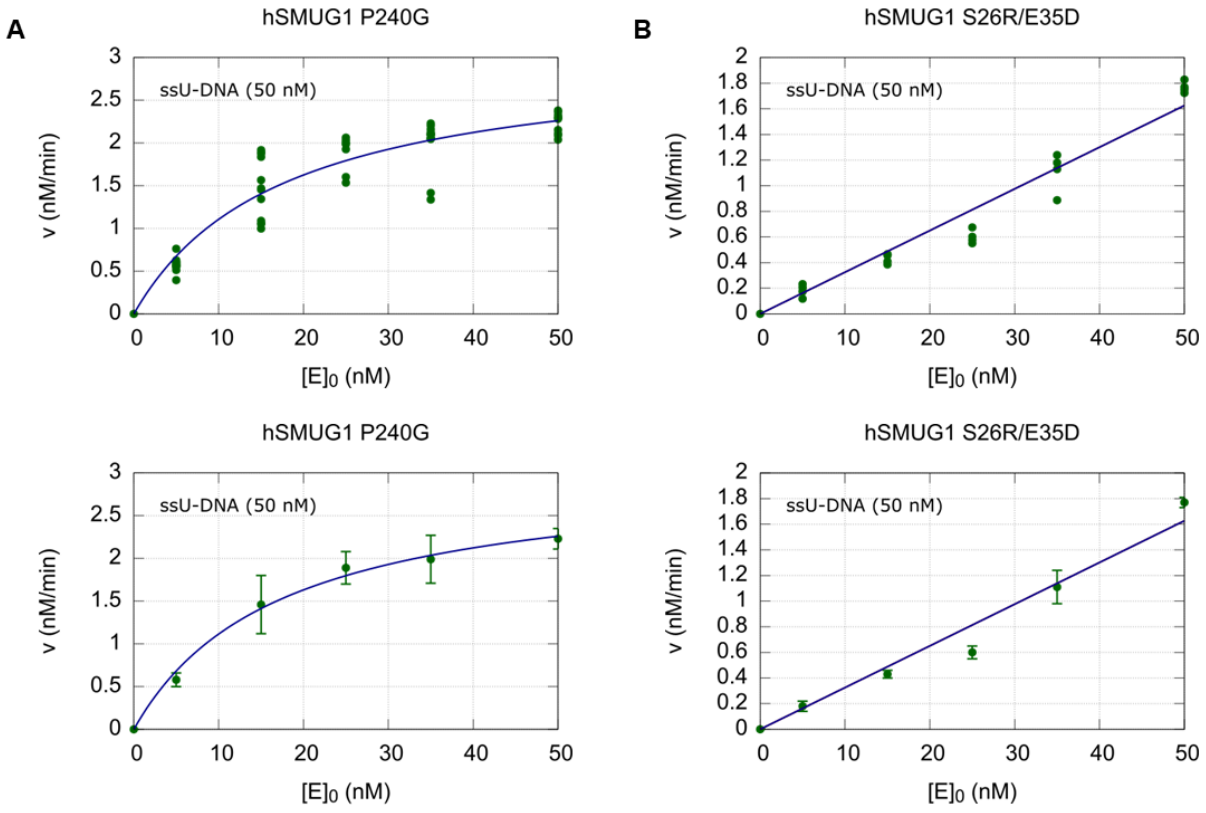


Figure 25. Concentration-dependent kinetics of hSMUG1 mutant proteins showing the excision velocity (v) in nM/min at ssU-DNA as a function of increasing (0–50 nM) initial enzyme concentrations, $[E]_0$ (nM). Upper panels; each value represents all independently measured data points. Lower panels; each value represents the average (±SD) of 4–10 independent measurements. ***(A)*** *Enzyme kinetics of hSMUG1 P240G adapted to saturation kinetics. All values exclusively represent the comparable data from the second trial.* ***(B)*** *Enzyme kinetics of hSMUG1 S26R/E35D adapted to linear fit curves. Abbreviations: ssU-DNA, single-stranded uracil-DNA; nM, nanomolar. Graphs generated using Gnuplot 5.2 patchlevel 8 via command prompt, two different scripts (see 3.5.1 Kinetic models), and two different functions (equation 3 and 4).*

4.3 Wild type hSMUG1 exhibited highest activity for uracil in ssDNA

The substrate (Figure 21A) was incubated with increasing amounts of wild type commercially purified hSMUG1(NEB) (Figure 26B) and in-house purified (by another member of the research group) (87) hSMUG1(TB) (Figure 26A). Repeated experiments of hSMUG1(NEB) demonstrates that this enzyme exhibited the highest glycosylase activity for ssU-DNA, removing the majority of all uracil residues at a very low enzyme concentration (Figure 26D) as compared to all other enzymes examined. At the highest hSMUG1(TB) concentration, nearly half of all uracil residues were removed from ssDNA (Figure 26C). No excision was observed without any enzyme (Figure 26A/B). All the experimental data (Figure 27, upper panels) and the average (±SD) (Figure 27B, lower panels) for both wild types were adapted into saturation kinetics (3.5.1 Kinetic model, equation 4), and their respective kinetic parameters are summarized in Table 3. The kinetic parameters show substantial variation and inactivated enzymes may have affected some measurements. The reduced UDG activity of hSMUG1(TB) is most likely explained by inactivation due to long-term storage of the unstable enzyme in non-optimal conditions (*i.e.* multiple freezing/thawing cycles and kept on ice during multiple experiments). Results from all repeated experiments are presented in Appendix B, Activity assays for excision of uracil-DNA by hSMUG1(TB) or hSMUG1(NEB).

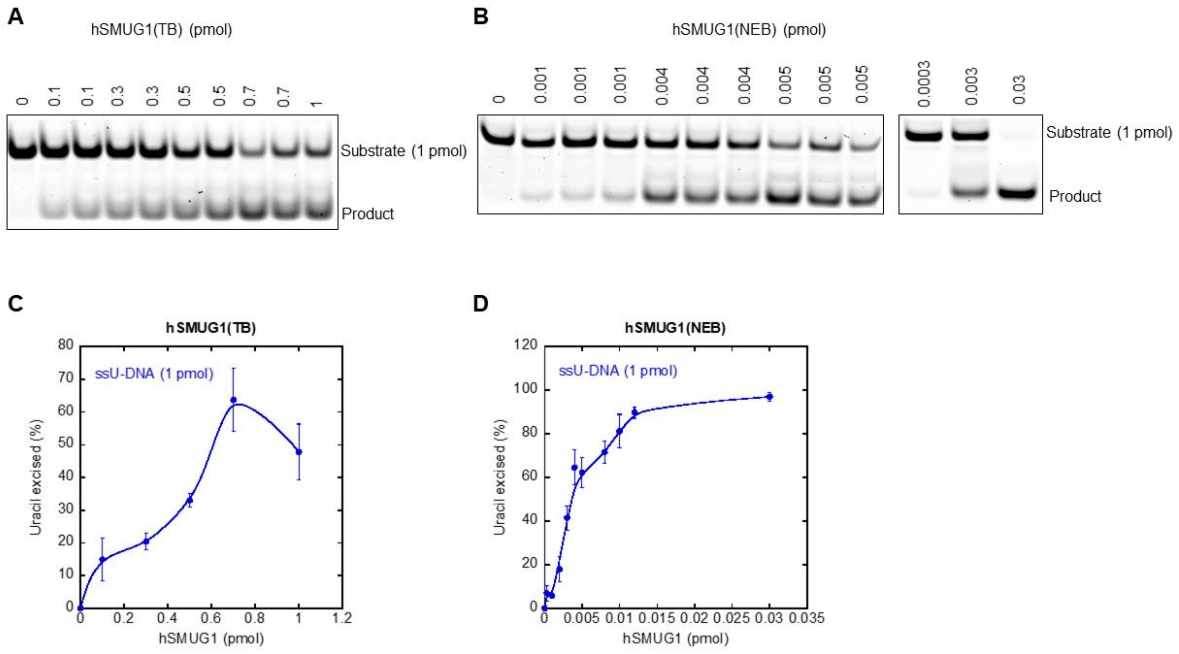


Figure 26. Excision of uracil in single-stranded DNA by wild type hSMUG1 purified in-house (TB) and by New England BioLabs (NEB) using the same conditions as described in Figure 21B. ***(A)*** *Increasing amounts of hSMUG1(TB) (0–1 pmol) were incubated with ssU-DNA.* ***(B)*** *Increasing amounts of hSMUG1(NEB) (0–0.03 pmol) were incubated with ssU-DNA.* ***(C)*** *Uracil excision (%) by hSMUG1(TB) is illustrated as a function of increasing enzyme amounts.* ***(D)*** *Uracil excision (%) by hSMUG1(NEB) as a function of increasing enzyme amounts. Each value represents the average (±SD) of 4–30 independent measurements. Abbreviation: ssU-DNA, single-stranded uracil-DNA.Graphs generated using KaleidaGraph (Courtesy of Svein Bjelland).*

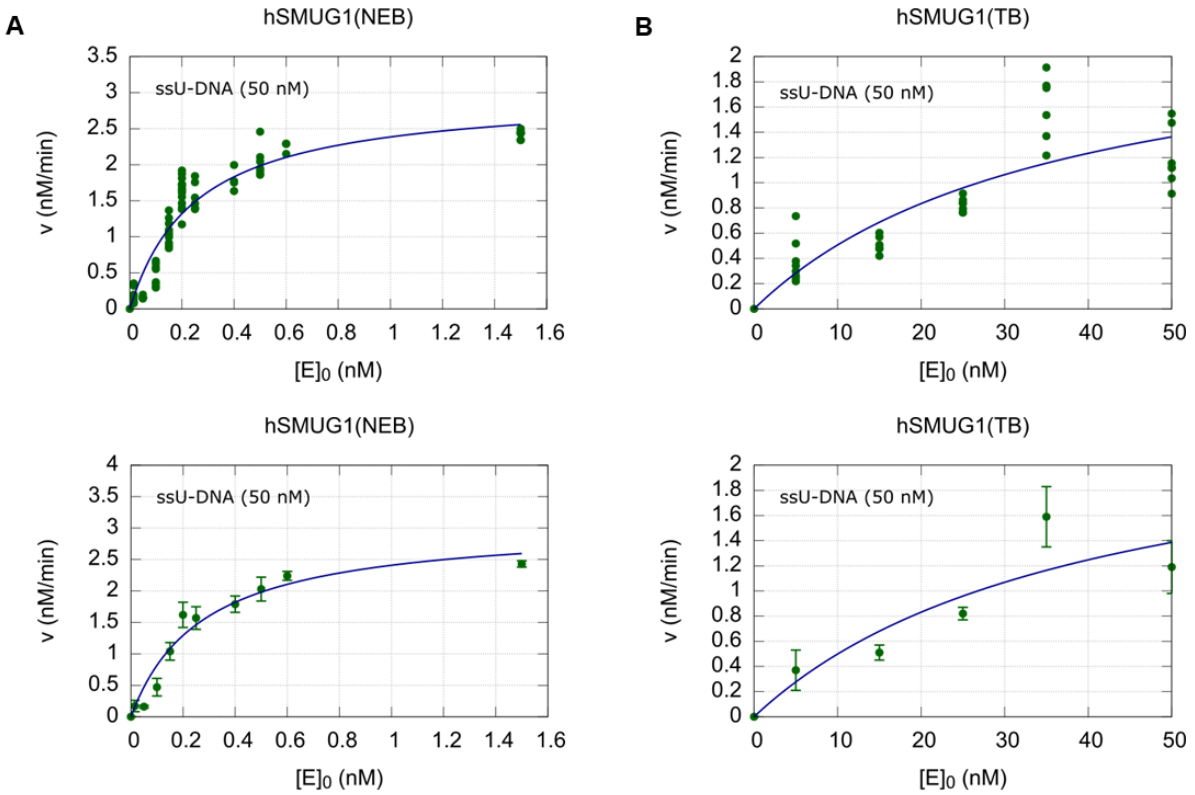


Figure 27. Concentration-dependent kinetics of wild type hSMUG1 illustrating the excision velocity (v) in nM/min at ssU-DNA as a function of the initial enzyme concentration, E_0 (nM). Bottom panels; each value represents the average (±SD) of 4–30 independent measurements. Top panels; each value represents all independently measured data points. ***(A)*** *Enzyme kinetics of hSMUG1(NEB) adapted to saturation kinetics.* ***(B)*** *Enzyme kinetics of hSMUG1(TB) adapted to both saturation and linear fit curves. Abbreviations: ssU-DNA, single-stranded uracil-DNA; nM, nanomolar; wt.TB, wild type purified in-house; wt.NEB, wild type purified by New England BioLabs. Graphs generated from Gnuplot 5.2 patchlevel 8 via command prompt, two different scripts (see 3.5.1 Kinetic model), and two different functions (equation 3 and 4).*

Table 3. Kinetic parameters of ssU-DNA excision activity by hSMUG1 P240G and wild type (NEB and TB), including asymptotic standard errors (±SD). Average; the value represents the average of multiple individual measurements. All; the value is calculated using all individual measurements. Abbreviations: K_D, dissociation constant; V_{max}, maximum velocity.

Parameter	Average	All	Parameter	Average	All
hSMUG1 P240G[c]			**hSMUG1(NEB)**		
K_D (nM)	17 ± 3	18 ± 3	K_D (nM)	0.28 ± 0.07	0.25 ± 0.03
V_{max} (nM/min)	3.0 ± 0.2	3.1 ± 0.2	V_{max} (nM/min)	3.1 ± 0.3	3.0 ± 0.1
V_{max}/K_D (min^{-1})	0.18	0.17	V_{max}/K_D (min^{-1})	11	12
hSMUG1 P240G[f]			**hSMUG1(TB)**		
K_D (nM)	40 ± 20	41 ± 9	K_D (nM)	40 ± 48	37 ± 18
V_{max} (nM/min)	4.0 ± 0.9	4.2 ± 0.5	V_{max} (nM/min)	2.5 ± 1.6	2.4 ± 0.6
V_{max}/K_D (min^{-1})	0.1	0.1	V_{max}/K_D (min^{-1})	0.06	0.06

[c], enzyme preparation compared/measured at a similar time; [f], the freshest enzyme preparation tested.

5 Discussion

5.1 hSMUG1 excision activity for hmU in ssDNA is destroyed by replacing Pro240 with Gly

Pro240 in hSMUG1 is a part of the intercalating loop (His239–Lys249) serving as a "wedge" penetrating the DNA in the region of the damaged base (66). Pro is a rigid residue not allowing rotations relative to adjacent residues (*i.e* His239 and Ser241) leading to stabilization of these residues in a specific position. In hSMUG1, His239 shares a strong ionic bond with the phosphate of the DNA backbone and a less strong H-bond to uracil O2 (Figure 10).

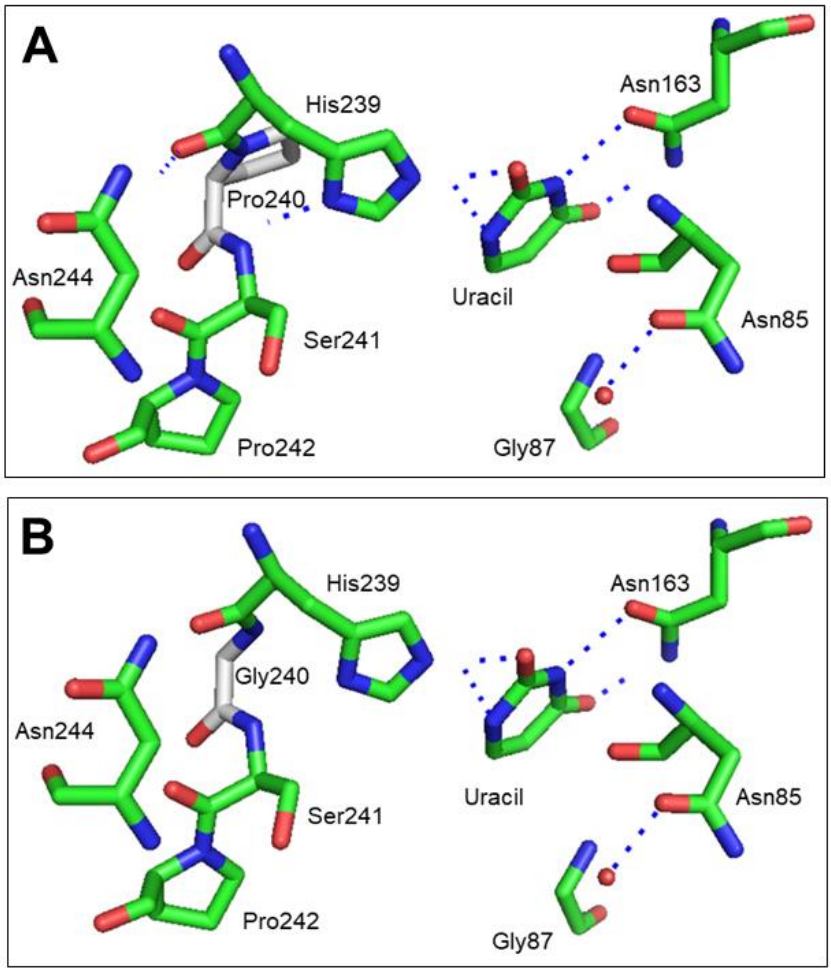


Figure 28. Close-up view of hSMUG1 residues in the region of site 240 (grey), partly showing the intercalating loop (His239–Lys249) and the catalytic residues (His239 and Asn85). Gly87–Met91 (not sown) contributes to C5-substituent recognition. The red sphere indicates a water molecule. Blue dotted lines indicate possible polar interactions. ***(A)*** *The native hSMUG1, showing the P240 residue (grey) and its interactions with Ser241, Asn244, and uracil.* ***(B)*** *The P240G variant of the native hSMUG1 showing the G240 residue (grey).* *The hSMUG1 residues were generated from the xSMUG1 crystal structure obtained from PDB (entry 1OE5) and constructed using PyMol Molecular Graphics System, Version 1.7.4, Schrödinger, LLC.*

Together with Asn85, these residues are thought to catalyze the hydrolysis of the N-glycosidic bond connecting the uracil DNA base to deoxyribose. When the rigid Pro (Figure 28A, grey residue) is replaced by the small and more flexible Gly (Figure 28B, grey residue), it might allow His239 to rotate away from the substrate resulting in rupture of (one of) its many hydrogen-bonds in SMUG1, including connections to Asn244 (amine group), Ser241 (backbone) or O2 of the uracil substrate base (Figure 28A), which might result in many conformational changes.

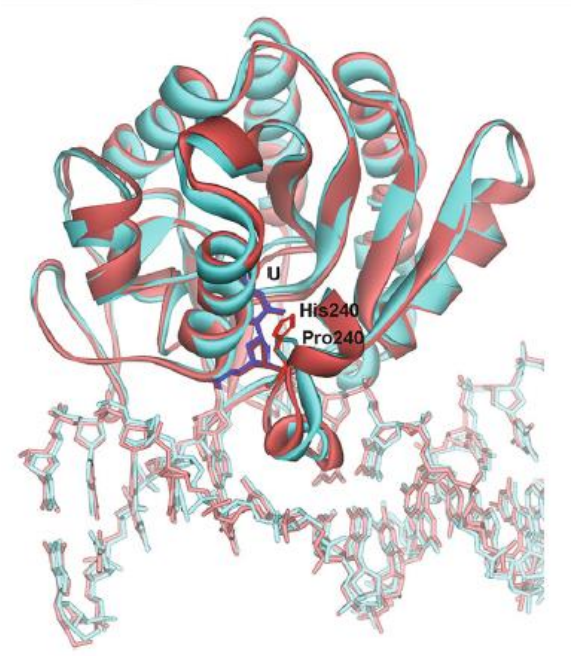


Figure 29. Enzyme-substrate complex for wild type (blue) and the polymorphic variant P240H (red) shows insignificant conformational changes when Pro is substituted for His, probably due to stabilization of the imidazole ring by nearby residue. The extra-helical position of uridine and the amino acid comprising a P240H substitution are shown (66).

Generally, Gly is believed to be necessary for conformational changes in the active sites of enzymes (84), supporting our suggestions above. Another possible explanation for the change in activity of the hSMUG1 P240G protein is that the Phe98 or/and the Asn163, both used to stabilize the flipped-out DNA base, are negatively affected by the possible conformational changes resulting from the substitution. Pro 240 also interacts with Asn244 (Figure 28A), where introduction of Gly instead might affect this interaction, allowing for an even larger conformational change. Recently, it was demonstrated that the hSMUG1 P240H mutant variant did not affect the DNA binding nor made a significant change in the enzyme-substrate structure (Figure 29), however, in contrast to Gly240, His240 imidazole ring can form stacking with the Trp251 indole ring and hydrogen bond with Arg140, leading to stabilization of the region. Interestingly, hSMUG1 P240H is a single nucleotide polymorphism (SNP) associated with melanoma (BioMuta database) (66).

An interesting phenomenon reported by the present study is that neither the P240G nor S26R/E35D variant reveals any activity for hmU (Figure 30A and B, red) as opposed to U (Figure 30A and B, dark blue) in ssDNA (87), even though the 25 nucleotide sequence is completely identical by only comprising a more bulky 5-substituent (hmU) at the uracil site. This indicates that these mutant proteins might not be able to recognize all the 5-substituents in SMUG1s broad substrate repertoire. The total obstruction of glycosylase activity for hmU-ssDNA might be due to the flexibility of Gly240 and its possible non-restricted rotation effect on its adjacent catalytic His239 residue, and that this (possibly together with the bulky 5-substituent of hmU) drives the His239 residue away from its substrate instead of facilitating recognition and hydrogen-bond interaction. Regarding the double mutant (S26R/E35D) where both substitutions are present far away from the active site of SMUG1, the total glycosylase activity obstruction for hmU-ssDNA is more difficult to explain. It has earlier been demonstrated that Glu29, Glu33, and Glu231 are necessary for proper DKC1-SMUG1 binding (44). The Ser26 and Glu35 that have been substituted (for Arg and Asp, respectively) in the double mutant of hSMUG1 are present in the same DKC1-binding region, in addition, Glu seems to be an important residue in this interaction, which tempts to speculate whether this substitution affects the DKC1 binding to hSMUG1 S26R/E35D and that this interaction might be crucial for glycosylase activity at some substrates (*i.e.* hmU-ssDNA, but not ssU-DNA) or alternatively, the double mutant contributes to conformational changes affecting the hmU recognition.

5.2 hSMUG1 mutant proteins P240G and S26R/E35D retain highest activity for uracil in ssDNA and R-loop DNA

The present study demonstrates that neither the P240G nor the S26R/E35D (86) amino acid replacements in hSMUG1 obstruct the glycosylase activity for uracil in ssDNA and both mutant proteins (Figure 30A and B). The only similarity shared by all four examined enzymes is the ssU-DNA preference as the most favorable substrate, which is also consistent with another study (18). Both hSMUG1 P240G and S26R/E35D proteins share the same relative substrate preference and show a shift in substrate preference for uracil in R-loop (Figure 30A and B, light blue) and bubble DNA (Figure 30A and B, purple) with higher relative activity for R-loop DNA, which is opposite to the wild type enzyme (Figure 30C and D).

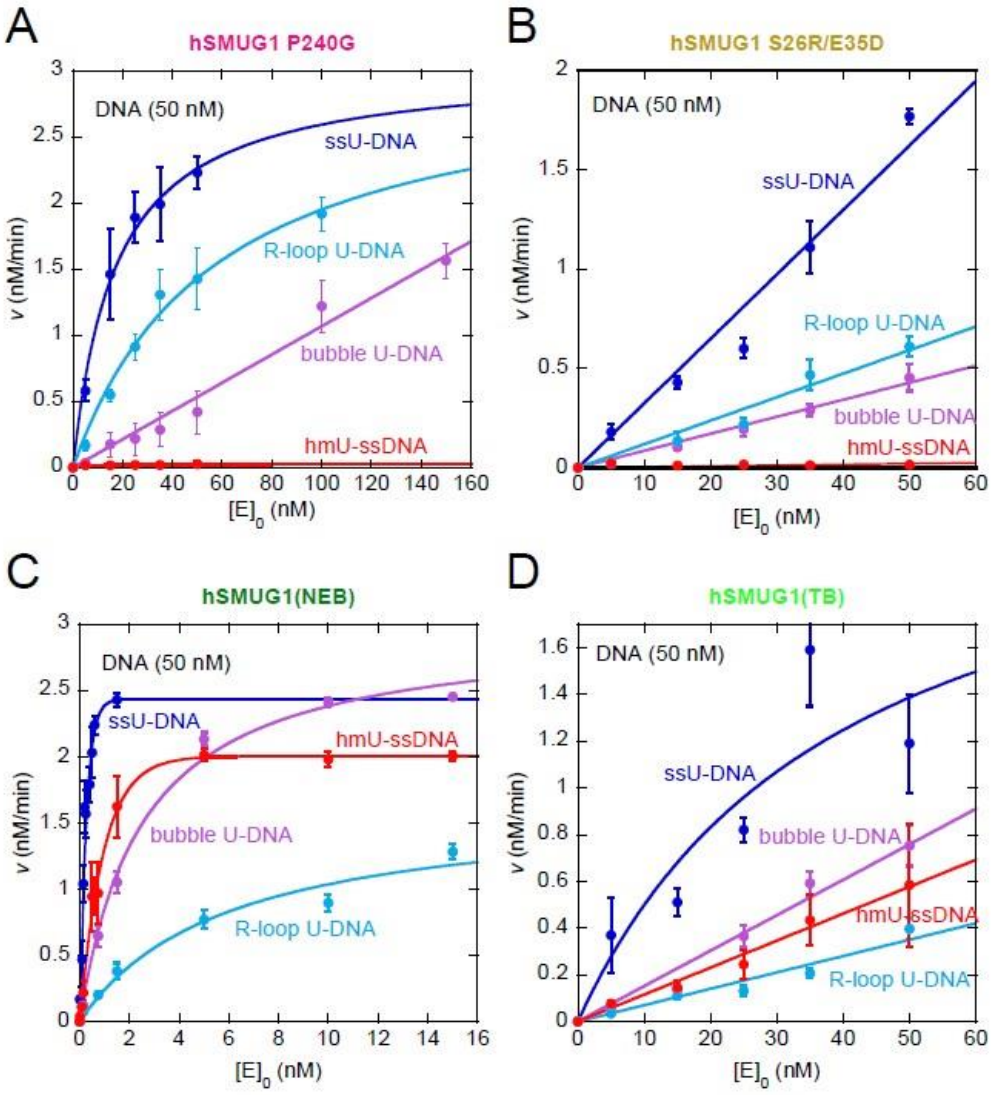


Figure 30. hSMUG1 substrate preferences of all enzymes examined by the research group, showing the increasing enzyme concentration, $[E]_0$ (nM), as a function of the velocity (nM/min). Each value represents the average (±SD) of multiple independent measurements. ***(A)*** *Substrate preference by the hSMUG1 P240G variant.* ***(B)*** *Substrate preference by the hSMUG1 S26R/E35D variant.* ***(C)*** *Substrate preference by the native hSMUG1(NEB).* ***(D)*** *Substrate preference by the native hSMUG1(TB). Abbreviations: v, velocity; nM, nanomolar, NEB, commercial wild type purified by New England Biolabs; TB, wild type purified by another member of the research group. Plots generated by KaleidaGraph (Courtesy of Svein Bjelland).*

R-loop is a three-stranded nucleic acid structure containing a DNA:RNA hybrid, which probably affects the way glycosylases bind to the substrate. The crystal structure of the xSMUG1-DNA complex reveals a wedge formed by a loop suggesting an invasive DNA penetration of the double-helix DNA structure and motifs that apparently interact with adjacent base pairs of the damaged base (47,53). It is also suggested that the C-terminal part of the wedge (243-RNPQANK-249) interacts with the distal strand of DNA, whereas the N-terminal part of the wedge (239-HPSPR-243) faces the uracil-containing strand. All these interactions, in addition to a more flexible residue (Gly instead of the rigid Pro) in the N-terminal wedge, might

explain the invasive interaction with triple-stranded R-loop U-DNA. Also, five out of seven mutant proteins regarding the C-terminal wedge region (close to P240) have previously demonstrated a significant increase in the hSMUG1 activity of dsDNA opposite G, where the P245A substitution increased the activity more than 2-fold compared to wild type (53). For more graphs, see Appendix B, Enzyme kinetics.

Table 4. Comparable kinetic variables of hSMUG1 (wild type(NEB) and P240G) excision activity from the present and previous studies showing reaction conditions. The K_D represents the average of multiple independent measurements (±SD). Abbreviations: wt, wild type; nt, nucleotides; mM, millimolar; K_D, dissociation constant; K_M, Michaelis-Menten constant.

hSMUG1	Substrate base	Opposite base	DNA (nt)	$MgCl_2$ (mM)	KCl (mM)	K_D *or* K_M (nM)	Reference
P240G[F]	U	none	25	0	70	40 ± 20[D]	
P240G[C]	U	none	25	0	70	17 ± 3[D]	
P240G	R-loop U	none	54	0	70	60 ± 10[D]	(86)
P240G	hmU	none	25	0	70	0[D]	(87)
NEB	U	none	25	0	70	0.28 ± 0.07[D]	
NEB	R-loop U	none	54	0	70	6 ± 2[D]	(86)
NEB	hmU	none	25	0	70	0.8 ± 0.1[D]	(87)
wt	U	none	50	7.5	70	1700±300[M]	(18)
wt	U	none	19	7.5	0	3000±500[M]	(53)
wt	U	G	50	7.5	70	1300±200[M]	(18)
wt	U	G	19	0	0	12[M]	(40)
wt	U	G	19	0	0	2.2[M]	(54)
wt	U	A	50	7.5	70	4000±900[M]	(18)
wt	U	A	19	0	0	96[M]	(40)
wt	hmU	G	50	7.5	70	2700±300[M]	(18)
wt	hmU	A	50	7.5	70	4600±400[M]	(18)
wt	hmU	none	50	7.5	70	3800±500[M]	(18)

[F], Fresh enzyme used in the first trial of activity assays; [C], Enzyme for comparison used in the second trial of activity assays; [D], Values representing K_D; [M], Values representing K_M.

Many kinetic values have been established for hSMUG1 excision activity and the majority of them reveals very high K_m values (Table 4), indicating a low affinity of the enzyme to its substrate. In the present study, K_D (dissociation constant) was measured (Table 3) from experimental data to be compared with previously measured K_m parameters, and the results revealed lower dissociation for almost all substrates examined (Table 4, top seven rows). As for the hSMUG1 P240G variant, it reveals K_D of 40 ± 20 nM (fresh enzyme preparation) and 17 ± 3 nM (later state enzyme preparation) for ssU-DNA, and 16 ± 10 nM for R-loop U-DNA. While the previously reported values for ssU-DNA containing 50 and 19 nucleotides (for

comparison with the 25 nucleotide oligo used in the present study), reveal K_m as high as 1700±300 and 3000±500 nM respectively. The wild type (NEB) showed, as expected, lower K_D compared to the P240G variant for all substrates except for hmU-ssDNA. Interestingly, previously K_m values for hmU-ssDNA are reported as 3800±500 nM revealing a much lower affinity compared to the values reported in the present study. A huge variation is previously displayed for dsU-DNA opposite G and A where K_m values of 2.2–1300 and 96–4000 nM respectively have been reported (Table 4).

References

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.
2. The National Academies Press. (2007) *The Limits of Organic Life in Planetary Systems*, Washington, DC.
3. Krokan, H.E. and Bjoras, M. (2013) Base excision repair. *Cold Spring Harb Perspect Biol*, **5**, a012583.
4. Krokan, H.E., Saetrom, P., Aas, P.A., Pettersen, H.S., Kavli, B. and Slupphaug, G. (2014) Error-free versus mutagenic processing of genomic uracil--relevance to cancer. *DNA Repair (Amst)*, **19**, 38-47.
5. Olinski, R., Zastawny, T., Budzbon, J., Skokowski, J., Zegarski, W. and Dizdaroglu, M. (1992) DNA base modifications in chromatin of human cancerous tissues. *FEBS Lett*, **309**, 193-198.
6. Xie, H., Gong, Y., Dai, J., Wu, X. and Gu, J. (2015) Genetic variations in base excision repair pathway and risk of bladder cancer: a case-control study in the United States. *Mol Carcinog*, **54**, 50–57.
7. Klungland, A. and Bjelland, S. (2007) Oxidative damage to purines in DNA: role of mammalian Ogg1. *DNA Repair (Amst)*, **6**, 481-488.
8. Bjelland, S. and Seeberg, E. (2003) Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. *Mutat Res*, **531**, 37-80.
9. Ide, H. and Kotera, M. (2004) Human DNA glycosylases involved in the repair of oxidatively damaged DNA. *Biol Pharm Bull*, **27**, 480-485.
10. Krokan, H.E., Standal, R. and Slupphaug, G. (1997) DNA glycosylases in the base excision repair of DNA. *Biochem J*, **325 (Pt 1)**, 1-16.
11. Krokan, H.E., Drablos, F. and Slupphaug, G. (2002) Uracil in DNA--occurrence, consequences and repair. *Oncogene*, **21**, 8935-8948.
12. Pettersen, H.S., Galashevskaya, A., Doseth, B., Sousa, M.M., Sarno, A., Visnes, T., Aas, P.A., Liabakk, N.B., Slupphaug, G., Saetrom, P. *et al.* (2015) AID expression in B-cell lymphomas causes accumulation of genomic uracil and a distinct AID mutational signature. *DNA Repair (Amst)*, **25**, 60-71.
13. Knaevelsrud, I., Ruoff, P., Anensen, H., Klungland, A., Bjelland, S. and Birkeland, N.K. (2001) Excision of uracil from DNA by the hyperthermophilic Afung protein is dependent on the opposite base and stimulated by heat-induced transition to a more open structure. *Mutat Res*, **487**, 173-190.
14. Nilsen, H. and Krokan, H.E. (2001) Base excision repair in a network of defence and tolerance. *Carcinogenesis*, **22**, 987-998.
15. Stavnezer, J., Guikema, J.E. and Schrader, C.E. (2008) Mechanism and regulation of class switch recombination. *Annu Rev Immunol*, **26**, 261-292.
16. Rada, C., Williams, G.T., Nilsen, H., Barnes, D.E., Lindahl, T. and Neuberger, M.S. (2002) Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr Biol*, **12**, 1748–1755.
17. Kavli, B., Andersen, S., Otterlei, M., Liabakk, N.B., Imai, K., Fischer, A., Durandy, A., Krokan, H.E. and Slupphaug, G. (2005) B cells from hyper-IgM patients carrying *UNG* mutations lack ability to remove uracil from ssDNA and have elevated genomic uracil. *J Exp Med*, **201**, 2011–2021.
18. Kavli, B., Sundheim, O., Akbari, M., Otterlei, M., Nilsen, H., Skorpen, F., Aas, P.A., Hagen, L., Krokan, H.E. and Slupphaug, G. (2002) hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup. *J Biol Chem*, **277**, 39926-39936.
19. Nilsen, H., An, Q. and Lindahl, T. (2005) Mutation frequencies and AID activation state in B-cell lymphomas from Ung-deficient mice. *Oncogene*, **24**, 3063–3066.
20. Akbari, M., Otterlei, M., Pena-Diaz, J., Aas, P.A., Kavli, B., Liabakk, N.B., Hagen, L., Imai, K., Durandy, A., Slupphaug, G. *et al.* (2004) Repair of U/G and U/A in DNA by UNG2-associated repair complexes takes place predominantly by short-patch repair both in proliferating and growth-arrested cells. *Nucleic Acids Res*, **32**, 5486–5498.

21. Slupphaug, G., Eftedal, I., Kavli, B., Bharati, S., Helle, N.M., Haug, T., Levine, D.W. and Krokan, H.E. (1995) Properties of a recombinant human uracil-DNA glycosylase from the *UNG* gene and evidence that UNG encodes the major uracil-DNA glycosylase. *Biochemistry*, **34**, 128–138.
22. Sarno, A., Lundbæk, M., Liabakk, N.B., Aas, P.A., Mjelle, R., Hagen, L., Sousa, M.M.L., Krokan, H.E. and Kavli, B. (2019) Uracil-DNA glycosylase UNG1 isoform variant supports class switch recombination and repairs nuclear genomic uracil. *Nucleic Acids Res*.
23. Guikema, J.E., Linehan, E.K., Tsuchimoto, D., Nakabeppu, Y., Strauss, P.R., Stavnezer, J. and Schrader, C.E. (2007) APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination. *J Exp Med*, **204**, 3017–3026.
24. Alexeeva, M., Moen, M.N., Xu, X.M., Rasmussen, A., Leiros, I., Kirpekar, F., Klungland, A., Alsøe, L., Nilsen, H.L. and Bjelland, S. (2020) Intrinsic strand-incision activity of human UNG predicts nick generation in antibody diversification.
25. Masani, S., Han, L. and Yu, K. (2013) Apurinic/apyrimidinic endonuclease 1 is the essential nuclease during immunoglobulin class switch recombination. *Mol Cell Biol*, **33**, 1468–1473.
26. Di Noia, J.M., Rada, C. and Neuberger, M.S. (2006) SMUG1 is able to excise uracil from immunoglobulin genes: insight into mutation versus repair. *EMBO J*, **25**, 585-595.
27. Durandy, A. (2003) Activation-induced cytidine deaminase: a dual role in class-switch recombination and somatic hypermutation. *Eur J Immunol*, **33**, 2069-2073.
28. Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y. and Honjo, T. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, **102**, 553–563.
29. Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.L. *et al.* (2013) Signatures of mutational processes in human cancer. *Nature*, **500**, 415–421.
30. Matsubara, M. (2004) Mutational analysis of the damage-recognition and catalytic mechanism of human SMUG1 DNA glycosylase. *Nucleic Acids Research*, **32**, 5291-5302.
31. Kemmerich, K., Dingler, F.A., Rada, C. and Neuberger, M.S. (2012) Germline ablation of SMUG1 DNA glycosylase causes loss of 5-hydroxymethyluracil- and UNG-backup uracil-excision activities and increases cancer predisposition of Ung-/-Msh2-/- mice. *Nucleic Acids Res*, **40**, 6016-6025.
32. Kaufman, E.R. (1986) Biochemical analysis of toxic effects of 5-hydroxymethyl-2'-deoxyuridine in mammalian cells. *Somat Cell Mol Genet*, **12**, 501-512.
33. Kahilainen Li Fau - Bergstrom, D.E., Bergstrom De Fau - Vilpo, J.A. and Vilpo, J.A. (1985) 5-Hydroxymethyl-2'-deoxyuridine. Cytotoxicity and DNA incorporation studied by using a novel [2-14C]-derivative with normal and leukemic human hematopoietic cells. **39**.
34. Mellac, S., Fazakerley, G.V. and Sowers, L.C. (1993) Structures of base pairs with 5-(hydroxymethyl)-2'-deoxyuridine in DNA determined by NMR spectroscopy. *Biochemistry*, **32**, 7779-7786.
35. Hashimoto, H., Hong, S., Bhagwat, A.S., Zhang, X. and Cheng, X. (2012) Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. *Nucleic Acids Res*, **40**, 10203–10214.
36. Rusmintratip, V. and Sowers, L.C. (2000) An unexpectedly high excision capacity for mispaired 5-hydroxymethyluracil in human cell extracts. *Proc Natl Acad Sci U S A*, **97**, 14183-14187.
37. Hollstein, M.C., Brooks, P., Linn, S. and Ames, B.N. (1984) Hydroxymethyluracil DNA glycosylase in mammalian cells. *Proc Natl Acad Sci U S A*, **81**, 4003-4007.
38. Boorstein, R.J., Cummings, A., Jr., Marenstein, D.R., Chan, M.K., Ma, Y., Neubert, T.A., Brown, S.M. and Teebor, G.W. (2001) Definitive identification of mammalian 5-hydroxymethyluracil DNA N-glycosylase activity as SMUG1. *J Biol Chem*, **276**, 41991-41997.
39. Matsubara, M., Masaoka, A., Tanaka, T., Miyano, T., Kato, N., Terato, H., Ohyama, Y., Iwai, S. and Ide, H. (2003) Mammalian 5-formyluracil-DNA glycosylase. 1. Identification and characterization of a novel activity that releases 5-formyluracil from DNA. *Biochemistry*, **42**, 4993-5002.
40. Masaoka, A., Matsubara, M., Hasegawa, R., Tanaka, T., Kurisu, S., Terato, H., Ohyama, Y., Karino, N., Matsuda, A. and Ide, H. (2003) Mammalian 5-formyluracil-DNA glycosylase. 2. Role of SMUG1 uracil-DNA glycosylase in repair of 5-formyluracil and other oxidized and deaminated base lesions. *Biochemistry*, **42**, 5003-5012.

41. Hong, H. and Wang, Y. (2007) Derivatization with Girard reagent T combined with LC-MS/MS for the sensitive detection of 5-formyl-2'-deoxyuridine in cellular DNA. *Anal Chem*, **79**, 322-326.
42. Darwanto, A., Theruvathu, J.A., Sowers, J.L., Rogstad, D.K., Pascal, T., Goddard, W., 3rd and Sowers, L.C. (2009) Mechanisms of base selection by human single-stranded selective monofunctional uracil-DNA glycosylase. *The Journal of biological chemistry*, **284**, 15835-15846.
43. Nilsen, H., Haushalter, K.A., Robins, P., Barnes, D.E., Verdine, G.L. and Lindahl, T. (2001) Excision of deaminated cytosine from the vertebrate genome: role of the SMUG1 uracil-DNA glycosylase. *EMBO J*, **20**, 4278-4286.
44. Jobert, L., Skjeldam, H.K., Dalhus, B., Galashevskaya, A., Vagbo, C.B., Bjoras, M. and Nilsen, H. (2013) The human base excision repair enzyme SMUG1 directly interacts with DKC1 and contributes to RNA quality control. *Mol Cell*, **49**, 339-345.
45. Knaevelsrud, I., Slupphaug, G., Leiros, I., Matsuda, A., Ruoff, P. and Bjelland, S. (2009) Opposite-base dependent excision of 5-formyluracil from DNA by hSMUG1. *Int J Radiat Biol*, **85**, 413-420.
46. An, Q., Robins, P., Lindahl, T. and Barnes, D.E. (2005) C --> T mutagenesis and gamma-radiation sensitivity due to deficiency in the Smug1 and Ung DNA glycosylases. *Embo j*, **24**, 2205-2213.
47. Wibley, J.E.A., Waters, T.R., Haushalter, K., Verdine, G.L. and Pearl, L.H. (2003) Structure and Specificity of the Vertebrate Anti-Mutator Uracil-DNA Glycosylase SMUG1. *Molecular Cell*, **11**, 1647-1659.
48. Mi, R., Dong, L., Kaulgud, T., Hackett, K.W., Dominy, B.N. and Cao, W. (2009) Insights from xanthine and uracil DNA glycosylase activities of bacterial and human SMUG1: switching SMUG1 to UDG. *J Mol Biol*, **385**, 761-778.
49. Dong, L., Meira, L.B., Hazra, T.K., Samson, L.D. and Cao, W. (2008) Oxanine DNA glycosylase activities in mammalian systems. *DNA Repair (Amst)*, **7**, 128-134.
50. An, Q., Robins, P., Lindahl, T. and Barnes, D.E. (2007) 5-Fluorouracil incorporated into DNA is excised by the Smug1 DNA glycosylase to reduce drug cytotoxicity. *Cancer Res*, **67**, 940-945.
51. Nagaria, P., Svilar, D., Brown, A.R., Wang, X.H., Sobol, R.W. and Wyatt, M.D. (2013) SMUG1 but not UNG DNA glycosylase contributes to the cellular response to recovery from 5-fluorouracil induced replication stress. *Mutat Res*, **743-744**, 26-32.
52. Alexeeva, M., Moen, M.N., Grøsvik, K., Tesfahun, A.N., Xu, X.M., Muruzábal-Lecumberri, I., Olsen, K.M., Rasmussen, A., Ruoff, P., Kirpekar, F. *et al.* (2019) Excision of uracil from DNA by hSMUG1 includes strand incision and processing. *Nucleic Acids Res*, **47**, 779–793.
53. Pettersen, H.S., Sundheim, O., Gilljam, K.M., Slupphaug, G., Krokan, H.E. and Kavli, B. (2007) Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms. *Nucleic Acids Res*, **35**, 3879-3892.
54. Matsubara, M., Tanaka, T., Terato, H., Ohmae, E., Izumi, S., Katayanagi, K. and Ide, H. (2004) Mutational analysis of the damage-recognition and catalytic mechanism of human SMUG1 DNA glycosylase. *Nucleic Acids Res*, **32**, 5291–5302.
55. Wiederhold, L., Leppard, J.B., Kedar, P., Karimi-Busheri, F., Rasouli-Nia, A., Weinfeld, M., Tomkinson, A.E., Izumi, T., Prasad, R., Wilson, S.H. *et al.* (2004) AP endonuclease-independent DNA base excision repair in human cells. *Mol Cell*, **15**, 209-220.
56. Vågbø, C.B. and Slupphaug, G. (2020) RNA in DNA repair. *DNA Repair*, **95**, 102927.
57. Masaoka, A., Matsubara, M., Tanaka, T., Terato, H., Ohyama, Y., Kubo, K. and Ide, H. (2003) Repair roles of hSMUG1 assessed by damage specificity and cellular activity. *Nucleic Acids Res Suppl*, 263-264.
58. Doseth, B., Ekre, C., Slupphaug, G., Krokan, H.E. and Kavli, B. (2012) Strikingly different properties of uracil-DNA glycosylases UNG2 and SMUG1 may explain divergent roles in processing of genomic uracil. *DNA Repair (Amst)*, **11**, 587-593.
59. Alsoe, L., Sarno, A., Carracedo, S., Domanska, D., Dingler, F., Lirussi, L., SenGupta, T., Tekin, N.B., Jobert, L., Alexandrov, L.B. *et al.* (2017) Uracil Accumulation and Mutagenesis Dominated by Cytosine Deamination in CpG Dinucleotides in Mice Lacking UNG and SMUG1. *Sci Rep*, **7**, 7199.
60. Dingler, F.A., Kemmerich, K., Neuberger, M.S. and Rada, C. (2014) Uracil excision by endogenous SMUG1 glycosylase promotes efficient Ig class switching and impacts on A:T substitutions during somatic mutation. *Eur J Immunol*, **44**, 1925–1935.

61. Kuznetsova, A.A., Iakovlev, D.A., Misovets, I.V., Ishchenko, A.A., Saparbaev, M.K., Kuznetsov, N.A. and Fedorova, O.S. (2017) Pre-steady-state kinetic analysis of damage recognition by human single-strand selective monofunctional uracil-DNA glycosylase SMUG1. *Mol Biosyst*, **13**, 2638-2649.
62. Zhang, Z., Shen, J., Yang, Y., Li, J., Cao, W. and Xie, W. (2016) Structural Basis of Substrate Specificity in Geobacter metallireducens SMUG1. *ACS Chem Biol*, **11**, 1729-1736.
63. Schormann, N., Ricciardi, R. and Chattopadhyay, D. (2014) Uracil-DNA glycosylases-structural and functional perspectives on an essential family of DNA repair enzymes. *Protein Sci*, **23**, 1667-1685.
64. Iakovlev, D.A., Alekseeva, I.V., Vorobjev, Y.N., Kuznetsov, N.A. and Fedorova, O.S. (2019) The Role of Active-Site Residues Phe98, His239, and Arg243 in DNA Binding and in the Catalysis of Human Uracil-DNA Glycosylase SMUG1. *Molecules*, **24**.
65. Slupphaug, G., Mol, C.D., Kavli, B., Arvai, A.S., Krokan, H.E. and Tainer, J.A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature*, **384**, 87–92.
66. Alekseeva, I.V., Bakman, A.S., Iakovlev, D.A., Kuznetsov, N.A. and Fedorova, O.S. (2021) [Comparative Analysis of the Activity of the Polymorphic Variants of Human Uracil-DNA-Glycosylases SMUG1 and MBD4]. *Mol Biol (Mosk)*, **55**, 277-288.
67. Scharer, O.D. and Campbell, A.J. (2009) Wedging out DNA damage. *Nat Struct Mol Biol*, **16**, 102-104.
68. Raja, S. and Van Houten, B. (2021) The Multiple Cellular Roles of SMUG1 in Genome Maintenance and Cancer. *Int J Mol Sci*, **22**.
69. Nilsen, H., Lindahl, T. and Verreault, A. (2002) DNA base excision repair of uracil residues in reconstituted nucleosome core particles. *EMBO J*, **21**, 5943-5952.
70. Otterlei, M., Warbrick, E., Nagelhus, T.A., Haug, T., Slupphaug, G., Akbari, M., Aas, P.A., Steinsbekk, K., Bakke, O. and Krokan, H.E. (1999) Post-replicative base exsision repair in replication foci. *EMBO J*, **18**, 3834-3844.
71. An, J.S., Huang, M.N., Song, Y.M., Li, N., Wu, L.Y. and Zhan, Q.M. (2013) A preliminary study of genes related to concomitant chemoradiotherapy resistance in advanced uterine cervical squamous cell carcinoma. *Chin Med J (Engl)*, **126**, 4109-4115.
72. Abdel-Fatah, T.M., Albarakati, N., Bowell, L., Agarwal, D., Moseley, P., Hawkes, C., Ball, G., Chan, S., Ellis, I.O. and Madhusudan, S. (2013) Single-strand selective monofunctional uracil-DNA glycosylase (SMUG1) deficiency is linked to aggressive breast cancer and predicts response to adjuvant therapy. *Breast Cancer Res Treat*, **142**, 515–527.
73. Zhong, J., Jermusyk, A., Wu, L., Hoskins, J.W., Collins, I., Mocci, E., Zhang, M., Song, L., Chung, C.C., Zhang, T. *et al.* (2020) A Transcriptome-Wide Association Study Identifies Novel Candidate Susceptibility Genes for Pancreatic Cancer. *J Natl Cancer Inst*, **112**, 1003-1012.
74. Oliveira, D.M., Laudanna, C., Migliozzi, S., Zoppoli, P., Santamaria, G., Grillone, K., Elia, L., Mignogna, C., Biamonte, F., Sacco, R. *et al.* (2018) Identification of different mutational profiles in cancers arising in specific colon segments by next generation sequencing. *Oncotarget*, **9**, 23960-23974.
75. Sun, H., Cao, D., Ma, X., Yang, J., Peng, P., Yu, M., Zhou, H., Zhang, Y., Li, L., Huo, X. *et al.* (2019) Identification of a Prognostic Signature Associated With DNA Repair Genes in Ovarian Cancer. *Front Genet*, **10**, 839.
76. Lirussi, L. and Nilsen, H. (2019) Telomere maintenance: regulating hTERC fate through RNA modifications. *Mol Cell Oncol*, **6**, e1670489.
77. Lirussi, L., Demir, Ö., You, P., Sarno, A., Amaro, R.E. and Nilsen, H. (2021) RNA Metabolism Guided by RNA Modifications: The Role of SMUG1 in rRNA Quality Control. *Biomolecules*, **11**.
78. Koliadenko, V. and Wilanowski, T. (2020) Additional functions of selected proteins involved in DNA repair. *Free Radical Biology and Medicine*, **146**, 1-15.
79. Aubert, M., O'Donohue, M.-F., Lebaron, S. and Gleizes, P.-E. (2018) Pre-Ribosomal RNA Processing in Human Cells: From Mechanisms to Congenital Diseases. *Biomolecules*, **8**, 123.
80. Cong, Y.-S., Wright, W.E. and Shay, J.W. (2002) Human telomerase and its regulation. *Microbiology and molecular biology reviews : MMBR*, **66**, 407-425.

81. Kroustallaki, P., Lirussi, L., Carracedo, S., You, P., Esbensen, Q.Y., Gotz, A., Jobert, L., Alsoe, L., Saetrom, P., Gagos, S. *et al.* (2019) SMUG1 Promotes Telomere Maintenance through Telomerase RNA Processing. *Cell Rep*, **28**, 1690-1702 e1610.
82. Studier, F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*, **41**, 207-234.
83. Riguero, V., Clifford, R., Dawley, M., Dickson, M., Gastfriend, B., Thompson, C., Wang, S.C. and O'Connor, E. (2020) Immobilized metal affinity chromatography optimization for poly-histidine tagged proteins. *J Chromatogr A*, **1629**, 461505.
84. Yan, B.X. and Sun, Y.Q. (1997) Glycine residues provide flexibility for enzyme active sites. *J Biol Chem*, **272**, 3190-3194.
85. Bailly, V. and Verly, W.G. (1987) *Escherichia coli* endonuclease III is not an endonuclease but a β-elimination catalyst. *Biochem J*, **242**, 565–572.
86. CA, R. (2021) *Uracil-excising activity of hSMUG1 in bubble and R-loop DNA*, University of Stavanger.
87. Bærheim, T. (2021) *5-Hydroxymethyluracil-excising activity of hSMUG1 in single-stranded DNA*, University of Stavanger.

Appendixes

Appendix A: Experimental protocols

Preparation of chemically competent cells and transformation

1. Inoculate 3 ml LB + appropriate antibiotic with a single colony of E. coli strain harboring the plasmid to be isolated. Grow at 37°C for overnight with vigorous shaking.
2. In the morning inoculate 25 ml LB + appropriate antibiotic with 200 µl of the overnight culture. Grow at 37°C overnight with vigorous shaking until the culture reaches an OD A_{600} of 0.3–0.5 (depending on the strain this takes 2–4 hours).
3. Place culture on ice for 10 minutes.
4. Split culture into round bottom falcon tubes 4 × 6 ml.
5. Centrifuge tubes for 10 min at 4000 rpm a 4°C.
6. Decant supernatant and resuspend in 3 ml ice cold sterile 100 mM $CaCl_2$. Let stand on ice for 30 minutes.
7. Centrifuge tubes for 10 min at 4000 rpm a 4°C.
8. Decant supernatant and resuspend in 400 µl ice cold sterile 100 mM $CaCl_2$. Split into Eppendorf tubes 200 µl cells per Eppendorf tube. Cells are now competent. The cells will remain competent for 24 hours with reducing transformation efficiency over time. It is possible to snap freeze cells and store for longer at –80°C; however, I do not recommend this.
9. Take 200 µl aliquot of cells from previous step and add appropriate quantity of DNA (20 ng for supercoiled plasmids). Incubate on ice for 30 minutes.
10. Place tube in a 42°C waterbath for exactly 45 sec (BL21) and 30 sec (Rosetta).
11. Remove from waterbath and place on ice.
12. Add 500 µl media (no antibiotics) to the tube and place the tube in a 37°C shaker (225 rpm) for an hour.
13. Centrifuge for 5 min at 6000 rpm in room temperature. Pour out the supernatant and resuspend in the drop that is left (app 50 µl).
14. Plate an appropriate volume of the transformation reaction on LB + appropriate antibiotic(s). Grow O/N at 37°C.

Media and buffers for preparation of competent cells and transformation

Antibiotics (for LB-agar plates)

	BL21(DE3)	Rosetta	pETM-11
Corresponding antibiotics	No resistance	Chloramphenicol	Kanamycin
Stock concentration		34 mg/ml	50 mg/ml
1:1000 dilution		34 µg/ml	50 µg/ml

$CaCl_2$

Make 100 mM

***E*xpression test**

1. Make o/n cultures: Take one colony from each plate (BL21 and Rosetta) in sterile LB-broth (3 ml) kanamycin (3 µl) and chloramphenicol (in Rosetta only, 3 µl). Incubate O/N at 37 °C, 225 rpm.
2. Upscale both cultures (1000 µl) in LB-broth (25 ml) with KAN. (25 µl) and CHL. (only in Rosetta, 25 µl). Incubate with shaking, 37 °C, until OD_{600} = 0.4-0.6.
3. The equivalent of 1 ml of cells at OD_{600} = 0.8 (*i.e.* 0.8/OD_{600} of sample = volume in ml) was aliquoted for each strain. Spin down at maximum capacity (13000 xg) for 5 min at 24 °C. Remove supernatant carefully and freeze the pellets at –20 °C. These are the two non-induced samples.
4. Transfer the rest of the sample to 5 round bottom tubes (3 ml in each) for each type of cell (Rosetta and BL21). Add 1 mM IPTG to all tubes (3 µl of 1 M). Mark the tubes with mutant, cell type, temperature, and hours to be incubated. Incubate at corresponding temp at 225 rpm.
5. Use the same calculations as above (0.8/OD) and transfer the corresponding volume to microtubes. Spin down at maximum capacity (11000 xg) for 5 min at 24 °C. Discard the supernatant carefully and freeze the pellet. These are the 8 induced samples.
6. Thaw the pellets.
7. Measure OD of the o/n culture (23 °C) and transfer the corresponding volume (0.8/OD_{600}) to a microtube, spin down at maximum capacity (11000 xg) for 5 min at 24° C and discard the supernatant.
8. Mix all pellets with 1x sample buffer (100 µl), boil for 10 min at 95 °C, cool down to RT, spin down at maximum capacity (11000 xg) for 5 min at 24 °C, and run SDS-PAGE at 220 V, 43 min on all samples (10 µl of supernatant).
9. Visualize the gels using Imaging ChemDoc.

Expression using autoinduction

1. Take one colony and inoculate in 3 ml of MDG (non-inducing minimal medium) with the 50 µg/ml of kan. Incubate o/n with shaking at 37 ºC.
2. Scale-up: 200 µl of o/n culture is inoculated in the 200 ml of ZYM-5052 (kan, complex auto-inducing medium). Incubate 28 ˚C at least 24 h (higher solubility of oxygen at the lower temperatures).
3. Harvest cells 6000 rpm for 20 min. Discard supernatant.
4. Freeze the pellet with liquid nitrogen and store at –80 ˚C.

Media for protein expression using autoinduction

MDG (keep in fridge)

	10 ml	40 ml
H2O (sterile MilliQ)	9.25 ml	37 ml
1M MgSO4	20 ul	80 ul
1000X metals	2 ul	8 ul
40% glucose	125 ul	500 ul
25% aspartate	100 ul	400 ul

20XM	500 ul	2 ml

ZY

	1 liter	500 ml
Tryptone (or peptone)	20 g	10 g
Yeast extract	10 g	5 g
H2O	to 1 liter	to 500 ml

ZYM-5052 (make fresh)

	10ml	100 ml	200 ml	500 ml	1000 ml
ZY (1×)	9.26 ml	92.6 ml	185.2 ml	463 ml	926 ml
1M MgSO4	20 ul	200 ul	400 ul	1 ml	2 ml
1M CitrateNa3	10 ul	100 ul	200 ul	500	1000 µl
1000X metals	10 ul	100 ul	200 ul	500	1000 µl
50X5052	200 ul	2 ml	4 ml	10 ml	20 ml
20XM	500 ul	5 ml	10 ml	25 ml	50 ml

Affinity purification using the Batch method

1. Take the frozen bacterial pellet after harvesting or frozen resuspended bacterial lysates and thaw it gradually at room temperature.
2. If you take the bacterial pallet then resuspend in the lysis buffer in respect to 7 ml per g of bacteria.
3. Add to the lysis buffer the lysozyme to concentration 1 mg/ml (final concentration), add a powder lysozyme, keep it on ice), add one EDTA-free protease inhibitor cocktail tablet incubated at 4 ℃ (shaker at cold room) for 30 min.
4. Supplement the cell lysate with 0.5 % NP-40 (tergitol), 5 mM MgCl2 and 40 µg/mL DNase (stock is 10 mg/ml) and 5 µg/ml RNAase (stock is 10 mg/ml). Incubate for 20 min at 4 ℃.
5. Sonicate the lysate on ice, 30 % amplitude for 10 min, use 5 sec on, 4 sec off, keep the lysate on ice with steering.
6. Remove insoluble debris by centrifugate at 20000 rpm for 45 min at 4 ˚C. Place the supernatant called “crude extract” on ice. Use special centrifuge tubes that are autoclaved. Leave the sample for SDS-PAGE analysis.
7. Precool the buffers, centrifuge with swing rotor in advance.
8. Talon beads are used for this batch purification. Use 2 ml of Talon beads for each lysate. Resuspend the beads and transfer them to the clean 50 ml Falcon tube.
9. Equilibration of beads include removal of storage buffer (20 % Ethanol), washing with water, and then with binding buffer or buffer A. Spin down the bead at 500xg for 5 min and discard the supernatant.
10. Add 10x of beads volume of MilliQ water and remove it by centrifugation 500xg for 5 min.
11. Add 10 ml of buffer A, incubate for 10 min at 4 ℃ and centrifuge 5 min at 500xg. Discard supernatant and repeat twice.
12. Add 10 ml of buffer A with 10 mM Imidazole and incubate for 5 min at 4 ℃. Centrifuge 5 min at 500xg, discard the buffer.

13. Incubate the crude extract from step 6 with the equilibrated bead for 30 min at cold room with gentle shaking. Spin down the crude extract with beads for 5 min at 500xg and transfer the flow through to the new 50 ml Falcon tube for SDS-PAGE analysis.
14. Add 10 ml of buffer A with 10 mM Imidazole, incubate for 5 min at cold room with gently shacking and spin down for 5 min at 500xg. Discard supernatant and repeat step 14 twice.
15. Add 2 ml of buffer A with 100 mM Imidazole, incubate for 10 min at 4 °C at gentle shaking. Spinn down the elution for 5 min at 500xg. Carefully transfer the elution fraction of your protein to the precooled 2 ml microtube. Take care of the sample for SDS-PAGE analysis and dialysis. Repeat this step twice.
16. Prepare 12 % SDS-PAGE gel and load molecular wight standard, crude extract sample, sample from flow through and samples for each elution. Prepare the 2x Laemeli sample buffer by mixing of 50 µl of βME with 950 µl of 2x Laemili buffer. Mix protein samples and 2xLaemili sample buffer in 1:2 ratio. Spin down samples on microcentrifuge and heat them for 5 min at 95 °C. Spin down again.
17. Load 3-5 µl of Protein Standard and 15-20 µl of protein sample. Run SDS-PAGE at 220 V for 40 min. Make sure that the gel runs until the bottom.
18. Remove gel from the glass (do not throw the glass and comb in the trash, they are reusable), and transfer them to the dish for staining. Add distilled water and boil in microwave. Discard water and repeat twice.
19. Add 15 ml of Simple blue and boil again in microwave. Leave for 5-10 min for staining.
20. Pool the fractions of target protein and dialyse it using the Pre-wetted RC tubing MWCO 16 kDa. Add 50 µl of AcTEV protease (Themo Fisher Scientific, 12575015), add fresh βME (final 2 mM). Dialyse O/N at 4 ˚C.
21. Run the second affinity chromatography using HisTrap Talon 1 ml crude column. Remember to collect the flow through (your protein might be there). Use the same Buffer A and Buffer B. Run the SDS-PAGE to analyse fractions and dialyse the pooled fractions o/n.
22. Estimate the protein concentration by Nanodrop using the buffer A as a blank solution.

Affinity purification using a peristaltic pump

1. Add 7 ml of lysis buffer to each g of pellet. Usually, you get 3-4 g pellet for 200 ml of media, use 25 ml of lysis buffer.
2. For 25 ml of lysis buffer, add 100 µl of 25 mg/ml Lysozyme (final 100 µg/ml), 62,5 µl 2mg/ml DNase (final 5µg/ml) and 12,5 µg 10 mg/ml RNase (final 5µg/ml). 178 µl of 70% Tergitol (final 0.5%), 125 µl of 1 M $MgCl_2$ and 1 tb of Complete EDTA-free protease inhibitor cocktail. Incubate for 30 min at RT. Extract should not be viscous at the end.
3. Sonicate the lysate on ice, 30% amplitude and 10 sec on. Alternatively, use 5 sec on, 4 sec off, 15 min total time, amp 30%, keep the lysate on ice with steering.

4. Remove insoluble debris by centuriation at 20000 rpm for 1 h at 4˚C. Place the supernatant called “crude extract” on ice. Use special centrifuge tubes that are autoclaved.
5. Prepare the Äkta Start by washing pumps and fractionation tube with water and then buffer A, wash pump B with buffer B.
6. Wash and Histrap HP 5ml (GE Heatlthcare) column with the water and then with the Buffer A using 3 CV.
7. Load the crude extract into the column using peristaltic pump. Collect the flow through for SDS-PAGE.
8. Set the flow rate 1 ml/min. Connect the column to the Äkta start and wash the buffer A with 3CV until the UV line will be stable.
9. Run the gradient with target concentration 100 % of Buffer B with the length 30 min. Set the fractionation with the size 1 ml.
10. Analyse the fraction from the crude extract, flow through and fractions from the peaks by the SDS-PAGE.
11. Mix 10 µL of sample with 10 µl 2x Laemmli Sample Buffer. Spin down and heat the mix for 5 min at 95 °C. Load 15–20 µl of each sample into the well of SDS-PAGE gel. Use the 1xTris-Glycine running buffer for SDS-PAGE. Run 200V for 30 min. Stain with the Simple blue protein stain or use Chemdoc Tough BioRad for stain-free gels. For simple blue staining, place the gel on the plate, rinse with dH2O and boil in microwave own, discard the water and repeat it two time to wash out the SDS. Add 15 ml of stain and boil it in microwave, leave on the orbital shaker for 1 h. Discard the stain and add water to remove the background. Discard water and add a new one several times until the background is clear. For the stain-free gels, use the protein plot in menu window, chose an activation to make the first picture.
12. Pool the fractions of target protein and dialyse it using the Pre-wetted RC tubing MWCO 16 kDa. Add 5o µl of AcTEV protease (Themo Fisher Scientific, 12575015), add fresh βME (final 2 mM). Dialyse O/N at 4 ˚C.
13. Run the second affinity chromatography using HiTrap Talon 1 ml crude column. Remember to collect the flow through (contains the protein of interests). Use the same Buffer A and Buffer B. Run the SDS-PAGE to analyse fractions and dialyse the pooled fractions o/n.
14. Estimatet the protein concentration by Nanodrop using the buffer A as a blank solution.

Media and buffers for affinity purification using the batch method, a peristaltic pump, and Äkta Start System

Lysis Buffer
50 mM HEPES, pH 8.0/TRIS pH 7.5
300 mM NaCl
5% glycerol
1 mM βME (add fresh)

Dialysis Buffer

50 mM HEPES, pH 8.0/TRIS pH 7.5
300 mM NaCl
2mM βME

Buffers for protein purification

	The Batch method	Äkta Start System
Buffer A (for equilibriating)	50 mM Tris pH 7.5 300 mM NaCl 2 mM βME (add fresh)	
Buffer A (for washing)	50 mM Tris pH 7.5 300 mM NaCl 10 mM Imizadole 2 mM βME (add fresh)	50 mM TRIS pH 7.5 300 mM NaCl
Buffer B (for elution)	50 mM Tris pH 7.5 300 mM NaCl 100 mM Imizadole 2 mM βME (add fresh)	50 mM TRIS pH 7.5 300 mM NaCl 500 mM Imidazle

Store buffer for –20 °C

Prepare a solution containing 50 % glycerol

20xM (1 M = 50 mM PO4, 50 mM NH4Cl, 5 mM Na2SO4)

	100 ml	400 ml	500 ml
NH4Cl	5.35 g	21.4 g	26.75 g
Na2SO4	1.42 g	5.68 g	7.1 g
KH2PO4	6.8 g	27.2 g	34 g
Na2HPO4	7.1 g	28.4 g	35.5 g

1 M MgSO4

24.65 g MgSO4.7 H2O
87 ml H2O

40 % glucose

To make 100 ml
74 ml H2O
40 g glucose

100 % glycerol (w/v)

100 g glycerol
20 ml H2O

50x5052 (1x5052 = 0.5% glycerol, 0.05% glucose, 0.2% a-lactose)

To make 100 ml
25 g glycerol (weigh in beaker)
73 ml H2O
2.5 g glucose
10 g α-lactose monohydrate

25 % aspartate (sodium salt)

84 ml H2O
25 g aspartic acid (1.88 M, MW: 133)
8 g NaOH (2.0 M, MW: 40)

1 M citrate (trisodium salt)

MW = 294.1
29.4 g trisodium citrate dihydrate
H2O to make 100 ml

0.1 M FeCl3 (in ~0.12 M HCl)

MW = 270.30
99 ml sterile H2O
1 ml conc HCl (~12 M)
2.7 g FeCl3-6H2O
Do not autoclave, as a large precipitate form.

1000x trace metal mixture (100 ml)

	Volume	MW	1× conc
Sterile H2O	39 ml		
0.1 M FeCl3 in ~0.12 M HCl	50 ml	270.3	50 uM Fe
1 M CaCl2	2 ml	110.99	50 uM Ca
1 M MnCl2-4H2O	1 ml	197.91	10 uM Mn
1 M ZnSO4-7H2O	1 ml	287.56	10 uM Zn
0.2 M CoCl2-6H2O	1 ml	237.95	2 uM Co
0.1 M CuCl2-2H2O	2 ml	170.486	2 uM Cu
0.1 M Na2MoO4-2H2O	2 ml	241.98	2 uM Mo
0.1 M H3BO3	2 ml	61.83	2 uM B

TB media

12 g Tryptone,
24 g Yeast extract,
4.7 ml of 85 % Glycerol, 0.9 L dH2O (4 ml of 100 %)

After autoclave, make a solution 1 L of phosphate solution for TB media
0.17 M KH2PO4, MW =136.08 g/mol, for 1 l is 23.13 g
0.72 M K2HPO4, MW = 174.18, for 1 l is 125.4
Add 100 ml of phosphate solution to the 900 ml of TB before use.

1 M $MgCl_2$ (100 ml)

1 M x 0.1 l = 0.1 mol, 0.1 mol x 95.22 g/mol = 9.522 g
Dilute to 100 ml.

Lysozyme (25 mg/mL)

Place 0.5 g lysozyme in 20 mL dH2O.
Aliquot and freeze at –20 °C.
To break down the cell wall for lysing bacteria, a final concentration of 0.25 mg/mL lysozyme (1:100 dilution) is added to the pelleted and resuspended cells in lysis buffer.

DNaseI (2 mg/mL)

Place 20 mg DNase I in 10 mL 20% glycerol, 75 mM NaCl.
Aliquot and freeze at –20 °C.
Add to a final concentration of 10 µg/mL in lysis buffer, with 5 mM MgCl2.

RNase stock solution

5 M NaCl (add 30 µl, final 15 mM)
1 M Tris-HCl, pH 8.0 (add 100 µl, final 10 mM)
H_2O to 10 ml
Dissolve 100 mg of pancreatic RNase A in the 10 ml of solution. Dispense 1-ml aliquots into 1.5-ml microcentrifuge tubes, and store them at −20°C. Just before use, boil the required amount for 10 minutes.

Affinity purification of TEV protease (including buffers)

The TEVsh protease is a construct of the tobacco etch virus protease produced by an in vitro directed evolution method (see van den Berg et al., Improved solubility of TEV protease by directed evolution, 2006). The TEV protease is a cysteine protease, so a reducing agent should be present at all times. It is easy and cheap to use 2mM beta-Mercaptoethanol (ß-ME; 142 ul per 1 liter is 2 mM). The protease is sensitive to an abrupt change in salt concentration. Changing from 500 mM NaCl to 200 mM NaCl, or from 200 mM NaCl to 100 mM NaCl, is fine. Changing from 500 mM directly to 100 mM NaCl will cause precipitation.

Buffers:

1. Break the cells (do not add protease inhibitors). Lysis buffer: (as needed) Same as IMAC buffer A.
2. IMAC capture. Buffer A (1.0 l): 50 mM Tris-Cl, pH 7.5, 500 mM NaCl, 10 % glycerol, 2 mM ß-ME. Buffer B (0.5 l): same plus 500 mM Imidazole (pH 7.5 final). Add Buffer B to the sample to be loaded on the column in order to make the sample 5 % buffer B. One 5 ml IMAC column is sufficient.
3. Dialysis (change from 500 mM NaCl to 200 mM) (1.0 l): 50 mM HEPES, 200 mM NaCl, 10 % glycerol, 2 mM ß-ME, pH 7.5.
4. Dilution of protein sample to less than 100 mM NaCl: Measure the volume of the dialyzed protein and add the same volume of cold Cation exchange buffer A. And, then you will need to centrifuge the sample to remove the precipitated protein (this step was skipped).
5. Cation exchange is nessessary (1 l of A and 0.5 l of B): Buffer A (1.0 l): 50 mM HEPES or Tris-Cl, 10 % glycerol, 2 mM β-ME, pH 7.0. Buffer B (0.5 l): 50 mM HEPES or Tris-Cl, 10 % glycerol, 1 M NaCl, 2 mM ß-ME, pH 7.0 (this step was skipped).

Storage buffer (50 ml or so): 50 mM HEPES, ~pH 7.0, 200 mM NaCl, 2 mM ß-ME, 1 mM EDTA, >80 % glycerol (used to dilute purified TEV protease to approx. 40 to 50% glycerol for storage at -80 degrees).

Day 1 – Cell lysis, His-tag capture, and dialysis or desalting:

1. Break the cells by sonication and centrifuge to remove cell debris. Centrifuge twice if necessary if still cloudy after 1 st. centrifugation (very important).
2. Load the supernatant onto a 5 ml IMAC column equilibrated in the lysis buffer. Then, wash the column with 10 volumes or more of lysis buffer. And finally, elute the TEV protease using a 15 column volume gradient from A to B.
3. Run a SDS-PAGE gel to see the purity (or can simply combine the large 2nd elution peak especially if doing the 2nd cation exchange purification step).
4. Combine the fractions containing TEV protease and dialyze against 50 mM HEPES, 200 mM NaCl, 10 % glycerol, 2 mM b-ME, pH 7.0. Or run a desalting column with the same buffer.

Day 2 (or a long Day 1) cation exchange plus storage:

1. Dilute the dialyzed (or desalted) protein using cold ion exchange buffer A to 3 time the original volume in two steps: First, measure the volume of dialyzed protein and then add the same volume of cold buffer A with mixing and wait and few minutes. Then add the same volume of cold Buffer A again with mixing and wait a few minutes. The concentration of NaCl should now be below 100 mM. Centrifuge the cloudy protein solution (20K for 20 minutes or as needed) and also filter if necessary.
2. Equilibrate the cation exchange column with 10 % B (100 mM NaCl). Load the protein sample onto the cation exchange column. Then wash the column with a few column volumes of 10 % buffer B. Finally, elute with a 10 column volume gradient from 10 % buffer B to 60 % buffer B. The TEV protease elutes early in the gradient. You may have to do several runs if you use a small column - no more than perhaps 15 mg or so on a 1ml column. A 10ml ResourceS column will bind >300 mg TEVsh and perhaps more.
3. Combine the fractions containing TEV and measure the concentration (1mg/ml = 1.1 absorption (1 cm path length, 280 nm). Dilute the purified TEV with the storage buffer

keeping the glycerol concentration >=40 %. Make 0.5 ml or smaller aliquots of the protein into 1.5ml Eppendorf tubes on ice.

4. An easy way to freeze the protein is to put a freezer box inside a larger styrofoam container and pour liquid nitrogen into the smaller freezer box. Then place the 1.5ml eppendorf tubes containing the aliquotted TEVsh directly into the freezer box. When the box is full, pour off the liquid nitrogen and place the box in the -80 degree freezer. Be careful working with liquid N2!

Buffers for SDS-PAGE (TRIS-glycine running System)

10× TRIS-Glysine running buffer for SDS-PAGE

Dissolve 30.0 g of Tris Base, 144.0 g of glycine, and 10.0 g of SDS in 1000 ml of H2O.
The pH of the buffer should be 8.3 and no pH adjustment is
required. Store the running
buffer at room temperature and dilute to 1× before use.

1× Tris-Glycine running buffer for SDS-PAGE

25 mM Tris Base (3.02 g)
0.192 M Glycine (14.4 g)
1 % SDS (1 g) in dH2O

Loading buffer

2x Laemmli buffer
2 mM β-ME

Expression using IPTG induction

1. Make o/n culture. Take one colony (Rosetta/BL21) in LB-broth (25 ml) kanamycin (25 µl) and chloramphenicol (Rosetta only, 25 µl), and make a control. Incubate o/n at 37 °C, 220 rpm.
2. Put the o/n culture on ice. Scaleup with LB-broth (1000 ml), kanamycin (1025 µl), chloramphenicol (Rosetta only, 1025 µl) and all the o/n culture (25 ml). Incubate at 220 rpm at 37 °C until OD_{600}≈0.6 in all samples. App. 1.5 h.
3. Aliquot the equivalent of 1 ml of cells at OD_{600} = 0.8 (*i.e.* 0.8/OD_{600} of sample = volume in ml) as the non-induced sample. Spin down at 11000 xg, for 5 min in RT. Discard the supernatant and freeze the non-induced pellet.
4. Induce the rest of the bacteria (1025 ml minus used volume) using 1 mM IPTG (1025 µl minus used volume). Incubate at 220 rpm at 37 °C for 2 h.
5. Harvest by spinning down in special centrifuge bottles 6000 rpm for 20 min at 24 °C. Discard the supernatant and freeze the pellet at -20 °C for later purification.

Media and buffers for protein expression using IPTG induction

IPTG

Make 1 M

Affinity purification using gravity column

1. Thaw the pellets (induced and non-induced) and add fresh 2 mM βME to the buffers.
2. Resuspend the induced pellets in lysis buffer (10 ml/l media used), then added lysozyme (1 mg/ml final gives 1 ml/10 ml lysis buffer). Incubate with gentle shaking at 4 °C for 30 min.
3. Equilibrated beads (1 ml, resuspended) using water (4 ml) and wash 2 buffer (4 ml), and pre-cool the centrifuge.
4. Add DNAse 1 (40 µg/ml final gives 40 µl/10 ml lysis buffer) and incubate with gentle shaking at 4 °C for 15-30 min.
5. Sonicate at 10 % altitude for 1 min (10 sec on/10 sec off 6 times).
6. Spin down in two small centrifuge tubes at 15000 rpm, 4 °C for 30 min.
7. Transfer the supernatant (crude extract) to a new tube. Check for clarity.
8. Add the crude extract (save 20 µl for SDS-PAGE) to the beads, resuspend and collect the flow through.
9. Wash with buffers: wash 2 (5 ml), wash 3 (5 ml) and wash 4 (5 ml) respectively (use wash 4 and elution buffer without PMSF).
10. Elute using elution buffer (2 ml) in 2 different fractions.
11. Regenerate the beads using 5x beads volume of MES and dH_2O. Store in 20 % ethanol (1 beads volume) at 4 °C to be reused.
12. Mix 2x Laemmli sample buffer and the samples in a 1:1 ratio. Resuspend the non-induced pellet in 100 µl 1x Laemmli sample buffer. Spin the samples, boil for 5 min at 95 °C, and again spin the samples.
13. Run SDS-PAGE by loading 20 µl of each sample (10 µl of the non-induced) in the following order: Ladder, non-induced, crude extract, flow-through, elution 1-2. Run with 300 V for 25 min. Visualize the gel.
14. Measure nanodrop in 3 parallels (blank=elution buffer).
15. Save 20 µl of elution 1 for SDS-PAGE before His-tag-cleavage by TEV. Mix sample in a 1:1 ratio with 2x Laemmli buffer with βME. Store in 4 °C.
16. Dialyze elution 1 and 2 (in two different membranes) with TEV (0.5 ml in each) using 2 l dialysis buffer (elution buffer minus imidazole) with βME (2 mM, 281 µl).

Performed gravity purification again, to remove the His-tag from the solution. SMUG should now be in flow-through, and the His-tag in the elutions, because the beads only bind His-tags.

17. Transfer the protein (elutions) from the dialysis pool to microtubes.
18. Equilibrate beads using water (5 ml) and wash 2 buffer (4 ml).
19. Make gels for SDS-PAGE according to the BioRad kit.
20. Add the dialyzed protein to the beads, incubate for 30 min in cold-room and collect the flow through.
21. Wash the beads with wash 2 (5 ml) and wash 4 (5 ml).

22. Elute the His-tag (after TEV cleavage) using elution buffer (4 ml) in 4 different fractions.
23. Regenerate the beads using 5x beads volume of MES and dH_2O. Store in 20 % ethanol (1x beads volume) at 4 °C to be reused.
24. Freeze the protein (present in flow-through after TEV-cleavage) in 50 % glycerol at –80 °C.
25. Measure Nanodrop (3 parallels, blank=dialysis buffer).
26. Mix all samples (protein before cleavage, protein after cleavage from the flow-through, and the His-tag protein from elution 1) with 2x Laemmili sample buffer in a 1:2 ratio. Spin down and heat the mix for 5 min at 95 °C. Load 15–20 µl of each sample into the well of SDS-PAGE gel together with a protein standard (3 µl). Use the 1xTris-Glycine running buffer for SDS-PAGE. Run at 200 V for 30 min.
27. Run SDS-PAGE by loading a protein standard, the protein sample before TEV-cleavage (SMUG1 w/His-tag), the protein from flow-through (SMUG1 w/o His-tag), and elution 1 after TEV-cleavage (containing only the His-tag protein).

Media and buffers for affinity purification using gravity column

Lysis buffer/wash 1

20 mM Tris pH 8.0 (used pH 7.5 last time) (1 ml of 1 M)
10 mM Imidazole (0.5 ml of 1 M)
150 mM NaCl (2.5 ml of 3 M)
0.2 % NP-40 (143 µl of 70 %)
1 µl PMSF
2 mM βME (add fresh, 7 µl/50 ml of 14.26 M)

Add Lysozyme 1 mg/ml after pellet is resuspended (200 µl).

Wash 2

Lysis buffer minus NP-40.

Wash 3

Wash 2 with 1 M NaCl (16.67 ml of 3 M)

Wash 4

Wash 2 with 20-50 mM imidazole (1 ml of 1 M), without PMSF.

Elution buffer

Wash 2 with 330 mM imidazole (16.5 ml of 1 M), without PMSF.

Western blotting

1. After SDS-PAGE, run Western blot (using Trans-Blot Turbo transfer pack) with 3 min transfer time. Block with 5 % BSA in PBST (0.1 % Twin 20) for 1 h at RT.
2. Cut the membrane into two parts (in the blocking solution) using sterile scissors and transfer one of the membranes to another container.

3. Discard the blocking solution and add the primary antibody solutions (primary rabbit anti-SMUG Ab 1:2000 in 5 % BSA for one membrane and anti-His Ab 1:2000 in 5 % BSA for the other membrane, 10 ml each), incubate at 4 ˚C o/n with gentle shaking.
4. Remove primary Ab (to be reused) and wash with PBST (0.1 % Tween 20) for 5 min, 3 times with shaking. Discard the washing solution after each use.
5. Add the secondary antibody solution: Added Precision Protein™ Strep Tactin-HRP Conjugate (1 µl) as an antibody for the ladder and secondary Goat anti rabbit-HRP IgG 1:2000 (5 µl) into 5 % BSA (10 ml). Incubate with shaking in RT for 1.5 h.
6. Remove secondary Ab (to be reused) and wash with PBST (0.1 % Tween 20) for 5 min, 3 times with shaking.
7. Prepare 2 ml of substrate solution (BioRad Clarity Western ECL Substrate) in a 1:1 ratio. Dry membrane by dipping edges on paper using a tweezer. Place on plastic with paper under (ladder to the left), resuspend solution, pipetted the solution on the whole membrane (in lines), and incubate open for 5 min. Check that the whole membrane is covered at all times. Close plastic and press carefully with paper to remove the excess substrate.
8. Image the membrane using a digital imager (BioRad ChemDoc).

Media and buffers for Western blotting

10 x PBS
Use PBS tablets to prepare 10x solution (recipe on the package), use MilliQ water, autoclave before use.

PBST (0.1 % Tween 20)
100 ml of 10xPBS 1 ml Tween 20 (detergent) 900 ml MilliQ H2O

5 % PBST/BSA
5 g milk/BSA 100 ml of PBST (0.1 % Tween 20)

hSMUG1 activity assay for excision of DNA-uracil

Notes:

- Put 96 % ethanol w/0.1 M NaAc in the freezer.
- Mark all tubes, remember control.
- Calculate the MW of the mutant using uniport (substitute one letter in the sequence). Then calculate the concentration from mg/ml to pmol/µl:
  (g/ml)/(g/mol) = mol/ml → (mol/ml)*10^{12}= pmol/ml → (pmol/ml)/1000= pmol/µl
- Make fresh 1 M DTT: Weight a small amount in a microtube (≈0.05). Calculate the mol and then the volume (µl) of MilliQ.
- Take 199 µl of 5x HEPES/Tris and add 1 µl of 1 M DTT (fresh).
- Make 1x HEPES: 100 µl 5x HEPES + 150 µl MilliQ + 250 µl 100 % glycerol.

- Dilute the enzyme from calculated conc. to 1 pmol/µl with 100 µl as final volume using 1x HEPES buffer (with DTT). Dilute in 3 serial dilutions (0.1, 0.01, 0.001 pmol/µl) using 90 µl 1x HEPES buffer (with DDT) and transferring 10 µl from tube to tube.
- ICE AND DARKNESS DURING THE ASSAY!

1. Mix reaction mixture:

Note for all sample tubes:

- First, add HEPES buffer, then add KCl, then add BSA, then add dH2O. Add the drops on 4 different sides of the tubes!
- Spin down the mixture (mini centrifuge).
- Get the substrate (Cy3-UG-subs3) **(remember darkness!)** and add 1 µl into the mixture.

2. Mix the reaction with a pipette and add the corresponding conc. of the enzyme at last (1x HEPES in N.E.).

	Stock	**Reaction mix**	**Reaction 1x (µl)**	**Reaction 2x (µl)**
HEPES buffer + DTT	**5×**	1x	4	8
KCl	**1 M**	70 mM	1.4	2.8
BSA	**10 mg/ml***	5x (0.5 mg/ml)	1	2
H_2O (MilliQ)			11.6	23.2
Labelled oligo	**1 pmol/ul**	1 pmol	1	2
Enzyme (protein)	**Varying**		1	2
Total volume			20	40

*BSA: 10 mg/ml = 10 ug/ul, dvs. 10 ug er tilsatt reaksjonen, dvs. 10 ug/20 ul = 0.5 ug/ul = 0.5 mg/ml

3. Spin down 4000 rpm for 1 min at RT.
4. Incubate at 37°C for 20 min.
5. Spin down, terminate the reaction with 45 µl of stop solution and 1 µl of Proteinase K (10 mg/ml), mix up and down by the pipet, incubate at 37 °C for 10 min.
6. Spin down and add 150 ml ice-cold 96 % ethanol w/0.1 M NaAc to each tube and then 1.6 µl tRNA (16 mg from Stock: 10 mg/ml). Invert the tubes several times.
7. Incubate the tubes at -70 for 2 h, or o/n at -20 in darkness.
8. Pre-cool centrifuge and turn on the heater.
9. Make fresh gel (and new 10 % APS if old).
10. Centrifuge the tubes at 13000 rpm at 4 °C. directly from the freezer, 15 min two times with different positions of the tubes. Discard the supernatant.
11. Add 300 ml cold 70 % ethanol (freezer). NO resuspending!
12. Centrifuge at 13000 rpm, 5 min at 4 °C. Discard the supernatant by tapping the tube on paper.

13. Centrifuge at 13000 rpm, 1 min at 4 °C. Discard the rest of the supernatant with a pipette.
14. Dry the pellet **in darkness** for 1-20 min on ice in the hood to completely evaporate the ethanol.
15. Dissolve the pellet in 10 µl 0.1 M NaOH and heat the samples for 10 min at 90 °C.
16. Fill the gel chamber with running buffer and carefully remove the comb. **Wash the gel wells carefully** using a pipette to remove the residual gel particles. This part is very critical in making the DNA bands sharper and straight.
17. Mix with 10 µl of denaturing loading buffer (DLB) and mix up and down with a pipette.
18. Load 5 µl of each sample into the wells and run the gel at 200 V (TBE-system) 120 V (Taurine-system) for 1.5 h **in darkness**.

Hybridization of Cy3-U:G substrate

ICE AND DARKNESS DURING THE ASSAY!

1. Mix in PCR-tube:
    a. 1 µl Cy3U (100 pmol/µl)
    b. 1 µl complementary strand (Gcomp, 100 pmol/µl)
    c. 8 µl 1x STE buffer
2. Spin down.
3. Incubate at 95 °C, 4 min (ug oligo) in the thermocycler.
4. Leave the tube in the thermocycler for appr. 2 hours to cool down.
5. Dilute with 90 µl 1x TE buffer to make 1 pmol/µl. Store at – 4 °C, in the dark.
6. Mark as Cy3-UG-subs1 (60), 2 (26) or 3 (60mod).

The ssDNA substrates were prepared by diluting 100 pmol/µl ssDNA stock (1 µl) in 1× TE buffer (99 µl) to make 1 pmol/µl substrate. The prepared substrates were stored at –4 °C in darkness.

Media and buffers for activity assay and hybridization of Cy3-U:G substrate

5x HEPES buffer	**Composition**	**Stock**	**Mix preparation**
Store in aliqouts (-20 °C)	225 mM HEPES, pH 7.5	238.3 g/mol	29.2 g
5 mM DTT (add fresh)	10 % glycerol	85 % (Merck, Cat #: 1.04094)	12 ml
	2 mM EDTA	0.5 M (lab stock)	400 ul
	Deionized H_2O		to 100 ml

1x buffer in glycerol	**Composition**	**Stock**	**Mix preparation**
5 mM DTT (add fresh)	5x HEPES	See as above	200 µl
	50 % glycerol	85 % (Merck, Cat #: 1.04094)	588 µl

	Deionized H_2O		212 µl
96 % ethanol w/ 0.1 M NaAc	**Composition**	**Stock**	**Mix preparation**
Store at room temperature	0.1 M NaAc	(Sigma, Cat #: S2889, 82.03 g/mol)	0.82 g
Store at -20 °C 1 h before use	96 % EtOH	96 %	225 ul
1M KCL	**Composition**	**Stock**	**Mix preparation**
Store at aliqouts (-20 °C)	1 M KCl	(Sigma, Cat #: 5405, 74.55 g/mol)	3,72 g
	Deionized H_2O		50 ml
Stop solution	**Composition**	**Stock**	**Mix preparation**
Store at room temperature	20 mM EDTA	0.5 M	2 ml
	0.5 % (w/v) SDS	99 % (Sigma, Cat #: L3771, 288.38 g/mol)	252 mg
	Deionized H_2O		To 50 ml
STE buffer (for oligo hybridization)	**Composition**	**Stock**	**Mix preparation**
Store in aliqouts (-20 °C)	10 mM Tris, pH 8.0	1 M	200 ul
	50 mM NaCl	(Sigma, Cat #: S5886, 58.44 g/mol)	58.44 mg
	1 mM EDTA	0.5 M (lab stock)	40 ul
	Deionized H_2O		To 20 ml
1x TE buffer	**Composition**	**Stock**	**Mix preparation**
Store in aliqouts (-20 °C)	10 mM Tris, pH 7.5	1 M	200 ul
Sterilfiltrated	1 mM EDTA, pH 8.0	0.5 M	1.6 ml
	Deionized H_2O		To 20 ml
20x taurine, running buffer	**Composition**	**Stock**	**Mix preparation**
Store in room temperature	1.78 M Tris base	(Sigma, Cat #: T6066, 121.14 g/mol)	216 g
	0.58 M Taurine	(Sigma, Cat #: T0625, 125.15 g/mol)	72 g
	$Na_2EDTAx2H_2O$		4 g
	Deionized H_2O		To 1000 ml
Denaturing 20 % PAGE gel	**Composition**	**Stock**	**Mix preparation**
	20 % acrylamide	40 % (Saveen Werner AB, Cat #: BIAC21)	7,5 ml
	1x Taurine	(Sigma, Cat #: T0625, 125.15 g/mol)	750 µl (from 20x prepared stock)
	3 % formamide	99.5 % (Sigma, Cat #: F9037)	450 µl
	Deionized H_2O		6,3 ml
	APS	(BioRad, Cat #: 161-0700, 228.2 g/mol)	150 ul (from 10% prepared stock)
	Temed	(Invitrogen Cat #: 15524-010)	15 ul

Denaturing Loading buffer	Composition	Stock	Mix preparation
Store in aliqouts (-20 °C)	80 % formamide	99.5 % (Sigma, Cat #: F9037)	40.20 ml
	1 mM EDTA	0.5 M	100 ul
	1 % (w/v) blue dextran	(Sigma, Cat #: D5751)	0.5 g
	Deionized H_2O		To 50 ml

Appendix B: Experiments

Production and purification of hSMUG1 P240G

Seven Lot's of hSMUG1 P240G were purified (Table 2), in addition, an expression test (Figure 15 and Table A 1) was performed to determine the optimal incubation time and check for possible differences in expression yield when the protein was produced in the strains *E. coli* BL21(DE3) and its derivative BL21(DE3) Rosetta respectively.

Table A 1. Expression test of hSMUG1 P240G. Temperature and time incubated after induction with 1 mM IPTG, the corresponding OD measured and the calculated volume used for the two different strains. Abbreviations: h, hour; RT, room temperature; OD, optical density.

Incubation time (h)	Temperature (°C)	*E. coli* strain	OD	Volume (µl)
0	–	BL21 Rosetta	0.714	1120
0	–	BL21	0.846	950*
1	37	BL21 Rosetta	1.351	592
1	37	BL21	1.288	621
2	37	BL21 Rosetta	1.626	492
2	37	BL21	1.603	499
3	37	BL21 Rosetta	1.715	467
3	37	BL21	1.900	421
4	37	BL21 Rosetta	1.783	449
4	37	BL21	2.072	386
O/N	RT	BL21 Rosetta	3.266	245
O/N	RT	BL21	3.157	253

*Should have taken 946 µl

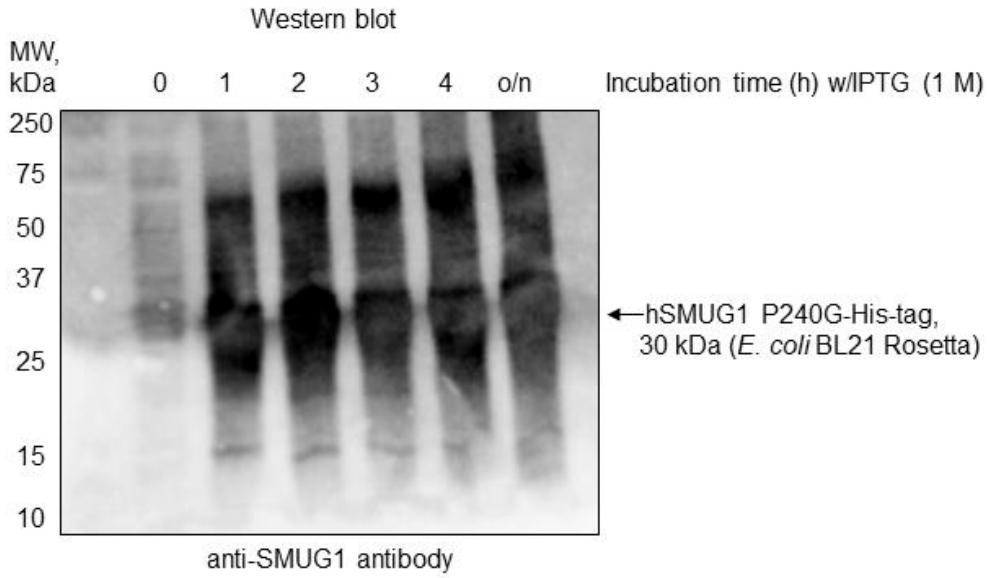


Figure A 1. Western blot from corresponding SDS-PAGE gel (Figure 15C) using monoclonal anti-SMUG1 antibody for identifying hSMUG1 P240G after induction by IPTG 0–4 h and o/n in E. coli BL21(DE3) Rosetta strain.

Lot No. 1

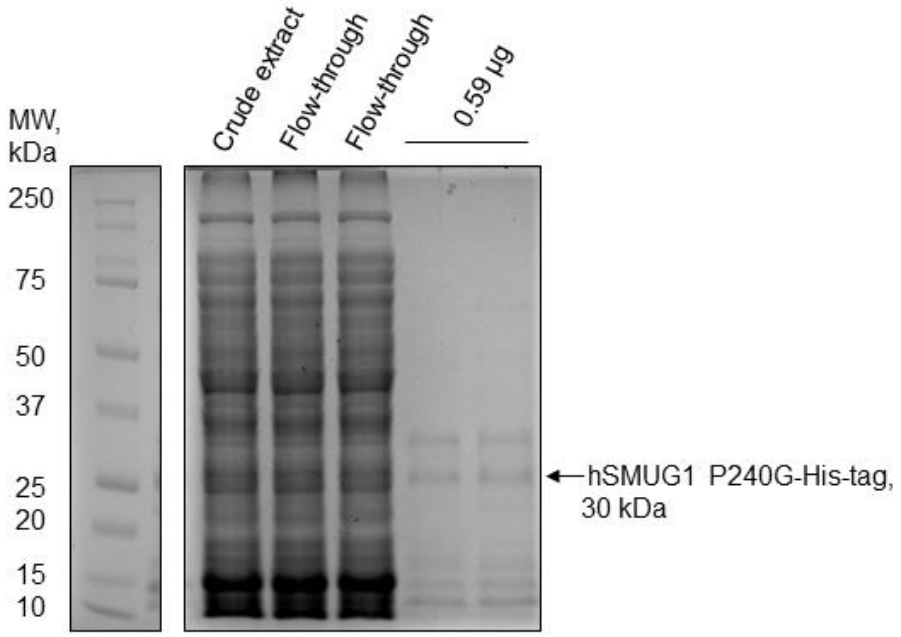


Figure A 2. SDS-PAGE of hSMUG1 P240G-His-tag produced in E. coli BL21(DE3) strain and purified by the Batch method.

Lot No. 2

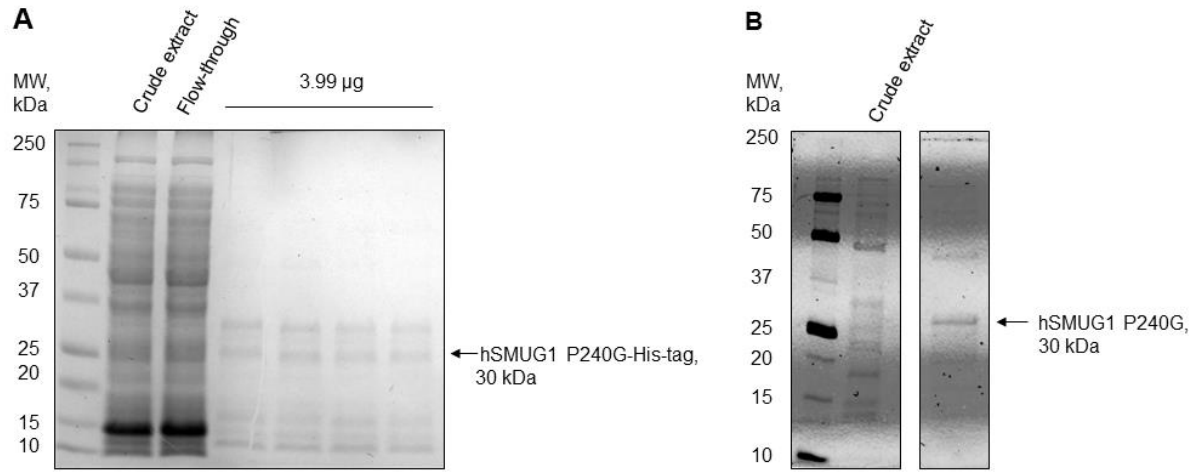


Figure A 3. SDS-PAGE of hSMUG1 P240G produced in E. coli BL21(DE3) strain and purified by the Batch method (A) and in addition by second affinity purification by Äkta Start System (B) in order to separate hSMUG1 P240G from His-tagged proteins and His-tag itself.

Lot No. 3

This particular enzyme revealed the highest UDG activity and was further used in all P240G experiments. The protein was produced in *E. coli* BL21(DE3) strain, and the steps that were followed are listed in Table A 2. See Result section 4.1.1 (Figure 16A and B) for SDS-PAGE gels.

Table A 2. Steps of protein production and purification of hSMUG1 P240G (Lot No. 3), showing the achieved volume/amount, corresponding buffer, and the possible activity detected. Nd, Not detected; O/n, overnight.

Step	**Production and purification step**	**Volume (ml)**	**Medium/buffer**	**Amount (g)**	**Detected**
1	O/n culture	3	MDG[2]	Nd	Visible turbidity
2	Autoinduction	1000	Zym-5052[2]	Nd	OD_{600} ~0.4–0.6
3	Harvesting	Nd	None	12.396	Nd
4	Cell-free extract	50	Lysis buffer[2]	Nd	Nd
5	Batch purification	2	Elution buffer w/imidazole[2]	Nd	SDS
6	Dialysis	~2.5	Elution buffer w/o imidazole[2]	Nd	~0.5 mg/ml[1]

7	Äkta purification	2	Elution buffer B[2]	Nd	Chromatogram ~0.1 mg/ml[1]
8	Ultrafiltration	0.2	Elution buffer B[2]	Nd	~0.2 mg/ml[1]

[1]Estimated by NanoDrop Spectrophotometer (Thermo Scientific)
[2]See Appendix A for details

Lot No. 4

See Results 4.1.1 (Figure 18) for SDS-PAGE of hSMUG1 P240G-His-tag produced in *E. coli* BL21(DE3) Rosetta. The protein was gravity purified and further analyzed by Western blot in 1:2 serial dilutions.

Lot No. 5

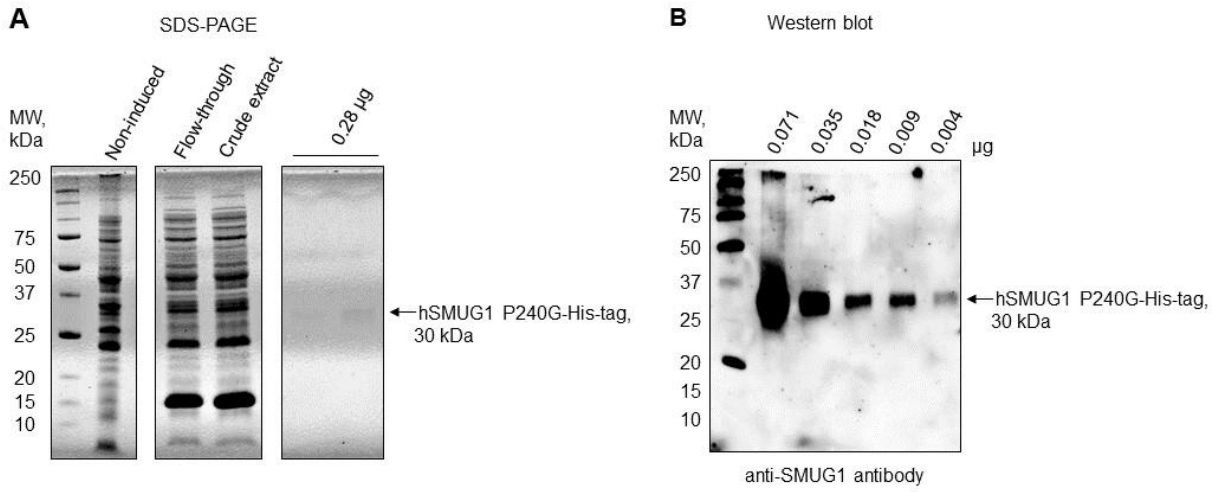


Figure A 4. ***(A)*** *SDS-PAGE of hSMUG1 P240G-His-tag produced in E. coli BL21(DE3) Rosetta and gravity purified.* ***(B)*** *The gravity purified hSMUG1 P240G-Histag was analyzed by Western blot in 1:2 serial dilutions.*

Lot No. 6

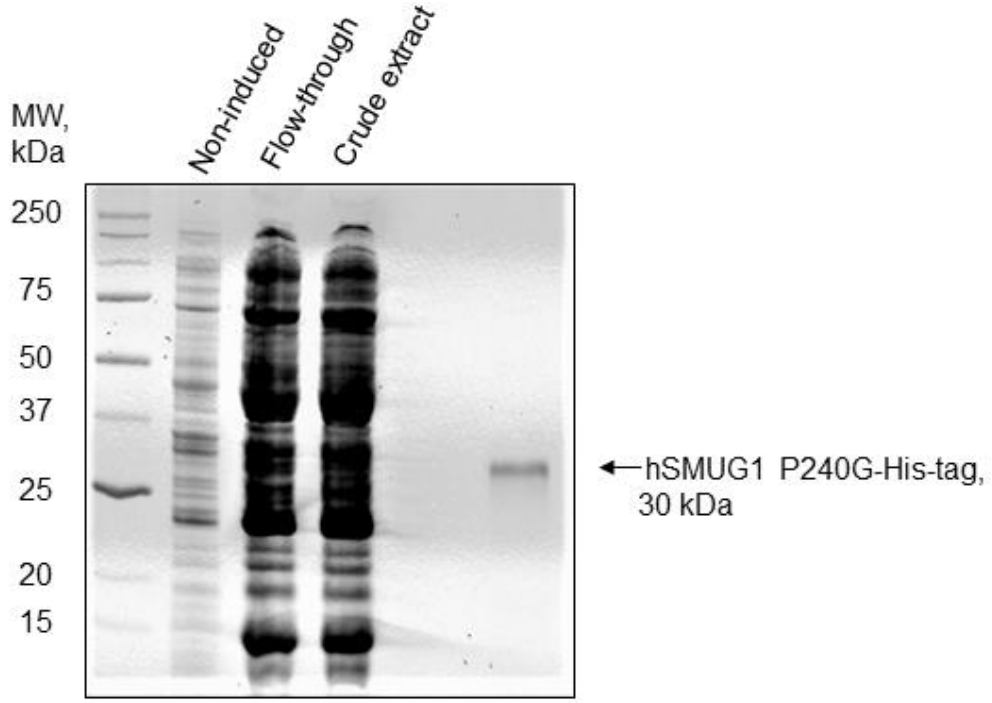


Figure A 5. SDS-PAGE of hSMUG1 P240G-His-tag produced in E. coli BL21(DE3) Rosetta and gravity purified.

Lot No. 7

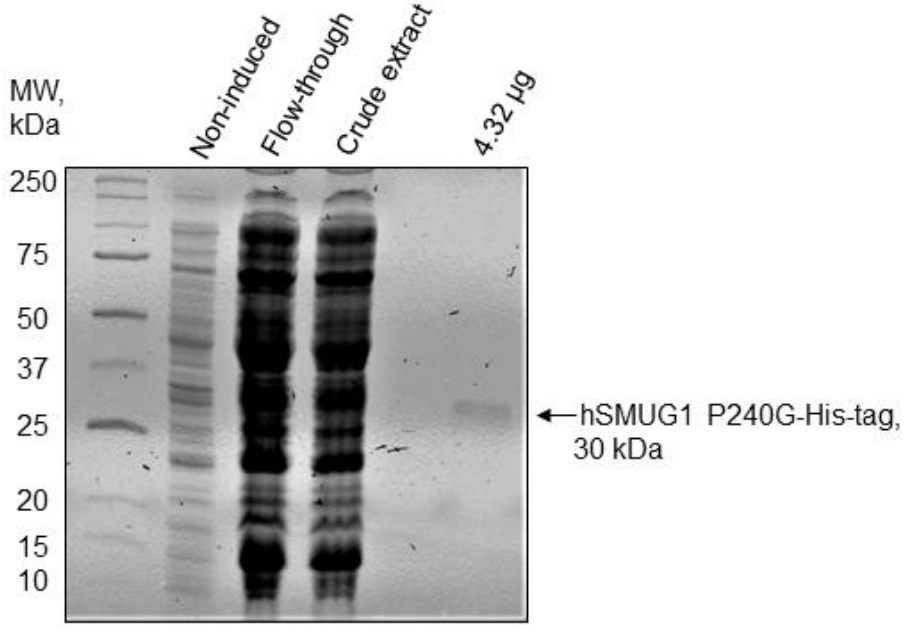


Figure A 6. SDS-PAGE of hSMUG1 P240G-His-tag produced in E. coli BL21(DE3) and gravity purified.

Production and purification of hSMUG1 S241A

One Lot of hSMUG1 S241A was purified and an expression test (Figure 20A and Table A 3) was performed to determine the optimal incubation time and check for possible differences in expression yield when the protein was produced in the strains *E. coli* BL21 and BL21(DE3) Rosetta respectively.

Table A 3. Expression test of hSMUG1 S241A. Temperature and time incubated after induction with 1 mM IPTG, the corresponding OD measured and the calculated volume used for the two different strains. Abbreviations: h, hour; RT, room temperature; OD, optical density.

Incubation time (h)	**Temperature (°C)**	***E. coli* strain**	**OD**	**Volume (µl)**
0	–	BL21 Rosetta	0.465	1720
0	–	BL21	0.677	1182
1	37	BL21 Rosetta	0.964	830
1	37	BL21	1.350	593
2	37	BL21 Rosetta	1.360	588
2	37	BL21	1.776	450
3	37	BL21 Rosetta	1.489	537
3	37	BL21	1.774	451
4	37	BL21 Rosetta	1.567	511
4	37	BL21	1.819	440
O/N	RT	BL21 Rosetta	2.483	322
O/N	RT	BL21	2.605	307

Lot No. 8

See Results 4.1.2 (Figure 20B) for SDS-PAGE of hSMUG1 S241A-His-tag produced in *E. coli* BL21(DE3 3) strain and gravity purified.

Purification of TEV protease

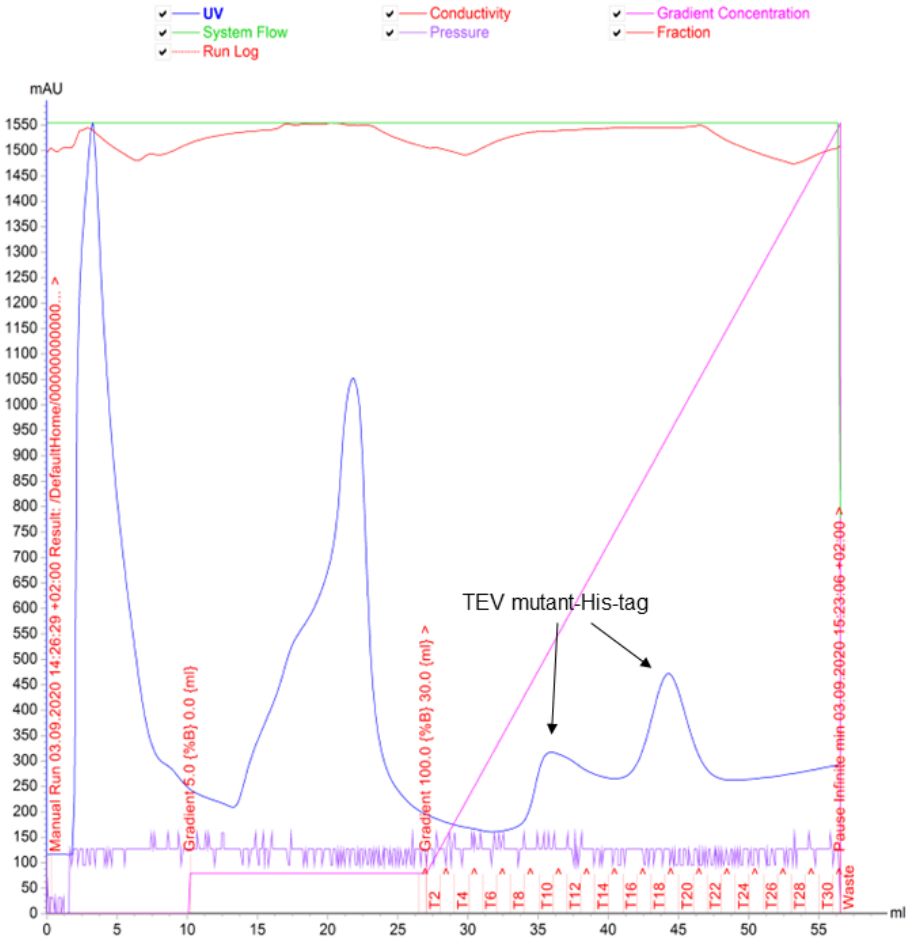


Figure A 7. TEV mutant-His-tag was purified using Äkta Start Protein Purification System followed by IPTG induction (see Materials and Methods 3.3.5 for IPTG induction and purification protocol in Appendix A, Affinity purification of TEV protease). A peristaltic pump was used to load the crude extract into a HisTrapHP column (5 ml). Samples 10, 11, 17, 18, 19, and 20 were further used.

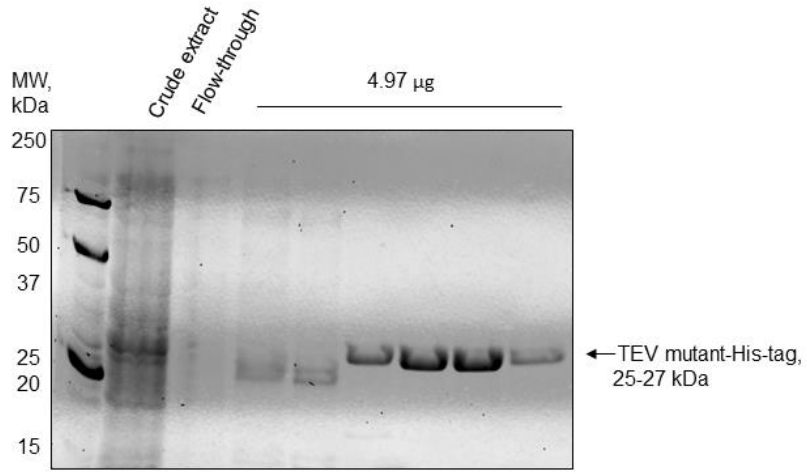
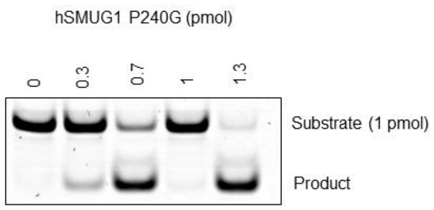
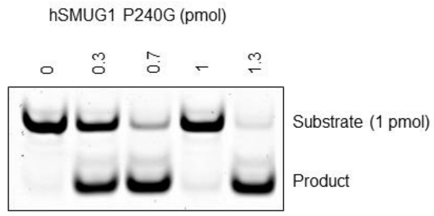


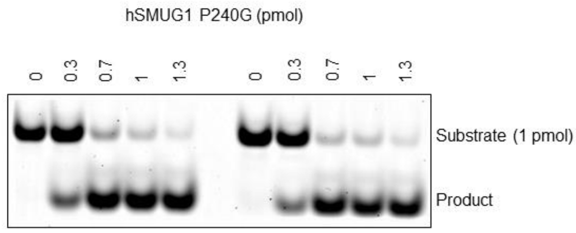
Figure A 8. SDS-PAGE analysis of purified mutated TEV protein. From the left: Molecular weight standard; crude extract; flow-through; sample 10; sample 11; sample 17; sample 18; sample 19; sample 20. The protein was further desalted by dialysis (all samples pooled together, 0.662 mg/ml total), up-concentrated (to 0.817 mg/ml), stored at –80 ºC in 50 % storage buffer (0.4085 mg/ml final TEV concentration) containing 50 mM HEPES ~pH 7.0, 200 mM NaCl, 2 mM β-mercaptoethanol, 1mM EDTA, and >80 % glycerol, and used to cleave the His-tag from hSMUG1 P240G.

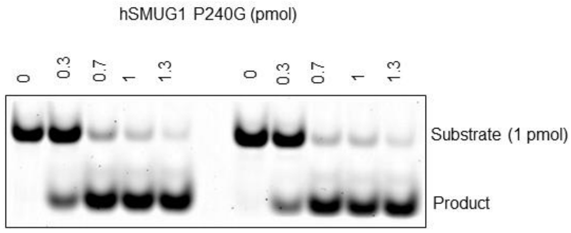
Activity assays for excision of uracil-DNA by hSMUG1 P240G

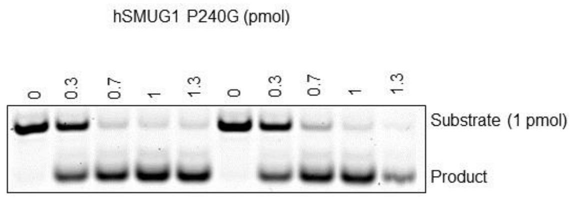
If not mentioned, all activity assays are performed on ssU-DNA and with NaOH.

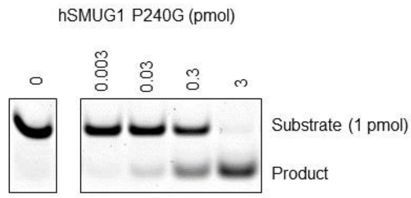


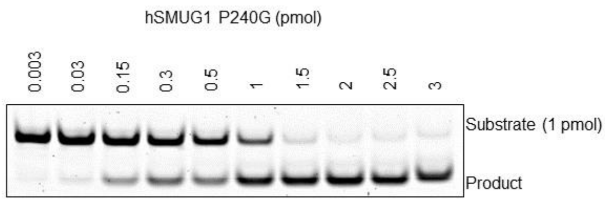


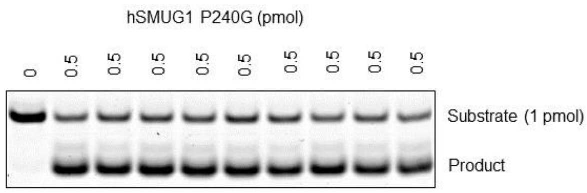


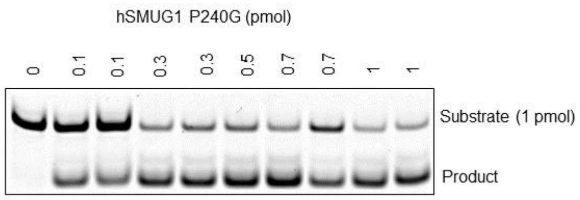


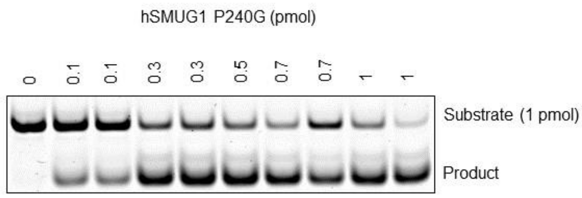


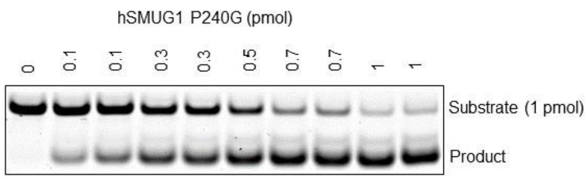


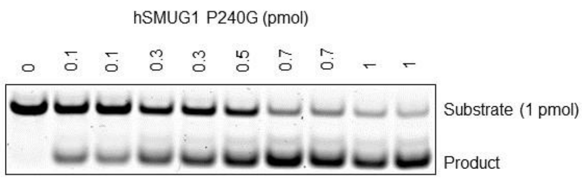


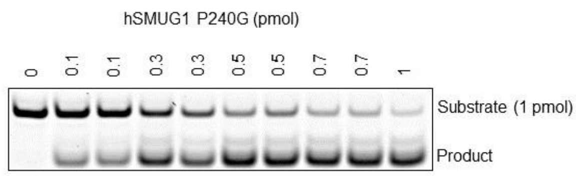


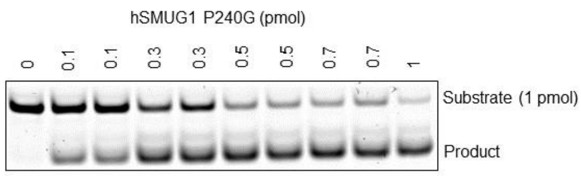


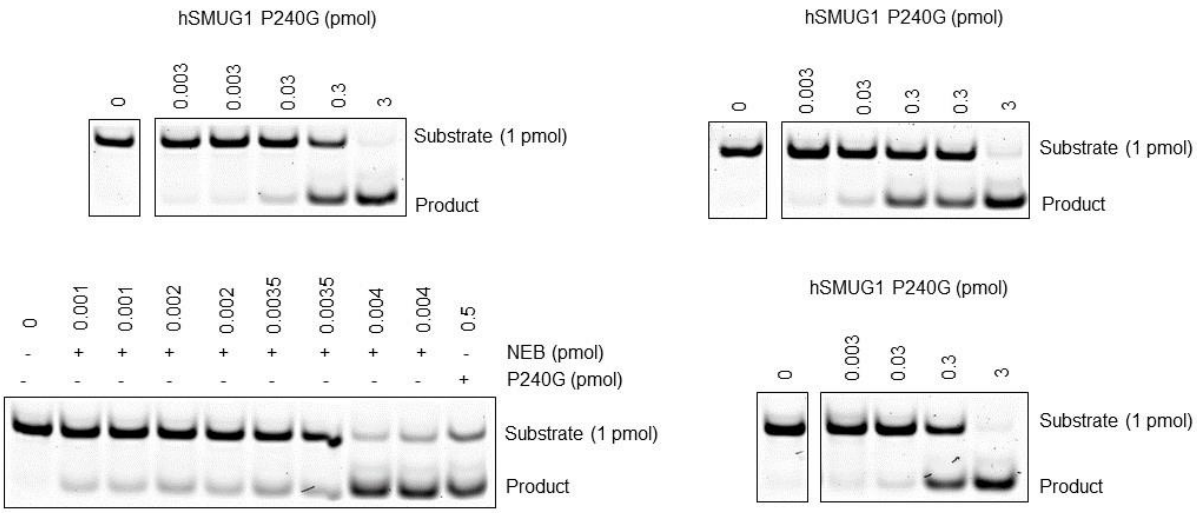




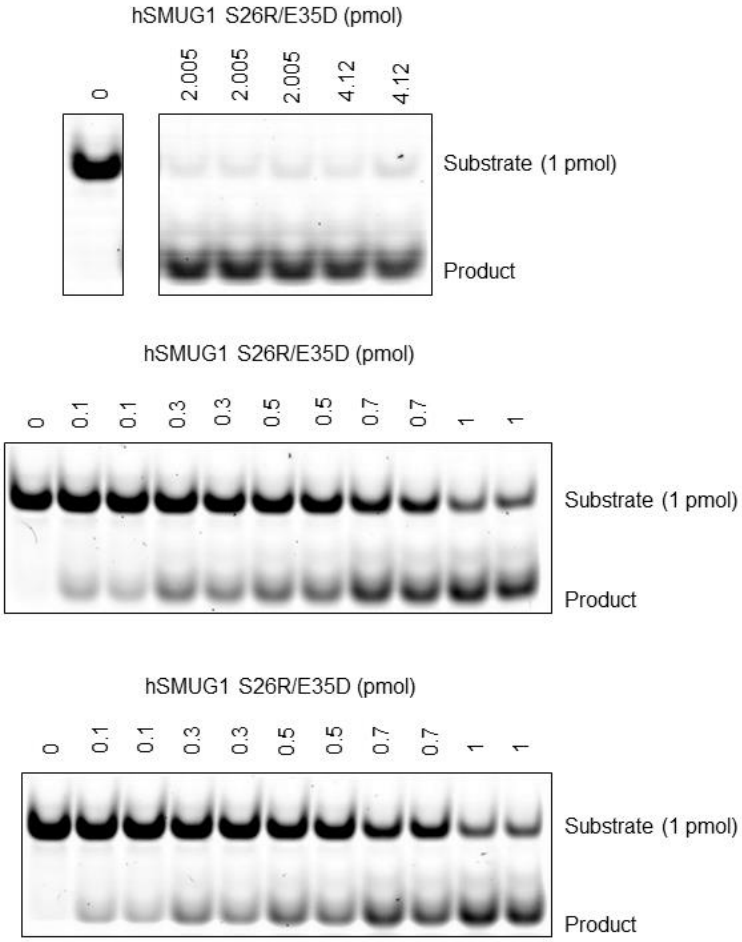




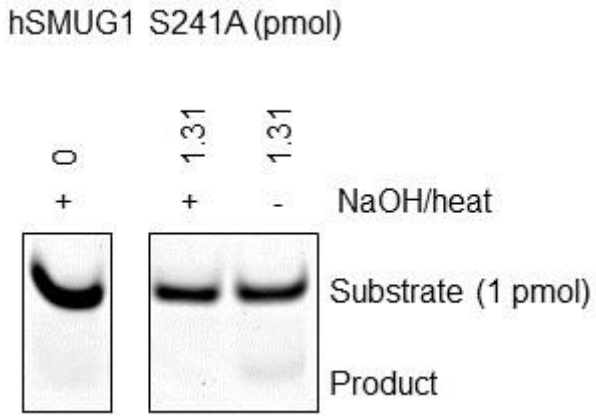




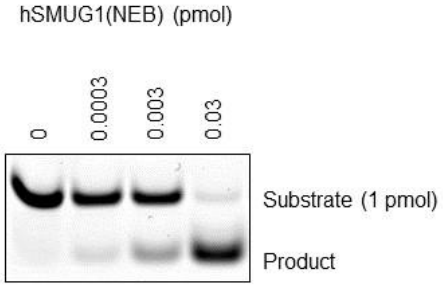
Activity assays for excision of uracil-DNA by hSMUG1 S26R/E35D

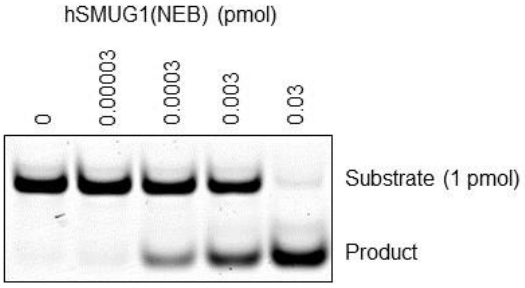


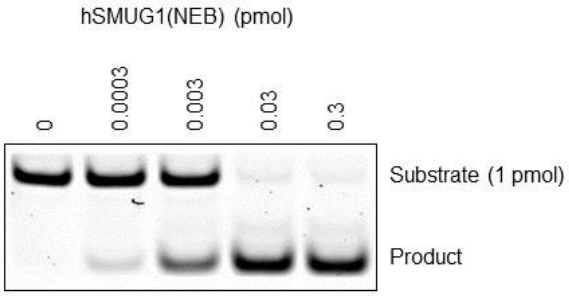
Activity assays for excision of uracil-DNA by hSMUG1 S241A

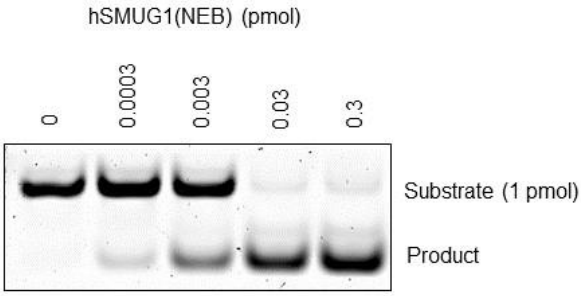


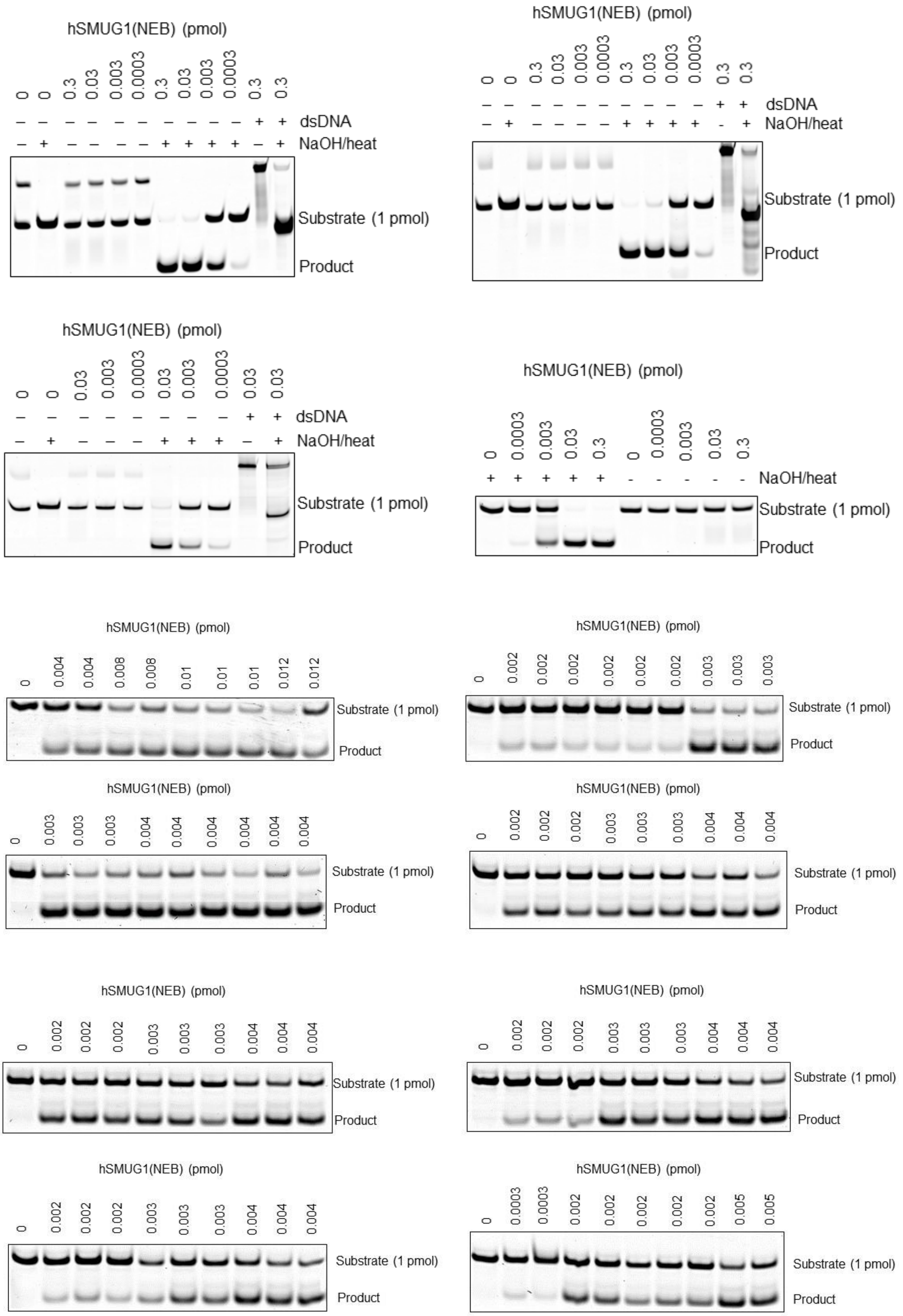
Activity assays for excision of uracil-DNA by hSMUG1(NEB)

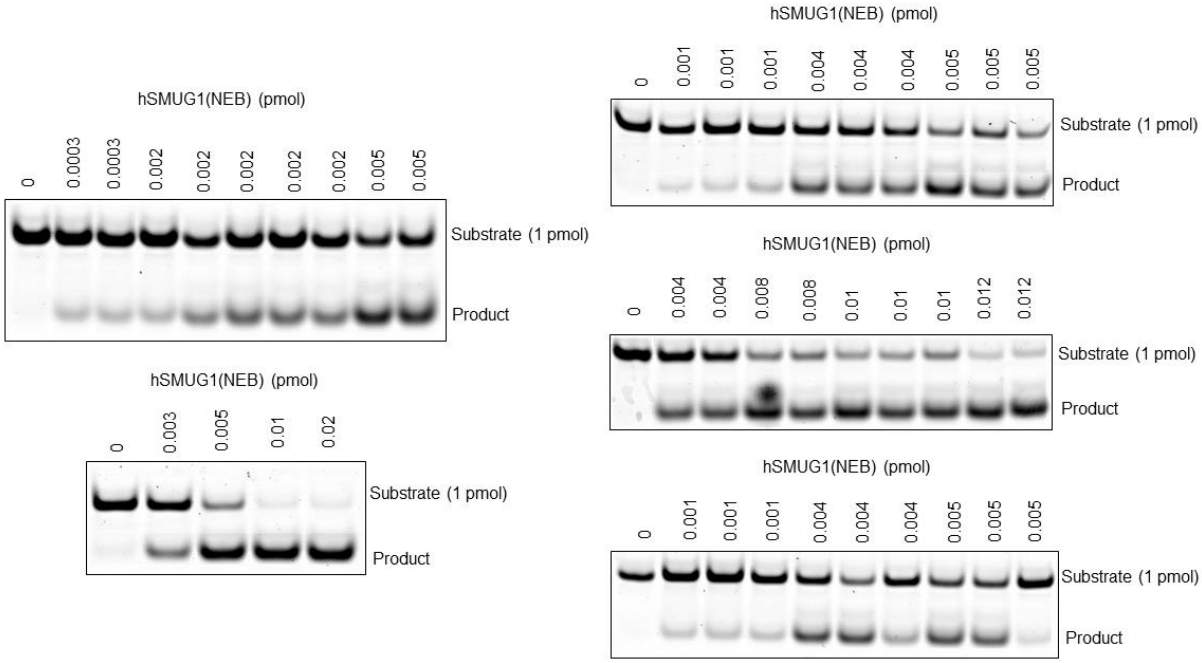




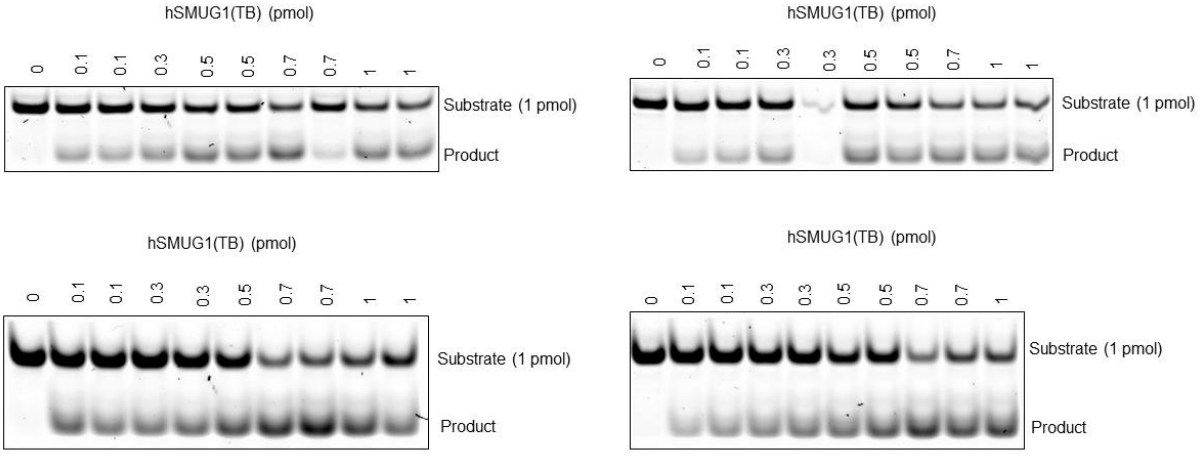





hSMUG1(NEB) (pmol)
0 0 0.3 0.03 0.003 0.0003 0.3 0.03 0.003 0.0003 0.3 0.3
dsDNA
NaOH/heat
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0 0.3 0.03 0.003 0.0003 0.3 0.03 0.003 0.0003 0.3 0.3
dsDNA
NaOH/heat
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0 0.03 0.003 0.0003 0.03 0.003 0.0003 0.03 0.03
dsDNA
NaOH/heat
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0.0003 0.003 0.03 0.3 0 0.0003 0.003 0.03 0.3
NaOH/heat
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0.004 0.004 0.008 0.008 0.01 0.01 0.01 0.012 0.012
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0.002 0.002 0.002 0.002 0.002 0.002 0.003 0.003 0.003
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0.003 0.003 0.003 0.004 0.004 0.004 0.004 0.004 0.004
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0.002 0.002 0.002 0.003 0.003 0.003 0.004 0.004 0.004
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0.002 0.002 0.002 0.003 0.003 0.003 0.004 0.004 0.004
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0.002 0.002 0.002 0.003 0.003 0.003 0.004 0.004 0.004
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0.002 0.002 0.002 0.003 0.003 0.003 0.004 0.004 0.004
Substrate (1 pmol)
Product
hSMUG1(NEB) (pmol)
0 0.0003 0.0003 0.002 0.002 0.002 0.002 0.002 0.005 0.005
Substrate (1 pmol)
Product



Activity assays for excision of uracil-DNA by hSMUG1(TB)



Enzyme kinetics

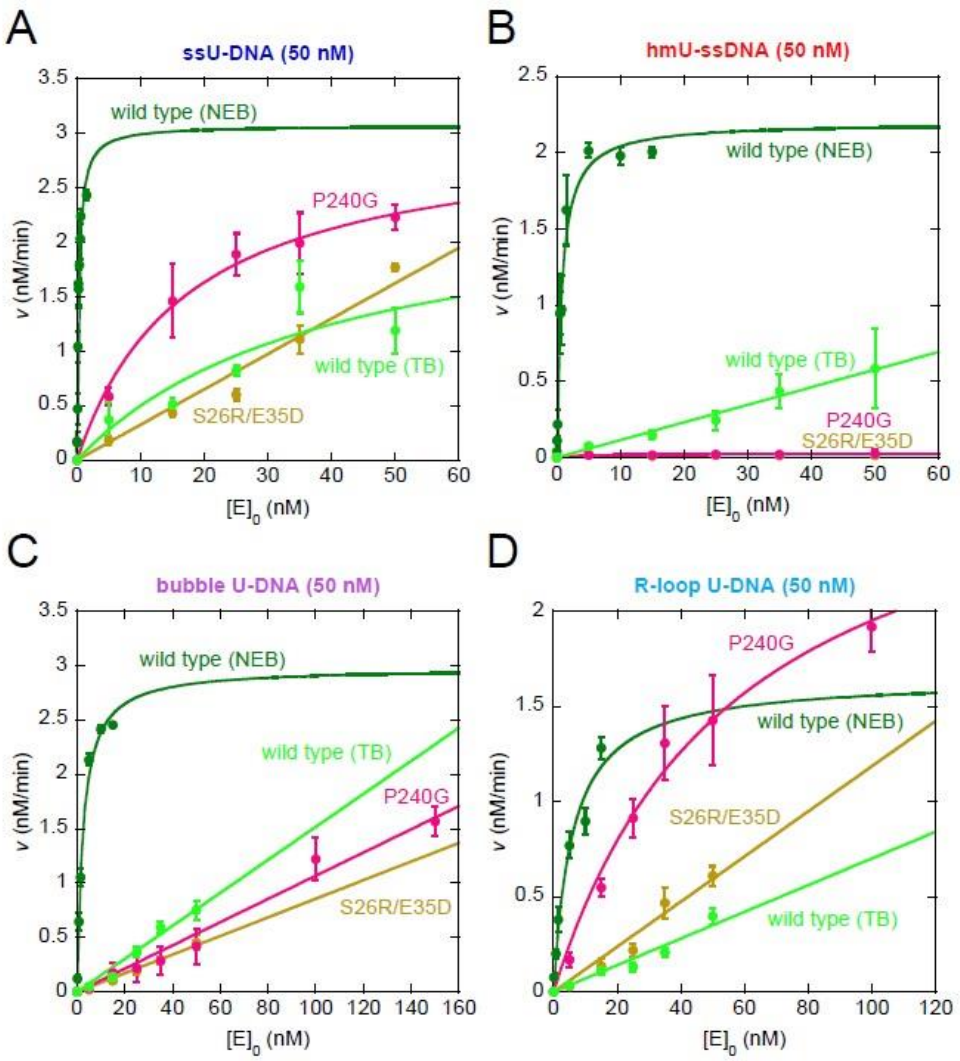


Figure A 9. Overview of all substrates examined by the research group, showing the increasing enzyme concentrations, $[E]_0$ (nM), as a function of the velocity (nM/min). Each value represents the average (±SD) of multiple independent measurements. ***(A)*** *Excision activity for ssU-DNA by hSMUG1 wild type and mutant variants.* ***(B)*** *Excision activity for hmU-ssDNA by hSMUG1 wild type and mutant variants.* ***(C)*** *Excision activity for bubble U-DNA by hSMUG1 wild type and mutant variants.* ***(D)*** *Excision activity for R-loop U-DNA by hSMUG1 wild type and mutant variants. Plots generated by KaleidaGraph (Courtesy of Svein Bjelland).*

Appendix C: Data collection and calculations

For Supplementary Data, see attached file below.

hSMUG1 mutant proteins for ssU,dsU-DNA

Conditions used for all activity assays:

Substrate, S: /5Cy3/CCACACAAAGGG**/dU/**AAAGCCGGGGCA (ssU-DNA), in TE buffer (10 mM Tris pH 7.5, 1 mM EDTA pH 8.0).

Enzyme, E: hSMUG1-FL-P240G, purified in centrifuge and Akta 16.09.2021, 1 µl (directly from stock).

Reaction buffer 45 mM HEPES, pH 7.5, 5 mM DTT, 0.4 mM EDTA, 70 mM KCl, and 0.5 mg/ml bovine serum albumin (BSA)

Incubation time: 20 min

Loading buffer: 80 % formamide, 1 mM EDTA pH 8, 1 % (w/v) blue dextran.

PAGE gel: Denaturing PAGE gel (20 % acrylamide, 3 % formamide, taurine, temed, APS); 5 µl added per well.

Electrophoresis: 120 V, 1.5 h

Date: 19.02.2021

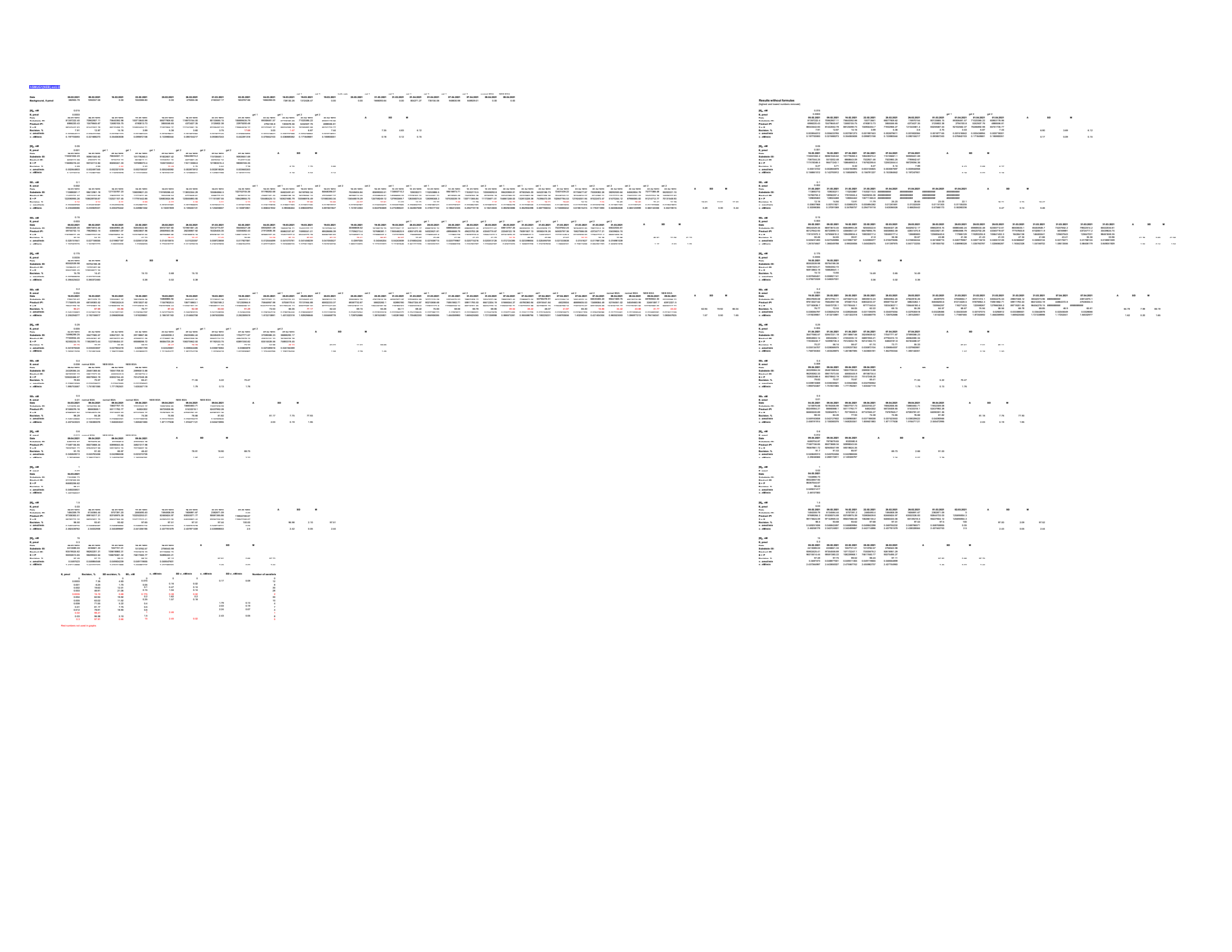
ssU-DNA			(Used for further activity assays)						
hSMUG1 mutant		P240G	P240G	P240G	P240G	P240G	P240G	S241A	NEB
E, pmol	**0**	**1.31**	**5.37**	**3.39**	**0.63**	**Nd**	**8.66**	**1.31**	**0.33**
Substrate	108360337.72	41695201.69	475694.41	96302865.95	64552678.16	93970034.51	73659031.2	90379326.70	3504467.17
Product	0.00	92752564.89	90854253.48	3431136.38	426197.24	6225732.87	8352503.81	993362.30	114728265.94
Excision, %	0.00	68.99	99.48	3.44	0.66	6.21	10.18	1.09	97.04

dsU-DNA					
hSMUG1 mutant	P240G	P240G	P240G	S241A	NEB
E, pmol	**1.31**	**3.39**	**0.63**	**1.31**	**0.33**
Substrate	66945923.84	73512416.97	79701584.60	82533226.72	11025047.27
Product	99609220.47	0.00	741042.38	0.00	159380480.89
Excision, %	59.81	0.00	0.92	0.00	93.53

Date: 01.06.2021

hSMUG1 S26R/E35D					
E, pmol	**2.06**	**2.06**	**2.06**	**4.12**	**4.12**
Substrate	0	1318628.92	1376415.36	857404.56	2256729.78
Product	47757429.31	49795788.55	44438664.55	41017696.12	40072536.7
Excision, %	100.00	97.42	97.00	97.95	94.67

hSMUG1 P240G His-tag						
E, pmol	**0**	**0.655**	**0.655**	**0.655**	**1.31**	**1.31**
Substrate	96924198.69	51010100.33	54783725.78	57224053.96	8196007.92	12474725.25
Product	0.00	48533538.84	33892222.18	31426093.76	51222398.64	43083276.41
Excision, %	0.00	48.76	38.22	35.45	86.21	77.55



hSMUG1(TB) ssU-

Date	**10.05.2021**	**11.05.2021**
Background, 0 pmol	0.00	0.00

$[E]_0$, nM 5
E, pmol 0.1

	gel1	gel1	gel2	gel2	gel1	gel1	gel2	gel2			
Date	**10.05.2021**	**10.05.2021**	**10.05.2021**	**10.05.2021**	**11.05.2021**	**11.05.2021**	**11.05.2021**	**11.05.2021**	**A**	**SD**	**M**
Substrate (S)	104438493.5	96416339.63	61399446.07	58583336.44	94101468.35	87609723.64	70853418.73	65449364.1			
Product (P)	16579847.94	11225913.16	25592841.6	15302517.04	10056569.55	11872637.53	6793789.7	11673795.3			
S + P	121018341.5	107642252.8	86992287.67	73885853.48	104158037.9	99482361.17	77647208.43	77123159.4			
Excision, %	13.70	10.43	29.42	20.71	9.66	11.93	8.75	15.14	14.97	6.52	12.82
***v*, pmol/min**	0.006850138	0.005214455	0.014709834	0.010355512	0.004827553	0.005967207	0.00437478	0.00756828			
***v*, nM/min**	0.342506924	0.260722738	0.735491682	0.517775607	0.241377664	0.298360367	0.21873902	0.37841406	0.37	0.16	0.32

$[E]_0$, nM 15
E, pmol 0.3

	gel1	gel2	gel2	gel1	gel2	gel2			
Date	**10.05.2021**	**10.05.2021**	**10.05.2021**	**11.05.2021**	**11.05.2021**	**11.05.2021**	**A**	**SD**	**M**
Substrate (S)	86348240.41	65274590.27	61356805.39	88112115.28	62014051.6	60565203.58			
Product (P)	17436051.62	15354424.1	18132293.53	22476841.78	14815434.07	19245924.2			
S + P	103784292	80629014.37	79489098.92	110588957.1	76829485.67	79811127.78			
Excision, %	16.80	19.04	22.81	20.32	19.28	24.11	20.40	2.44	19.80
***v*, pmol/min**	0.00840014	0.009521649	0.011405522	0.010162336	0.009641763	0.012057168			
***v*, nM/min**	0.420006999	0.476082469	0.570276106	0.508116777	0.482088157	0.602858421	0.51	0.06	0.50

$[E]_0$, nM 25
E, pmol 0.5

	gel1	gel1	gel2	gel1	gel1	gel2	gel2			
Date	**10.05.2021**	**10.05.2021**	**10.05.2021**	**11.05.2021**	**11.05.2021**	**11.05.2021**	**11.05.2021**	**A**	**SD**	**M**
Substrate (S)	58919072.75	63726975.34	60989403.84	70018566.49	52504000.23	35709591.7	48690616.83			
Product (P)	29723195.7	28006753.08	28370970.82	35295285.26	23089288.53	20610703.33	25678314.77			
S + P	88642268.45	91733728.42	89360374.66	105313851.8	75593288.76	56320295.03	74368931.6			
Excision, %	33.53	30.53	31.75	33.51	30.54	36.60	34.53	33.00	2.05	33.51
***v*, pmol/min**	0.016765814	0.015265243	0.015874469	0.01675719	0.015272049	0.018297759	0.017264141			
***v*, nM/min**	0.8382907	0.763262149	0.793723474	0.837859519	0.763602461	0.914887933	0.863207062	0.82	0.05	0.84

$[E]_0$, nM 35
E, pmol 0.7

	gel1	gel2	gel2	gel1	gel2	gel2			
Date	**10.05.2021**	**10.05.2021**	**10.05.2021**	**11.05.2021**	**11.05.2021**	**11.05.2021**	**A**	**SD**	**M**
Substrate (S)	35751752.74	15680083.72	22746345.14	30849783.31	11862890.43	19912203.21			
Product (P)	43321822.76	36653027.45	55110285.51	29212017.53	38725160.44	31745262.15			
S + P	79073575.5	52333111.17	77856630.65	60061800.84	50588050.87	51657465.36			
Excision, %	54.79	70.04	70.78	48.64	76.55	61.45	63.71	9.72	65.75
***v*, pmol/min**	0.027393363	0.035018965	0.035392159	0.0243183	0.038275007	0.030726694			
***v*, nM/min**	1.369668138	1.750948235	1.769607966	1.215914988	1.913750371	1.536334677	1.59	0.24	1.64

$[E]_0$, nM 50
E, pmol 1

	gel1	gel1	gel2	gel2	gel1	gel1	gel2			
Date	**10.05.2021**	**10.05.2021**	**10.05.2021**	**10.05.2021**	**11.05.2021**	**11.05.2021**	**11.05.2021**	**A**	**SD**	**M**
Substrate (S)	41435008.99	32848337.92	24594330.47	40033051.72	31094200.83	32924218.38	24518864.58			
Product (P)	29299192.18	26616938.57	35384304.2	23043322.68	26662511.09	26632777.75	39839481.44			
S + P	70734201.17	59465276.49	59978634.67	63076374.4	57756711.92	59556996.13	64358346.02			
Excision, %	41.42	44.76	58.99	36.53	46.16	44.72	61.90	47.78	8.56	44.76
***v*, pmol/min**	0.020710768	0.022380236	0.029497424	0.018266207	0.023081743	0.022359067	0.030951294			
***v*, nM/min**	1.035538385	1.119011806	1.474871193	0.913310368	1.154087127	1.117953367	1.547564687	1.19	0.21	1.12

E, pmol	**Excision, %**	**SD excision, %**	**$[E]_0$, nM**	***v*, nM/min**	**SD *v*, nM/min**	**Number of parallels**
0	0	0	0	0	0	0
0.1	14.97	6.52	5	0.37	0.16	8
0.3	20.4	2.44	15	0.51	0.06	6
0.5	33	2.05	25	0.82	0.05	7
0.7	63.71	9.72	35	1.59	0.24	6
1	47.78	8.56	50	1.19	0.21	7

hSMUG1 P240G ssU-D

						gel 1	gel 2	gel 1	gel 2	gel 1
Date	25.02.2021	26.02.2021	01.03.2021	02.03.2021	04.03.2021	08.03.2021	08.03.2021	09.03.2021	09.03.2021	10.03.2021
Background, 0 pmol	0.00	479255.06	2163347.17	1653767.66	1684288.05	882142.81	689409.94	0.00	0.00	0.00

[E]$_0$, nM	1.5								
E, pmol	0.03								
Date	25.02.2021	26.02.2021	01.03.2021	02.03.2021	04.03.2021		A	SD	M
Substrate (S)	77420044.88	93486870.07	78301395.79	109925719.17	111418776.23				
Product (P)	5050028.01	2392326.10	3020420.83	551795.53	833638.79				
S + P	82470072.89	95879196.17	81321816.62	110477514.70	112252415.02				
Excision, %	6.12	2.50	3.71	0.50	0.74		2.71	2.07	2.50
v, pmol/min	0.003061734	0.001247573	0.001857079	0.000249732	0.000371323				
v, nM/min	0.153086684	0.062378654	0.092853952	0.012486603	0.018566166		0.07	0.05	0.06

[E]$_0$, nM	15												
E, pmol	0.3					gel 1	gel 1	gel 2	gel 1	gel 2			
Date	25.02.2021	26.02.2021	01.03.2021	02.03.2021	04.03.2021	08.03.2021	09.03.2021	09.03.2021	10.03.2021	10.03.2021	A	SD	M
Substrate (S)	44702913.79	83849766.23	49406097.12	65120632.50	90020046.11	86089225.21	91556374.83	70385401.51	85809886.34	75447615.72			
Product (P)	43793106.71	38400670.66	26017705.39	43085948.13	26817782.23	88153786.52	30257765.31	24781041.63	27581980.94	19900074.89			
S + P	88496020.50	122250436.89	75423802.51	108206580.63	116837828.34	174243011.7	121814140.1	95166443.14	113391867.3	95347690.61			
Excision, %	49.49	31.41	34.50	39.82	22.95	50.59	24.84	26.04	24.32	20.87	32.48	10.34	28.73
v, pmol/min	0.024742981	0.01570574	0.017247675	0.019909116	0.011476498	0.025296219	0.012419644	0.013019842	0.012162239	0.010435531			
v, nM/min	1.23714904	0.785286982	0.862383774	0.995455819	0.573824904	1.264810933	0.620982204	0.650992115	0.608111975	0.521776531	0.81	0.26	0.72

[E]$_0$, nM	25												
E, pmol	0.5												
Date	16.03.2021	16.03.2021	16.03.2021	16.03.2021	16.03.2021	16.03.2021	16.03.2021	16.03.2021	16.03.2021		A	SD	M
Substrate (S)	29145087.50	36746018.96	36335838.12	33273551.73	40182487.22	32718356.09	25996983.61	29961065.37	25109937.84				
Product (P)	108249374.3	93403016.1	109618410.3	98316059.58	91761913.25	66659333.45	88765023.19	69906810.16	74093344.74				
S + P	137394461.79	130149035.06	145954248.43	131589611.31	131944400.47	99377689.54	114762006.80	99867875.53	99203282.58				
Excision, %	78.79	71.77	75.10	74.71	69.55	67.08	77.35	70.00	74.69		73.23	3.65	74.69
v, pmol/min	0.039393645	0.0358831	0.037552319	0.037357075	0.034772947	0.03353838	0.038673523	0.034999648	0.0373442				
v, nM/min	1.969682272	1.794154987	1.877615957	1.86785375	1.738647357	1.676918978	1.933676172	1.749982409	1.867210006		1.83	0.09	1.87

[E]$_0$, nM	35									
E, pmol	0.7	gel 2	gel 1	gel 2	gel 1	gel 2				
Date	08.03.2021	08.03.2021	09.03.2021	09.03.2021	10.03.2021	10.03.2021		A	SD	M
Substrate (S)	13026305.07	20784119.66	11671247.97	9016876.37	9541751.47	3866976.83				
Product (P)	116943293.5	80133252.39	80814846.32	68944735.07	82079037.73	67667184.65				
S + P	129969598.54	100917372.05	92486094.29	77961611.44	91620789.20	71534161.48				
Excision, %	89.98	79.40	87.38	88.43	89.59	94.59		88.23	4.55	89.01
v, pmol/min	0.044988711	0.039702407	0.043690269	0.044217105	0.044792802	0.047297112				
v, nM/min	2.249435537	1.98512037	2.184513438	2.210855246	2.239640109	2.36485559		2.21	0.11	2.23

[E]$_0$, nM	50							
E, pmol	1	gel 2	gel 1	gel 2				
Date	09.03.2021	09.03.2021	10.03.2021	10.03.2021		A	SD	M
Substrate (S)	5096818.54	7828229.18	5039461.55	4748513.48				
Product (P)	70306609.58	60512874.72	77310414.82	48624600.29				
S + P	75403428.12	68341103.90	82349876.37	53373113.77				
Excision, %	93.24	88.55	93.88	91.10		91.69	2.09	92.17
v, pmol/min	0.046620301	0.044272679	0.046940213	0.045551587				
v, nM/min	2.331015026	2.213633936	2.347010652	2.27757933		2.29	0.05	2.30

[E]$_0$, nM	65									
E, pmol	1.3	gel 1	gel 1	gel 2	gel 1	gel 2				
Date	08.03.2021	08.03.2021	09.03.2021	09.03.2021	10.03.2021	10.03.2021		A	SD	M
Substrate (S)	3253168.72	2438791.01	3035676.25	5585436.13	2358181.21	2443519.48				
Product (P)	102630845.3	86577106.78	88100481.20	67231282.94	83270444.70	61430775.79				
S + P	105884013.97	89015897.79	91136157.45	72816719.07	85628625.91	63874295.27				
Excision, %	96.93	97.26	96.67	92.33	97.25	96.17		96.10	1.73	96.80
v, pmol/min	0.048463806	0.048630137	0.048334538	0.04616473	0.048623018	0.048087243				
v, nM/min	2.423190277	2.431506869	2.416726897	2.308236481	2.431150909	2.404362175		2.40	0.04	2.42

[E]$_0$, nM	150								
E, pmol	3								
Date	25.02.2021	26.02.2021	01.03.2021	02.03.2021	04.03.2021		A	SD	M
Substrate (S)	1608121.91	2451037.69	791779.48	0.00	1802767.34				
Product (P)	87734301.30	93406985.61	67226155.28	84992539.50	53963881.39				
S + P	89342423.21	95858023.30	68017934.76	84992539.50	55766648.73				
Excision, %	98.20	97.44	98.84	100.00	96.77		98.25	1.12	98.20
v, pmol/min	0.049100023	0.048721527	0.049417963	0.05	0.048383651				
v, nM/min	2.455001167	2.436076355	2.470898136	2.5	2.41918255		2.46	0.03	2.46

E, pmol	Excision, %	SD excision, %	[E]$_0$, nM	*v*, nM/min	SD *v*, nM/min	Number of parallels
0	0	0	0	0	0	0
0.03	2.71	2.07	1.5	0.07	0.05	5
0.3	32.5	10.38	15	0.81	0.26	10
0.5	73.23	3.65	25	1.83	0.09	9
0.7	88.26	4.51	35	2.21	0.11	4
1	91.69	2.09	50	2.29	0.05	4
1.30	96.11	1.73	65	2.4	0.04	4
3	98.25	1.12	150	2.46	0.03	5

Red numbers not used in graphs

hSMUG1 P240G ssU

Results of this sheet have been performed at same days as the following substrates: bubbleU-DNA and R-loopU-DNA (CAR), sshm5U-DNA (TB)

	gel 1	gel 2	gel 1	gel 2	gel 1	gel 2
Date	**12.04.2021**	**12.04.2021**	**14.04.2021**	**14.04.2021**	**15.04.2021**	**15.04.2021**
Background, 0 pmol	1081739.21	614762.58	1495817.72	0	0	0

$[E]_0$, nM 5
E, pmol 0.1

	gel 1	gel 1	gel 2	gel 2	gel 1	gel 1	gel 2	gel 2	gel 1	gel 1	gel 2	gel 2						
Date	**12.04.2021**	**12.04.2021**	**12.04.2021**	**12.04.2021**	**14.04.2021**	**14.04.2021**	**14.04.2021**	**14.04.2021**	**15.04.2021**	**15.04.2021**	**15.04.2021**	**15.04.2021**	**A**		**SD**		**M**	
Substrate (S)	98567801.40	120699500.01	114741233.95	105015096.16	117667755.81	114375226.47	105869184.43	112267679.99	81890833.21	77373189.37	92979790.67	87667802.76						
Product (P)	43221616.69	38685213.59	35314457.95	29436775.63	22070875.63	36698887.77	35458750.18	28985480.18	23718831.22	23528074.69	26831815.60	25857736.13						
S + P	141789418.09	159384713.60	150055691.90	134451871.79	139738631.44	151074114.24	141327934.61	141253160.17	105609664.43	100901264.06	119811606.27	113525538.89						
Excision, %	30.48	24.27	23.53	21.89	15.79	24.29	25.09	20.52	22.46	23.32	22.40	22.78		23.07		3.21		23.05
***v*, pmol/min**	0.015241482	0.012135798	0.011767117	0.010946956	0.007897199	0.012145988	0.012544848	0.010260117	0.011229479	0.011658959	0.011197503	0.011388511						
***v*, nM/min**	0.76207409	0.606789897	0.588355855	0.547347821	0.39485995	0.607299403	0.627242418	0.513005871	0.561473975	0.582947967	0.559875133	0.569425532		0.58		0.08		0.58

$[E]_0$, nM 15
E, pmol 0.3

	gel 1	gel 1	gel 2	gel 2	gel 1	gel 1	gel 2	gel 2	gel 1	gel 1	gel 2	gel 2						
Date	**12.04.2021**	**12.04.2021**	**12.04.2021**	**12.04.2021**	**14.04.2021**	**14.04.2021**	**14.04.2021**	**14.04.2021**	**15.04.2021**	**15.04.2021**	**15.04.2021**	**15.04.2021**	**A**		**SD**		**M**	
Substrate (S)	19298357.55	23733637.85	35880110.74	40938047.17	89694087.47	82379279.27	70884736.72	81015848.05	39243717.3	26386107.95	35489216.76	48290220.79						
Product (P)	63440089.39	65784991.39	111419208.7	123458393.1	59604195.12	63571700.63	51639168.6	59566096.76	54619392.24	37576252.81	59490841.58	56177875.9						
S + P	82738446.94	89518629.24	147299319.5	164396440.3	149298282.6	145950979.9	122523905.3	140581944.8	93863109.54	63962360.76	94980058.34	104468096.7						
Excision, %	76.68	73.49	75.64	75.10	39.92	43.56	42.15	42.37	58.19	58.75	62.64	53.78		58.52		13.72		58.47
***v*, pmol/min**	0.038337733	0.036743744	0.03782068	0.037548986	0.019961447	0.021778443	0.021073099	0.021185543	0.029095239	0.029373723	0.031317543	0.026887575						
***v*, nM/min**	1.916886639	1.837187185	1.891034004	1.877449306	0.99807235	1.088922128	1.053654968	1.05927715	1.454761954	1.468686129	1.565877159	1.344378755		1.46		0.34		1.46

$[E]_0$, nM 25
E, pmol 0.5

	gel 1	gel 1	gel 2	gel 2	gel 1	gel 1	gel 2	gel 2				
Date	**12.04.2021**	**12.04.2021**	**12.04.2021**	**14.04.2021**	**14.04.2021**	**15.04.2021**	**15.04.2021**	**15.04.2021**	**15.04.2021**	**A**	**SD**	**M**
Substrate (S)	25190110.05	28893312.63	47739806.37	52970955.17	15355440.16	17607344.78	15784165.22	12918987.97				
Product (P)	84663826.65	115355770.9	85219104.86	84226105.55	72399768.44	68220898.01	62760352.6	55845153.81				
S + P	109853936.7	144249083.6	132958911.2	137197060.7	87755208.6	85828242.79	78544517.82	68764141.78				
Excision, %	77.07	79.97	64.09	61.39	82.50	79.49	79.90	81.21		75.70	7.65	79.69
***v*, pmol/min**	0.038534726	0.039984923	0.032047158	0.030695302	0.041250981	0.039742686	0.03995209	0.040606305				
***v*, nM/min**	1.926736292	1.999246166	1.602357903	1.534765124	2.062549038	1.987134298	1.99760449	2.030315233		1.89	0.19	1.99

$[E]_0$, nM 35
E, pmol 0.7

	gel 1	gel 1	gel 2	gel 2	gel 1	gel 1	gel 2	gel 2	gel 1	gel 1	gel 2	gel 2						
Date	**12.04.2021**	**12.04.2021**	**12.04.2021**	**12.04.2021**	**14.04.2021**	**14.04.2021**	**14.04.2021**	**14.04.2021**	**15.04.2021**	**15.04.2021**	**15.04.2021**	**15.04.2021**	**A**		**SD**		**M**	
Substrate (S)	15734732.41	37358096.11	16643616.84	48449039.42	21745155.62	20501145.05	23636259.49	20276869.36	9192222.71	7995666.69	10787076.83	12246066.02						
Product (P)	100334752	42985562.89	94760046.62	63218736.07	98377293.95	93721019.64	125500020.2	103938732.8	68066172.22	65711060.8	56981304.87	57253701.11						
S + P	116069484.4	80343659	111403663.5	111667775.5	120122449.6	114222164.7	149136279.7	124215602.1	77258394.93	73706727.49	67768381.7	69499767.13						
Excision, %	86.44	53.50	85.06	56.61	81.90	82.05	84.15	83.68	88.10	89.15	84.08	82.38		79.76		11.28		83.88
***v*, pmol/min**	0.043221848	0.026751061	0.042530041	0.028306616	0.040948755	0.041025759	0.042075617	0.041838034	0.044050988	0.044576027	0.042041217	0.041189851						
***v*, nM/min**	2.161092395	1.337553063	2.12650203	1.415330783	2.047437725	2.051287942	2.103780858	2.091901721	2.202549389	2.228801326	2.102060852	2.059492552		1.99		0.28		2.10

$[E]_0$, nM 50
E, pmol 1

	gel 1	gel 1	gel 2	gel 2	gel 1	gel 1	gel 2	gel 2	gel 1	gel 2								
Date	**12.04.2021**	**12.04.2021**	**12.04.2021**	**12.04.2021**	**14.04.2021**	**14.04.2021**	**14.04.2021**	**14.04.2021**	**15.04.2021**	**15.04.2021**	**A**		**SD**		**M**			
Substrate (S)	11301444.91	12694097.43	23083556.77	7043877.57	8073942.36	9791586.7	11498378.29	10030088.99	3766271.85	2672040.02								
Product (P)	61101977.2	64734216.51	102203874.9	94776227.83	91238245.02	112487086.3	71278828.31	105137796.4	65843073.79	53252606.19								
S + P	72403422.11	77428313.94	125287431.7	101820105.4	99312187.38	122278673	82777206.6	115167885.4	69609345.64	55924646.21								
Excision, %	84.39	83.61	81.58	93.08	91.87	91.99	86.11	91.29	94.59	95.22		89.37		4.71		91.58		
***v*, pmol/min**	0.042195504	0.041802677	0.04078776	0.046541018	0.04593507	0.0459962	0.043054623	0.045645449	0.047294708	0.047611035								
***v*, nM/min**	2.109775181	2.090133868	2.03938802	2.327050916	2.296753486	2.299810007	2.152731145	2.282272442	2.36473541	2.380551769		2.23		0.12		2.29		

E, pmol	**Excision, %**	**SD excision, %**	**$[E]_0$, nM**	***v*, nM/min**	**SD *v*, nM/min**	**Number of parallels**
0	0	0	0	0	0	0
0.1	23.07	3.21	5	0.58	0.08	12
0.3	58.52	13.72	15	1.46	0.34	12
0.5	75.7	7.65	25	1.89	0.19	8
0.7	79.76	11.28	35	1.99	0.28	12
1	89.37	4.71	50	2.23	0.12	10

S26RE35D ssU-DNA c

Results of this sheet have been performed at same days as the following substrates: bubbleU-DNA and R-loopU-DNA (CAR), sshm5U-DNA (TB)

	gel 1	gel 2					
Date	**09.06.2021**	**09.06.2021**					
Background, 0 pmol	0	0					
$[E]_0$, nM	5						
E, pmol	0.1						
	gel 1	gel 1	gel 2	gel 2			
Date	**09.06.2021**	**09.06.2021**	**09.06.2021**	**09.06.2021**	**A**	**SD**	**M**
Substrate (S)	90072238.21	92389966.72	91137866.41	83447658.19			
Product (P)	9241482.67	6943514.68	8155910.46	4125197.04			
S + P	99313720.88	99333481.4	99293776.87	87572855.23			
Excision, %	9.31	6.99	8.21	4.71	7.30	1.71	7.60
***v*, pmol/min**	0.004652672	0.003495053	0.00410696	0.002355294			
***v*, nM/min**	0.232633582	0.174752626	0.205347976	0.117764718	0.18	0.04	0.19
$[E]_0$, nM	15						
E, pmol	0.3						
	gel 1	gel 1	gel 2	gel 2			
Date	**09.06.2021**	**09.06.2021**	**09.06.2021**	**09.06.2021**	**A**	**SD**	**M**
Substrate (S)	86939292.51	84659407.12	79734866.84	71528477.25			
Product (P)	19922280.31	15447605.27	17807759.06	13974694.66			
S + P	106861572.8	100107012.4	97542625.9	85503171.91			
Excision, %	18.64	15.43	18.26	16.34	17.17	1.33	17.30
***v*, pmol/min**	0.009321536	0.007715546	0.009128193	0.008172033			
***v*, nM/min**	0.466076808	0.385777302	0.45640967	0.408601644	0.43	0.03	0.43
$[E]_0$, nM	25						
E, pmol	0.5						
	gel 1	gel 1	gel 2	gel 2			
Date	**09.06.2021**	**09.06.2021**	**09.06.2021**	**09.06.2021**	**A**	**SD**	**M**
Substrate (S)	77252227.56	76420126.87	66219991.22	65880895.17			
Product (P)	28553908.12	21607455.82	20989196.62	19735861.40			
S + P	105806135.7	98027582.69	87209187.84	85616756.57			
Excision, %	26.99	22.04	24.07	23.05	24.04	1.85	23.56
***v*, pmol/min**	0.013493503	0.01102111	0.012033822	0.0115257			
***v*, nM/min**	0.674675148	0.55105551	0.601691093	0.576285011	0.60	0.05	0.59
$[E]_0$, nM	35						
E, pmol	0.7						
	gel 1	gel 1	gel 2	gel 2			
Date	**09.06.2021**	**09.06.2021**	**09.06.2021**	**09.06.2021**	**A**	**SD**	**M**
Substrate (S)	44951064.31	58859536.56	50064176.43	39237283.95			
Product (P)	44174207.43	32380720.84	41248915.54	35043056.96			
S + P	89125271.74	91240257.4	91313091.97	74280340.91			
Excision, %	49.56	35.49	45.17	47.18	44.35	5.35	46.17
***v*, pmol/min**	0.024782088	0.017744755	0.022586529	0.023588379			
***v*, nM/min**	1.239104425	0.887237766	1.129326437	1.17941896	1.11	0.13	1.15
$[E]_0$, nM	50						
E, pmol	1						
	gel 1	gel 1	gel 2	gel 2			
Date	**09.06.2021**	**09.06.2021**	**09.06.2021**	**09.06.2021**	**A**	**SD**	**M**
Substrate (S)	21488472.92	22613820.45	16789251.16	19431849.41			
Product (P)	52030107.09	50250350.76	45670933.06	45340687.7			
S + P	73518580.01	72864171.21	62460184.22	64772537.11			
Excision, %	70.77	68.96	73.12	70.00	70.71	1.53	70.39
***v*, pmol/min**	0.035385686	0.034482208	0.036560037	0.034999932			
***v*, nM/min**	1.769284278	1.724110421	1.828001856	1.749996593	1.77	0.04	1.76

E, pmol	**Excision, %**	**SD excision, %**	**$[E]_0$, nM**	***v*, nM/min**	**SD *v*, nM/min**	**Number of parallels**
0	0	0	0	0	0	0
0.1	7.30	1.71	5	0.18	0.04	4
0.3	17.17	1.33	15	0.43	0.03	4
0.5	24.04	1.85	25	0.6	0.05	4
0.7	44.35	5.35	35	1.11	0.13	4
1	70.71	1.53	50	1.77	0.04	4