Use of sequence microdivergence in mycobacterial ortholog to analyze contributions of the water-activating loop histidine of *Escherichia coli* uracil–DNA glycosylase in reactant binding and catalysis

Narottam Acharya, Ramappa K. Talawar, Kedar Purnapatre,1 and Umesh Varshney*

Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560 012, India

Received 5 June 2004
Available online 25 June 2004

Abstract

Uracil–DNA glycosylase (Ung), a DNA repair enzyme, pioneers uracil excision repair pathway. Structural determinations and mutational analyses of the Ung class of proteins have greatly facilitated our understanding of the mechanism of uracil excision from DNA. More recently, a hybrid quantum-mechanical/molecular mechanical analysis revealed that while the histidine (H67 in *Eco*Ung) of the GQDPYH motif (o loop) in the active site pocket is important in positioning the reactants, it makes an unfavorable energetic contribution (penalty) in achieving the transition state intermediate. Mutational analysis of this histidine is unavailable from any of the Ung class of proteins. A complication in demonstrating negative role of a residue, especially when located within the active site pocket, is that the mutants with enhanced activity are rarely obtained. Interestingly, unlike the most Ung proteins, the H67 equivalent in the o loop in mycobacterial Ung is represented by P67. Exploiting this natural diversity to maintain structural integrity of the active site, we transplanted an H67P mutation in *Eco*Ung. Uracil inhibition assays and binding of a proteinaceous inhibitor, Ugi (a transition state substrate mimic), with the mutant (H67P) revealed that its active site pocket was not perturbed. The catalytic efficiency (*V*<sub>max</sub>/*K*<sub>m</sub>) of the mutant was similar to that of the wild type Ung. However, the mutant showed increased *K*<sub>m</sub> and *V*<sub>max</sub>. Together with the data from a double mutation H67P/G68T, these observations provide the first biochemical evidence for the proposed diverse roles of H67 in catalysis by Ung.

© 2004 Elsevier Inc. All rights reserved.

Spontaneous deamination of cytosines and/or erroneous incorporation of dUMP by DNA polymerase(s) are chiefly responsible for the occurrence of uracil in DNA as G:U or A:U pairs, respectively [1]. Appearance of uracil in the genome in either form hampers cellular processes. If not repaired prior to replication, the presence of G:U pairs leads to GC → AT mutations in one of the progeny duplexes. And, the A:U base pairs can impede recognition by the cognate DNA binding proteins [2]. Therefore, to maintain error free genome and safeguard the physiological functions, cells possess a superfamily of uracil–DNA glycosylases (UDGs) which initiate uracil excision repair [3,4]. Ung proteins, which efficiently utilize both the single- and double-stranded DNAs as substrates [5], are a highly conserved class of UDGs found in all organisms including many viruses. The activity of Ung on single-stranded DNA is physiologically relevant as such DNAs arise during transcription or replication. In fact, the cytosines in single-stranded DNA are inherently prone to deamination.

Crystal structure analyses have revealed an amazing conservation of the overall architecture and active site geometry of various Ung proteins. Further, the cocrystal structures of the protein with an uncleavable substrate analog or the reaction products have provided a wealth of information on the mechanism of substrate recognition and catalysis [6–12]. Several hydrogen bonds are established between the conserved Ung residues such as histidine of HPSLS motif (DNA intercalation loop) and asparagine of GVLLLN motif (specificity pocket) and the 2, 3, and 4 positions of uracil. The asparagine
hydrogen bonds with O4 and N3 of uracil and restricts cytosine from entering the pocket. Likewise, the tyrosine of the GQDPYPY motif (o loop), which is in van der Waals’ contact with the C5 position of the uracil, excludes thymine with a methyl group at this position or the purines with bulky rings. A number of mutational analyses have confirmed these predictions and yielded further insights into the additional roles of many of these residues [13–22].

More recently, the mechanism of catalysis by human Ung was analyzed by a hybrid quantum-mechanical/molecular mechanical (QM/MM) approach [23]. The findings of this elegant computer analysis provide strong evidence to the dissociative mode of catalysis by Ung [12]. Importantly, the QM/MM study has highlighted the crucial contribution of the phosphates in the DNA backbone, and the role of the key residues in Ung in transformation of the substrate to transition state intermediate(s).

A residue of particular interest to us, in this study, is the histidine of the GQDPYPY motif (loop). The QM/MM study revealed that while this histidine is important in positioning the reactants, it slows down the reaction by making an unfavorable energetic contribution (penalty) to achieving the transition state intermediate. Unfortunately, no mutational analyses of this histidine (H67 in EcoUng) have been reported so far. A complication in demonstrating the role of a residue in proteins is that the desired mutants with enhanced activity are rarely obtained. To add to this challenge, the histidine in question is located within the active site pocket. Therefore, any mutational designs to test the function of this residue must be carefully planned.

To address the issue of H67, we took advantage of the natural diversity within the available Ung sequences. Unlike the most Ung proteins, the o loop in mycobacterial Ung is represented by the GQDPYP sequence [24]. Considering that mycobacterial Ung is an efficient UDG [25], we transplanted an H67P mutation in Escherichia coli Ung. We show that the biochemical parameters (K_a and V_max) of uracil excision by H67P and an additional mutant, H67P/G68T, are consistent with the proposed diverse roles of H67 in reactant binding and the actual catalysis by Ung.

Materials and methods

DNA oligomers and 5’-end labeling of SSU9. DNA oligomers were obtained from Microsynth, Switzerland. The oligomer, SSU9 [5’-d(cagatcttatcaggtagagact)-3’] containing dU, was [5’-32P]-end labeled using [y-32P]ATP (6000 Ci/mmol, DuPont NEN, USA) and T4 polynucleotide kinase, and purified on Sephadex G-50 minicolumns [26].

Generation of EcoUng mutants. To generate H67P mutation in EcoUng, 5’-cagatcttatcaggtagagact-3’ and 5’-tggcggatatcggtagtctg-3’ primers were employed in a quick change mutagenesis protocol using Pfu DNA polymerase along with ~50 ng pTrcEcoUng. To generate a double mutation (H67P and G68T), the H67P template [pTrcEcoUng(H67P)] was further mutated using 5’-cctatcaccagcagggc-agc-3’ and 5’-gtgcttggtggtaggaaag-3’ primers [19]. The mutations were verified by DNA sequence analysis [26].

Generation of overexpression constructs for Ung and Ung–Ugi complexes. The Nru–HindIII fragment carrying a part of Ung ORF (downstream of the site of mutation) and the Ugi ORF from pTrcEcoUng–Ugi was subcloned into the similar sites of pTrcEcoUng(H67P) and pTrcEcoUng(H67P/G68T) to generate pTrcEcoUng(H67P–Ugi and pTrcEcoUng(H67P/G68T–Ugi bicistronic constructs, respectively [27]. The Ncol–HindIII fragments from these constructs were then subcloned into the same sites of pET11d and referred to as pETEcoUng(H67P)–Ugi and pETEcoUng(H67P/G68T)–Ugi. The bicistronic constructs were digested with EcoRI to excise Ugi and self-ligated to generate pETEcoUng(H67P) and pETEcoUng(H67P/G68T) for overproduction of Ung proteins.

Analysis of the in vivo formed Ung–Ugi complexes. E. coli BL26(DE3) was transformed with the pET11d based expression constructs and grown in 2YT medium containing ampicillin (100 µg/ml). At mid-log phase, the cells were induced with 0.5 mM IPTG for 3–4 h, harvested by centrifugation, sonicated in TME (25 mM Tris-HCl, pH 8.0, 2 mM β-mercaptoethanol, and 1 mM Na2EDTA), and clarified by centrifugation at 20,000g for 10 min. The cell-free extracts thus obtained were analyzed on 15% polyacrylamide (19:1 crosslinking) gels without or with 2–8 M urea [19].

Purification of Ung mutants and determination of K_a and V_max. The pET11d based Ung constructs were introduced into E. coli BL26 (DE3) and the transformants were inoculated into 1 L of 2YT medium [26]. At mid-log phase, the cells were induced with 0.5 mM IPTG for 3–4 h. The cells were harvested and Ung proteins were purified [19]. To determine kinetic parameters of uracil excision, reactions (15 µl) containing varying amounts of SSU9 along with 20,000 cpm of the [5’-32P]-end labeled substrate and appropriate amounts of Ung proteins in the reaction buffer were incubated at 37°C for 10 min and stopped by adding 5 µl of 0.2 N NaOH. The reaction mixture was dried in vacuo, taken up in 10 µl of loading dye, and half of the contents were electrophoresed on 15% polyacrylamide–8 M urea gels. The bands corresponding to the protein and the remaining substrate were quantified by using a BioImage Analyser. The percentage values of product and the remaining substrate were quantified.

Inhibition of Ung by uracil. The reaction mixture (20 µl) was set up in 2× Ung buffer with 5 pmol of SSU9 containing 20,000 cpm of the [5’-32P]-end labeled substrate as tracer, in the absence or presence of 5 mM uracil at 37°C. The reaction was started by addition of 5 µl of appropriate dilutions of Ung in the reaction buffer and incubated at 37°C for 10 min. The reactions were terminated by adding 5 µl of 0.2 N NaOH and heating at 90°C for 30 min, and analyzed on 15% polyacrylamide–8 M urea gel. The bands corresponding to substrate and product were quantified by a BioImage Analyser (Fuji, FLA 2000) and pmol uracil released per min per µg of Ung was determined.

Results

Comparison of mycobacterial Ung proteins with EcoUng and design of the Ung mutants

Mycobacterial Ung is an efficient enzyme [25]. More recently, the open reading frame of Ung from Mycobacterium smegmatis (MsmUng) was cloned and sequenced [24]. The MsmUng showed an overall identity of 39% and 51% similarity with EcoUng. Interestingly, this comparison revealed that the GQDPYPY motif, a
highly conserved sequence in the o loop of various Ung proteins, is represented by GQDPYP in MsmUn
(g Fig. 1). Taking advantage of this natural microdiversity in the conserved motif sequence, we transplanted H67P in the sequence context of EcoUng to analyze the signiﬁcance of the role of H67 in Ung. In addition, as a control, because the GQDPYP sequence is followed by a Thr in MsmUng (as opposed to Gly in other Ung sequences), we generated yet another construct of EcoUng which included both the H67P/G68T mutations.

**Purification of EcoUng mutants**

The H67P and the H67P/G68T mutants were overproduced in E. coli BL26 (DE3) by using the T7 RNA polymerase based constructs (Fig. 2) and puriﬁed [19] to apparent homogeneity (Fig. 3).

**Inhibition of Ung proteins by uracil**

Uracil, one of the reaction product, is a known inhibitor of Ung with a $K_i$ of ~0.2–5 mM. Presence of uracil in the reaction leads to an inhibition which varies from 40% to 70% [29–34]. In our earlier study we have demonstrated that structural perturbation in the uridine-binding pocket due to the Y66W mutation resulted in complete loss of inhibition of the Y66W protein by uracil [22]. Therefore, to probe the active site pockets of the H67P and H67P/G68T mutants, we examined the effects of inclusion of uracil in the reaction. As shown in Fig. 4, inclusion of uracil in the reactions resulted in inhibition of the wild type, H67P, and H67P/G68T proteins to roughly the same extent. As neither H67 nor G68 make any contacts with the uracil, a direct effect of the mutations was not anticipated. However, these
observations are significant in supporting that the mutations did not result in structural perturbation of the residues that contact uracil in the active site pocket.

**Analysis of Ung–Ugi complexes in urea PAGE**

Ugi is a small (9.4 kDa), highly acidic (pI of 4.2) and thermostable protein of the *Bacillus subtilis* bacteriophage (PBS-1/-2) origin [2]. It forms an extremely specific and exclusively stable complex with the Ung class of UDGs. The co-crystal structures of Ung–Ugi have shown that Ugi is a transition state mimic for substrate DNA [10]. The conserved sequence motifs of the Ung lie at the interface with Ugi to establish an intricate network of interactions. The histidine of the GQDPYH sequence (H67 of *Eco*Ung), establishes two hydrogen bonds with the Q19 and S21 of Ugi [9,10].

Hence, to further ensure proper folding as well as to probe for structural alterations, if any, in the active site face of the mutants (H67P and H67P/G68T), we analyzed the complexes of Ugi with the Ung mutants using a polyacrylamide gel system that we developed [19] to analyze the stability of the in vivo formed Ung–Ugi complexes. In these gels, the prominent band in the middle of the gel corresponds to Ung–Ugi complex. As shown in Fig. 5, electrophoretic analysis on the native gel did not reveal any differences between the complexes of Ugi with the wild type or the mutant Ung proteins (panel i). Therefore, to study the effect of mutations at H67 and H67/G68, we electrophoresed the complexes on urea containing gels. Similar to the complex of Ugi with the wild type Ung, its complexes with the two mutants were fully stable to electrophoresis in 2 or 4 M urea gels (panels ii and iii), suggesting that the mutants suffered no major structural changes. However, as evidenced from the presence of smear and the separated Ugi, some dissociation of the complex involving H67P/G68T Ung was detectable in 6 M urea (panel iv). Upon further increase in urea concentration (8 M), a minimal dissociation of the Ung complex with H67P was also seen (panel v). However, as the H67 makes H-bonds with Ugi, some dissociation of H67P mutant, in 8 M urea gel, was indeed expected [35]. Therefore, these observations are consistent with the notion that other than a direct effect of loss of H-bonds with Ugi, the H67P mutant did not cause a detectable structural change. On the other hand, because G68 is not known to make any interactions with Ugi, a greater degree of dissociation of H67P/G68T over that of H67P in 8 M urea gel, might be argued to point to a subtle structural perturbation in the active site pocket of the H67P/G68T mutant.

**Steady state kinetics of uracil release by Ung mutants**

The steady state kinetic parameters of uracil excision from SSU9, by the wild type and the two mutant Ung proteins carried out in duplicates, are shown in Table 1. The catalytic efficiencies (\(V_{\text{max}}/K_m\)) of uracil excision by the H67P mutant are very similar to those of the wild type Ung. Interestingly, a closer examination of the \(K_m\) and \(V_{\text{max}}\) numbers shows that with respect to the wild type protein, there is an increase in the average values of both the \(K_m\) and the \(V_{\text{max}}\) of uracil excision by the H67P Ung protein. These observations are fully consistent...
with the QM/MM prediction that H67 contributes positively to reactant binding but adversely to transformation of substrate into the transition state complex. In the H67P mutant, the loss of contact from H67 results in an increase in $K_m$ but at the same time removal of the adverse energetic factor towards the substrate transformation results in increased $V_{\text{max}}$. While the introduction of a double mutation, H67P/G68T, in Eco Ung, which arguably suffered a small structural perturbation in the active site pocket (Fig. 5) resulted in an increase in $K_m$, it did not reveal a significant increase in $V_{\text{max}}$ comparable to that seen for the H67P mutant.

Importantly, although the structural perturbation in H67P/G68T, as deduced from the data shown in Fig. 5, appears only a subtle one, the observations with this mutant underscore the challenge posed by such structural perturbations while carrying out mutational analysis of the active site residues.

### Discussion

The Ung proteins are amongst the most active DNA repair enzymes. High-resolution X-ray crystallographic studies on these conserved proteins from various sources have been extremely useful in identifying the key amino acid residues involved in the catalysis [13,16,17]. These studies have provided a strong basis for a large number of mutational designs and biochemical studies to probe the role of the Ung residues identified from the structural studies. Furthermore, such studies provided the basis for proposals on actual mechanisms of the glycosidic bond cleavage based on the general acid base catalysis, or only the general base catalysis involving the Asp of the water-activating loop and the His of the DNA intercalation loop [6,7,14,15,36,37]. However, neither the general acid base nor the general base mechanisms are consistent with the observations made from more recent co-crystal structures of human UNG with uncleavable substrate harboring pseudouridine analog (C–C glycosidic bond) in place of uridine (N–C glycosidic bond). Instead, these structures and the subsequent QM/MM study suggest that the glycosidic bond cleavage occurs by substrate autocatalysis [12,23]. The substrate phosphates and the Pro–Ser loops in Ung (PPS and the HPSPLS motifs, Fig. 1) make crucial contributions to lowering the activation energy needed to attain the transition state before the substrate decomposes into products. Interestingly, the QM/MM study also makes a prediction that although the His of the GQDPYH motif is important in binding and positioning the reactants as it bridges the $5'$ phosphate of the $-1$ base to the water molecule that attacks the C1' of the deoxyribose, it actually slows down the transition of

![Fig. 5. Stability of Ung–Ugi complexes. Cell free extracts (~15 μg) of the transformants expressing Ung–Ugi complexes from the bicistronic constructs, as shown, were subjected to electrophoresis on 15% polyacrylamide gels containing different concentrations of urea (as indicated) and stained with Coomassie brilliant blue.](image)

<table>
<thead>
<tr>
<th>Ung</th>
<th>$K_m \times 10^{-7}$ (M)</th>
<th>$V_{\text{max}} \times 10^3$ pmol/min/μg</th>
<th>$V_{\text{max}}/K_m \times 10^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>W.T.</td>
<td>4.8</td>
<td>228</td>
<td>45.5</td>
</tr>
<tr>
<td>H67P</td>
<td>10.0</td>
<td>416</td>
<td>43.1</td>
</tr>
<tr>
<td>H67P/G68T</td>
<td>9.9</td>
<td>238</td>
<td>28.6</td>
</tr>
</tbody>
</table>

Table 1

Kinetic parameters of uracil excision from SSU9
the substrate into the products. In fact, this residue has also been shown to be important in making hydrogen bonds with the Q19 and S21 of Ugi, a transition state substrate mimic [9,10]. Despite such an involvement of this residue, there have been no reports on the effects of substitution of this residue with any other amino acid.

We have been interested in studies on EcoUng. And, although the QM/MM studies have been performed using human UNG, the fact that the human UNG and EcoUng possess identical sequences within the conserved motifs, and share incredibly similar three dimensional structure [9,10], in this study we chose to carry out mutational analysis of H67 in EcoUng. We were guided by the sequence of the loop of the mycobacterial Ung, to tailor an H67 to P67 mutation. Further, because in mycobacterial Ung, P67 is followed by a single-stranded substrate, SSU9. However, a and the Indian Council of Medical Research, New Delhi, India.

References

[20] P. Handa, N. Acharya, U. Varshney, Effects of mutations at tyrosine 66 and asparagine 123 in the active site pocket of Escherichia coli uracil-DNA glycosylase on uracil excision from synthetic DNA oligomers: evidence for the occurrence of

Acknowledgments

We thank our laboratory colleagues for their suggestions. This work was supported in part by research grants from the Department of Biotechnology, Council of Scientific and Industrial Research (CSIR), and the Indian Council of Medical Research, New Delhi, India.
long-range interactions between the enzyme and substrate, Nucleic Acids Res. 30 (2002) 3086–3095.


