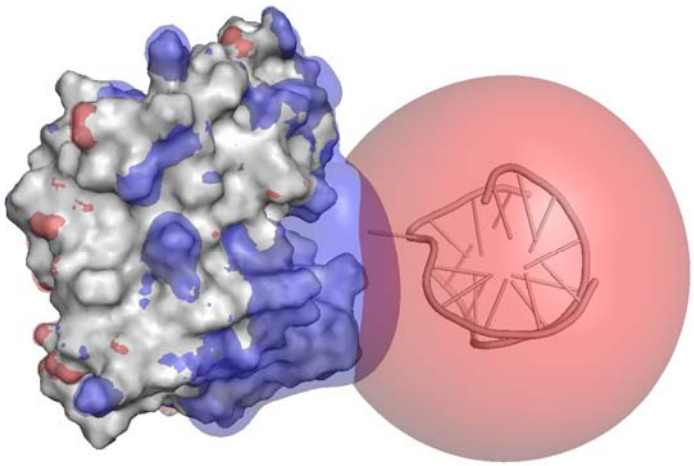
A DISSERTATION FOR THE DEGREE OF PHILOSOPHIAE DOCTOR

The molecular origin of cold adaptation:

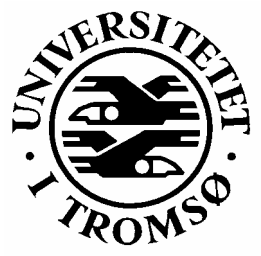
A comparative study of cold- and warm-active uracil DNA glycosylase

Magne Olufsen

April 2007



Department of Chemistry, Faculty of Science,
University of Tromsø, Norway



A DISSERTATION FOR THE DEGREE OF PHILOSOPHIAE DOCTOR

The molecular origin of cold adaptation:

A comparative study of cold- and warm-active uracil DNA glycosylase

Magne Olufsen

April 2007

Department of Chemistry, Faculty of Science,
University of Tromsø, Norway

Acknowledgements

The work leading to this thesis has been carried out at the Norwegian Structural Biology Centre (NorStruct), Department of Chemistry, University of Tromsø, Norway, from January 2003 to April 2007. Financial support was given from the Research Council of Norway (NFR), which is greatly acknowledged.

I wish to express my deepest gratitude to my supervisors Professor Arne O. Smalås and Dr. Bjørn Olav Brandsdal. Arne, your structured leadership, pleasant personality, supportive nature and your wise guidance has made it a pleasure to work for you. Bjørn Olav, your optimistic nature, wise guidance, expertise in the field and our scientific discussions have been very inspiring and of great help through these years.

I would also like to thank all the former and present group members both at Department of Chemistry and at NorStruct for creating a stimulating working environment.

I am grateful to my mother Gunn Robertsen and Odd Valdermo for endless hours of child care, which has made it possibly for me to focus more on my research. A warm gratitude goes to my girlfriend Vibeke Aune. Your love, care and encouragement, especially in down periods, have meant a lot to me. Finally, I want to thank my daughter Martine. You have shown exceptional patience when I had to work late, travel abroad or when you had to stay at the University with me.

Magne Olufsen

Tromsø, April 2007

Abbreviations

Throughout the text the standard 3-letter abbreviations for amino acids are used. Other abbreviations are listed below:

UDG	Uracil DNA glycosylase
cUDG	Cod uracil DNA glycosylase
hUDG	Human uracil DNA glycosylase
MD	Molecular dynamics
MM	Molecular mechanics
MM-PBSA	Molecular-mechanics/Poisson-Boltzmann/surface area
r.m.s.d.	Root mean squared deviations (starting structure as reference)
r.m.s.f.	Root mean squared fluctuations (average structure as reference)
ε_p	Dielectric constant of the protein
PB	Poisson-Boltzmann
GB	Generalized-Born
LIE	Linear interaction energy
E. coli	Escherichia coli

List of papers

I. Moe, E., Leiros, I., Riise, E.K., **Olufsen, M**., Lanes, O., Smalås, A.O. & Willassen, N.P. Optimisation of the surface electrostatics as a strategy for cold adaptation of uracil-DNA N-glycosylase (UNG) from atlantic cod (Gadus morhua). J. Mol. Biol. (2004). **343**, 1221-1230

Comments: My contribution to this paper has been to crystallize and solve the structure of the hUDG-E171V mutant.

II. **Olufsen, M**., Smalås, A.O., Moe, E. & Brandsdal, B.O. Increased flexibility as a strategy for cold adaptation – A comparative molecular dynamics study of cold- and warm-active uracil DNA glycosylase. J. Biol. Chem. (2005). **280**, 18042-18048

III. **Olufsen, M**., Brandsdal, B.O. & Smalås, A.O. Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme. J. Mol. Graph. Model. (2007), in press.

IV. **Olufsen, M**., Papaleo, E., Smalås, A.O. & Brandsdal, B.O. Ion pairs and their role in modulating stability of cold- and warm-active uracil DNA glycosylase. Submitted to Proteins.

V. **Olufsen, M**., Smalås, A.O. & Brandsdal, B.O. Electrostatic interactions play an essential role in DNA repair and cold-adaptation of Uracil DNA Glycosylase. Manuscript.

Reprints were made with permission from the copyright holders

Related publications

i. Mekonnen, S.M., **Olufsen M**., Smalås, A.O. & Brandsdal, B.O. Predicting proteinase specificities from free energy calculations. J. Mol. Graph. Model. (2006). **25**, 176-185.

ii. Papaleo, E., **Olufsen, M**., De Gioia, L. & Brandsdal, B.O. Optimization of electrostatics as a strategy for cold-adaptation: A case study of cold- and warm-active elastases. J. Mol. Graph. Model. (2007), in press.

Table of contents

Background

During the last decade or so, scientists at the Norwegian Structural Biology Centre (NorStruct) have used the enzyme uracil DNA glycosylase (UDG) as a model system in the study of cold-adaptation, protein-DNA recognition and enzyme specificity. The cod UDG (cUDG) and human UDG (hUDG) are thoroughly biologically characterized and there are established well-functioning recombinant expression systems to produce mutants. More than 30 mutants of cUDG and hUDG have been expressed, purified and characterized. The kinetic constants (k_{cat} and K_m) have been obtained for the majority of these mutants. In addition, the crystal structure has been determined for some of them. This Ph.D. project is an extension of earlier work done at NorStruct.

Enzyme kinetics

Enzymes are protein catalysts which speed up rates of chemical reactions by temporarily binding to the substrate (target molecule) and lower the activation energy needed to convert the substrate to product. Michaelis and Menten developed in 1913 the typical scheme for enzymatic reactions [1]:

$$(I) \qquad E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \overset{k_2}{\rightarrow} E + P$$

where E, S and P are enzyme, substrate and product, respectively. In the simple Michaelis-Menten kinetics with only one enzyme-substrate complex and all binding steps are fast, then k_2 is equal to k_{cat}. Enzymatic reactions are usually monitored in terms of the substrate concentration ([S]) and are analyzed in terms of velocity (v):

$$(II) \qquad v = \frac{-d[S]}{dt}$$

The steady-state approximation gives the velocity as a function of substrate consentration:

$$(III) \qquad v = \frac{[E]_0 \cdot [S] \cdot k_{cat}}{K_m + [S]}$$

where k_{cat} and K_m are the turnover number and the Michaelis constant, respectively. The catalytic efficiency of an enzymatic reaction is defined as k_{cat}/K_m, and in order for an enzyme to increase its catalytic efficiency it can either increase k_{cat}, decrease K_m or adjust both parameters [2]. The Michaelis constant is linked to the binding of the enzyme-substrate:

$$(IV) \qquad K_m = K_S + \frac{k_2}{k_1}$$

where K_S is the dissociation constant (inverse of the binding constant). When k_2 is the rate limiting step ($k_2 << k_{-1}$), then K_m is equal to K_s. Equation IV is also valid for Briggs-Haldane kinetics (when $k_2 >> k_{-1}$) [1]. The Michaelis-Menten scheme may be extended to cover a wide range of different cases with additional intermediates in the reaction pathway, but the Michaelis-Menten still applies [1]. In these cases K_m is always less or

equal to K_S. There is still thus an inverse relationship between the binding energy of the enzyme-substrate complex and the K_m value. k_{cat}/K_m is also called the second order rate constant, and with low substrate concentration ([S] << K_m), equation III can be reduced to [1]:

$$(V) \qquad v = \frac{k_{cat}}{K_m} \cdot [E]_0 \cdot [S]$$

Enzymes can also use binding energy to lower the activation energy of an enzymatic reactions, instead of lowering the K_m [1,3]. If the enzyme binds stronger to the transition state relative to the ground state then the binding energy will lower the activation energy. In such cases it is much better to compare k_{cat}/K_m for an enzymatic reaction with two different substrates, because k_{cat}/K_m includes both the activation energy and the binding energy. By using different substrates and comparing the k_{cat}/K_m values, it is possibly to calculate the difference in both activation and binding energy associated with reaction with different substrates:

$$(VI) \qquad \Delta\Delta G = -R \cdot T \cdot \frac{(k_{cat}/K_m)_A}{(k_{cat}/K_m)_B}$$

Extremophiles

On our planet there are many harsh habitats that are considered extreme. Examples of such habitats include: the deep sea with high pressure, salt lakes with high salt concentrations, areas close to volcanoes with high temperatures, dry desert regions and polar regions with low temperatures. Organisms have populated all these different habitats, and in order to survive in the harsh environments the organisms have adapted to different extreme conditions. In some cases they have adapted to more than just one of the extreme conditions. Like for example organisms living in the deep sea, they have adapted to high pressure, low temperatures and salt. Extremophiles can be classified according to their environmental requirements for optimal growth [4], and Table 1 lists the most common subclasses.

Table 1
Examples of extremophiles and the conditions they are adapted to.

Name	Adapted to
Thermophiles	High temperatures
Psychrophiles	Low temperatures
Barophilies/piezophiles	High pressure
Alkaliephiles	High pH
Acidophiles	Low pH
Xerophiles	Dry conditions
Osmophiles	High sugar concentrations

2.1 Cold adaptation

A vast amount of our planet consists of cold environments, like for example the Arctic, Antarctic, mountain regions, glaciers and deep sea waters. Microorganisms and ectotherms (cold-blooded animals, whose body temperature is regulated by their behavior or surroundings) that live in cold areas have to adapt to the surrounding environment. Ectotherms and micro organisms living in cold environments have many physiological adaptations which help them to cope with low temperature. Lipids in the cell membrane of Arctic fish species are for example less saturated than those of southern fish (a

chemical exchange that is equivalent of replacing butter with olive oil), making the lipids more liquid at low temperatures. The cold also affects the metabolism, the complete set of chemical reactions that occur in a living organism to maintain its body. Enzymes are crucial to metabolism as their main task is to catalyze unfavorably chemical reactions in the cells. Enzymes from microorganisms and ectotherms living in cold regions are often referred to as cold-adapted or psychrophilic enzymes (Greek; psychro means cold, cold loving enzymes) [5]. The temperature is one of the most important factors for enzyme activity, and enzymatic reaction rates can be reduced 30-80 times when the temperature decreases from 37 to 0°C [6]. To deal with this temperature dependency, cold-adapted enzymes usually have higher catalytic efficiency at moderate and low temperatures compared to their mesophilic and thermophilic homologues [6]. The temperature dependence of chemical reactions, including enzymatic reactions can be described by the Arrhenius equation [6]:

$$(VII) \qquad k_{cat} = Zp \cdot e^{-E_a / RT}$$

where k_{cat} is the turnover number, E_a is the activation energy, R is the gas constant, T is the temperature, Z is the collision frequency and p is a steric factor. Increased k_{cat} can be achieved by lowering the activation energy for the reaction [7].

Several strategies have been postulated to explain how enzymes adapt to cold environments, but the most widely accepted hypothesis is that increased structural flexibility of components involved in the catalytic cycle in psychrophilic enzymes enhances the catalytic efficiency [5,8,9]. Psychrophilic enzymes are less heat stabile than their mesophilic homologues, and this is thought to be a result of the above mentioned increase in flexibility [5,7]. The stability/flexibility relationship is controversial because cold-adapted organisms are under no selective pressure to stabilize their proteins at elevated temperatures, and it is believed that the stability property has slowly vanished due to genetic drift [10]. Although the flexibility hypothesis has been the dominating theory to explain the increased catalytic efficiency of cold-adapted enzymes, other adaptational strategies should not be ruled out. For example, cold-adapted enzymes can increase their catalytic activity by optimizing the electrostatic properties at and around the active site [5].

2.2 Structural adaptation to low temperatures

Considerable effort has been directed towards defining structural features important for thermal adaptation. Numerous protein homologues from cold and warm environments have been analyzed in order to pinpoint structural and sequential differences to explain thermal adaptation. Unfortunately, differences that are critical for thermal adaptation are often hidden among other differences produced by genetic drift and other adaptational effects, which make it difficult to identify features specific for thermal adaptation [7]. The overall fold and the active site of cold- and warm-active protein homologues are very similar, indicating that the catalytic mechanism and reaction pathway are the same [8]. There are, however, also differences between cold- and warm-active enzymes, and progress has been made, both to explain adaptation to cold and to warm environments.

Weakening of intramolecular non-bonded interactions are often referred to in order to explain the higher molecular flexibility and lower thermal stability of psychrophilic proteins compared to their mesophilic homologues. Several comparative studies have reported that the number of salt-bridges are lower in cold-adapted enzymes [7,11,12]. Psychrophilic enzymes have also shown to lack surface salt-bridges and ion pairs between secondary structure elements and domains compared to mesophilic homologues [13,14]. In most cases salt-bridges are thought to stabilize the protein structure, thus making the psychrophilic enzymes, with fewer salt-bridges, more flexible and thermally unstable compared to mesophilic homologues. The picture is further complicated as the dielectric constant of water changes with temperature (from 55.5 at 100°C to 88.0 at 0°C [15]), consequently, the strength of ionic interactions will decrease with decreasing temperature due to increased screening.

In other cases, psychrophilic enzymes have a general lack of aromatic interactions [13] or fewer hydrogen bonds than their mesophilic counterparts [11,16]. Another trend is that psychrophilic enzymes have smaller and less hydrophobic residues in the core of the protein compared to both mesophilic and thermophilic enzymes [7,9]. This will lead to cavities in the core of the protein, and such structural differences will probably make the psychrophilic enzymes more unstable and more flexible.

The molecular surface of enzymes seems to be important for cold adaptations, and psychrophilic enzymes tend to have a higher proportion of hydrophobic residues at the surface [7]. Hydrophobic residues at the surface will destabilize the protein structure due to a decrease in entropy of the water molecules surrounding the protein [7]. For some cold-adapted enzymes an increase in charged residues at the protein surface has been observed, especially for negatively charged residues [17,18]. For example, charge-charge repulsions is thought to increase flexibility of the linker region of the psychrophilic cellulase [19].

Cold-adapted enzymes have been shown to possess an increased number of glycine residues and a reduced number proline residues in their sequences, particularly in loop regions when compared to their warm-active homologues [2,5]. Glycine, which lacks a side chain, can make the main chain more flexible. The side chain of Pro, on the other hand, forms a five member ring with the Cα atom and the main chain nitrogen, leading to a more rigid main chain. Enzymes form cold environments have also in some cases low relative Arg content [Arg/(Arg+Lys)] compared to the mesophilic counterparts [2,5,7]. The Arg side chain has a much higher hydrogen bond potential compared to Lys and these hydrogen bonds are able to stabilize the protein structure. In addition, it has been reported that psychrophilic enzymes have larger accessibility to the catalytic cavity compared to its warm-active homologues [20,21], giving rise to higher specific activity at low temperatures.

There are also other structural features which are thought to be important for cold adaptation, but the above section covers the most frequently discussed features in the literature. It is important to note that not all cold-adapted enzymes possess all the features listed above, but seem to usually use a few of them in order to achieve efficient catalysis at low temperatures. In an extensive study of structural differences between thermophilic, mesophilic and psychrophilic enzymes it was found that different protein families use different strategies to adapt to low temperatures [11].

Uracil DNA glycosylase (UDG) as a model system

In this study UDG has been used as a model system to study cold adaptation and protein-DNA recognition. The biological function of UDG is well known. In contrast to other nucleotide excision repair systems, UDG does not require any additional co-factors for activity, which makes it an excellent candidate for studies of cold adaptation and protein-DNA recognition.

3.1 DNA repair

DNA can be damaged by a variety of agents and processes, such as spontaneous deamination of bases, radiation, oxidative stress, alkylating agents and replication errors [22]. Faithful maintenance of the genome is crucial to the individual and to species [23]. In humans, DNA damage can be repaired by four major repair pathways, and several proteins are involved in each of the pathways [24]. If the DNA repair system is defect and the DNA is not repaired, severe diseases such as cancer can occur. Uracil, which does not normally occur in DNA, can appear in DNA either if deoxyuridine triphosphate is misincorporated in DNA instead of thymine or as a result of deamination of cytosine [23]. There are at least 11 different mammalian DNA glycosylases which initiate the base excision repair (BER) and removes damaged or inappropriate bases [24]. DNA glycosylases cleave the N-glycosylic bond between the target base and deoxyribose, and releases a free base and leaves an apurinic/apyrimidine site [22].

3.2 Uracil DNA glycosylase

Uracil DNA glycosylase is a DNA-repair enzyme in the base excision repair (BER) pathway and removes uracil from both single and double stranded DNA. All free living organisms express uracil DNA glycosylase, indicating that this is a highly important enzyme [25]. Six members in the uracil DNA glycosylase family are known: Uracil DNA N-glycosylase (Family-1), mismatch-specific DNA glycosylase, single-stranded selective monofunctional uracil DNA glycosylase, thermostable uracil DNA glycosylase, uracil

DNA glycosylase B and MIG protein/endonuclease III/*Methanococcus jannaschii* uracil DNA glycosylase family [26-28]. The different families of uracil DNA glycosylase have very limited sequence similarity and differ also in the makeup of the active site [28,29]. Surprisingly, there are also a lack in conservation of catalytic residues among the different families of uracil DNA glycosylase [27].

Human UDG (hUDG) and cod UDG (cUDG), both from Family-1, have been used as the model system to study cold adaptation. Throughout this thesis the abbreviation UDG will be used for enzymes from the Family-1 uracil DNA glycosylase superfamily. In humans this enzyme occurs both in nuclei and mitochondria [23], and the two enzymes have different sequence in the N-terminal. UDG from the nuclei consists of 313 amino acids while the mitochondria UDG only consists of 304 amino acids, but the catalytic domain is identical for the two variants of the enzyme [30]. Even if the N-terminal part of the enzyme is removed and only the catalytic domain is left, the enzyme is still fully active [30]. Some results also indicate that UDG is able to interact with other proteins [31-33], this interaction seems to be through the presequence (the N-terminal end) of UDG, which is not necessary for catalytic activity [32]. UDG is highly specific for uracil and show negligible activity towards the natural DNA bases or uracil in RNA (Fig. 1). UDG binds the uracil base in a specificity pocket, and the residues forming this pocket is highly conserved in this family [27].

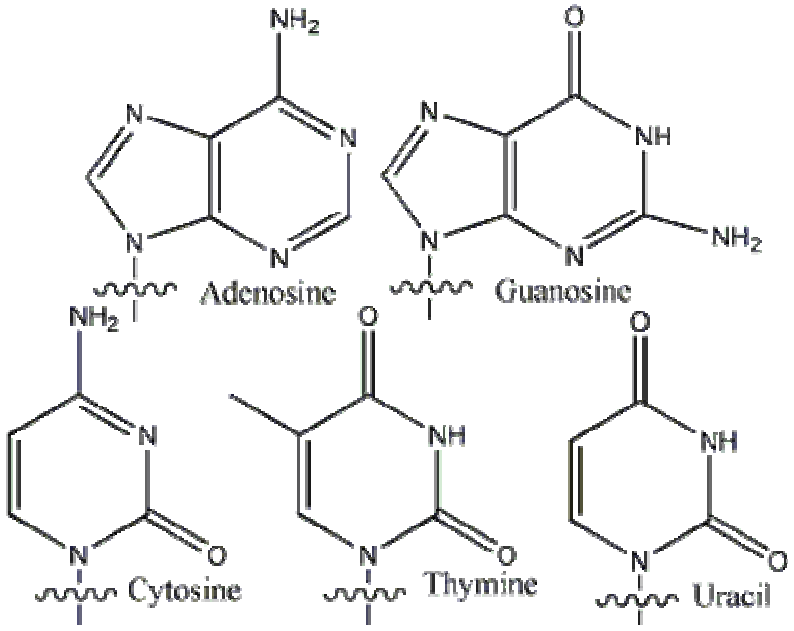


Figure 1
The four natural occurring bases in the DNA and the uracil base.

3.2.1 Recognition of DNA and catalytic mechanism of UDG

Perhaps the least understood stage in the processing of uracil bases in DNA is how UDG recognize the damaged DNA sites within vast stretches of DNA [34]. UDG can search for damaged bases either in a distributive or a processive mechanism. In the distributive mechanism the UDG dissociates from the DNA after removing the uracil base. While in a processive mechanism the UDG locates sequential uracil prior to dissociation [35]. In vitro experiments have indicated that UDG slides along the DNA and scans the strand for uracil residues in a processive manner [35,36]. This mechanism is highly affected by the salt concentration, and already at around 50 mM salt the search is shifted to a distributive search mechanism [36]. There are two different views of how UDG recognize the uracil in the DNA strand. The *base sampling model* suggests that UDG localizes uracil by breaking base pairs in the double stranded DNA and flip them out to test them against the interactions offered in the specificity pocket [27]. Another view is the *inherent extrahelicity model*, suggesting that the base pairs involving uracil are inherently weak and that the uracil will spontaneous flip-out to an extrahelical conformation,

complementary to the binding interactions offered by UDG [27,37]. The UDG then trap the extrahelical uracil base.

UDG catalyses the removal of uracil from both double and single stranded DNA by cleaving the bond between the uracil base and the sugar ring (glycosylic bond) [22]. Studies of the catalytic mechanism of human and *Escherichia coli* (*E. coli*) UDG have shown that UDG removes uracil in a stepwise reaction mechanism (Fig. 2) [38,39]. The first step is the cleavage of the glycosylic bond, where His268 and Asp145 stabilize the first transition state. In the next step, a water molecule bound to His148 and Pro146 attacks the carbon bound to uracil base and one of the hydrogens on the water molecule is transferred to the Asp145 (Fig. 2) [38]. Four negatively charged phosphate groups on the DNA repel the anionic leaving group and stabilize the positive charge on the sugar ring. These four phosphate groups stabilize the rate-determining transition state with more than 20 kcal/mol [38]

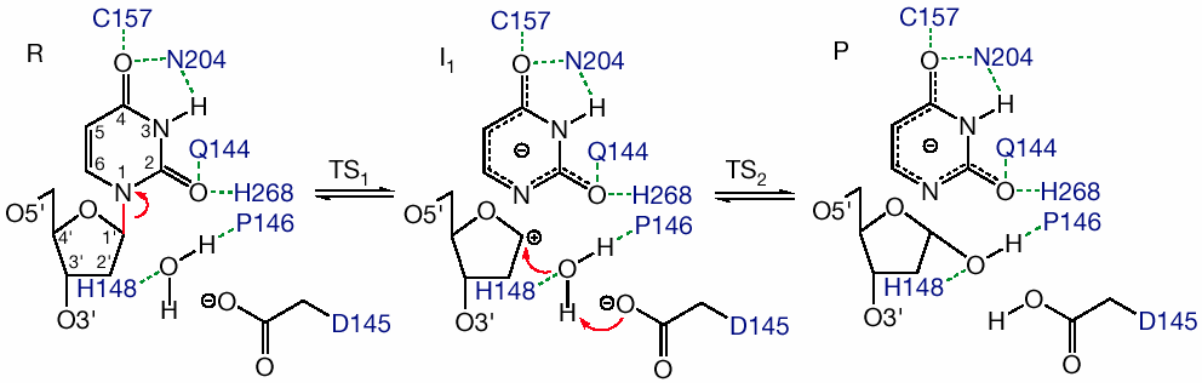


Figure 2
Catalytic mechanism for UDG. Figure adopted from Dinner *et al.* [38].

3.2.2 Comparison of mesophilic hUDG and psychrophilic cUDG

In order to improve our understanding of cold adaptation, cUDG and hUDG are used as a comparative model system. High sequence and structural similarity makes UDG a very good choice as a model system to explore adaptational features. The high amount of characterized mutants and crystal structures available for both cUDG and hUDG also make it an excellent system for theoretical studies. Even though the two enzymes have very similar three-dimensional structure, the cold-adapted enzyme is up to ten times more catalytically active in the temperature range from 288-310 K and is less heat stable compared to the warm-active representative [40]. Thus, cUDG has the typical features of cold adaptation. cUDG and hUDG have a sequence identity of 75%, an overall displacement of main chain atoms of 0.63 Å and similar secondary and tertiary structure [41]. The catalytic domain of both enzymes consists of 223 amino acids and the active site is situated in the C-terminal region of the enzyme. The tertiary structure of both enzymes consists of a β-sheet with 4 parallel strands and 11 helices (Fig. 3A). The side of the enzyme facing the DNA upon binding has a positive electrostatic potential for both enzymes (Fig. 3B). The loops: 4-Pro loop (^{165}PPPPS^{169}), the Gly-Ser loop (^{246}GS^{247}), the Leu272 loop (^{268}HPSPLSVYR^{276}) and the water-activating loop (^{145}DPYH^{148}) are in close contact with the DNA upon binding and are thought to be important for detection and catalysis [42]. The sequences in the loops are from hUDG, and cUDG has two mutations in the Leu272 loop. Residue 274 and 275 are Ala and His in cUDG, respectively.

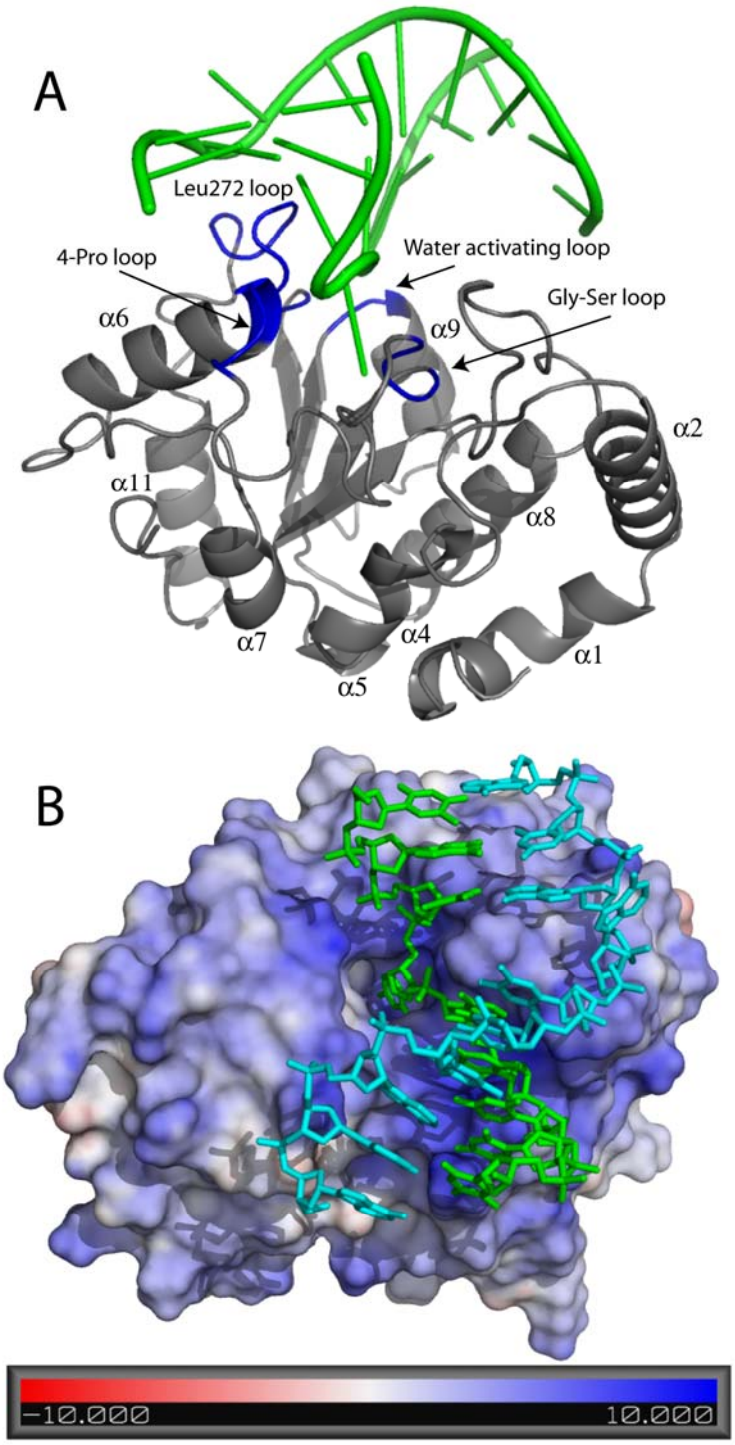


Figure 3
Two different representations of cUDG in complex with DNA. **A:** Ribbon style representation of cUDG bound to DNA. The four loops important for recognition and catalysis are shown in blue. The helices are also numbered. **B:** The electrostatic potential is projected to the surface of cUDG, and colored in kT/q. The electrostatic potential was calculated in DelPhi. The figure was generated using PyMol [43].

3.2.3 Structural adaptations to cold in psychrophilic cUDG

Psychrophilic cUDG has been thoroughly studied to find structural features which could explain the increase in catalytic efficiency and reduced thermal stability compared to its mesophilic homologue. A comparative study of the crystal structure of mesophilic hUDG and cold-adapted cUDG show that the mesophilic enzyme has 11 salt-bridges while the psychrophilic variant of the enzyme has only 5 salt-bridges [41]. The cold-adapted homologue has also three more hydrogen bonds compared to its mesophilic cousin [41]. There is a slight decrease in the size of the hydrophobic core residues in cUDG compared to hUDG, which leads to an increase in cavities in the core [40,41]. Some of the additional cavities are situated close to the active site, which could give rise to increased flexibility of the active site. A more flexible DNA recognition loop (Leu272 loop) has also been proposed to increase the catalytic efficiency and reduce the thermal stability of cUDG compared to hUDG [41].

Theoretical methods

Today computational methods play a central role in many applications in the study of biophysical properties of macromolecules. Since only a limited number of properties of biomolecular systems are actually accessible to measurement by experimental techniques, computer simulations can complement experiments by providing time series, distributions or average values of any definable quantity [44]. For example, the study of folding/unfolding pathways, conformational distributions and interactions between parts of systems are properties that computational methods are well suited for.

4.1 Molecular mechanics

Molecular mechanics (MM) refer to computational approaches that adopt classical mechanics to analyze the structure and energetic of molecular systems [45]. MM ignores the electronic motions and the energy of the system is calculated as a function of the nuclear positions only. The atoms are treated as a set of soft spheres with point charges. These energy functions along with the parameters make up a force field. Today several force fields are available for the study of biological systems, for example: AMBER [46], GROMOS [47], CHARMM [48] and OPLS-AA [49]. Equation VIII shows a typical form of a potential energy function (see also Fig. 4):

$$(VIII) \qquad V\left(r^N\right) = \sum_{bonds} \tfrac{1}{2} k_b \left(b_i - b_{i,0}\right)^2 + \sum_{angles} \tfrac{1}{2} k_\theta \left(\theta_i - \theta_{i,0}\right)^2 + \sum_{torsions} \tfrac{1}{2} V_n \left(1 + \cos(n\phi - \delta)\right)$$
$$+ \sum_{\substack{inproper \\ torsions}} \tfrac{1}{2} k_\zeta \left(\zeta_i - \zeta_{i,0}\right)^2 + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right)$$

$V(\boldsymbol{r}^N)$ is the potential energy as a function of the positions ($\boldsymbol{r}$) of all N atoms. In equation VIII bond lengths, angles and improper torsions are evaluated in terms of deviations from their equilibrium values ($b_{i,0}$, $\theta_{i,0}$ *and* $\zeta_{i,0}$), while torsions are evaluated from the minimal value ($cos(n\phi\text{-}\delta)$). The last term in equation VIII is the non-bonded potential, which describes the energy between atoms separated by more than three bonds or atoms in separate molecules. The van der Waals contribution is typically evaluated with the 6-12

Lennard-Jones potential, while the electrostatic energy is obtained by applying Coulombs law. However, more sophisticated force fields contain additional terms than those mentioned above.

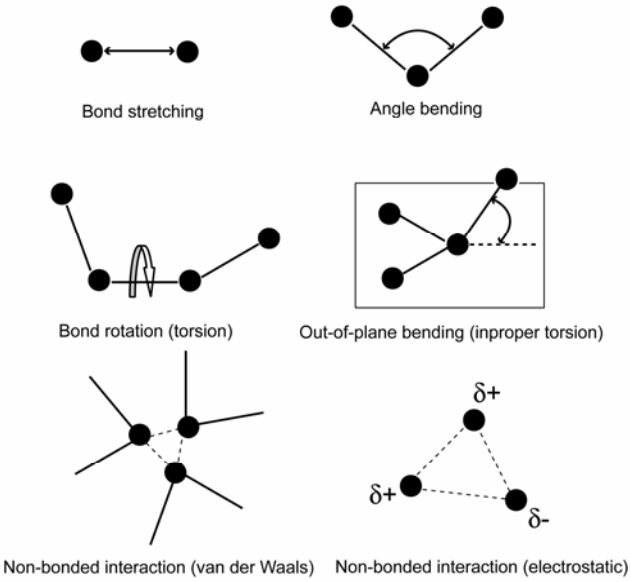


Figure 4
Schematic representations of the main contributors to a typical potential energy function in force fields.

4.2 Molecular dynamics

Molecular dynamics (MD) simulations generate a trajectory or a set of conformations that describes how the system varies over time [45] by integrating Newton`s laws of motions. The resulting trajectory contains information about positions and velocities of all atoms and how they change over time. Prior to a MD simulation the initial velocities and positions of all atoms are needed. The velocities are normally obtained from a random Maxwell distribution, while the initial positions are often taken from X-ray or NMR structures, but could also be computed from a homology model at atomic resolution. In a force field the motions of all particles are coupled, giving rise to a many-body problem that cannot be solved analytically. But by applying the finite difference method, the problem can be solved by breaking the integration into many small steps, all separated in time by a fixed time step (δt). The potential energy function with respect to the atomic

positions give the force acting on the atoms and this force is used to calculate the accelerations of all particles. The accelerations of the particles combined with positions and velocities are used to calculate the new positions and velocities at time (t + δt). The force is assumed to be constant during the time step. In order to keep the temperature stable during the MD simulation, the system is normally coupled to an external bath [50].

In a MD simulation a combination of 3N positions and 3N momenta defines the points in the 6N-dimential phase space, where N is the number of atoms [45]. If the sampling of the phase space is sufficient, the trajectory from the MD simulation can give good estimates of thermodynamic quantities and dynamic properties. MD simulations are often used to sample the conformational space prior to calculations of binding energies [51-54], pK_a values [55,56] and strengths of salt-bridges [57]. In addition, MD simulations are an important tool in the study of the dynamic nature of biological macromolecules and folding/unfolding of proteins [58-60]. They are also crucial in the structure refinement of macromolecules using NMR or X-ray crystallography.

4.3 Free energy calculation methods

The ability to calculate the strength of non-covalent interaction has been an important objective in computational chemistry. There are several methods that are currently used to predict the strength of binding energies. The methods are ranging from computationally expensive methods like free energy perturbation (FEP) and thermodynamic integration (TI) (see e.g. Brandsdal *et al.* [61] for a review of the methods) to various empirical or knowledge-based scoring approaches [62-64] where sampling of the conformational space is neglected. The linear interaction energy (LIE) method [65] and the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method [66,67] are commonly used for calculation of association energies. These approaches are based on analysis of the molecular dynamics or Monte Carlo trajectories.

4.4 MM-PBSA method

The MM-PBSA method [66-68] was initially used to study the stability of various DNA and RNA fragments, but the method has in later years also been applied in the calculation of binding free energies of proteins and small ligand [53,54], protein-protein [69-71] and protein-DNA complexes [72]. This method estimates the free energy of every single conformation according to:

$$(IX) \qquad G = H_{MM} + G_{sol} - TS_{MM}$$

where H_{MM} is the molecular mechanical energy and can be divided into several energy terms:

$$(X) \qquad \mathrm{H}_{MM} = E_{bond} + E_{angle} + E_{torsion} + E_{elec} + E_{vdW}$$

where E is the bond, angle, torsion, electrostatic and the van der Waals term in the molecular mechanical force field. The term $-TS_{MM}$ is the solute entropy and is usually calculated with normal mode analysis. The G_{sol} is the solvation free energy and can be divided into two terms:

$$(XI) \qquad G_{\mathrm{sol}} = G_{pol} + G_{np}$$

where G_{pol} is the electrostatic contributions to the solvation free energy and can be obtained by either solving the Poisson-Boltzmann (PB) equation [73] or by the generalized-Born (GB) method [74,75]. The non-polar solvation free energy, G_{np}, is determined with a solvent-accessible-surface-area (SASA) dependent term [76]. The binding free energy is finally calculated according to:

$$(XII) \qquad \Delta\mathrm{G} = \left\langle \Delta\mathrm{G} \right\rangle_{\mathrm{complex}} - \left\langle \Delta\mathrm{G} \right\rangle_{\mathrm{receptor}} - \left\langle \Delta\mathrm{G} \right\rangle_{\mathrm{ligand}}$$

where $\left\langle \Delta\mathrm{G} \right\rangle_{\mathrm{complex}}$, $\left\langle \Delta\mathrm{G} \right\rangle_{\mathrm{receptor}}$ and $\left\langle \Delta\mathrm{G} \right\rangle_{\mathrm{ligand}}$ are the free energies of the complex, the protein and the ligand, respectively, averaged over a set of snapshots extracted from a MD simulation. The energy contributions to the binding energy can either be extracted from a single trajectory simulation of the complex or it is possible to use separate trajectories for the complex, receptor and the ligand. A drawback with the separate trajectory method is that it is very difficult to get the H_{MM} energy to converge for large molecules within reasonable computing time [61]. The single trajectory method assumes

that there are no changes in the structure of the receptor and ligand upon binding, which is not always the case. Studies using linear interaction energy (LIE) calculations have shown that the ligand, if not very small and rigid, often adopt different conformations when free in solution compared to the conformations in the complex [61]. But in contrast to the LIE method, the MM-PBSA method does not require optimization of any parameters for the calculation of the total energy. In addition, the MM-PBSA method estimates the entropy of the solute.

4.5 Continuum electrostatics

Explicit solvation energies of large macromolecules are very time consuming to calculate with computational methods. An alternative approach is to apply continuum electrostatics, which describe the solvent properties as average values [73]. The advantages of continuum electrostatic models are their low computational cost and the possibility of visualization of the electrostatic potential. In these methods the solvent is represented as a high dielectric medium, while the solute is described as a low dielectric medium, and the molecular surface area is often used as the boundary (Fig. 5). The classical treatment of electrostatics in solution is based on the Poisson-Boltzmann (PB) equation:

$$(XIII) \qquad \nabla\cdot\left[\varepsilon(r)\nabla\cdot\phi(r)\right]-\varepsilon(r)\kappa(r)^2\sinh\left[\phi(r)\right]+\frac{4\pi\rho^f(r)}{kT}=0$$

where ε is the dielectric constant, ϕ(r) is the electrostatic potential in units of kT/q, k is the Boltzmann constant, T is the temperature, q is the charge of a proton, r is the position vector and ρ^f is the fixed charge density (in proton charge units). The term $\kappa^2 = 1/\lambda^2 = 8\pi q^2 I/ekT$, where λ is the Debye length and I is the ionic strength of the bulk solution. The salt effect is described by the second term in equation XIII, and if there are no mobile ions in the system, this term is absent. With no mobile ions present and with a uniform dielectric constant for the entire system, the PB equation is reduced to the Coulomb's law.

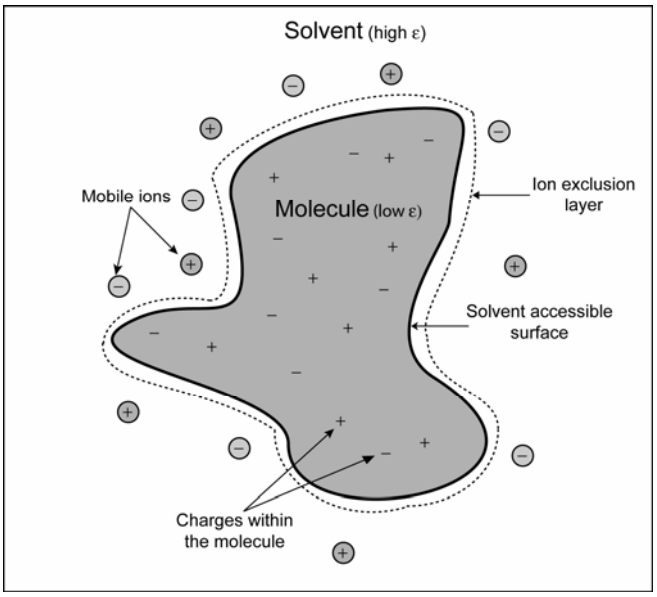


Figure 5
A molecule in a heterogeneous dielectric medium.

The PB equation is solved by numerical methods like finite difference methods, and today several programs have implemented an algorithm to solve the PB equation. Examples of programs that are widely used for this purpose are DelPhi [77] and UHBD [78]. Even though the continuum approaches have provided a wealth of information in different fields, they also have some limitations. Continuum models do not tend to work well when short range effects become important, for example when individual waters molecules bridging functional groups [79]. Such models are also very dependent on the dielectric constant, and it can be very challenging to choose the optimal dielectric constant of the solute [79,80]. The protein dielectric constant (ε_p) is not an universal constant, but is a parameter which depend on the model used [81]. ε_p represents the electrostatic interactions which are not represented explicitly in the model. Such factors can be the effect of water penetration into the protein core or fluctuations of polar groups in the protein [80-82]. Because proteins are structurally heterogeneous and possess flexible areas, the ε_p can also have different optimal values in different regions of the protein [80]. All the above mentioned factors make it difficult to choose the optimal protein dielectric constant. Early calculations used a low dielectric constant for the protein, typically in the range 1-4, while more recent studies indicate that a larger constant is needed to reproduce experimental shifts in pK_a values [56,81].

Aims of study

The main focus of this thesis is to explore the structural features responsible for cold adaptation using uracil DNA glycosylase (UDG) as a model system. Since UDG is a DNA binding protein, general aspects of protein-DNA recognition will also be studied. These objectives will mainly be studied with computational methods.

Subgoal 1:

Is improved flexibility responsible for increased catalytic efficiency of cold-adapted UDG?

In general, cold-adapted enzymes are believed to possess higher flexibility in structural components involved in the catalytic cycle compared to their warm-active homologues. In UDG, increased flexibility in the DNA recognition loop is thought to explain the high catalytic efficiency of cold-adapted cod UDG (cUDG). Molecular flexibility is very difficult to study by experimental methods, thus MD simulations will be used to study differences in flexibility between the cold-adapted cUDG and the warm-active human UDG (hUDG).

Subgoal 2:

Which forces are important for thermal stability of UDG?

The general view is that cold-adapted enzymes are less stable than their warm-active homologues. Experimental results have also shown that psychrophilic cUDG has lower kinetically derived stability than mesophilic hUDG. Because of the reduced stability of cUDG, our hypothesis is that the psychrophilic enzyme will unfold at lower temperatures or have higher unfolding rate than the mesophilic homologue. High temperature MD simulations will be used to analyze the unfolding pathway and find important molecular contacts for stability of cUDG and hUDG.

Comparative studies of cold- and warm-active enzymes have shown that warm-active enzymes often have increased number of salt-bridges compared to cold-adapted homologues. Analysis of the crystal structure of cUDG and hUDG showed that the

mesophilic enzyme has increased number of salt-bridges. Salt-bridges are generally thought to have favorable effect on the protein stability. Continuum electrostatics will be used to calculate the strength of ion pairs in both cold- and warm-active UDG to investigate if the warm-active hUDG are more stabilized by salt-bridges than the cold-adapted homologue.

Subgoal 3:

Is the binding to DNA different for cold- and warm-active UDG?

Kinetic studies have shown that cold-adapted cUDG has lower K_m values compared to the mesophilic counterpart. If the enzyme follows the simple Michaelis-Menten mechanism, K_m will be lowered by a more favorable substrate binding. So far, the binding energy between cold-adapted UDG and DNA has not been studied. In this study MD simulations and free energy calculations methods will be used to calculate the strength of the binding for hUDG and cUDG in complex with double stranded DNA.

Results and discussion

UDG is an excellent model system for the study of enzymatic adaptation to low temperature. Its biological function is well-known, and several crystal structures of both cUDG and hUDG are available, including mutant and native structures, which have been characterized in terms of stability and kinetics [40,41,83-85]. Even though cUDG and hUDG have highly similar secondary structure and 75% sequence identity, cUDG has been shown to be up to 10 times more catalytically efficient (k_{cat}/K_m) in the temperature range from 288-310 K [40]. cUDG has also been found to be much more pH and temperature labile than hUDG [83], indicating that cUDG is generally less stable compared to hUDG. In this study the psychrophilic cUDG and the mesophilic hUDG enzymes have been explored with various methods in order to describe how UDG adapt to cold environments at a molecular level.

Life in the cold has to cope with reduction in chemical reaction rates and lower fluid viscosity. In order to maintain sufficient metabolic fluxes at low temperatures, the psychrophilic organisms produce "cold-adapted" enzymes, which are able to maintain high catalytic efficiency even at low temperature [86]. The relationship between activity-flexibility-stability is probably the leading hypothesis to explain enzyme properties in temperature adaptation [2,5,7]. Increased molecular flexibility will lead to higher activity and less stable enzymes, and on the other hand, a very stable enzyme will be too rigid to perform its catalytic function at high rate, and yielding a less efficient enzyme. Flexibility can be considered either as a static or a dynamic property [87,88]. Static flexibility refers to the number and structural diversity of the different conformers in the equilibrium ensemble. Dynamic flexibility is how quickly the structure can interconvert between the conformers and is a measure of the energy barriers between the conformers in the equilibrium ensemble [87,88]. Static flexibility can be gained from B-factors from crystal structures, from hydrogen/deuterium (H/D) exchange experiments or from conformers from MD simulations. Dynamic flexibility can be measured by dynamic fluorescence quenching or proteolytic nicking, for more information on these methods see e.g. *Siddiqui and Cavicchioli* [7].

5.1 Molecular flexibility

Enzymes are not static structures, but possess a certain dynamic property. Flexibility of regions which are directly involved in the enzyme catalysis is believed to be essential for enzyme activity [89]. Thermophilic enzymes have decreased catalytic efficiency at low temperatures, which is thought to be caused by a too rigid structure at low temperatures [90-92]. The dynamic properties of protein structures decrease with decreasing temperature and cold-adapted enzymes are thought to have improved flexibility in structural components involved in the catalytic cycle in order to maintain high catalytic activity at low temperatures [2]. The increased flexibility would increase the specific activity (k_{cat}), as the increased flexibility will enable good complementary with the substrate [2]. Psychrophilic cUDG has higher specific activity compared to mesophilic hUDG (paper I). Previous studies have postulated that the increased catalytic efficiency of cUDG could be explained by increased flexibility of the DNA recognition loop also called the Leu272 loop [41]. MD simulations has been used with success to reveal flexibility in engineered mesophilic subtilisin [93]. The engineered subtilisin showed typical cold-adapted features and also increased local and global flexibility.

MD simulations on cold- and warm-active UDG and on mutants of the enzymes were performed to get insight into the flexibility of the two homologues enzymes (paper II). The simulations were run for both the UDG-DNA complex and for the uncomplexed enzyme solvated by water. The MD simulations show that the cold-adapted cUDG has higher overall flexibility per residue compared to the mesophilic homologue. This indicates that the structure of the mesophilic enzyme is more rigid than the psychrophilic counterpart. The plot of root-mean-squared fluctuations (r.m.s.f.) per residue shows that especially in the DNA recognition loop, there are large differences in flexibility between the two enzymes (paper II). The psychrophilic enzyme has much higher flexibility in this loop compared to the mesophilic counterpart. The finding fits well with the emerged picture that cold-adapted enzymes have higher flexibility in the active site cleft or in loop structures around the active site in order to maintain sufficient activity at low temperatures [7,93].

The results from the MD simulations of all six different UDG variants show that there seem to be a correlation between catalytic efficiency and flexibility of this loop (paper II). This is especially evident for the cUDG-V171E mutant, which has the lowest flexibility in the DNA recognition loop and lowest catalytic efficiency of the variants in this study. It is interesting to note that the mutation of the 171 residue does not have any effect on the flexibility on nearby residues in the sequence even though it has large effect on the flexibility of the DNA recognition loop. Introducing the negatively charged Glu171 reduces the positive electrostatic potential in the DNA recognition loop area. The Leu272 loop consists of neutral residues and one positively charged Arg residue, and for these residues it would probably be destabilizing to be in a positive electrostatic potential. Reducing the positive electrostatic potential near this loop will maybe stabilize the loop and make it less flexible. Even if the largest differences in flexibility between cUDG and hUDG are seen in the Leu272 loop, there is also increased flexibility of the Gly-Ser loop (residue 246-247) in the psychrophilic enzyme. The Ser in the Gly-Ser loop forms a hydrogen bond to the DNA in the enzyme-DNA complex and this loop is probably involved in orienting the enzyme prior to the DNA scan [42].

5.1.1 Structural features responsible for increased Leu272 loop flexibility in cUDG

If we compare the amino acid sequence of the Leu272 loop of cUDG and hUDG, Val274 and Tyr275 in hUDG are mutated to Ala274 and His275 in cUDG. In addition, Phe279 in hUDG is mutated to Leu in the cUDG structure. Residues Val274, Tyr275 and Phe279 in the mesophilic enzyme are believed to form a hydrophobic cluster which could restrict the motion of the Leu272 loop [41]. In the psychrophilic enzyme the three residues are mutated to smaller amino acids that are not able to form a hydrophobic cluster. The cUDG-H275Y mutant has much lower flexibility in the Leu272 loop compared to cUDG, but similar flexibility to the human variant of the enzyme (paper II). This indicates that His275 in cUDG is the main contributor to the increased flexibility of the Leu272 loop, relative to the human enzyme.

5.2 Thermal stability

Proteins are usually only marginally stable at their physiologically relevant temperature, and their free energy of unfolding typically varies between 5 to 15 kcal/mol [1]. Large opposing energies are involved in the stabilization of proteins, and all contributions, both favorable and unfavorable, are therefore important when considering protein stability. How proteins achieve their stability do not seem to follow any general rules [94], but the emerging picture is, for globular proteins, that the hydrophobic effect and burial of nonpolar side chains stabilizes the native state [95,96]. Disulfide bonds, electrostatic interactions and hydrogen bonds are, however, also important for structural stability and contribute favorably to protein stability [97-99].

Kinetically derived stability of cUDG and hUDG has been measured as half-life times (the time it takes for the enzyme to lose 50% of its activity). The psychrophilic enzyme has shorter half-life for all temperatures examined [40,83], and has also lower temperature optimum (the temperature with highest relative activity) than its mesophilic counterpart. The temperature optimum for cUDG and hUDG is 314 K and 318 K, respectively [83,100]. These results show that the psychrophilic variant loses activity at lower temperature than its mesophilic homologue. Even though this is not a direct measure of thermodynamic stability, it indicates that mesophilic hUDG might be more thermostable than the psychrophilic variant of the enzyme.

In order to investigate the stability of UDG, high temperature MD simulations of cUDG and hUDG were performed, to investigate the unfolding process of these enzymes (paper III). The MD simulations were carried out at three different temperatures: 375 K, 400 K and 425 K. The root-mean-squared deviations (r.m.s.d.) were used to measure unfolding during the MD simulations. For the two lowest temperatures, the psychrophilic enzyme unfolds more rapidly than the mesophilic variant, indicating reduced structural stability. If the time taken to reach r.m.s.d. of 15 Å, is used as a measure of unfolding rate, there are large differences in the unfolding rate at 400 K. The r.m.s.d. of the psychrophilic enzyme reaches a value of 15 Å after ~0.5 ns at 400 K while the mesophilic variant needs 4.2 ns to reach the same degree of unfolding at the same temperature (paper III). At 425

K the psychrophilic enzyme has the same high unfolding rate as observed for the 400 K simulations, while at this temperature hUDG has also reached high unfolding rate, similar to the rate of cUDG. This indicates that the cold-adapted cUDG needs lower temperature to reach a fast unfolding rate compared to the warm-active hUDG. However, once the thermal energy of the systems is sufficiently high (i.e. 425 K in this case), rapid unfolding is observed for both enzymes.

Salt-bridges are also expected to be important for stability of proteins [99], and it is believed that the warm-active enzymes are more stabilized by salt-bridges than the cold-adapted homologues [7]. The crystal structure of the psychrophilic cUDG (pdb code 1OKB) has a reduced number of salt-bridges compared to the mesophilic hUDG (pdb code 1AKZ) [41]. In our study we have analyzed 5000 structures from MD simulations to identify putative salt-bridges in cUDG and hUDG (paper IV). The results from this study show that the two enzymes have similar amount of salt-bridges, cUDG and hUDG have 11 and 12 salt-bridges, respectively. Continuum electrostatics was used to calculate the electrostatic contributions to the stability of each salt-bridge (paper IV). Virtually all ion-pairs present in both cUDG and hUDG have a favorable electrostatic contribution, which probably lead to increased structural stability. This is in accordance with other studies of salt-bridges in other proteins [99,101]. Without taking the physiological temperature of the organisms into considerations, we find comparable electrostatic stability of ion-pairs in cUDG and hUDG. However, if we look at their respective environmental temperatures, the ion-pairs are more stabilizing in hUDG when compared to cUDG (paper IV). When it comes to their net contribution to protein stability, entropy must also be considered. Salt-bridge formations will reduce the available conformational space and the conformational entropy of the folded state will consequently decrease. Hence, ion-pairs destabilize the native state from an entropic point of view, whereas they are enthalpically stabilizing. The entropic effect is intrinsically difficult to estimate through computer simulations, and the fact that it changes with temperature makes it even more challenging to calculate.

5.2.1 Structural features responsible for stability in UDG

Comparative analysis of the ion pairs and unfolding of cold- and warm-active UDG have shown that certain molecular contacts seem to explain the differences in stability between the two enzymes. One of the first events in the unfolding process of cUDG is the melting of the N-terminal, while in hUDG both terminals unfold as an early step in the unfolding pathway. Thus, molecular contacts which stabilize the terminals are probably very important for stability of UDG. In the N-terminal there are especially three hydrogen bonds which are thought to be important for the stability of both enzymes. When the hydrogen bonds: Ser88:Oγ-Asp133:Oδ2, Trp89:Nε1-Cys132:O and Trp89:Nε1-Thr129:O are lost, the N-terminal of both enzymes rapidly unfolds (paper III). These hydrogen bonds connect the N-terminal of helix α1 and the helix α4 and the loop between helix α4 and helix α5 (for helix numbering see Fig 3A). In addition, the side chain of Trp89 packs into a hydrophobic area between helices α4 and α8. Interestingly, the crystal structure of a thermophilic uracil DNA glycosylase from *Thermus thermophilus* HB8 (TthUDG) has a stabilizing [4Fe-4S] cluster close to the N-terminal [102]. TthUDG belongs to the family-4 uracil DNA glycosylases while hUDG and cUDG are from family-1. Even though the amino acid sequence homology is low between TthUDG and hUDG or cUDG, the topology and order of the secondary structure elements are similar between the two uracil DNA glycosylase families [102]. Temperature adaptation of UDG seems to involve additional stabilization of the N-terminal part.

The high temperature MD simulations show that there are large differences in stability in the C-terminal between the two enzymes (paper III). hUDG unfolds in the C-terminal at the two lowest temperatures, while at the highest temperature both enzymes unfold. The C-terminal of both enzymes is situated at the surface, and the side facing away from the solvent forms a hydrophobic cluster with helix α7 and α11. In addition, the Trp301 side chain forms stacking interactions with the side chains of Phe284 and Trp195. When the hydrophobic packing/stacking interactions are lost, the C-terminal starts to melt. The hydrogen bond between Gln198 and residue 300 (Asp in hUDG and Asn in cUDG) might also be responsible for stabilizing the C-terminal. But even if the interactions between Gln198 and Asn300 are lost, the C-terminal seems to resist unfolding. Structural analysis

shows that the ionic interaction between Lys138 and residue 297 is important for stability of the C-terminal. cUDG has a Glu in this position and the distance between the atoms Lys138:Nζ and Glu297:Oε1 is 5.4 Å. Thus, this ionic interaction will stabilize the C-terminal of cUDG. hUDG, on the other hand, has a Lys in position 297 and the positive charges will repel each other and destabilize the C-terminal of hUDG (paper III). To verify the importance of this ionic contact, the Glu297 in cUDG was mutated to Lys and the Lys297 in hUDG was mutated to Glu. MD simulations were performed on these two mutants as well. The cUDG-E297K did unfold in the C-terminal, and this shows that the Lys138-Glu297 ionic interaction is important for C-terminal stability. The C-terminal of the hUDG-K297E mutant did also unfold, indicating that there are also other molecular features than the former mentioned ionic contact that is important for stabilizing the C-terminal. From residue 293 to the C-terminal end, there are four residues in the sequence that make cUDG more hydrophobic than hUDG. The substitutions are K293L, K296T, D300N and E303A (mutation from hUDG to cUDG), all charged residues in hUDG, substituted to hydrophobic or uncharged in cUDG (for a full alignment see Leiros *et al.* [41]). It will be more unfavorable for cUDG to unfold in the C-terminal and expose its more hydrophobic residues to water compared to hUDG which has more hydrophilic residues close to the C-terminal. This could also explain why cUDG have a more stable C-terminal compared to hUDG.

In the comparative salt-bridge study, we identified 12 and 11 salt-bridges in hUDG and cUDG, respectively (paper IV). The strength of the different salt-bridges varies for both enzymes. In the mesophilic enzyme we found 3 very strong salt-bridges at its physiological temperature (ΔG_{tot} > 2.0 kcal/mol), Asp180-Arg282, Asp183-Lys302 and Asp300-Lys302. In the psychrophilic enzyme only one strong salt-bridge was found, Asp180-Lys282. But the mesophilic enzyme has also the three weakest salt-bridges (paper IV). The Asp180-Lys/Arg282 salt-bridge is probably especially important for UDG stability. In the unfolding study of the warm-active UDG this specific salt-bridge seems to be responsible for keeping the 165-190 region folded (paper III).

The warm- and the cold-active UDG have 5 and 3 global salt-bridges, respectively. The global salt-bridges will connect different part of the structure and this probably leads to a more rigid overall structure. Local salt-bridges, on the other hand, can still be intact even if the protein is highly unfolded. Thus, global salt-bridges are most likely more important for protein stability compared to local ones. The warm-active UDG has also a strong ionic network (Asp183, Lys302 and Asp300), which is not observed in the cold-adapted UDG. These factors could probably stabilize the mesophilic UDG more than its cold-adapted homologue.

5.3 UDG-DNA binding and electrostatics

The ability of proteins to interact with other macromolecules in a highly specific manner is an important feature for a variety of biological processes, including DNA repair, antigen-antibody interactions, signal transduction, enzymatic catalysis, drug design among others. It has been proposed that cUDG has increased substrate affinity compared to hUDG due to enhanced positive electrostatic potential at surface areas central to formation of the enzyme-substrate complex [41]. Continuum electrostatics was applied to investigate the difference in electrostatics between cold- and warm-active UDG. Both cUDG and hUDG has a highly positive electrostatic potential in the specificity pocket and in nearby areas that are know to interact directly with DNA. As indicated (paper I), there are differences in the electrostatic surface potential between cUDG and hUDG, as the psychrophilic enzyme has a more positive electrostatic potential near the active site. The 171 residue seems to be a key residue for explaining difference in electrostatic potential (paper V).

The binding studies of cUDG and hUDG with the MM-PBSA method show that there are large energies involved in the binding between UDG and DNA. Even if the standard deviations are very high, especially for the separate trajectory, the energies calculated with both single and the separate trajectories indicate that the psychrophilic enzyme associates stronger to DNA than the mesophilic variant (paper V). All the four loops which are important for detection and catalysis in UDG have favorable interactions with

the DNA, but the DNA recognition loop seems to be especially important for binding, and is responsible for 34.1 % and 44.2 % of the enthalpic contribution to the binding energy in the warm- and cold-active UDG, respectively. Residues 275 and 276 in particular, have a much stronger binding energy per residue in cUDG. Residue 276 is an Arg in both enzymes, and 275 is a Tyr in hUDG and the simulations show that this residue bends away from the DNA in the complex, while cUDG has a His in this position which hydrogen bonds to the DNA.

Concluding remarks

Several structural features important for cold-adaptation have been pinpointed in this study. This study shows that electrostatic surface properties near the active site and flexibility of the DNA recognition loop seem to be important for adaptation to cold for psychrophilic cUDG. A link between high positive electrostatic potential at and around the active site and flexibility of the DNA recognition loop is proposed, but further studies are needed to fully investigate this hypothesis. The UDG-DNA binding study identified several residues which are important for binding and especially the DNA recognition loop form strong interactions with the DNA.

Summary of the thesis

In this study uracil DNA glycosylase (UDG) has been used as a model system for the study of cold adaptation and protein-DNA recognition with computational methods. Cold-adapted, or so-called psychrophilic, enzymes are attractive as targets for commercialization due to their reduced thermal stability which is usually accompanied with an increased catalytic efficiency. UDG removes uracil from the DNA strand as the first step in the base excision repair pathway (DNA repair system). A deficient DNA repair system is associated with serious diseases, such as development of cancer in humans. Uracilation of DNA represents a constant threat to the survival of many organisms, and since UDG is the most efficient of all the enzymes in the UDG superfamily and found in all free living organisms its a vital enzyme.

MD simulations, continuum electrostatics, X-ray crystallography and free energy calculations have been used to study the structural and energetic differences between warm-active human UDG and cold-active cod UDG at atomic level. Analyses of the MD simulations show that the psychrophilic UDG has a highly flexible DNA recognition loop compared to its warm-active homologue. This is thought to explain the observed high catalytic efficiency for cold-adapted UDG. This is in accordance with the emerged picture that cold-adapted enzymes have higher flexibility of components involved in the catalytic cycle. Analyses of several mutants of cod and human UDG indicate that there is a correlation between catalytic efficiency and flexibility in this DNA recognition loop.

Continuum electrostatics calculations has been applied to analyze all ionic contacts in the warm- and cold-active UDG. The electrostatic contribution of the ion-pairs is slightly more favorable in cod UDG at 298 K. This is primarily attributed to more optimized interactions between the ion-pairs and nearby dipoles/charges in cod UDG. When we take the environmental temperatures into account, the electrostatic stability becomes more favorable for the ion-pairs in the mesophilic enzyme. Comparative studies of the electrostatic potential of cod and human UDG show that both enzymes have a positive electrostatic potential near the active site, but the potential is even higher for the cold-

adapted enzyme. This difference in electrostatic potential could probably explain the different K_m values for these two enzymes.

Comparative high temperature MD simulations were used to study the unfolding and structural stability of cod and human UDG. The simulations showed that there are distinct structural differences in the unfolding pathway between the warm- and cold-active UDG, particularly evident in the N- and C-terminals. The results from these MD simulations also showed that at certain temperatures the psychrophilic enzyme has a higher unfolding rate compared to its mesophilic homologue. The MM-PBSA method was used to analyze the binding energy of the UDG-DNA complex. Cod UDG possesses a slightly more favorable DNA binding energy compared to human UDG. Decomposition of the binding energy per residue made it possible to pin-point residues that were important to the binding energy. The DNA recognition loop is responsibly for 34.1 % and 44.2 % of the enthalpic contribution to the binding energy in the warm- and cold- active UDG, respectively. Thus, this loop is very important for UDG-DNA association, and especially important for cod UDG.

References

[1] Fersht, A. Structure and mechanism in protein science. W.H. Freeman and Company, 1999.

[2] Georlette, D., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Hoyoux, A., Marx, J. C., Sonan, G., Feller, G., Gerday, C. Some like it cold: Biocatalysis at low temperatures. FEMS Microbiol. Rev. 2004, 28, 25-42.

[3] Avis, J. M., Fersht, A. R. Use of binding-energy in catalysis - optimization of rate in a multistep reaction. Biochemistry. 1993, 32, 5321-5326.

[4] van den Burg, B. Extremophiles as a source for novel enzymes. Curr Opin Microbiol. 2003, 6, 213-218.

[5] Smalås, A. O., Leiros, H. K. S., Os, V., Willassen, N. P., 2000. Cold adapted enzymes. In. Biotechnol annu rev, 6, 1-57

[6] Lonhienne, T., Gerday, C., Feller, G. Psychrophilic enzymes: Revisiting the thermodynamic parameters of activation may explain local flexibility. Bba-Protein Struct M. 2000, 1543, 1-10.

[7] Siddiqui, K. S., Cavicchioli, R. Cold-adapted enzymes. Annu Rev Biochem. 2006, 75, 403-433.

[8] Feller, G. Molecular adaptations to cold in psychrophilic enzymes. Cell Mol Life Sci. 2003, 60, 648-662.

[9] Feller, G., Gerday, C. Psychrophilic enzymes: Molecular basis of cold adaptation. Cell. Mol. Life Sci. 1997, 53, 830-841.

[10] Miyazaki, K., Wintrode, P. L., Grayling, R. A., Rubingh, D. N., Arnold, F. H. Directed evolution study of temperature adaptation in a psychrophilic enzyme. J. Mol. Biol. 2000, 297, 1015-1026.

[11] Gianese, G., Bossa, F., Pascarella, S. Comparative structural analysis of psychrophilic and meso- and thermophilic enzymes. Proteins. 2002, 47, 236-249.

[12] Riise, E. K., Lorentzen, M. S., Helland, R., Smalas, A. O., Leiros, H. K. S., Willassen, N. P. The first structure of a cold-active catalase from vibrio salmonicida at 1.96 angstrom reveals structural aspects of cold adaptation. Acta Crystallogr. D. 2007, 63, 135-148.

[13] Davail, S., Feller, G., Narinx, E., Gerday, C. Cold adaptation of proteins - purification, characterization, and sequence of the heat-labile subtilisin from the antarctic psychrophile bacillus ta41. J. Biol. Chem. 1994, 269, 17448-17453.

[14] Feller, G., Payan, F., Theys, F., Qian, M. X., Haser, R., Gerday, C. Stability and structural-analysis of alpha-amylase from the antarctic psychrophile alteromonas-haloplanctis-a23. Eur. J. Biochem. 1994, 222, 441-447.

[15] Kumar, S., Nussinov, R. Different roles of electrostatics in heat and in cold: Adaptation by citrate synthase. Chembiochem. 2004, 5, 280-290.

[16] Smalås, A. O., Heimstad, E. S., Hordvik, A., Willassen, N. P., Male, R. Cold adaption of enzymes - structural comparison between salmon and bovine trypsins. Proteins. 1994, 20, 149-166.

[17] Feller, G., Zekhnini, Z., LamotteBrasseur, J., Gerday, C. Enzymes from cold-adapted microorganisms - the class c beta-lactamase from the antarctic psychrophile psychrobacter immobilis a5. Eur. J. Biochem. 1997, 244, 186-191.

[18] Leiros, H. K. S., Willassen, N. P., Smalas, A. O. Structural comparison of psychrophilic and mesophilic trypsins - elucidating the molecular basis of cold-adaptation. Eur. J. Biochem. 2000, 267, 1039-1049.
[19] Violot, S., Aghajari, N., Czjzek, M., Feller, G., Sonan, G. K., Gouet, P., Gerday, C., Haser, R., Receveur-Brechot, V. Structure of a full length psychrophilic cellulase from pseudoalteromonas haloplanktis revealed by x-ray diffraction and small angle x-ray scattering. J. Mol. Biol. 2005, 348, 1211-1224.
[20] Aghajari, N., Van Petegem, F., Villeret, V., Chessa, J. P., Gerday, C., Haser, R., Van Beeumen, J. Crystal structures of a psychrophilic metalloprotease reveal new insights into catalysis by cold-adapted proteases. Proteins. 2003, 50, 636-647.
[21] Russell, R. J. M., Gerike, U., Danson, M. J., Hough, D. W., Taylor, G. L. Structural adaptations of the cold-active citrate synthase from an antarctic bacterium. Structure. 1998, 6, 351-361.
[22] Krokan, H. E., Standal, R., Slupphaug, G. DNA glycosylases in the base excision repair of DNA. Biochem J. 1997, 325, 1-16.
[23] Lindahl, T., Wood, R. D. Quality control by DNA repair. Science. 1999, 286, 1897-1905.
[24] Krokan, H. E., Kavli, B., Slupphaug, G. Novel aspects of macromolecular repair and relationship to human disease. J. Mol. Med. 2004, 82, 280-297.
[25] Priet, S., Gros, N., Navarro, J. M., Boretto, J., Canard, B., Querat, G., Sire, J. Hiv-1-associated uracil DNA glycosylase activity controls dutp misincorporation in viral DNA and is essential to the hiv-1 life cycle. Mol. Cell. 2005, 17, 479-490.
[26] Moe, E., Leiros, I., Smalas, A. O., McSweeney, S. The crystal structure of mismatch-specific uracil-DNA glycosylase (mug) from deinococcus radiodurans reveals a novel catalytic residue and broad substrate specificity. J. Biol. Chem. 2006, 281, 569-577.
[27] Pearl, L. H. Structure and function in the uracil-DNA glycosylase superfamily. Mutat. Res-DNA Rep. 2000, 460, 165-181.
[28] Sartori, A. A., Fitz-Gibbon, S., Yang, H. J., Miller, J. H., Jiricny, J. A novel uracil-DNA glycosylase with broad substrate specificity and an unusual active site. EMBO J. 2002, 21, 3182-3191.
[29] Aravind, L., Koonin, E. V. The alpha/beta fold uracil DNA glycosylases: A common origin with diverse fates. Genome Biol. 2000, 1, RESEARCH0007.
[30] Nilsen, H., Otterlei, M., Haug, T., Solum, K., Nagelhus, T. A., Skorpen, F., Krokan, H. E. Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the ung gene. Nucleic Acids Res. 1997, 25, 750-755.
[31] Handa, P., Acharya, N., Varshney, U. Chimeras between single-stranded DNA-binding proteins from escherichia coli and mycobacterium tuberculosis reveal that their c-terminal domains interact with uracil DNA glycosylases. J. Biol. Chem. 2001, 276, 16992-16997.
[32] Nagelhus, T. A., Haug, T., Singh, K. K., Keshav, K. F., Skorpen, F., Otterlei, M., Bharati, S., Lindmo, T., Benichou, S., Benarous, R., Krokan, H. E. A sequence in the n-terminal region of human uracil-DNA glycosylase with homology to xpa interacts with the c-terminal part of the 34-kda subunit of replication protein a. J. Biol. Chem. 1997, 272, 6561-6566.

[33] Otterlei, M., Warbrick, E., Nagelhus, T. A., Haug, T., Slupphaug, G., Akbari, M., Aas, P. A., Steinsbekk, K., Bakke, O., Krokan, H. E. Post-replicative base excision repair in replication foci. EMBO J. 1999, 18, 3834-3844.
[34] Krokan, H. E., Drabløs, F., Slupphaug, G. Uracil in DNA - occurrence, consequences and repair. Oncogene. 2002, 21, 8935-8948.
[35] Bennett, S. E., Sanderson, R. J., Mosbaugh, D. W. Processivity of escherichia-coli and rat-liver mitochondrial uracil-DNA glycosylase is affected by nacl concentration. Biochemistry. 1995, 34, 6109-6119.
[36] Higley, M., Lloyd, R. S. Processivity of uracil DNA glycosylase. Mutat. Res. 1993, 294, 109-116.
[37] Cao, C. Y., Jiang, Y. L., Stivers, J. T., Song, F. H. Dynamic opening of DNA during the enzymatic search for a damaged base. Nat. Struct. Mol. Biol. 2004, 11, 1230-1236.
[38] Dinner, A. R., Blackburn, G. M., Karplus, M. Uracil-DNA glycosylase acts by substrate autocatalysis. Nature. 2001, 413, 752-755.
[39] Werner, R. M., Stivers, J. T. Kinetic isotope effect studies of the reaction catalyzed by uracil DNA glycosylase: Evidence for an oxocarbenium ion-uracil anion intermediate. Biochemistry. 2000, 39, 14054-14064.
[40] Lanes, O., Leiros, I., Smalås, A. O., Willassen, N. P. Identification, cloning, and expression of uracil-DNA glycosylase from atlantic cod (gadus morhua): Characterization and homology modeling of the cold-active catalytic domain. Extremophiles. 2002, 6, 73-86.
[41] Leiros, I., Moe, E., Lanes, O., Smalås, A. O., Willassen, N. P. The structure of uracil-DNA glycosylase from atlantic cod (gadus morhua) reveals cold-adaptation features. Acta Crystallogr. D. 2003, 59, 1357-1365.
[42] Parikh, S. S., Mol, C. D., Slupphaug, G., Bharati, S., Krokan, H. E., Tainer, J. A. Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. EMBO J. 1998, 17, 5214-5226.
[43] DeLano, W. L., The pymol molecular graphics system, 2002, DeLano Scientific, San Carlos, CA, USA
[44] van Gunsteren, W. F., Bakowies, D., Baron, R., Chandrasekhar, I., Christen, M., Daura, X., Gee, P., Geerke, D. P., Glattli, A., Hunenberger, P. H., Kastenholz, M. A., Ostenbrink, C., Schenk, M., Trzesniak, D., van der Vegt, N. F. A., Yu, H. B. Biomolecular modeling: Goals, problems, perspectives. Angew. Chem. Int. Edit. 2006, 45, 4064-4092.
[45] Leach, A. R. Molecular modelling. second Ed., Person Education Limited, 2001.
[46] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., Kollman, P. A. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. J. Am. Chem. Soc. 1995, 117, 5179-5197.
[47] van Gunsteren, W., Berendsen, H. J. C., Molecular simulation (gromos) library manual, 1987, Groningen
[48] MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R.,

Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B. 1998, 102, 3586-3616.

[49] Weiner, S. J., Kollman, P. A., Nguyen, D. T., Case, D. A. An all atom force-field for simulations of proteins and nucleic-acids. J. Comput. Chem. 1986, 7, 230-252.

[50] Berendsen, H. J. C., Postma, J. P. M., Vangunsteren, W. F., Dinola, A., Haak, J. R. Molecular-dynamics with coupling to an external bath. J. Chem. Phys. 1984, 81, 3684-3690.

[51] Almlof, M., Brandsdal, B., Aqvist, J. Accurate predictions of protein-protein binding free energies. Abstr. Pap. Am. Chem. Soc. 2005, 229, U784-U784.

[52] Almlof, M., Brandsdal, B. O., Aqvist, J. Binding affinity prediction with different force fields: Examination of the linear interaction energy method. J. Comput. Chem. 2004, 25, 1242-1254.

[53] Brigo, A., Lee, K. W., Fogolari, F., Mustata, G. L., Briggs, J. M. Comparative molecular dynamics simulations of hiv-1 integrase and the t66i/m154i mutant: Binding modes and drug resistance to a diketo acid inhibitor. Proteins. 2005, 59, 723-741.

[54] Kuhn, B., Kollman, P. A. A ligand that is predicted to bind better to avidin than biotin: Insights from computational fluorine scanning. J. Am. Chem. Soc. 2000, 122, 3909-3916.

[55] Tarus, B., Straub, J. E., Thirumalai, D. Dynamics of asp23-lys28 salt-bridge formation in a beta(10-35) monomers. J. Am. Chem. Soc. 2006, 128, 16159-16168.

[56] van Vlijmen, H. W. T., Schaefer, M., Karplus, M. Improving the accuracy of protein pk(a) calculations: Conformational averaging versus the average structure. Proteins. 1998, 33, 145-158.

[57] Papaleo, E., Olufsen, M., De Gioia, L., Brandsdal, B. O. Optimization of electrostatics as a strategy for cold-adaptation: A case study of cold- and warm-active elastases. J. Mol. Graph. Model. 2006,

[58] Caflisch, A., Karplus, M. Acid and thermal-denaturation of barnase investigated by molecular-dynamics simulations. J. Mol. Biol. 1995, 252, 672-708.

[59] Li, A. J., Daggett, V. Characterization of the transition-state of protein unfolding by use of molecular-dynamics - chymotrypsin inhibitor-2. Proc. Natl. Acad. Sci. U. S. A. 1994, 91, 10430-10434.

[60] Li, A. J., Daggett, V. Molecular dynamics simulation of the unfolding of barnase: Characterization of the major intermediate. J. Mol. Biol. 1998, 275, 677-694.

[61] Brandsdal, B. O., Osterberg, F., Almlof, M., Feierberg, I., Luzhkov, V. B., Aqvist, J. Free energy calculations and ligand binding. Adv. Prot. Chem. 2003, 66, 123-158.

[62] Bohm, H. J. The development of a simple empirical scoring function to estimate the binding constant for a protein ligand complex of known 3-dimensional structure. J. Comput. Aid. Mol. Des. 1994, 8, 243-256.

[63] Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V., Mee, R. P. Empirical scoring functions.1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J. Comput. Aid. Mol. Des. 1997, 11, 425-445.

[64] Jain, A. N. Scoring noncovalent protein-ligand interactions: A continuous differentiable function tuned to compute binding affinities. J. Comput. Aid. Mol. Des. 1996, 10, 427-440.
[65] Åqvist, J., Medina, C., Samuelsson, J. E. New method for predicting binding-affinity in computer-aided drug design. Protein Eng. 1994, 7, 385-391.
[66] Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S. H., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., Cheatham, T. E. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. Acc. Chem. Res. 2000, 33, 889-897.
[67] Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A., Case, D. A. Continuum solvent studies of the stability of DNA, rna, and phosphoramidate - DNA helices. J. Am. Chem. Soc. 1998, 120, 9401-9409.
[68] Massova, I., Kollman, P. A. Computational alanine scanning to probe protein-protein interactions: A novel approach to evaluate binding free energies. J. Am. Chem. Soc. 1999, 121, 8133-8143.
[69] Adekoya, O. A., Willassen, N. P., Sylte, I. The protein-protein interactions between smpi and thermolysin studied by molecular dynamics and mm/pbsa calculations. J. Biomol. Struct. Dyn. 2005, 22, 521-531.
[70] Luo, C., Xu, L. F., Zheng, S. X., Luo, Z., Jiang, X. M., Shen, J. H., Jiang, H. L., Liu, X. F., Zhou, M. D. Computational analysis of molecular basis of 1: 1 interactions of nrg-1 beta wild-type and variants with erbb3 and erbb4. Proteins. 2005, 59, 742-756.
[71] Wang, W., Kollman, P. A. Free energy calculations on dimer stability of the hiv protease using molecular dynamics and a continuum solvent model. J. Mol. Biol. 2000, 303, 567-582.
[72] Zhang, Q., Schlick, T. Stereochemistry and position-dependent effects of carcinogens on tata/tbp binding. Biophys. J. 2006, 90, 1865-1877.
[73] Honig, B., Nicholls, A. Classical electrostatics in biology and chemistry. Science. 1995, 268, 1144-1149.
[74] Onufriev, A., Bashford, D., Case, D. A. Modification of the generalized born model suitable for macromolecules. J. Phys. Chem. B. 2000, 104, 3712-3720.
[75] Onufriev, A., Bashford, D., Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins. 2004, 55, 383-394.
[76] Sitkoff, D., Sharp, K. A., Honig, B. Accurate calculation of hydration free-energies using macroscopic solvent models. J. Phys. Chem. 1994, 98, 1978-1988.
[77] Nicholls, A., Honig, B. A rapid finite-difference algorithm, utilizing successive over-relaxation to solve the poisson-boltzmann equation. J. Comput. Chem. 1991, 12, 435-445.
[78] Davis, M. E., Mccammon, J. A. Solving the finite-difference linearized poisson-boltzmann equation - a comparison of relaxation and conjugate-gradient methods. J. Comput. Chem. 1989, 10, 386-391.
[79] Simonson, T. Macromolecular electrostatics: Continuum models and their growing pains. Curr. Opin. Struc. Biol. 2001, 11, 243-252.

[80] Demchuk, E., Wade, R. C. Improving the continuum dielectric approach to calculating pk(a)s of ionizable groups in proteins. J. Phys. Chem. 1996, 100, 17373-17387.
[81] Schutz, C. N., Warshel, A. What are the dielectric "constants" of proteins and how to validate electrostatic models? Proteins. 2001, 44, 400-417.
[82] Warshel, A., Sharma, P. K., Kato, M., Parson, W. W. Modeling electrostatic effects in proteins. Bba-Proteins Proteom. 2006, 1764, 1647-1676.
[83] Lanes, O., Guddal, P. H., Gjellesvik, D. R., Willassen, N. P. Purification and characterization of a cold-adapted uracil-DNA glycosylase from atlantic cod (gadus morhua). Comp. Biochem. Physiol. 2000, 127, 399-410.
[84] Leiros, I., Lanes, O., Sundheim, O., Helland, R., Smalås, A. O., Willassen, N. P. Crystallization and preliminary x-ray diffraction analysis of a cold-adapted uracil-DNA glycosylase from atlantic cod (gadus morhua). Acta Crystallogr. D. 2001, 57, 1706-1708.
[85] Mol, C. D., Arvai, A. S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H. E., Tainer, J. A. Crystal-structure and mutational analysis of human uracil-DNA glycosylase - structural basis for specificity and catalysis. Cell. 1995, 80, 869-878.
[86] Hoyoux, A., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Huston, A. L., Marx, J. C., Sonan, G., Zeng, Y. X., Feller, G., Gerday, C. Extreme catalysts from low-temperature environments. J. Biosci. Bioeng. 2004, 98, 317-330.
[87] Tang, K. E. S., Dill, K. A. Native protein fluctuations: The conformational-motion temperature and the inverse correlation of protein flexibility with protein stability. J. Biomol. Struct. Dyn. 1998, 16, 397-411.
[88] Zecchinon, L., Claverie, P., Collins, T., D'Amico, S., Delille, D., Feller, G., Georlette, D., Gratia, E., Hoyoux, A., Meuwis, M. A., Sonan, G., Gerday, C. Did psychrophilic enzymes really win the challenge? Extremophiles. 2001, 5, 313-321.
[89] Tsou, C. L. The role of active site flexibility in enzyme catalysis. Biochemistry-Moscow. 1998, 63, 253-258.
[90] Jaenicke, R. Stability and stabilization of globular proteins in solution. J. Biotechnol. 2000, 79, 193-203.
[91] Li, W. F., Zhou, X. X., Lu, P. Structural features of thermozymes. Biotechnol. Adv. 2005, 23, 271-281.
[92] Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z., Kern, D. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. Nat. Struct. Mol. Biol. 2004, 11, 945-949.
[93] Tindbaek, N., Svendsen, A., Oestergaard, P. R., Draborg, H. Engineering a substrate-specific cold-adapted subtilisin. Protein. Eng. Des. Sel. 2004, 17, 149-156.
[94] Scandurra, R., Consalvi, V., Chiaraluce, R., Politi, L., Engel, P. C. Protein stability in extremophilic archaea. Front. Biosci. 2000, 5, D787-D795.
[95] Dill, K. A. Dominant forces in protein folding. Biochemistry. 1990, 29, 7133-7155.
[96] Kumar, S., Tsai, C. J., Nussinov, R. Maximal stabilities of reversible two-state proteins. Biochemistry. 2002, 41, 5359-5374.

[97] Pace, C. N. Polar group burial contributes more to protein stability than nonpolar group burial. Biochemistry. 2001, 40, 310-313.
[98] Pace, C. N., Alston, R. W., Shaw, K. L. Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. Protein Sci. 2000, 9, 1395-1398.
[99] Kumar, S., Nussinov, R. Salt bridge stability in monomeric proteins. J. Mol. Biol. 1999, 293, 1241-1255.
[100] Slupphaug, G., Eftedal, I., Kavli, B., Bharati, S., Helle, N. M., Haug, T., Levine, D. W., Krokan, H. E. Properties of a recombinant human uracil-DNA glycosylase from the ung gene and evidence that ung encodes the major uracil-DNA glycosylase. Biochemistry. 1995, 34, 128-138.
[101] Kumar, S., Nussinov, R. Relationship between ion pair geometries and electrostatic strengths in proteins. Biophys. J. 2002, 83, 1595-1612.
[102] Hoseki, J., Okamoto, A., Masui, R., Shibata, T., Inoue, Y., Yokoyama, S., Kuramitsu, S. Crystal structure of a family 4 uracil-DNA glycosylase from thermus thermophilus hb8. J. Mol. Biol. 2003, 333, 515-526.

Paper I

doi:10.1016/j.jmb.2004.09.004 J. Mol. Biol. (2004) xx, 1–10

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Optimisation of the Surface Electrostatics as a Strategy for Cold Adaptation of Uracil-DNA *N*-glycosylase (UNG) from Atlantic Cod (*Gadus morhua*)

Elin Moe[1], Ingar Leiros[2], Ellen Kristin Riise[3], Magne Olufsen[3] Olav Lanes[4], Arne O. Smalås[3] and Nils Peder Willassen[1]*

[1]*Department of Molecular Biotechnology, Institute of Medical Biology, Faculty of Medicine, University of Tromsø N-9037 Tromsø, Norway*

[2]*The European Synchrotron Radiation Facility, 38043 Grenoble Cedex, France*

[3]*The Norwegian Structural Biology Centre, University of Tromsø, N-9037 Tromsø Norway*

[4]*Biotec Pharmacon ASA N-9008 Tromsø, Norway*

**Corresponding author*

Cold-adapted enzymes are characterised by an increased catalytic efficiency and reduced temperature stability compared to their mesophilic counterparts. Lately, it has been suggested that an optimisation of the electrostatic surface potential is a strategy for cold adaptation for some enzymes. A visualisation of the electrostatic surface potential of cold-adapted uracil-DNA *N*-glycosylase (cUNG) from Atlantic cod indicates a more positively charged surface near the active site compared to human UNG (hUNG). In order to investigate the importance of the altered surface potential for the cold-adapted features of cod UNG, six mutants have been characterised and compared to cUNG and hUNG. The cUNG quadruple mutant (V171E, K185V, H250Q and H275Y) and four corresponding single mutants all comprise substitutions of residues present in the human enzyme. A human UNG mutant, E171V, comprises the equivalent residue found in cod UNG. In addition, crystal structures of the single mutants V171E and E171V have been determined. Results from the study show that a more negative electrostatic surface potential reduces the activity and increases the stability of cod UNG, and suggest an optimisation of the surface potential as a strategy for cold-adaptation of this enzyme. Val171 in cod UNG is especially important in this respect.

Keywords: uracil-DNA *N*-glycosylase; cold-adapted enzymes; electrostatics; stability; crystal structure

Introduction

Cold-adapted enzymes possess lower thermal stability and higher catalytic efficiency than their mesophilic homologues. The dominating hypothesis is that these features originate from an increased structural flexibility.[1] Comparative studies have shown that cold-adapted enzymes have fewer cross-linking interactions, in particular in regions associated with the catalytic site. Although not proven, these findings indicate that increased local flexibility in key regions of the structure is a major determinant for cold activity.[2] However, other strategies for cold adaptation might exist, and an optimisation of the electrostatic surface potential near the active site is observed for some cold-adapted enzymes. Salmon trypsin is reported to have a more negatively charged surface near the catalytic area compared to bovine trypsin.[3] The increased negative potential is believed to cause the increased substrate affinity for positively charged substrates, and the optimised surface potential is suggested to be one of the strategies for adaptation of trypsin to cold environments.[4] Malate dehydrogenase (MD) from *Aquaspirillium arcticum* is reported to have a more positive potential on the surface around the substrate-binding site and a decreased negative potential around the co-factor-binding site compared to thermophilic MD.[5] It is suggested that the altered potentials may facilitate the interaction with the negatively charged substrate (oxaloacetate) and increase the catalytic efficiency at low temperatures. Citrate synthase (CS) from an Antarctic bacterium

Abbreviations used: UNG, uracil-DNA *N*-glycosylase; cUNG, recombinant cod UNG; hUNG, recombinant human UNG.

E-mail address of the corresponding author: nilspw@fagmed.uit.no

is reported to have a more negatively charged surface potential near the positively charged active site of the enzyme compared to thermophilic CS.[6] The hypothesis is that the negative potential serves to focus the negatively charged substrate (oxaloacetate) towards the positive potential of the active site.

Uracil-DNA *N*-glycosylase (UNG) is a DNA repair enzyme involved in the base excision repair pathway (BER) for removal of uracil in DNA.[7] The structure of the catalytic domain is known from human UNG[8] herpes simplex virus 1 (HSV-1) UNG,[9] *Escherichia coli* UNG[10] and Atlantic cod UNG.[11] The four homologues have similar structures and consist of a classic single-domain α/β-fold with a central four-stranded parallel and twisted β-sheet surrounded by 11 α-helices. The active site is located within a positively charged groove at the C-terminal end of the parallel β-sheet. Four loops are involved in the detection and removal of uracil in DNA: The 4-Pro loop (165-PPPPS-169), the Gly-Ser loop (246-GS-247), the Leu272-loop (268-HPSPLSVY-275) and the water-activating loop (145-DPYH-148). Some of the most important residues for catalysis are the conserved His268 and Asp145. His268 is located in the Leu272-loop, while Asp145 is located in the water-activating loop.[12]

The gene encoding the catalytic domain of cold-adapted UNG from Atlantic cod (*Gadus morhua*) has been identified.[13] Biochemical analysis of recombinant cod UNG (cUNG) revealed typical cold-adapted features such as low temperature stability, low temperature optimum and increased catalytic efficiency compared to recombinant human UNG (hUNG). In addition, the kinetic constants of cUNG showed a reduced K_M value, thus an increased substrate affinity compared to hUNG. A visualisation of the electrostatic surface potential of the three-dimensional structure of cUNG indicates that it possesses a more positive potential near the active site compared to hUNG.[11] The increased positive charge on the surface of cod UNG might explain the increased affinity for the negatively charged substrate DNA.

Further analyses suggested that four amino acid residues on the surface of cUNG were responsible for the differences in surface potentials. In order to study how the surface electrostatics affect the cold-adapted features of cod UNG, one mutant with four substituted amino acids (V171E, K185V, H250Q and H275Y) and four corresponding single mutants were constructed and characterised, and the 3D-structure of the single mutant V171E was determined. In addition, a mutant of human UNG, E171V, was investigated. The results show that the mutations affect both the catalytic efficiency and the stability of cod UNG. In particular, Val171 is shown to be essential in order to maintain the cold-adapted features of cUNG. The substitution of Val with Glu reduced the substrate affinity, with a 50% reduction in activity as a result, and increased the stability by adding one ion-pair.

Results

Construction of mutants

A visualisation of the electrostatic surface potential of the crystal structure of cUNG using SwissPDBViewer[14] and GRASP[15] reveals a more positive surface potential near the DNA binding site compared to hUNG (Figure 1(a) and (b)). Five mutants of cUNG were constructed on the basis of sequence and structural analysis, one mutant with four substitutions, V171E, K185V, H250Q and H275Y (quadruple), and four corresponding single mutants. In addition, one mutant of recombinant hUNG, E171V, was constructed. A model of cUNG with the four amino acid substitutions is shown in Figure 1(c).

Large-scale expression and purification

The cUNG mutants and cUNG were produced in a 15 l Chemap 3000 fermenter and purified by ion-exchange chromatography and gel-filtration to apparent purity as shown in Figure 2.

The molecular mass of the mutants of cod UNG was determined to be approximately 28 kDa, which is the same as observed previously.[13] The hUNG migrates slower than cUNG, with a molecular mass estimated as 29 kDa. The yield varied, but was typically between 2 mg and 4 mg per fermentation. The hUNG E171V mutant was produced and purified using the methods described above (results not shown).

Kinetics

The kinetic constants k_{cat} and K_M were determined for cUNG, hUNG and the mutants at three different temperatures, 15 °C, 22 °C and 37 °C at the optimal condition for cUNG (pH 7.5 and 50 mM NaCl). The results are summarised in Table 1.

The results show that k_{cat} of the quadruple mutant is reduced by approximately 40%, while k_{cat} of the V171E and H275Y mutants is reduced by more than 50% compared to cUNG. k_{cat} of K185V is increased at 37 °C and 15 °C, while it is decreased at 22 °C compared to cUNG. The k_{cat} value of H250Q is increased at 37 °C and reduced at 22 °C relative to cUNG, while it is approximately the same as of cUNG at 15 °C.

The K_M values of the V171E mutant are approximately on the same level as hUNG, indicating a reduced substrate affinity compared to cUNG. Quadruple and H275Y possess slightly lower K_M values than cUNG, while the K185V and H250Q mutants have approximately the same K_M values as cUNG.

Alterations of the kinetic constants of the quadruple, V171E and H275Y mutants result in reduced catalytic efficiencies (k_{cat}/K_M) compared to cUNG. Cod UNG K185V possesses an increased catalytic efficiency at 37 °C and 22 °C and decreased efficiency at 15 °C, while the H250Q mutant

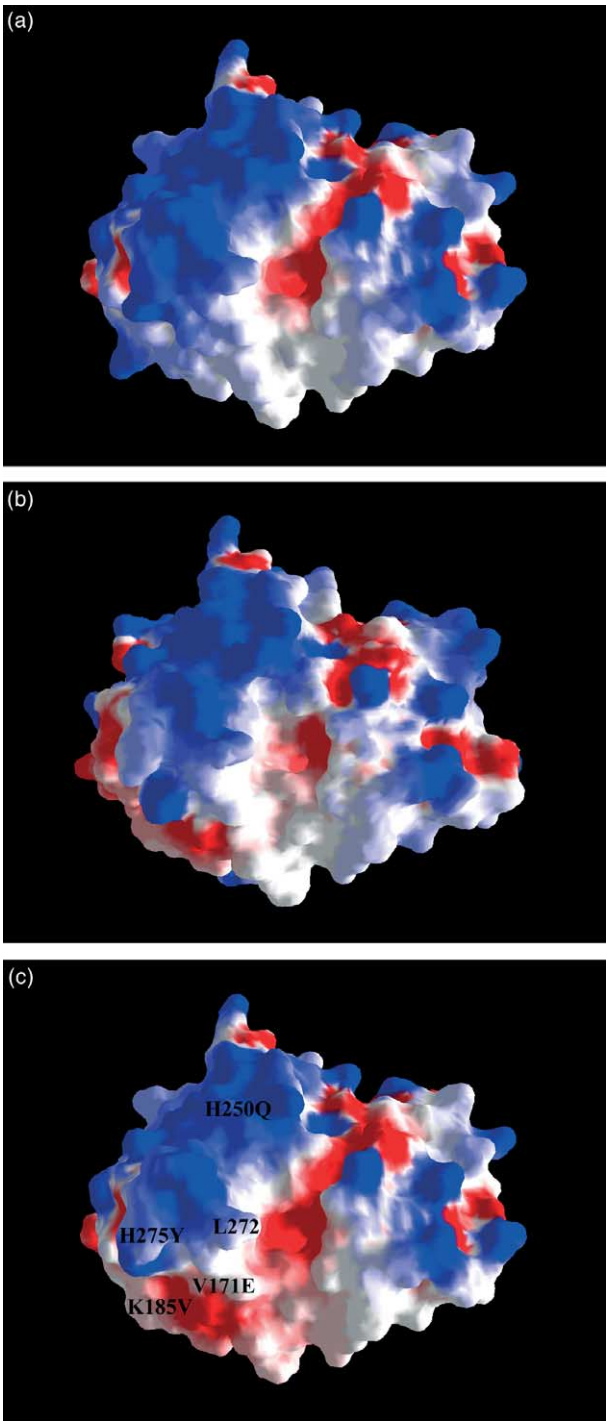


Figure 1. Difference in electrostatic surface potential of cUNG and hUNG. The Figure was made using the SwissPDBViewer,[14] with an electrostatic surface potential imported from GRASP[15] and contoured at ± 3 kT/e, where red describes a negative potential and blue describes a positive potential: (a) cUNG, (b) hUNG, (c) cUNG with substituted residues believed to increase the surface potential.

shows higher catalytic efficiency at 37 °C and lower catalytic efficiency at 22 °C and 15 °C, relative to cUNG. The hUNG E171V mutant possesses an increased catalytic efficiency at all three

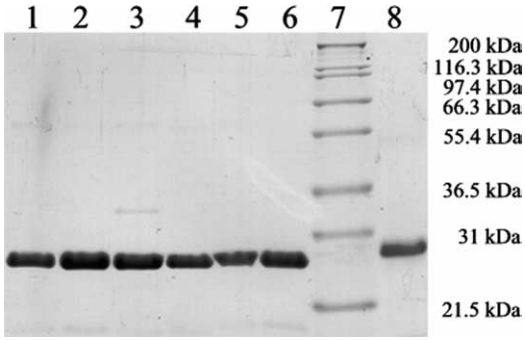


Figure 2. SDS-PAGE of purified cUNG and mutants. Approximately 3 μg of each enzyme was applied to the gel. Lane 1, cUNG; lane 2, rcV171E, lane 3, rcK185V, lane 4, rcH250Q, lane 5, rcH275Y, lane 6, quadruple, lane 7, Mark 12™ wide-range protein standard (NOVEX, San Diego, CA); lane 8, hUNG.

temperatures compared to hUNG due to reduced K_M values.

Temperature stability

The thermal stability of cUNG, hUNG and the mutants were determined at 37 °C. The results are shown in Figure 3. There is an apparent increase in the stability of all the mutants at 37 °C compared to cUNG, and the quadruple and the V171E mutants seem to be as stable as hUNG. Analysis of the temperature stability of the hUNG E171V mutant at 37 °C shows reduced stability compared to hUNG.

Crystal structure of the cUNG V171E and hUNG E171V mutants

The crystal structure of the catalytic domain of the cUNG V171E and hUNG E171V mutants has been determined at 1.7 Å and 1.95 Å resolution, respectively. The overall structures of the mutants are very similar to the catalytic domains of cUNG[11] and hUNG[8] with an overall root-mean-square (r.m.s.) difference of 0.27 Å and 0.32 Å, based on all main-chain atoms.

Electron density near residue 171 in the structure of cUNG V171E confirms that the valine to glutamic acid mutation introduces the expected ion pair interaction between Glu171 and His186. The orientation of Glu171 is slightly different from the corresponding residue in hUNG (Figure 4), but the ion pair distances are comparable, indicating similar strength of the interactions. Distances between the carboxylate oxygen atoms of Glu171 and the His nitrogen atom are 3.05 Å and 3.65 Å in hUNG, and 3.12 Å and 3.33 Å in the cUNG V171E mutant, respectively. The structure of the hUNG E171V mutant, correspondingly show no dramatic changes compared to hUNG at the mutation site other than the expected loss of the 171–186 interactions (Figure 5). The structure becomes more flexible in this region and the local

Table 1. Kinetic constants for cUNG, hUNG and mutants at 37 °C, 22 °C and 15 °C

UNG	Temperature (°C)	V_{max} (nmol min^{-1} mg^{-1})	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (min^{-1} μM^{-1})
cUNG	37	49,135±1626	1241	0.8±0.09	1499
	22	26,807±555	677	0.8±0.06	892
	15	11,064±374	279	0.5±0.07	607
Quadruple (V171E, K185V, H250Q, H275Y)	37	29,204±2210	737	0.6±0.22	1219
	22	13,244±282	334	0.5±0.05	612
	15	6753±213	171	0.5±0.07	335
cUNGV171E	37	20,189±553	510	2.0±0.12	260
	22	8070±239	204	1.7±0.12	117
	15	4041±143	102	1.4±0.13	72
cUNGK185V	37	63,997±1607	1616	0.8±0.07	1948
	22	25,865±617	653	0.6±0.06	1166
	15	16,159±750	408	0.8±0.13	510
cUNGH250Q	37	52,888±1079	1335	0.8±0.06	1637
	22	20,063±512	507	0.6±0.06	844
	15	10,879±282	275	0.7±0.07	419
cUNGH275Y	37	23,590±859	596	0.6±0.09	1017
	22	11,974±553	302	0.6±0.11	509
	15	6166±384	156	0.7±0.16	224
hUNG	37	25,347±1256	647	2.1±0.23	309
	22	12,530±740	320	2.4±0.29	135
	15	5580±399	142	2.2±0.34	66
hUNGE171V	37	26,529±885	677	1.0±0.1	684
	22	10,256±285	262	0.7±0.07	396
	15	6545±253	167	1.0±0.12	164

environment becomes similar to that of cUNG, although the conformation of the Val171 side-chain is slightly different in the two structures.

Discussion

An optimisation of the electrostatic surface potential is suggested to be a strategy for adaptation of enzymes to cold environments. A qualitative evaluation of the electrostatic surface potential of the 3D structure of cod UNG revealed more positive potential near the active site compared to human UNG. The presence of residues giving rise to more positive electrostatic potential in the active-site region might explain the increased substrate affinity (reduced K_M) and thus, in part, the enhanced catalytic efficiency of this enzyme. In order to investigate this hypothesis, five mutants of cUNG and one mutant of hUNG were produced, characterised and compared to cUNG and hUNG. In addition, the 3D X-ray structure of the cUNG V171E and the hUNG E171V mutants were determined

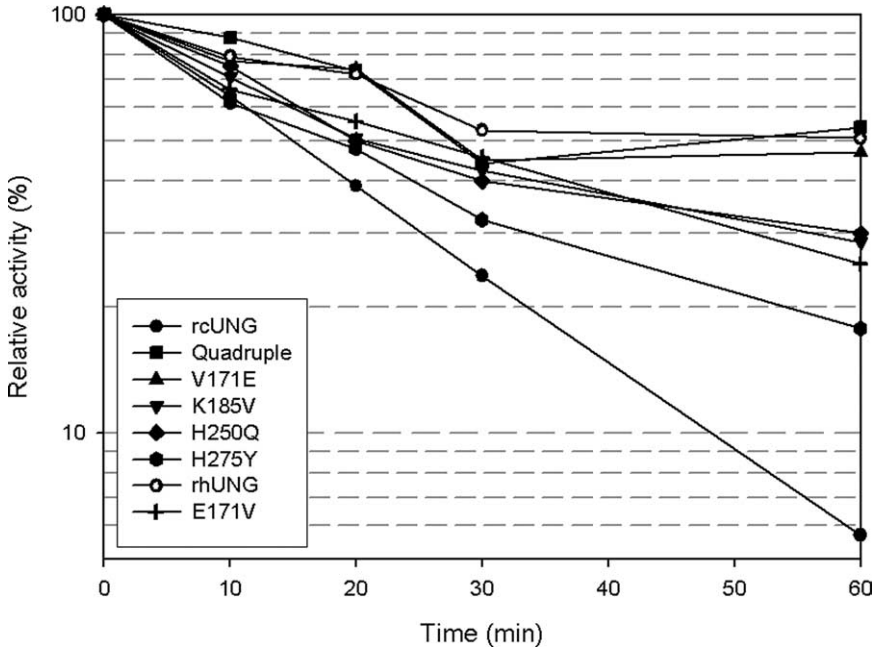


Figure 3. Temperature stability of cUNG, hUNG and six mutants at 37 °C. Stability was measured after incubation at 37 °C in 10, 20, 30 and 60 minutes. Standard deviations are not indicated, but were typically found in the range of 1–5%.

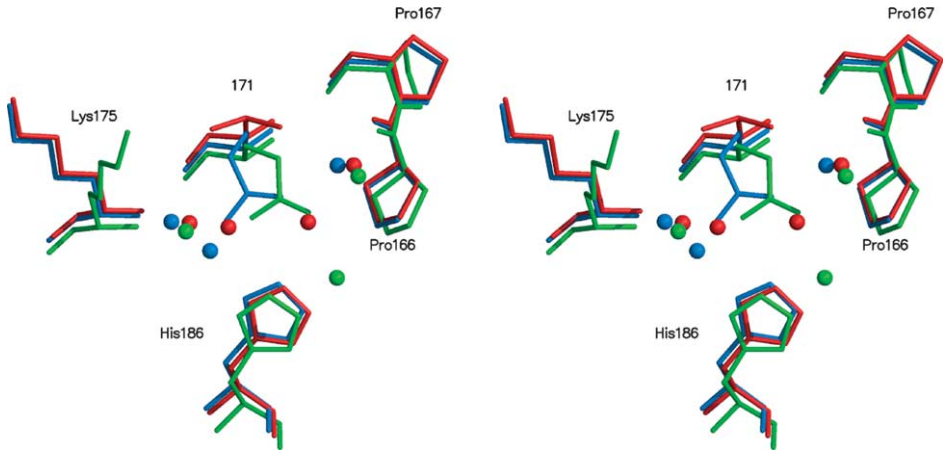


Figure 4. Stereo view of residue 166, 167, 171, 175 and 186 of cUNG V171E (blue), cUNG (red) and hUNG (green). The water molecules near residue 171 are the same colour as the enzyme. One ion-pair, less than 4 Å, is created between Glu171 and His186 in the mutant.

and analysed. The structure of cUNG with locations of the substituted amino acids is illustrated in Figure 6.

The cUNG V171E mutant seems to closely resemble hUNG in its behaviour. Compared to cUNG, the K_M values are increased at all temperatures tested, resulting in a significantly reduced catalytic efficiency. Substituting Val with Glu reduces the surface potential near the active site and leads to a weaker interaction between UNG and the negatively charged substrate DNA. The V171E mutant also shows increased temperature stability at 37 °C. According to the structure of hUNG,[8] Glu171 is involved in two salt-bridges, one weak interaction to Lys175 (5.7 Å) and one to His186 (3.05 Å and 3.65 Å). Residue 171 is located in helix α6, and the ion pair Glu171–Lys175 is intrahelical, while Glu171–His186 runs from α6 to the loop between α6 and α7. The helix α6 is located close to the DNA-binding area and moves upon DNA-binding. The corresponding salt-bridges are not present in cUNG. However, the crystal structure of the cUNG V171E mutant shows that the substitution of Val171 to Glu introduces one salt-bridge between Glu171 and His186 (3.12 Å and 3.33 Å) (Figure 4). This additional salt-bridge probably results in a more rigid structure of this area of the mutant. On the basis of the hypothesis about an inverse relationship between rigidity and activity, it is possible to explain the reduced k_{cat} by a decreased flexibility of this mutant compared to cUNG. The reduced flexibility also explains the increased temperature stability of this mutant.

The hUNG E171V mutant has correspondingly obtained reduced temperature stability and increased catalytic efficiency as cold-adapted features. Substituting Glu171 with Val leads to a removal of the previously described salt-bridges between Glu171 and Lys175 and His186 (Figure 5), giving rise to reduced stability of structural elements close to the active site. The increased catalytic efficiency is caused mainly by an increased substrate affinity (reduced K_M) and might be explained by an enhanced interaction between the

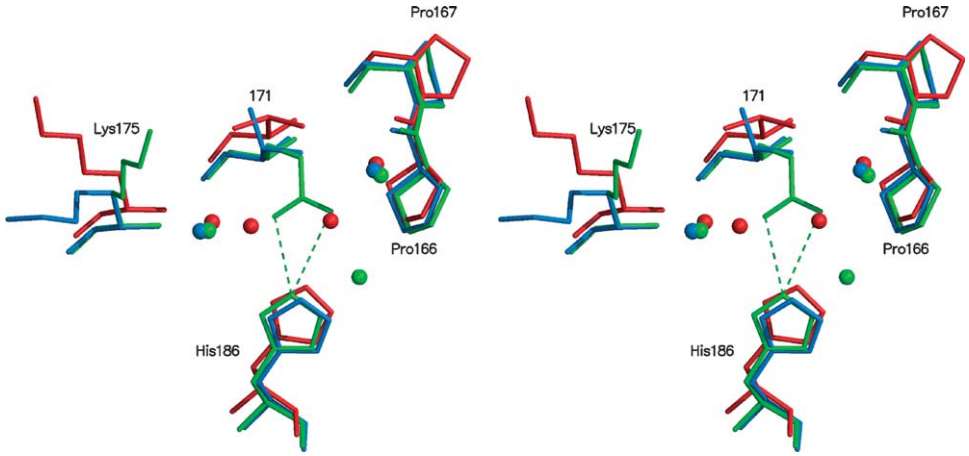


Figure 5. Stereo view of residues 166, 167, 171, 175 and 186 of hUNG E171V (blue), cUNG (red) and hUNG (green). The water molecules near residue 171 are the same colour as the enzyme. The broken line indicates the salt-bridge between 171 OE1 and OE2 to 186 NE2 in hUNG.

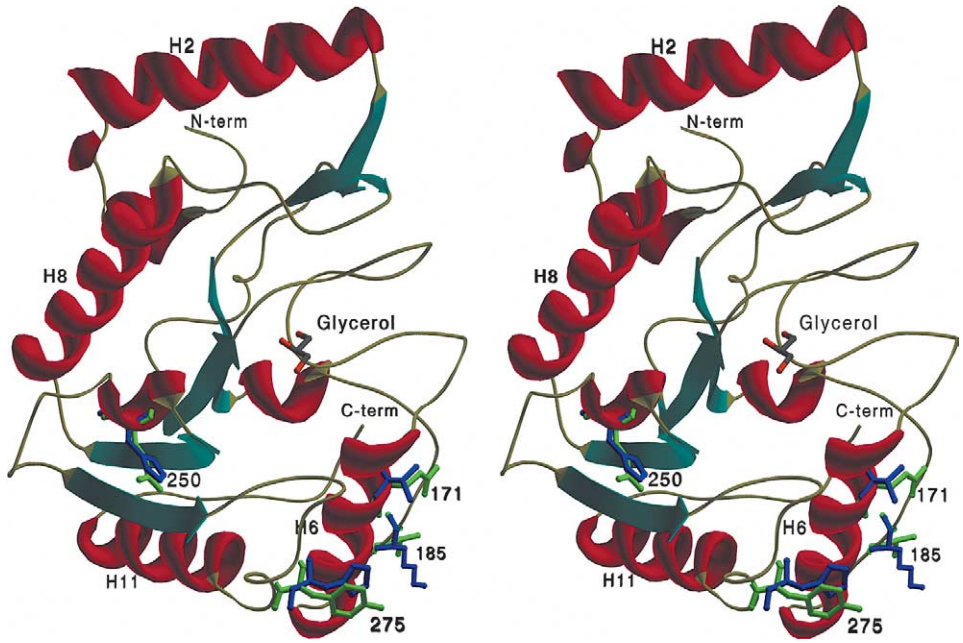


Figure 6. Stereo view of the crystal structure of cUNG with residues believed to affect the electrostatic surface potential (V171, K185, H250 and H275). cUNG residues are shown in blue, while hUNG residues are shown in green. The glycerol molecule found in the active site of cUNG is included for clarity. Secondary structure elements are as defined by Leiros *et al*.[11]

enzyme and the substrate due to the removal of the negatively charged Glu, and the increased flexibility resulting from the removal of an ion-pair interaction. On the basis of the results from analysis of the V171E and the E171V mutants, it seems as if residue 171 is a structural determinant for catalytic activity and stability in both hUNG and cUNG.

The catalytic efficiency of the cUNG K185V mutant is increased at 37 °C and 22 °C, and reduced slightly at 15 °C relative to cUNG, i.e. displays a slight loss of cold-adaptation properties. The structures of cUNG and hUNG reveal that this amino acid is located at the surface of the enzyme, in the loop between α6 and α7 in cUNG and has no interaction with other amino acids (Figure 6). According to the theory about a reduced substrate interaction caused by a more negatively charged surface, one would expect the catalytic efficiency to be reduced for this mutant. However, it is possible that the distance between residue 185 and the substrate is too large to affect the enzyme activity. The results from the temperature stability measurement show an increased stability compared to cUNG, thus Val185 has a stabilising effect at 37 °C.

The catalytic efficiency of the cUNG H250Q mutant is increased at 37 °C, while it is similar at 22 °C and reduced at 15 °C compared to cUNG, and as for the cUNG K185V mutant, this indicates loss of cold-adaptation features. In the structure of cUNG, His250 is located in α9, and forms hydrogen bonds with Ala266 and Leu244. Ala266 is located at the terminus of the β4 strand near the catalytically important Leu272 loop, and Leu244 is located in the β3 strand, near the highly conserved Gly-Ser loop (residues 246–247) (Figure 6). Thus, His250 forms hydrogen bonds to residues close to regions involved in DNA-binding and catalysis. In hUNG, Gln250 forms a hydrogen bond to Thr266, but not to residue 244. The substitution of His250 to Gln in cUNG could remove the hydrogen bond between residues 250 and 244 in cUNG and introduce a more flexible Gly-Ser loop, thereby explaining the increased catalytic efficiency of the H250Q mutant at 37 °C. However, the suggested increased flexibility does not seem to improve the catalytic efficiency at low temperatures.

The catalytic efficiency of the cUNG H275Y mutant is decreased compared to cUNG at all temperatures measured, while the temperature stability is increased slightly. His275 is located in the proximity of the so-called Leu272-loop, or the minor groove intercalation loop. The movement of this loop is important for bringing the catalytically important His268 within hydrogen-bonding distance of uracil O2, and for accomplishing the formation of the uracil-recognition pocket.[12] The substitution of His275 to Tyr probably reduces the flexibility of the Leu272 loop, and hence reduces the catalytic efficiency and increases the temperature stability of cUNG.

The quadruple mutant possesses all the four substitutions as described above, and adopts features such as reduced catalytic efficiency and increased temperature stability compared to cUNG. It has obtained a stronger substrate affinity (lower K_M) than cUNG, while all the single mutants have a weaker or similar substrate affinity than cUNG. The temperature stability of the quadruple mutant (Figure 3) seems to be additive, as it is more stable than any of the mutants, but the activity for the

quadruple mutant appears to be an average of the four individual mutations (Table 1). However, the total picture indicates that the quadruple mutant displays a greater loss in catalytic efficiency at lower temperatures, whereas the temperature stability is increased, again indicative of a loss of cold-adaptation features.

The difference in catalytic efficiency between most cold-adapted enzymes and homologous mesophilic enzymes increases with decreasing temperatures. This is true also for cUNG and hUNG, where the difference in catalytic efficiency is 80% at 37 °C, and 89% at 15 °C (Table 1). On the basis of this observation it should be possible to evaluate the effect of amino acid substitution on the cold-adapted features of cUNG, by calculating the difference in catalytic efficiency at high and low temperatures between the mutants and the native cold-adapted enzyme as shown in Table 2.

Calculations show that the reduction in catalytic efficiency is most severe for the V171E mutant, but there is also a considerable effect for the quadruple and H275Y mutants, and the difference in catalytic efficiency increases with decreasing temperature for these mutants compared to cUNG. There seems to be no reduction in the catalytic efficiency of the K185V or the H250Q mutants at 37 °C or 22 °C, while some effect is observed at 15 °C. On the basis of this information and the results from the characterisation, it seems as if Val171 is the single most important residue among those investigated in order to maintain the cold-adapted features of cUNG, while the other residues are of a somewhat more moderate importance. However, all the mutants showed a decrease in catalytic efficiency that was most pronounced at low temperature, i.e. there appears to be a rational relationship between charge complementarity and cold adaptation.

Conclusion

The results from the present study show that the increased electrostatic surface potential is important for the catalytic efficiency of cUNG, hence indicating that an optimisation of the potential in the substrate-binding area is a strategy for cold adaptation of this enzyme. Val171 is shown to be particularly important in this respect. Substitution of Val171 to Glu in cUNG resulted in a reduced substrate affinity (increased K_M) that contributes strongly to the reduced catalytic efficiency. An additional salt-bridge was introduced by the mutation and led to an increased temperature stability compared to cUNG. The opposite features were observed for the hUNG E171V mutant, caused by an enhanced substrate interaction (reduced K_M) and removal of a salt-bridge. Thus the presence of Val171 in cUNG seems to optimise the interaction between the enzyme and the substrate, and increase the flexibility of the structure close to the active site, and by that means the enzyme performs a more efficient catalysis at low temperatures. Characterisations of the other surface charge mutants tested in this study (K185V, H250Q, H275Y and the quadruple mutant) confirm the relation between surface charge and cold activity.

The cross-mutation between Val and Glu in cUNG and hUNG at residue 171 demonstrates that local flexibility features and electrostatic environments, especially near the catalytic region of an enzyme, is of vital importance for the catalytic efficiency. The clear correlation between mutations imposing rigidity and poorer electrostatic complementarity between enzyme and substrate, with a reduced catalytic efficiency, an effect that is increasingly notable at reduced temperature, shows that cold activity of enzymes can in fact be manipulated. Additionally, the data on the mutants show a clear correlation between catalytic efficiency and stability. This is very strong support for the classical hypothesis on a close relation between stability and flexibility on one hand, and catalytic efficiency on the other.

Table 2. Reduction (%) in catalytic efficiency at different temperatures for cUNG mutants relative to cUNG

Construct	Temperature (°C)		
	37	22	15
cUNG	0	0	0
Quadruple	20	30	45
V171E	83	87	88
K185V	0	0	16
H250Q	0	0	31
H275Y	32	43	63

Materials and Methods

Materials

Vent DNA polymerase and EcoRI restriction enzyme were obtained from New England Biolabs (Beverly, MA). SalI restriction enzyme, dNTPs and phage T4 DNA ligase were purchased from Promega (Madison, WI). Primers were obtained from MedProbe (Eurogentech SA, Belgium), and the Qiaquick gel extraction kit was purchased from Qiagen (Germany). The GeneAmp 9700 Thermocycler was obtained from Perkin Elmer (Foster City, CA). Expression vector pTRC99A, Äkta Explorer Purification System, purification columns and deoxy [5-^{3}H]uridine 5′-triphosphate (16.2 Ci/mmol) were obtained from Amersham Biosciences (Uppsala, Sweden). The Bio-Rad Protein Assay Dye Reagent Concentrate and the Mini-Protean® 3 electrophoresis system were purchased from Bio-Rad (Hercules, CA), and SigmaPlot was obtained from SPSS Inc. (Chicago, IL). The QuikChange® site-directed mutagenesis kit was purchased from Stratagene (La Jolla, CA).

E. coli TOP10 [F-*mc*rA (*mrr-hs*dRMS-*mc*rBC) 80*lac*Z M15 *lac*X74 *de*oR *rec*A1 *ara*D139 (*ara-le*u) 7697 *ga*lU *ga*lK *rps*L (Str R) *en*dA1 *nu*pG] was purchased from Invitrogen (Carlsbad, CA) and *E. coli* NR8052 [Δ(*pro-lac, thi, ara, trpE9777, ung1*)] and purified recombinant human UNG (UNGΔ84) were kindly provided by Dr Hans E. Krokan, Institute for Cancer Research and Molecular Biology, Norwegian University of Science and Technology.

Construction of mutants

The mutants of cUNG were constructed as described.[16] In short, two PCR fragments were amplified using the gene upstream primer and mutation downstream primer, and the mutation upstream primer and the gene downstream primer. The fragments were purified from gel using the Qiaquick gel extraction kit, and eluted in 20 μl of elution buffer supplied by the manufacturer. Then 1 μl of each of the purified fragments was applied in a new PCR reaction and ligated using the gene upstream and downstream primers.

PCR was performed in a GeneAmp 9700 Thermocycler. Each reaction had a volume of 50 μl and contained 0.4 unit of Vent DNA polymerase and buffer supplied by the manufacturer, 0.125 mM dNTPs and each of the upstream and downstream primers at a concentration of 1 μM, in addition to the template. The amplification was carried out at 94 °C for two minutes, 80 °C for two minutes, followed by 30 cycles at 94 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 30 seconds, and a final extension step at 72 °C for seven minutes. The polymerase was added to the PCR mixture during the step at 80 °C. The fragments containing the genes encoding mutated cUNG were purified from 1% (w/v) agarose gel, digested with EcoRI and SalI restriction enzymes, and ligated into the pTrc99A expression vector using 200 units of T4 DNA ligase. Competent *E. coli* TOP10 was transformed with the plasmid and screened for positive clones. Plasmid DNA was isolated from the TOP10 cells using the Qiagen plasmid purification kit, and used for transformation of competent *E. coli* NR8052 and sequencing.

The following mutants were constructed: V171E, K185V, H250Q, H275Y, and the quadruple mutant V171E, K185V, H250Q and H275Y. The template for construction of the mutants was the gene encoding the catalytic domain of recombinant cod UNG with three additional residues (MEF) and two altered Arg codons (from AGA to CGT) in the N-terminal part of the sequence (cUNGΔ81o).

The gene upstream and downstream primers for the cod mutants are OPUNG (5′-ATG GAA TTC TTC GGA GAG AC-3′) and NPUNG (5′-CTG CAG GTC GAC TTA GAG TGC-3′) or NPpTrc (5′-GAT GCC TGG CAG TTC CCT AC-3′). OPUNG posesses an EcoRI restriction enzyme site and NPUNG posesses a SalI restriction enzyme site. Upstream primers for the mutants are (the mutation codons are underlined): OPV171E (5′-C AGT CTC GAA AAC ATA TAC-3′), OPK185V (5′-GAT GGC TTC GTG CAT CCT GG-3′), OPH250Q (5′-ca tac gcc CAG aag aag gga g-3′), and OPH275Y (5′-G TCT GCT TAT CGT GGG TTC-3′). The downstream primers were complementary and reverse of the primers listed above.

The human UNG mutant E171V was constructed using the QuikChange® site-directed mutagenesis kit and according to the manual from the manufacturer. The primers used were as follows: 5′-GTT CCG CCT CCG CCC AGT TTG GTG AAC ATT TAT AAA G-3′; 5′-C TTT ATA AAT GTT CAC CAA ACT GGG CGG AGG CGG AAC-3′.

DNA sequencing

DNA sequencing was performed using the Amersham Biosciences Thermo Sequenase Cy5 Dye Terminator Kit, ALFexpress™ DNA sequencer and ALFwin Sequence Analyser version 2.10, and according to the protocol supplied by the manufacturer. Gels were made with Reprogel™ Long Read and Reproset UV-polymeriser. All items were purchased from Amersham Biosciences (Uppsala, Sweden).

Expression

Large-scale expression was performed using a 15 l Chemap 3000 fermenter (Switzerland). A 200 ml preculture of *E. coli* NR8052 transformed with pTrc99A containing genes encoding the cUNG or mutants were used for inoculation of 7 l of 2×LB-medium supplemented with 20 mM glucose and 100 μg/ml of ampicillin. The cells were grown at 30 °C and the expression induced with 1 mM IPTG at A_{600}=2.0: 50 ml of 20% (w/v) glucose was added to the growth medium every hour after induction to avoid glucose starvation. The cells were harvested six hours after induction by centrifugation at 10,000*g* for ten minutes and stored at −20 °C.

Protein purification

The cUNG and the H275Y mutant were purified as described.[13] The quadruple, V171E, K185V, H250Q and E171V mutants were purified without using the Blue Sepharose FF column (1.6/5.0). This purification step was exchanged with purification on a Source 15 S column (2.6/3.0), followed by concentration of the sample to 3 ml and further purification on a Superdex 75 column (2.6/60). Fractions containing UNG after gel-filtration were pooled, concentrated and stored at −20 °C in 50% (v/v) glycerol. All solutions containing enzyme were kept on ice except during the purification steps, which were performed at room temperature. The centrifugations were performed at 4 °C.

Activity measurements

Preparation of substrate ([^{3}H]dUMP DNA made by nick translation) and measurements of UNG activity was performed as described.[17] One unit is defined as the amount of enzyme required for releasing 1 nmol of acid-soluble uracil per minute at 37 °C.

Protein determination

Protein concentration were determined with Bio-Rad Protein Assay Dye Reagent Concentrate based on the Bradford dye-binding procedure[18] according to the microtiter plate protocol described by the manufacturer (Bio-Rad, Hercules, CA), using bovine serum albumine (BSA) as standard. SDS-PAGE was performed as described,[19] with a 12% (w/v) polyacrylamide separating gel and a 5% (w/v) polyacrylamide stacking gel using the Mini-Protean® 3 electrophoresis system (BioRad, Hercules, CA). After electrophoresis, gels were stained with Coomassie brilliant blue G250.

Kinetics

K_M and k_{cat} were measured in the presence of eight different concentrations of [^{3}H]dUMP substrate in the range of 0.56–4.5 μM at 15 °C, 22 °C and 37 °C. The amounts of UNG used were different for each mutant at the different temperatures, giving cpm values between 500 and 5000. Assays were performed in 25 mM Tris–HCl (pH 7.5 at 37 °C), 50 mM NaCl, 1 mM EDTA and 100 μg/ml of BSA. Calculation of the kinetic constants was performed by the enzyme kinetics module in SigmaPlot (SPSS Inc., Chicago, IL).

Temperature stability measurements

The temperature stability was examined by incubation of diluted enzyme in 10 mM Tris–HCl (pH 8.0), 50 mM NaCl, 1 mM EDTA, 1% (v/v) glycerol in a total volume of 50 µl at 37 °C (pH was adjusted at 37 °C). After different time intervals, as indicated in the Figure legends, the tubes were transferred onto ice. From each tube, 5 µl was used in assay mixtures for measuring residual activity under standard assay conditions.

X-ray structure determination

Crystals of the cod UNG V171E mutant were obtained by the vapour-diffusion, hanging-drop method at room temperature with a protein concentration of 10 mg/ml prior to mixing with equal volumes of reservoir solution (2 µl) of 1.4 M sodium citrate in a 0.1 M Hepes-buffer at pH 7.5. Diffraction quality crystals of maximal size of 0.4 mm appeared in the drop after approximately three days. The X-ray data were collected at 100 K at ID14 EH4 at the European Synchrotron Radiation Facility (ESRF) at a wavelength of 0.933 Å. Reservoir solution containing 20% glycerol was used as cryo-protectant. The selected crystal of the V171E mutant (0.2 mm × 0.15 mm × 0.02 mm) diffracted to 1.7 Å resolution and belongs to the $P2_1$ space group with cell dimensions $a=69.01$ Å, $b=67.77$ Å, and $c=70.42$ Å, $\beta=119.4°$ ($V=286{,}926$ Å^3) and with two molecules in the asymmetric unit. The cell is slightly larger than for the cUNG ($V=274{,}324$ Å^3).[20]

Crystals of the hUNG E171V mutant were obtained by hanging-drop methods from a protein solution of 12.5 mg/ml prior to mixing with equal amounts (1.5 µl) of a reservoir solution made from 0.1 M sodium cacodylate buffer (pH 6.5), 25% (w/v) PEG 8000 and 0.1 M sodium acetate buffer (pH 4.9). X-ray data were collected to 1.95 Å resolution at the Swiss-Norwegian beam line, ESRF, at a wavelength of 0.933 Å. The crystals belong to space group $P2_12_12_1$ with cell dimensions $a=47.51$ Å, $b=54.77$ Å, and $c=78.07$ Å with one molecule in the asymmetric unit.

Both data sets were integrated using DENZO,[21] and scaling and merging was carried out using SCALA of the CCP4 program suite.[22] Due to the difference in cell parameters in cUNG V171 compared to the cUNG, the initial phases were obtained from molecular replacement searches in CNS[23] using cUNG as search model. The search gave two possible rotation-solutions, 180° apart, with the first peak 7.46 times higher than the mean of rotation function, and the other 5.26 times higher than peak number 3. These two peaks correspond to the two molecules in the asymmetric unit. There is 7 Å difference in the *c*-axis between the hUNG E171V and native hUNG[8] unit cells; thus, initial phases were obtained from molecular replacement searches also in this case.

Alternating cycles of model building using the program O[24] and refinement with CNS brought the final *R*-factor to 19.8% and R_{free} to 22.6% for cUNG V171E. The corresponding final *R*-factors for hUNG E171V were 21.0% and 25.4%. For both models, the refinement data were gradually extended to maximum resolution, and 5% of the reflections were kept out of refinement for cross validation.[25] Solvent molecules were added to the models where the difference density exceeded 4σ and within reasonable hydrogen bonding distances (3.4 Å). Details from the data collection and refinement characteristics are given in Table 3.

The quality of the final models is good. The electron density is generally well defined and about 90% of the main-chain dihedral angles are in the most favoured regions of the ϕ/ψ plot as judged from PROCHECK.[26]

Table 3. Data collection and refinement summary

	cUNG V171E	hUNG E171V
A. *Data collection*		
Resolution range (Å)	25–1.70	25–1.95
No. unique reflections	64,933	15,383
Redundancy (multiplicity)	3.2	2.3
R-merge (%)	4.6	13.2
Completeness (%)	94.3	93.2
Average $I/\sigma(I)$	9.7	4.0
B. *Refinement statistics*		
R-value (%)	19.82	20.98
Free *R*-value (%)	22.56	25.44
Deviation from ideal geometry		
Bond lengths (Å)	0.005	0.006
Bond angles (deg.)	1.189	1.233
Average *B*-values (Å^2)		
Main-chain atoms (number)	25.80 (1778)	15.41 (892)
Side-chain atoms (number)	27.99 (1800)	17.47 (914)
Glycerol (2)	38.30[a]	–
Chloride ion (2)	49.26[a]	–
Water molecules (number)	37.75 (365)	21.11 (126)
All atoms	28.01	16.76
Ramachandran plot (%)		
Most-favoured regions	89.4	89.9
Additionally allowed regions	10.1	10.1
Generously allowed regions	0.5[b]	0

[a] The values are calculated from the *B*-values of each atom.
[b] Phe158.

Acknowledgements

The Norwegian research council supported this work. We thank Eva Ryeng and Marit Sjo Lorentzen for technical assistance, and the organisers of the Swiss-Norwegian beam line (SNBL) and ID14 at the European Synchrotron Radiation Facility, Grenoble, France, for providing beam time.

References

1. Hochachka, P. W. & Somero, G. N. (1984). *Biochemical Adaptation*. Princeton University Press, Princeton, NJ.
2. Smalas, A. O., Leiros, H. K., Os, V. & Willassen, N. P. (2000). Cold adapted enzymes. *Biotechnol. Annu. Rev.* **6**, 1–57.
3. Gorfe, A. A., Brandsdal, B. O., Leiros, H. K., Helland, R. & Smalas, A. O. (2000). Electrostatics of mesophilic and psychrophilic trypsin isoenzymes: qualitative evaluation of electrostatic differences at the substrate binding site. *Proteins: Struct. Funct. Genet.* **40**, 207–217.
4. Brandsdal, B. O., Smalas, A. O. & Aqvist, J. (2001). Electrostatic effects play a central role in cold adaptation of trypsin. *FEBS Letters*, **499**, 171–175.
5. Kim, S. Y., Hwang, K. Y., Kim, S. H., Sung, H. C., Han,

Y. S. & Cho, Y. (1999). Structural basis for cold adaptation. Sequence, biochemical properties, and crystal structure of malate dehydrogenase from a psychrophile *Aquaspirillium arcticum*. *J. Biol. Chem.* **274**, 11761–11767.
6. Russell, R. J., Gerike, U., Danson, M. J., Hough, D. W. & Taylor, G. L. (1998). Structural adaptations of the cold-active citrate synthase from an Antarctic bacterium. *Structure*, **6**, 351–361.
7. Lindahl, T. & Nyberg, B. (1974). Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry*, **13**, 3405–3410.
8. Mol, C. D., Arvai, A. S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H. E. & Tainer, J. A. (1995). Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis. *Cell*, **80**, 869–878.
9. Savva, R., McAuley-Hecht, K., Brown, T. & Pearl, L. (1995). The structural basis of specific base-excision repair by uracil-DNA glycosylase. *Nature*, **373**, 487–493.
10. Ravishankar, R., Bidya Sagar, M., Roy, S., Purnapatre, K., Handa, P., Varshney, U. & Vijayan, M. (1998). X-ray analysis of a complex of *Escherichia coli* uracil DNA glycosylase (EcUDG) with a proteinaceous inhibitor. The structure elucidation of a prokaryotic UDG. *Nucl. Acids Res.* **26**, 4880–4887.
11. Leiros, I., Moe, E., Lanes, O., Smalas, A. O. & Willassen, N. P. (2003). The crystal structure of uracil-DNA *N*-glycosylase from Atlantic cod (*Gadus morhua*) reveals cold-adaptation features. *Acta Crystallog. sect. D*, **59**, 1357–1365.
12. Parikh, S. S., Mol, C. D., Slupphaug, G., Bharati, S., Krokan, H. E. & Tainer, J. A. (1998). Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO J.* **17**, 5214–5226.
13. Lanes, O., Leiros, I., Smalas, A. O. & Willassen, N. P. (2002). Identification, cloning, and expression of uracil-DNA glycosylase from Atlantic cod (*Gadus morhua*): characterization and homology modeling of the cold-active catalytic domain. *Extremophiles*, **6**, 73–86.
14. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
15. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.
16. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51–59.
17. Lanes, O., Guddal, P. H., Gjellesvik, D. R. & Willassen, N. P. (2000). Purification and characterization of a cold-adapted uracil-DNA glycosylase from Atlantic cod (*Gadus morhua*). *Comp. Biochem. Physiol. B*, **127**, 399–410.
18. Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.
19. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
20. Leiros, I., Lanes, O., Sundheim, O., Helland, R., Smalas, A. O. & Willassen, N. P. (2001). Crystallization and preliminary X-ray diffraction analysis of a cold-adapted uracil-DNA glycosylase from Atlantic cod (*Gadus morhua*). *Acta Crystallog. sect. D*, **57**, 1706–1708.
21. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
22. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
23. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al*. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
24. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
25. Kleywegt, G. J. & Brunger, A. T. (1996). Checking your imagination: applications of the free *R* value. *Structure*, **4**, 897–904.
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

Edited by R. Huber

(Received 6 May 2004; received in revised form 1 September 2004; accepted 3 September 2004)

Paper II

Paper III

Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling xxx (2006) xxx–xxx
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme

Magne Olufsen, Bjørn Olav Brandsdal, Arne Oskar Smalås *

The Norwegian Structural Biology Centre, Department of Chemistry, University of Tromsø, N-9037 Tromsø, Norway

Received 14 June 2006; received in revised form 17 October 2006; accepted 18 October 2006

Abstract

Uracil DNA glycosylase (UDG) is a DNA repair enzyme involved in the base excision repair (BER) pathway, removing misincorporated uracil from the DNA strand. The native and mutant forms of Atlantic cod and human UDG have previously been characterized in terms of kinetic and thermodynamic properties as well as the determination of several crystal structures. This data shows that the cold-adapted enzyme is more catalytically efficient but at the same time less resistant to heat compared to its warm-active counterpart. In this study, the structure–function relationship is further explored by means of comparative molecular dynamics (MD) simulations at three different temperatures (375, 400 and 425 K) to gain a deeper insight into the structural features responsible for the reduced thermostability of the cold-active enzyme. The simulations show that there are distinct structural differences in the unfolding pathway between the two homologues, particularly evident in the N- and C-terminals. Distortion of the mesophilic enzyme is initiated simultaneously in the N- and C-terminal, while the C-terminal part plays a key role for the stability of the psychrophilic enzyme. The simulations also show that at certain temperatures the cold-adapted enzyme unfolds faster than the warm-active homologues in accordance with the lower thermal stability found experimentally.

Keywords: Molecular dynamics; Uracil DNA glycosylase; Protein unfolding; Cold adaptation; Psychrophilic

1. Introduction

Several cold-adapted enzymes from different organisms living in constantly cold environments have been identified and characterized in the last few decades. Such enzymes possess unique properties, attractive not only for the academic research but also to the biotechnological market. The emerging picture suggests that cold-adapted enzymes have higher catalytic efficiency, lower thermostability and improved flexibility in the structural parts directly involved in the catalytic cycle [1,2]. Indeed, most psychrophilic enzymes studied so far display reduced overall stability when compared to their mesophilic counterparts. The increased catalytic efficiency is usually accompanied by a decrease in thermal stability, which is believed to be a consequence of higher structural flexibility in the psychrophilic enzymes. When the temperature decreases from 37 to 0 °C the reaction rates are generally reduced by 30- to 80-fold [2], and in order to maintain a sufficient metabolic flux at low temperatures it has been suggested that psychrophilic enzymes have higher flexibility in the structural parts directly involved in the catalytic cycle [1]. However, enzymes are shown to use different structural adaptation strategies to achieve the above-mentioned features [3], and psychrophilic enzymes often possess decreased number of proline residues, increased number of glycine residues, low relative arginine content [arg/(arg + lys)] and lower numbers of ion pairs, aromatic interactions and/or hydrogen bonds compared to the mesophilic counterparts [3–5].

Cold-adapted uracil DNA glycosylase from cod (cUDG) has been shown to be up to 10 times more catalytic efficient (k_{cat}/K_m) in the temperature range from 15 to 37 °C compared to the warm-active human counterpart (hUDG) [6]. UDG is the first enzyme in the BER pathway, and catalyzes the hydrolysis of promutagenic uracil residues from single- or double-stranded DNA [7]. The crystal structure of the catalytic domain of UDG from several species are known: human (hUDG) [8], Atlantic

* Corresponding author. Tel.: +47 77644070; fax: +47 77644765.
E-mail address: Arne.Smalas@chem.uit.no (A.O. Smalås).

1093-3263/$ – see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2006.10.003

cod (cUDG) [9], virus 1 [10] and *Escherichia coli* [11]. The catalytic domain of hUDG and cUDG consists of 223 amino acids with a sequence identity of 75%, and the overall topology is a typical α/β protein [8]. The r.m.s.d. between the two crystal structures is 0.492 Å (all atoms were included). The higher catalytic efficiency of cUDG compared to hUDG has been postulated to arise from a more flexible DNA-recognition loop [12] and optimization of the electrostatic surface potential near the specificity pocket [9] in the cold-adapted enzyme.

cUDG has also been found to be much more pH and temperature labile than hUDG [13], indicating that cUDG is generally less stable compared to hUDG. Large opposing energies are involved when proteins go from the unfolded to the folded state, but the overall change in energy is small, typically ranging from 1 to 15 kcal/mol [14]. All contributions, both favorable and unfavorable, are therefore important when considering protein stability [14]. How proteins achieve their stability does not seem to follow any general rules [15], but the emerging picture is, for globular proteins, that the hydrophobic effect and burial of non-polar side chains stabilizes the native state [16,17]. Disulfide bonds, electrostatic interactions and hydrogen bonds are, however, also important for structural stability and can contribute favorably to protein stability [18–20]. The majority of reversible two-state folding proteins, including thermophilic, mesophilic and psychrophilic enzymes, show maximum stability around room temperature [17].

High temperature MD simulations have been shown to provide valuable insights into the unfolding process of several proteins: chymotrypsin inhibitor 2 (CI2) [21], barnase [22,23] and engrailed homeodomain (En-HD) [24]. Protein unfolding studies using MD simulations are carried out by increasing the temperature necessary to overcome the enthalpic forces stabilizing the 3D structure. Because of the differences in accessible time scales by experiments and computer studies, higher temperatures are needed when performing computer experiments relative to that observed in experimental unfolding studies. A MD study of CI2 has shown that the unfolding pathway is not widely affected by the simulation temperature, and thus raising the temperature only increases the reaction rate [25].

In this study, we have employed MD simulations to investigate the structure–function relationship further with focus on possible differences in the thermal unfolding pathway between a psychrophilic and a mesophilic UDG. We have especially analyzed the structures to see if there are important molecular contacts, which could explain the observed difference in thermal stability between these enzymes. Comparative high temperature (375, 400, 425 K) simulations for 6 and 9 ns have been carried out for the two enzymes, and the unfolding process has been carefully monitored both at an overall and a detailed structural level.

2. Methods

All MD simulations were carried out using the AMBER7 program package [26] with the recently developed force field parm99 [27]. The generalized Born methodology [28] was applied to describe the electrostatics from the solvent. Implicit solvent has been chosen in this study. The advantage of using implicit solvent models is the increase in computational efficiency. MD simulations with implicit solvent have successfully been used to investigate protein folding [29,30], and recent development in the implicit solvent methodology now provide a level of realism that challenge the results from the explicit solvent simulations [31]. The simulations were initiated with 100 cycles of energy minimization and the protein atoms constrained with a 5 kcal/mol $Å^2$ force constant. The temperature was slowly raised to the final temperature (375, 400 or 425 K) during a heating phase of approximately 20 ps. In this phase, the protein atoms were weakly restrained to their initial positions. SHAKE [32] was employed to constrain all bonds involving hydrogen atoms. A cutoff of 15 Å was used for the non-bonded interactions and every 1000 step the rotational and translational motion was removed. The time step was set to 2 fs, and the temperature was maintained by the Berendsen coupling algorithm [33]. The hydrophobic effect was taken into account by adding a surface energy term to the total energy [34].

Crystal structures were used as starting structures in the MD simulations for both hUDG [8] and cUDG [35] (Protein Data Bank entries 1AKZ and 1OKB, respectively). The crystal structures were recombinant enzymes with three mutations in the N-terminal end: P82M, V83E and G84F. UDG contains several histidines, and in the simulation the His 148 was charged, while the rest of the histidines were considered as neutral with the hydrogen atom placed on the Nε atom. These choices were based on data from NMR and continuum electrostatics calculations [36].

Hydrogen bonds and r.m.s.d. values of the Cα atoms were calculated with the Amber program package. In the hydrogen bond calculations a distance cutoff of 3.4 Å and an angle cutoff of 60° (the angle between: hydrogen-donor–acceptor) were applied. These calculations were also tested with no angle cutoff, but these calculations gave similar results, and the angle cutoff method was used in this study. The DSSP program [37] was used to find the secondary structure elements of each snapshot. The secondary structure elements were calculated for snapshots every 10th ps and for smoothening the plots, the number of residues in secondary elements was averaged every 10 snapshots. The surface area was calculated by the MSMS program [38], and a probe radius of 1.5 Å was applied to calculate the solvent accessible surface area (SASA).

3. Results and discussion

Organisms have adapted to extreme environments and survive at different temperatures by optimization of their proteins at some level and thereby maintain sufficient metabolic fluxes. Increased availability of sequences, kinetic, thermodynamic and structural information of proteins from psychrophilic, mesophilic and thermophilic enzymes have facilitated a deeper understanding of features specific for adaptation at the molecular level. Presently, much deeper insight has been

Please cite this article in press as: M. Olufsen et al., Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme, J. Mol. Graph. Model. (2006), doi:10.1016/j.jmgm.2006.10.003

achieved on how proteins have adapted to cope with high temperatures without denaturation, as compared to the low temperature adaptation. Higher temperatures involve increased reaction rates and atomic mobility, while low temperature slows down the chemical reaction rates and reduced solvent viscosity. Hence, it may appear that adaptation to high and low temperature requires opposite changes, but several studies have shown that the picture is far more complicated than simply reversing the features responsible for high temperature adaptation when explaining how the challenges are met at low temperature.

Uracil DNA glycosylase is an excellent model system for the study of the mechanisms of enzymatic adaptation to low temperature. Its biological function is well-known, and several crystal structures of both cUDG and hUDG are available, including mutant and native structures, which have also been characterized in terms of stability and kinetics [6,8,9,13,35,39]. It is well-established that cold-adapted enzymes are less resistant to heat as compared to their warm-active homologues and this has also been shown to be the case for cUDG [13]. In light of the small differences revealed by structural analysis, computational simulation of the thermal unfolding process was a natural next step.

In order to gain further insights into the thermal stability of cUDG and hUDG eight comparative MD simulations at three different temperatures (375, 400 and 425 K) have been carried out. Molecular interactions and intramolecular contacts (like salt-bridges, hydrogen bonds and hydrophobic contacts) have been thoroughly monitored in order to search for specific features and structurally important regions that could explain the difference in thermal stability. Even higher temperatures were tested, but increasing the temperature beyond 425 K yielded complete loss of ordered structure within just a few ps. The resulting trajectories were too difficult to analyze as most of the sampling were on the fully unfolded polypeptide chains.

3.1. Structural fluctuation in the MD simulations

Computer aided thermal unfolding of proteins are usually performed at temperatures between 200 and 225 °C (473–498 K) [40], and a commonly used criterion for an unfolded state has been a Cα r.m.s.d. exceeding 10 Å [41]. According to this, the proteins studied here are all fully denaturated after approximately 5–6 ns. There are still some fluctuations in the r.m.s.d. values in the simulations at the two highest temperatures (Fig. 1A and B). However, the protein unfolding proceeds slowly in the 375 K simulations (Fig. 1A), and the structural ensembles show less fluctuations when compared to the 400 and 425 K simulations. The 375 K simulations were therefore extended to 9 ns to reach this threshold. cUDG and hUDG show similar unfolding rates for the first 5 ns of the simulations. However, the psychrophilic enzyme displays

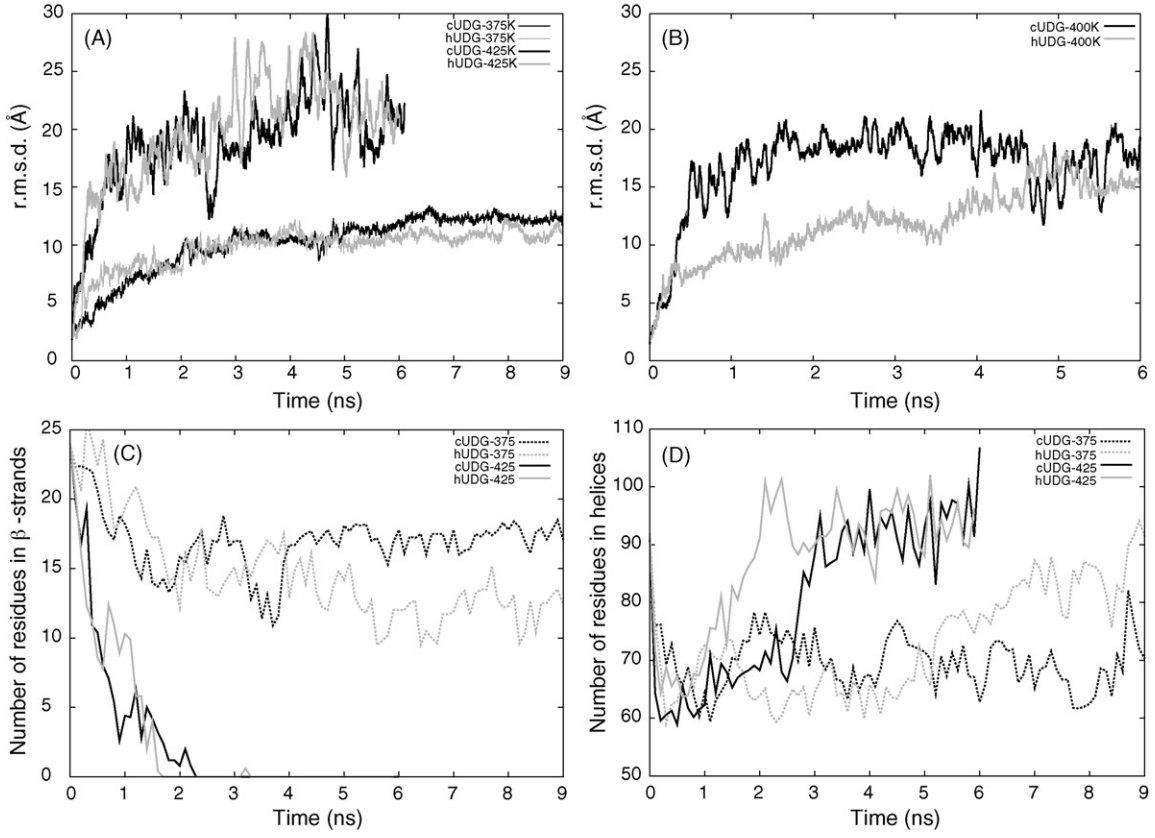


Fig. 1. Various plots from the simulations. (A) and (B) are average r.m.s.d. plots as a function of time during the simulation. The calculated r.m.s.d. values are based on all Cα atoms in the starting structure and from every snapshot during the simulation. (C) and (D) are number of residues in secondary structure elements during the simulation. The secondary structure elements are calculated by DSSP [37].

Please cite this article in press as: M. Olufsen et al., Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme, J. Mol. Graph. Model. (2006), doi:10.1016/j.jmgm.2006.10.003

higher r.m.s.d. values in the last 4 ns (Fig. 1A), indicating greater structural changes in the cold-adapted enzyme. The unfolding rate is faster when the temperature is raised to 400 K, and the simulations show that the enzymes unfolded to a higher degree at this temperature when compared to the 375 K results. Interestingly, large differences in the unfolding rates between the psychrophilic and the mesophilic enzyme are observed in the 400 K simulations. The psychrophilic enzyme unfolds much more rapidly than the mesophilic counterpart, indicating reduced structural stability (Fig. 1B). Destruction of the native protein structures occurs very fast when the temperature is raised to 425 K, and unfolding of cUDG and hUDG proceeds at a similar rate. Moreover, the simulations show that the mesophilic enzyme requires higher temperatures to achieve similar degree of structural distortions as observed for the psychrophilic enzyme. This already indicates differences in the interactions stabilizing the 3D structure, and is also in agreement with the observed reduced stability of the cold-adapted enzyme.

4. Structural investigations

4.1. Secondary structure

The starting structure of both cUDG and hUDG contain 92 and 24 residues in helices (α-helices, π-helices and 3–10 helices) and β-strands, respectively. The amount of residues in β-strands decreases at approximately the same rate in the simulation at 375 and 400 K for both enzymes (Fig. 1C), while for the simulation at the highest temperature, all β-strands are lost at around 2 ns for both enzymes (Fig. 1C). Both enzymes also have similar extent of residues involved in helices in the 375 and 400 K simulations, while it actually increases for both enzymes the 425 K simulation as the structure get distorted (Fig. 1D) (residues in helices for the 400 K simulation is not shown). Visual inspection of the simulated trajectories shows that the unfolding process generates new helices, originating from areas that were initially loops and β-strands in their respective crystal structures. The catalytic domain of both cUDG and hUDG is distorted already in the beginning of the 425 K simulations, and after 600 ps both enzymes have r.m.s.d. values of more than 15 Å (Figs. 1A and 7). When the domain becomes distorted and the β-sheets are destroyed, solvent gains access to the core β-strands yielding rapid unfolding. On the contrary, helices are stable even after an opening of the catalytic domain. However, both NMR and MD studies have shown that the denatured state may comprise significant amount of secondary structure [42]. There are also indications that the unfolded protein can increase the helical content during the simulation and that helices can be formed in water in the absence of significant tertiary interactions [22].

4.2. Hydrogen bonding pattern

The starting structures of cUDG and hUDG contained 240 and 223 intramolecular hydrogen bonds (shorter than 3.4 Å), respectively, and many are lost early in the simulations. However, after approximately 1 ns the number of hydrogen bonds reaches a stabile plateau, which is steadily maintained throughout the simulations. The average number of hydrogen bonds for the last 5 ns of the cUDG simulations is 198, 179 and 165 for the 375, 400 and the 425 K simulation, respectively, while the corresponding numbers for the mesophilic enzyme are 191, 174 and 156. Thus, as the simulation temperature is increased, there is a concomitant decrease in number of intact hydrogen bonds. This is reasonable as the structures become more distorted as the simulation temperature is raised.

4.3. Accessible surface area

The solvent accessible surface area (SASA) is an important parameter for mapping unfolding. The SASA has been calculated for the crystal structures and of snapshots from the MD simulations. The initial SASA for cUDG and hUDG is 10259 and 10376 $Å^2$, respectively. SASA increases rapidly for the 375 K simulations by approximately 20%, and stays roughly at the same level throughout the simulations (Fig. 2A). In the 400 K simulation the SASA increase to roughly 40% at the end of the simulation (Fig. 2B). While at 425 K the enzyme rapidly unfolds, and already after only 100–200 ps the SASA

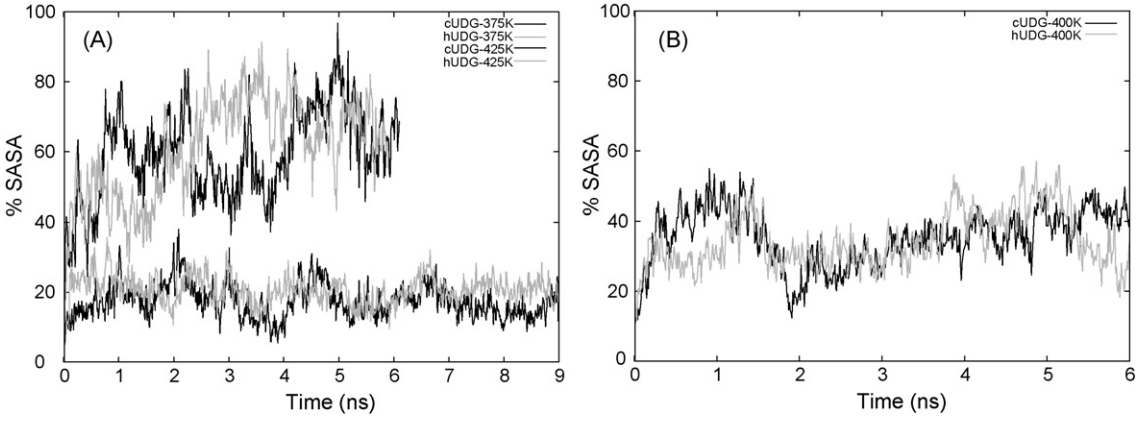


Fig. 2. Solvent-accessible surface area (SASA) as a function of simulation time. (A) shows percent increase in SASA from the 375 and the 425 K simulation. (B) shows percent increase in SASA from the 400 K simulation. The reference is the SASA from the crystal structures. The SASA are calculated by MSMS [38].

has increased by 40% (Fig. 2A). There are only minor differences in SASA among the enzymes as they unfold at different temperatures.

5. Stability of the N- and C-terminals

The N-terminal of cUDG unfolds at all three temperatures and as expected to a greater extent as the temperature is raised (Fig. 3A, C and E). While the N-terminal of the mesophilic variant of the enzyme also unfolds at all three temperatures (Fig. 3B, D and F), difference is observed when compared to cUDG. The N-terminal is more stable in the hUDG simulations as observed in the 375 and 400 K simulations. Structural investigations of the interactions involving the N-terminal show that there are several stabilizing contacts present. Numerous hydrophobic contacts between residues in the N-terminal helices (α1 and α2, α-helix 1 and 2, numbering according to Leiros et al. [9]) and residues in α6 and the loop between α6 and β1 are found. However, three hydrogen bonds (Ser88:Oγ-Asp133:Oδ2, Trp89:Nε1–Cys132:O and Trp89:Nε1–Thr129:O), which are conserved in the two enzymes, appear to be important for the N-terminal stability (Fig. 4). These hydrogen bonds connect the

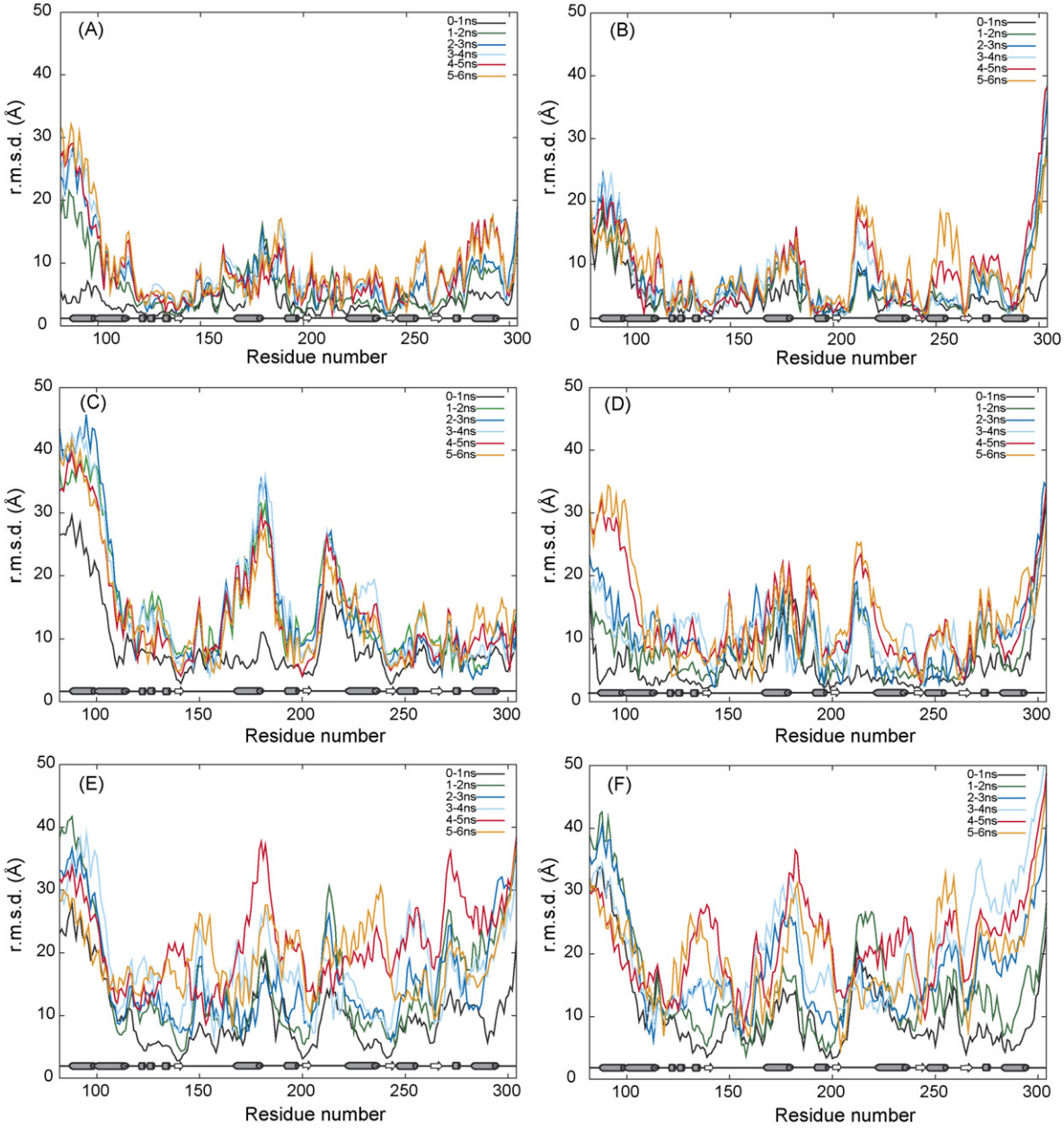


Fig. 3. Flexibility in the structure (r.m.s.d. per residue). The figure shows r.m.s.d. per residue. values in different part of the simulations, 0–1 ns is the r.m.s.d. values from the simulations from 0 to 1 ns, and so on. All the Cα atoms are used to calculate the r.m.s.d. values and the starting structure is used as a reference. The gray cylinder and the white arrow at the bottom of each plot is where helices and β-strands are located in the starting structure, respectively. (A), (C) and (E) is the r.m.s.d. values from the cUDG simulations at 375, 400, and 425 K, respectively. While (B), (D) and (F) is the r.m.s.d. values from the hUDG simulations at 375, 400 and 425 K, respectively.

Please cite this article in press as: M. Olufsen et al., Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme, J. Mol. Graph. Model. (2006), doi:10.1016/j.jmgm.2006.10.003

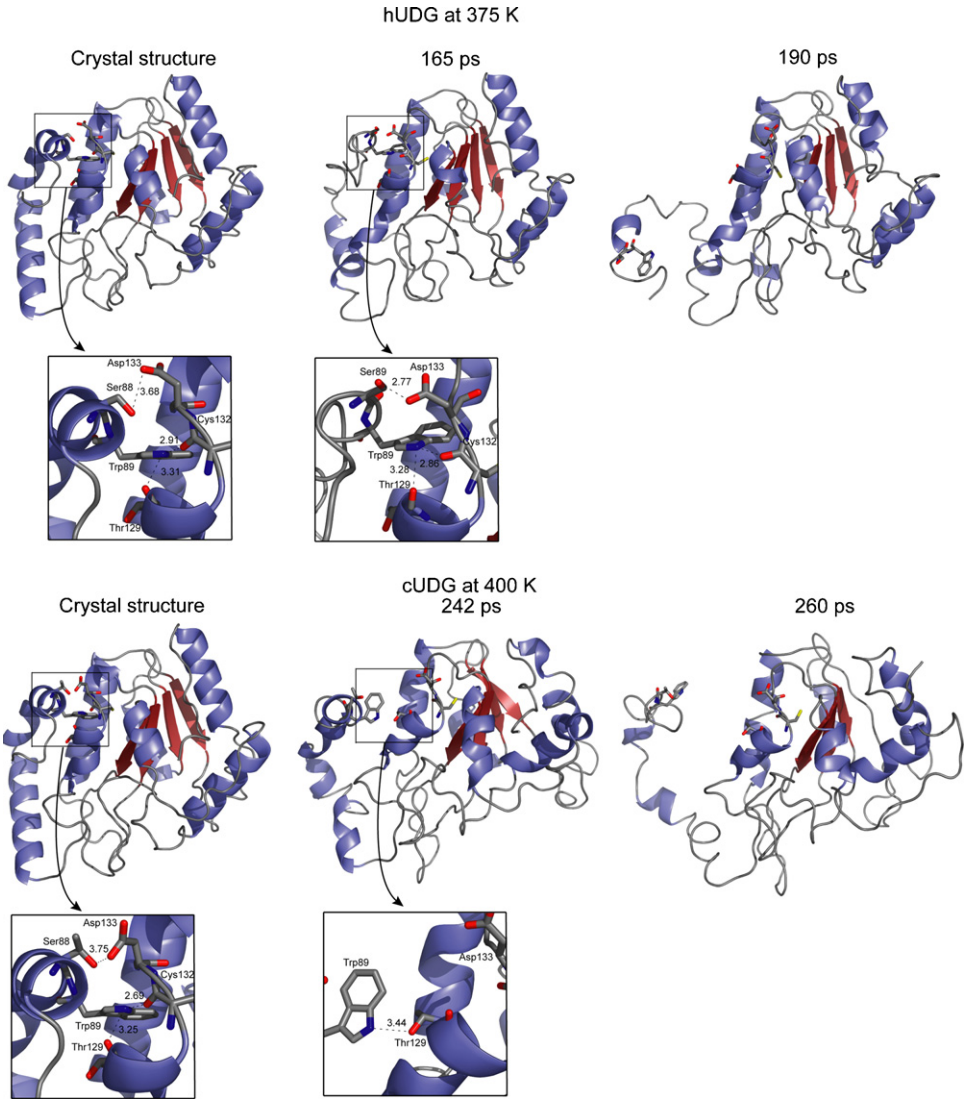


Fig. 4. Snapshots from the unfolding of hUDG at 375 K and cUDG at 400 K. The figure shows how the N-terminal unfolds from the rest of the structure. Some important residues are shown in ball-and-stick. The figure is generated in PyMOL [45].

N-terminal of helix α1 and the helix α4 and the loop between helix α4 and helix α5. In addition, the side chain of Trp89 packs into a hydrophobic area between helices α4 and α8.

Subtle differences in the hydrogen-bonding pattern observed during the simulations of cUDG and hUDG result in strikingly different behavior of their respective N-terminals. In the 375 K cUDG simulation the N-terminal starts to unfold at around 470 ps (Figs. 3A, 5A and 5B), as the previously mentioned hydrogen bond, Trp89:Nε1–Thr129:O, is lost. In contrast, the N-terminal starts to unfold at 177 ps in the 375 K simulation of hUDG, here the hydrogen bond between Ser88:Oγ and Asp133:Oδ2 is lost after 169 ps, and 8 ps later the Trp89:Nε1–Thr129:O hydrogen bond also disappear. Together, an immediate unfolding of the α1 and α2 helices is observed (Fig. 5A and B) in hUDG. Fig. 4 shows a detailed picture of the unfolding events occurring in the N-terminal of hUDG. The mesophilic enzyme starts to unfold earlier than the psychrophilic homologue at this temperature, but at around 2 ns the psychrophilic enzyme is more severely unfolded in the N-terminal.

When the temperature is increased to 400 K, the simulations show that the N-terminal of cUDG becomes highly flexible after only 70 ps. This is again a result of the loss of the Ser88:Oγ-Asp133:Oδ2 hydrogen bond. After the distortion of an additional hydrogen bond, Trp89:Nε1–Thr129:O, complete unfolding of the terminal is observed (Fig. 5D). Figs. 4 and 7 show the unfolding events observed for the N-terminal. Analysis of the simulation of hUDG at 400 K reveals that the hydrogen bond between the Ser88:Oγ and the Asp133:Oδ2 is lost after about 1025 ps (Fig. 5E), while the hydrogen bond between the Trp89:Nε1 and the Thr129:O disappear after 1075 ps (Fig. 5D). These events lead to a melting of the N-terminal in a similar manner as observed for cUDG.

In the 425 K simulation, molecular contacts are lost much faster than at lower temperatures, for example in the hUDG simulation important hydrogen bond in the N-terminal (Trp89:Nε1–Thr129:O) are lost after 110 ps and then the N-terminal starts to unfold (Fig. 5G). The corresponding in hydrogen bond in the psychrophilic enzyme is lost already after

Please cite this article in press as: M. Olufsen et al., Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme, J. Mol. Graph. Model. (2006), doi:10.1016/j.jmgm.2006.10.003

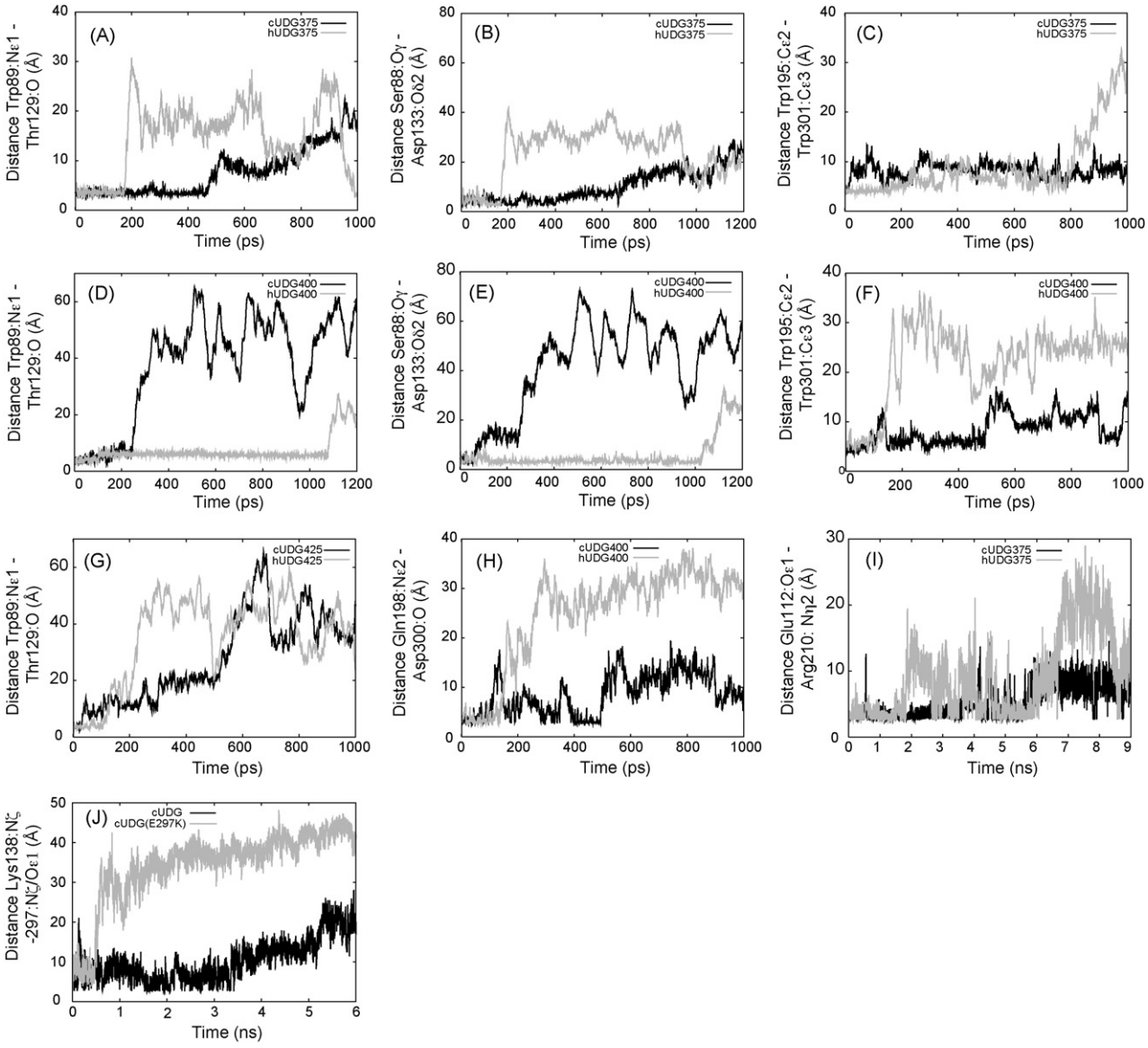


Fig. 5. Atom distance plots. Distances between selected atoms during simulation. (A), (D) and (G): distanced between the atoms Trp89:Nε1 and Thr129:O for the 375, 400 and 425 K simulation, respectively. (B) and (E): distance between the atoms Ser88:Oγ and Asp133:Oδ2 for the 375 and 400 K simulation, respectively. These above-mentioned atom pairs form hydrogen bonds close to the N-terminal. (C) and (F): distance between the atoms Trp195:Cε2 and Trp301:Cε3, for the 375 and 400 K simulation, respectively, which form a hydrophobic contact in the C-terminal. (H): distance between the atoms Gln198:Nε2 and Asp300:O in the 400 K simulation, these residues form a hydrogen bond in the C-terminal. (I): distance between atoms Glu112:Oε1 and Arg 210:Nη2. (J): distance between the atoms Lys138:Nζ and atom (Lys297:Nζ or Glu297:Oε1) in cUDG(E297K) and cUDG, respectively. Both simulations are done at 400 K simulation. There are different scales on the axes in the different plots.

25 ps (Fig. 5G). Many molecular contacts and interactions have been monitored in these simulations as well, and the unfolding of the two enzymes is very similar compared to the lower simulations temperatures. However, it is apparent that the kinetic energy of the system is so high that too many contacts are lost simultaneously making analysis of the simulations more difficult.

The different behavior of the N-terminal of cUDG compared to the hUDG is not obvious, but it is likely that subtle differences in the interactions contribute differently to the stability of the two N-terminals. In particular, the Trp89:Nε1–Thr129:O hydrogen bond seems to be important for N-terminal stability. Interestingly, the recently reported crystal structure of a thermophilic UDG from *Thermus thermophilus* HB8 (TthUDG) has a stabilizing [4Fe-4S] cluster close to the N-terminal [43]. TthUDG belongs to the family 4 UDG while hUDG and cUDG are from family 1. Even though the amino acid sequence homology is low between TthUDG and hUDG or cUDG, the topology and order of the secondary structure elements are very similar between the two UDG families [43]. High temperature adaptation thus seems to involve additional stabilization of the N-terminal part. Hence, increasing the stability of the N-terminal might make the mesophilic and the psychrophilic enzyme more thermostable. MD studies of Chymotrypsin inhibitor 2 have also shown that the N-terminal unfolds from the rest of the structure as one of the first events in

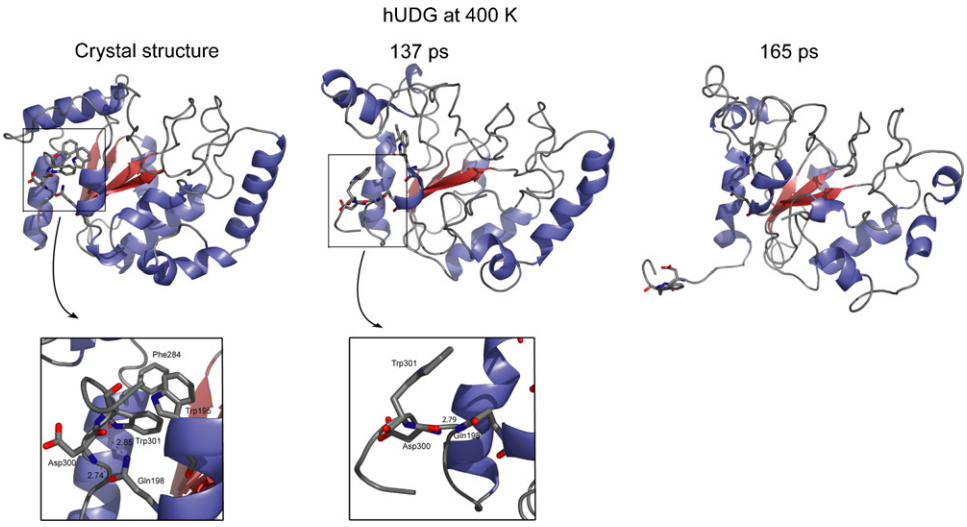


Fig. 6. Snapshots from the unfolding of hUDG at 400 K. The figure shows how the C-terminal unfolds from the rest of the structure. Some important residues are shown in ball-and-stick. The figure is generated in PyMOL [45].

the unfolding process [25,44], acting as an initiator of unfolding.

The very end of the C-terminal part (residue 293–304) is situated at the surface of the protein, and the side of the C-terminal facing away from the solvent is involved in a hydrophobic cluster with helices α7 and α11. In addition, the Trp301 side chain forms stacking interactions with the side chains of Phe284 and Trp195. There are also several hydrogen bonds that could be important for C-terminal stability, especially involving residue Gln198, forming hydrogen bond interaction with residue 300 (Asp in hUDG and Asn in cUDG) (Fig. 6). The fact that Trp301, Phe284, Trp195 and Gln198 are all conserved in the sequences studied by Leiros et al. [9], additionally indicates structural importance. The simulation shows that the C-terminal is far more stable in the cold-adapted enzyme when compared to its warm-active counterpart. The C-terminal of hUDG unfolds at all three simulations temperatures, while the C-terminal of cUDG only melts at 425 K.

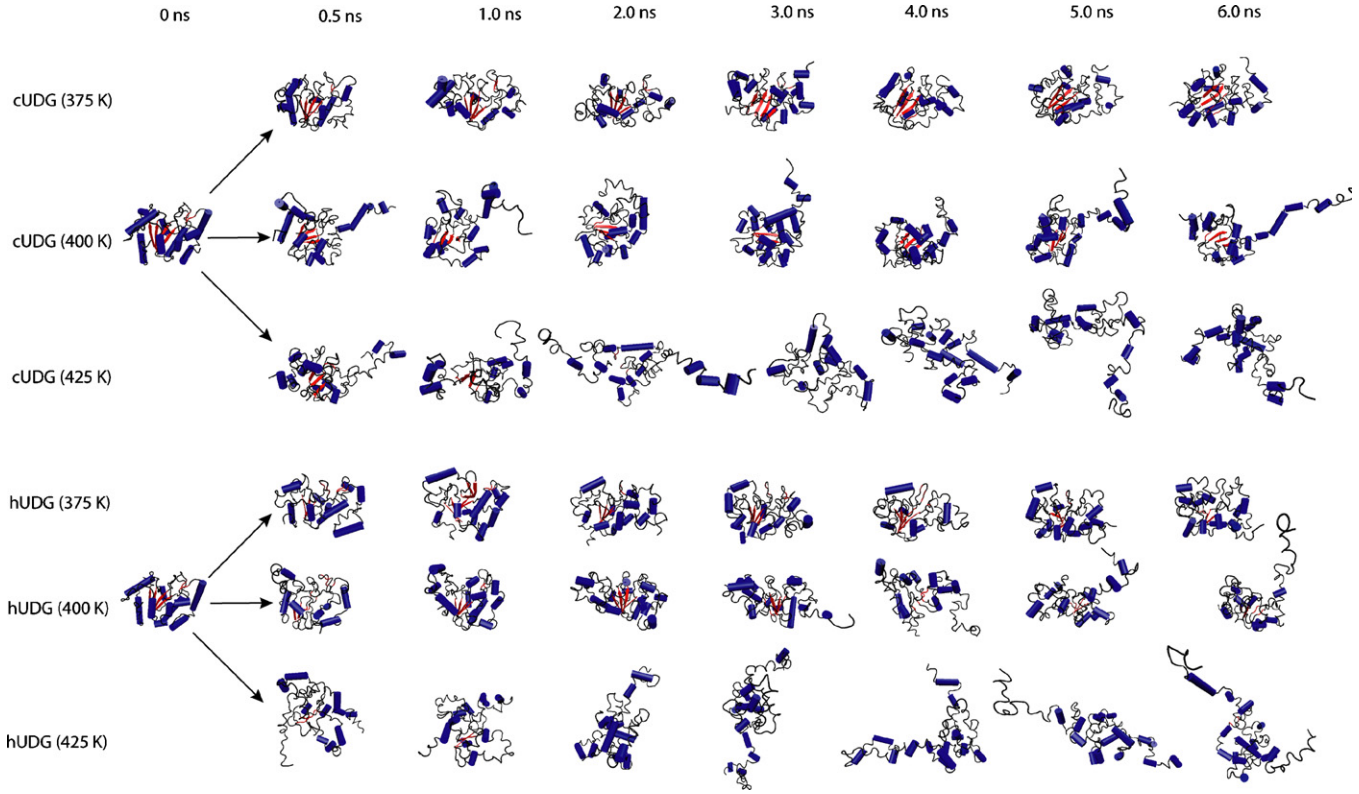


Fig. 7. Snapshots from the thermal unfolding simulations of cUDG and hUDG. The structures are made with the VMD program [46], helices, β-strands and loops are colored blue, red and gray, respectively.

Please cite this article in press as: M. Olufsen et al., Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme, J. Mol. Graph. Model. (2006), doi:10.1016/j.jmgm.2006.10.003

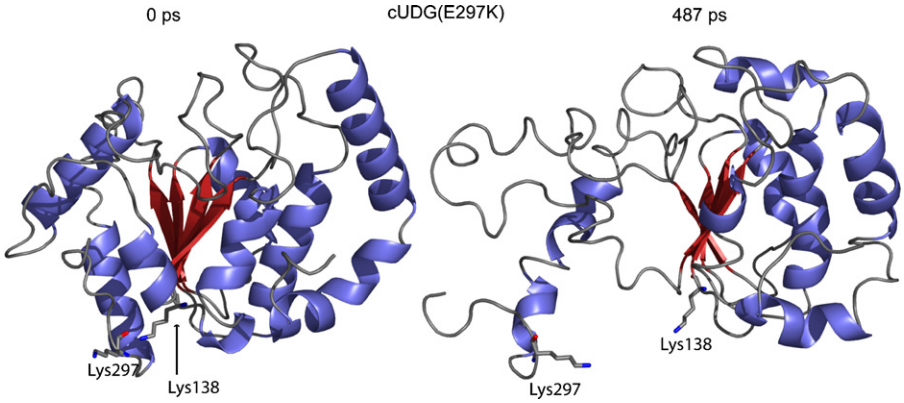


Fig. 8. Snapshots from the unfolding of cUDG(E297K) mutant at 400 K. The figure shows how the C-terminal unfolds from the rest of the structure. Some important residues are shown in ball-and-stick. The figure is generated in PyMOL [45].

Again, slightly different intramolecular interactions are responsible for the observed behavior in the C-terminal of cUDG and hUDG. The 375 K simulation of hUDG shows that around 450–500 ps, the hydrogen bonds between Gln198 and Asp300 disappear. Additional hydrophobic contacts in the C-terminal are also disrupted at around 800 ps, among others the Trp195–Trp301 side chain contact (Fig. 5C). These events seem to initiate melting of the C-terminal. The C-terminal of cUDG does not unfold at all at this temperature, but the last three residues have large fluctuations (Fig. 3A). Increasing the temperature to 400 K result in similar observations of the stability of the C-terminals. After only 140 ps several intermolecular contacts are lost, among others the hydrogen bond between Gln198 and residue 300 (Fig. 5H) and the hydrophobic contact between Trp195 and Trp301 (Fig. 5F). Together, these events lead to a rapid unfolding of the C-terminal as shown in Fig. 6. Again, the C-terminal is stable in the simulation of the psychrophilic enzyme, even though the interactions between residue Gln198 and Asn300 are not present after 500 ps. Favorable hydrophobic packing interactions between several residues in the C-terminal, among others the Trp195–Trp301 residues (Fig. 5F), is a possible explanation of this observation. The 425 K simulations show that both C-terminals unfold, but again with a higher stability in hUDG (Fig. 3E, F and 7).

Structural differences between the two enzymes may explain the differences observed in the stability of the C-terminal. First of all, residue 297 is a Lys in hUDG and a Glu in cUDG. In the crystal structure of cUDG this residue form a weak salt-bridge to Lys138 (5.4 Å between the Lys Nζ atom and the Glu Oε1 atom), while an unfavorable electrostatic interaction between two positively charged residues is observed in hUDG. To further investigate whether this explains the differences in unfolding of the C-terminal, two mutants of the enzymes were modeled from the crystal structures. The Lys297 residue in hUDG was mutated to a Glu and the Glu297 residue in cUDG was mutated to a Lys. Then 6 ns MD simulations of these two mutants were performed at 400 K. The C-terminal of the cUDG(E297K) mutant unfolds after ∼500 ps (Figs. 5J and 8), while the C-terminal of cUDG does not unfold at this temperature at all. This indicates that the weak salt bridge between Lys138 and Glu297 stabilizes the C-terminal in cUDG. The simulation of the hUDG (K297E) mutant shows that the enzyme still unfolds even when the mutations are introduced. Thus, other elements are also likely contributing to the observed difference in the unfolding of the C-terminal. From residue 293 to the C-terminal end, there are four residues in the sequence that make cUDG more hydrophobic than hUDG. The substitutions are K293L, K296T, D300N and E303A (mutation from hUDG to cUDG), all charged residues in hUDG, substituted to hydrophobic or uncharged in cUDG (for a full alignment see Leiros et al. [9]). It seems likely that the hydrophobic contacts are particularly important in the C-terminal, especially the hydrophobic/stacking contact between residue Trp195 and Trp301 (Fig. 5C and F). When this interaction is destroyed the C-terminal easily unfolds (Fig. 6).

6. Other differences in stability

Even if the largest differences in unfolding pathway among the two enzymes are in the terminals, there are also other parts of the structure possessing differences in unfolding pathways. In the 375 K simulation, the mesophilic enzyme unfolds to a high extent in the residues 207–226 region, whereas, the psychrophilic enzyme has a much lower r.m.s.d. per residue in this part of the structure (Fig. 3A and B). One possible reason for this difference is that the salt-bridge between residue Glu112 and Arg210 is not present for most of the simulation of hUDG whereas, it is intact for the first 5.5 ns of the simulation of cUDG (Fig. 5I).

In the 400 K simulation there are also other parts than the terminals of the structure that unfolds. Especially the segment from residue 205 to 220 unfolds rapidly for both enzymes, while in the 165–190 region the cUDG enzyme unfolds to a much higher extent than the mesophilic counterpart (Fig. 3C and D). There are several intermolecular interactions that may explain the difference in unfolding of this area. By analyzing atomic interactions trajectories it seems like residue 183, situated in a sharp turn, could explain much of the observed difference in this area. This residue is an Asp in hUDG and forms a salt bridge with Lys302 that is lost after 80 ps and does therefore probably not explain the observed differences. cUDG has a Gly at position 183 and one can therefore expect that the psychrophilic enzyme

Please cite this article in press as: M. Olufsen et al., Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme, J. Mol. Graph. Model. (2006), doi:10.1016/j.jmgm.2006.10.003

becomes more flexible in this area of the structure. In addition, there is a salt-bridge between Asp180 and Lys/Arg282 (Lys in cUDG and Arg in hUDG), which is lost at around 85 ps for both enzymes. This interaction is however reformed again in the hUDG structure, and is actually intact for large parts of the simulation. Hence, these specific interactions can explain the differences observed in the 165–190 area between the two enzymes. In the 425 K simulation both enzymes unfolds rapidly. After a short period of simulation, both cUDG and hUDG are highly distorted (Fig. 3E, F and 7).

7. Concluding remarks

The high sequence identity and structural similarity between cUDG and hUDG allows for the construction of various hypotheses that can be tested experimentally. Such studies (covering cloning, protein production, purification, characterization, etc.) are demanding in terms of costs and labor. Integrating computational approaches with different levels of complexity with the biochemical and/or biophysical studies, therefore allows in many cases for a more rationalized experimental process. In this study we have used molecular dynamics simulations at high temperatures to investigate not only the specific features of the unfolding mechanism of two homologous DNA repair enzymes, but also to gain a deeper understanding of the structurally important regions modulating thermal stability of the two. It is important to note that the objective of the present study is not by any means to try to solve the problem of protein folding. Instead, these simulations are used as tools to increase our knowledge of the structure–function relationship in uracil DNA glycosylase.

The simulations presented here (at 375 and 400 K) show that the psychrophilic UDG has lower thermal stability compared to the mesophilic UDG, which is in accordance with experiments [13]. The mesophilic enzyme needs higher temperature to reach the same rate of unfolding compared to cUDG. However, once the thermal energy of the systems is sufficiently high (i.e. 425 K in this case), rapid unfolding is observed for both enzymes. Subsequent analysis of the resulting trajectories provides indication of the structural regions important for protein stability. It is particularly the terminals that have different behavior between the cold- and warm-adapted UDG enzymes. Important molecular contacts, probably decisive for the stability of both terminals have been pin-pointed.

Acknowledgements

This work has been supported by a grant from the Research Council of Norway (RCN) (NFR 154197/432), and the Norwegian Structural Biology Centre (NorStruct) is supported by the National program in Functional Genomics (FUGE) in the Research Council of Norway.

References

[1] G. Feller, Molecular adaptations to cold in psychrophilic enzymes, Cell. Mol. Life Sci. 60 (2003) 648–662.
[2] T. Lonhienne, C. Gerday, G. Feller, Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility, Biochim. Biophys. Acta 1543 (2000) 1–10.
[3] D. Georlette, V. Blaise, T. Collins, S. D'Amico, E. Gratia, A. Hoyoux, J.C. Marx, G. Sonan, G. Feller, C. Gerday, Some like it cold: biocatalysis at low temperatures, FEMS Microbiol. Rev. 28 (2004) 25–42.
[4] C. Gerday, M. Aittaleb, M. Bentahir, J.P. Chessa, P. Claverie, T. Collins, S. D'Amico, J. Dumont, G. Garsoux, D. Georlette, A. Hoyoux, T. Lonhienne, M.A. Meuwis, G. Feller, Cold-adapted enzymes: from fundamentals to biotechnology, Trends Biotechnol. 18 (2000) 103–107.
[5] A.O. Smalås, H.K.S. Leiros, V. Os, N.P. Willassen, Cold Adapted Enzymes, vol. 6, Elsevier Science B.V., Amsterdam, 2000, pp. 1–57.
[6] O. Lanes, I. Leiros, A.O. Smalås, N.P. Willassen, Identification, cloning, and expression of uracil-DNA glycosylase from Atlantic cod (gadus morhua): Characterization and homology modeling of the cold-active catalytic domain, Extremophiles 6 (2002) 73–86.
[7] T. Lindahl, B. Nyberg, Heat-induced deamination of cytosine residues in deoxyribonucleic-acid, Biochemistry 13 (1974) 3405–3410.
[8] C.D. Mol, A.S. Arvai, G. Slupphaug, B. Kavli, I. Alseth, H.E. Krokan, J.A. Tainer, Crystal-structure and mutational analysis of human uracil-DNA glycosylase—structural basis for specificity and catalysis, Cell 80 (1995) 869–878.
[9] I. Leiros, E. Moe, O. Lanes, A.O. Smalås, N.P. Willassen, The structure of uracil-DNA glycosylase from Atlantic cod (gadus morhua) reveals cold-adaptation features, Acta Crystallogr. D 59 (2003) 1357–1365.
[10] R. Savva, K. Mcauleyhecht, T. Brown, L. Pearl, The structural basis of specific base-excision repair by uracil-DNA glycosylase, Nature 373 (1995) 487–493.
[11] R. Ravishankar, M.B. Sagar, S. Roy, K. Purnapatre, P. Handa, U. Varshney, M. Vijayan, X-ray analysis of a complex of *Escherichia coli* uracil DNA glycosylase (ecudg) with a proteinaceous inhibitor. The structure elucidation of a prokaryotic UDG, Nucl. Acids Res. 26 (1998) 4880–4887.
[12] M. Olufsen, A.O. Smalås, E. Moe, B.O. Brandsdal, Increased flexibility as a strategy for cold adaptation—a comparative molecular dynamics study of cold- and warm-active uracil DNA glycosylase, J. Biol. Chem. 280 (2005) 18042–18048.
[13] O. Lanes, P.H. Guddal, D.R. Gjellesvik, N.P. Willassen, Purification and characterization of a cold-adapted uracil-DNA glycosylase from Atlantic cod (gadus morhua), Comp. Biochem. Physiol. 127 (2000) 399–410.
[14] A.R. Fersht, V. Daggett, Protein folding and unfolding at atomic resolution, Cell 108 (2002) 573–582.
[15] R. Scandurra, V. Consalvi, R. Chiaraluce, L. Politi, P.C. Engel, Protein stability in extremophilic archaea, Front. Biosci. 5 (2000) D787–D795.
[16] K.A. Dill, Dominant forces in protein folding, Biochemistry 29 (1990) 7133–7155.
[17] S. Kumar, C.J. Tsai, R. Nussinov, Maximal stabilities of reversible two-state proteins, Biochemistry 41 (2002) 5359–5374.
[18] C.N. Pace, Polar group burial contributes more to protein stability than non-polar group burial, Biochemistry 40 (2001) 310–313.
[19] C.N. Pace, R.W. Alston, K.L. Shaw, Charge-charge interactions influence the denatured state ensemble and contribute to protein stability, Protein Sci. 9 (2000) 1395–1398.
[20] S. Kumar, R. Nussinov, Salt bridge stability in monomeric proteins, J. Mol. Biol. 293 (1999) 1241–1255.
[21] A.J. Li, V. Daggett, Characterization of the transition-state of protein unfolding by use of molecular-dynamics—chymotrypsin inhibitor-2, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 10430–10434.
[22] A.J. Li, V. Daggett, Molecular dynamics simulation of the unfolding of barnase: characterization of the major intermediate, J. Mol. Biol. 275 (1998) 677–694.
[23] A. Caflisch, M. Karplus, Acid and thermal-denaturation of barnase investigated by molecular-dynamics simulations, J. Mol. Biol. 252 (1995) 672–708.
[24] U. Mayor, N.R. Guydosh, C.M. Johnson, J.G. Grossmann, S. Sato, G.S. Jas, S.M.V. Freund, D.O.V. Alonso, V. Daggett, A.R. Fersht, The complete folding pathway of a protein from nanoseconds to microseconds, Nature 421 (2003) 863–867.

[25] R. Day, B.J. Bennion, S. Ham, V. Daggett, Increasing temperature accelerates protein unfolding without changing the pathway of unfolding, J. Mol. Biol. 322 (2002) 189–203.
[26] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham, S. Debolt, D. Ferguson, G. Seibel, P. Kollman, Amber a package of computer-programs for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules, Comput. Phys. Commun. 91 (1995) 1–41.
[27] J.M. Wang, P. Cieplak, P.A. Kollman, How well does a restrained electrostatic potential (resp) model perform in calculating conformational energies of organic and biological molecules? J Comput. Chem. 21 (2000) 1049–1074.
[28] A. Onufriev, D. Bashford, D.A. Case, Modification of the generalized born model suitable for macromolecules, J. Phys. Chem. B. 104 (2000) 3712–3720.
[29] J. Karanicolas, C.L. Brooks, Integrating folding kinetics and protein function: Biphasic kinetics and dual binding specificity in a ww domain, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 3432–3437.
[30] Y.Z. Ohkubo, C.L. Brooks, Exploring flory's isolated-pair hypothesis: statistical mechanics of helix-coil transitions in polyalanine and the c-peptide from rnase a, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 13916–13921.
[31] M. Feig, C.L. Brooks, Recent advances in the development and application of implicit solvent models in biomolecule simulations, Curr. Opin. Struc. Biol. 14 (2004) 217–224.
[32] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, Numerical-integration of cartesian equations of motion of a system with constraints—molecular-dynamics of *n*-alkanes, J. Comput. Phys. 23 (1977) 327–341.
[33] H.J.C. Berendsen, J.P.M. Postma, W.F. Vangunsteren, A. Dinola, J.R. Haak, Molecular-dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[34] D. Sitkoff, K.A. Sharp, B. Honig, Accurate calculation of hydration free-energies using macroscopic solvent models, J. Phys. Chem. 98 (1994) 1978–1988.
[35] I. Leiros, O. Lanes, O. Sundheim, R. Helland, A.O. Smalås, N.P. Willassen, Crystallization and preliminary X-ray diffraction analysis of a cold-adapted uracil-DNA glycosylase from Atlantic cod (gadus morhua), Acta Crystallogr. D 57 (2001) 1706–1708.
[36] A.R. Dinner, G.M. Blackburn, M. Karplus, Uracil-DNA glycosylase acts by substrate autocatalysis, Nature 413 (2001) 752–755.
[37] W. Kabsch, C. Sander, Dictionary of protein secondary structure—pattern-recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[38] M.F. Sanner, A.J. Olson, J.C. Spehner, Reduced surface: an efficient way to compute molecular surfaces, Biopolymers 38 (1996) 305–320.
[39] E. Moe, I. Leiros, E.K. Riise, M. Olufsen, O. Lanes, A. Smalas, N.P. Willassen, Optimisation of the surface electrostatics as a strategy for cold adaptation of uracil-DNA *n*-glycosylase (UNG) from Atlantic cod (gadus morhua), J. Mol. Biol. 343 (2004) 1221–1230.
[40] V. Daggett, A. Fersht, The present view of the mechanism of protein folding, Nat. Rev. Mol. Cell Biol. 4 (2003) 497–502.
[41] V. Daggett, Molecular dynamics simulations of the protein unfolding/folding reaction, Acc. Chem. Res. 35 (2002) 422–429.
[42] C.J. Bond, K.B. Wong, J. Clarke, A.R. Fersht, V. Daggett, Characterization of residual structure in the thermally denatured state of barnase by simulation and experiment: description of the folding pathway, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 13409–13413.
[43] J. Hoseki, A. Okamoto, R. Masui, T. Shibata, Y. Inoue, S. Yokoyama, S. Kuramitsu, Crystal structure of a family 4 uracil-DNA glycosylase from *Thermus thermophilus* hb8, J. Mol. Biol. 333 (2003) 515–526.
[44] A.J. Li, V. Daggett, Identification and characterization of the unfolding transition state of chymotrypsin inhibitor 2 by molecular dynamics simulations, J. Mol. Biol. 257 (1996) 412–429.
[45] W.L. DeLano, The Pymol Molecular Graphics System, DeLano Scientific, San Carlos, CA, USA, 2002.
[46] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.

Please cite this article in press as: M. Olufsen et al., Comparative unfolding studies of psychrophilic and mesophilic uracil DNA glycosylase: MD simulations show reduced thermal stability of the cold-adapted enzyme, J. Mol. Graph. Model. (2006), doi:10.1016/j.jmgm.2006.10.003

Paper IV

Paper V