

## Molecular dynamics simulations as a tool for improving protein stability

Mariël G. Pikkemaat<sup>1</sup>, Antonius B.M. Linsen<sup>2</sup>,  
Herman J.C. Berendsen<sup>2</sup> and Dick B. Janssen<sup>1,3</sup>

<sup>1</sup>Laboratory of Biochemistry and <sup>2</sup>Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands

<sup>3</sup>To whom correspondence should be addressed.  
E-mail: D.B.Janssen@chem.rug.nl

**Haloalkane dehalogenase (DhlA) was used as a model protein to explore the possibility to use molecular dynamics (MD) simulations as a tool to identify flexible regions in proteins that can serve as a target for stability enhancement by introduction of a disulfide bond. DhlA consists of two domains: an  $\alpha/\beta$ -hydrolase fold main domain and a cap domain composed of five  $\alpha$ -helices. MD simulations of DhlA showed high mobility in a helix–loop–helix region in the cap domain, involving residues 184–211. A disulfide cross-link was engineered between residue 201 of this flexible region and residue 16 of the main domain. The mutant enzyme showed substantial changes in both thermal and urea denaturation. The oxidized form of the mutant enzyme showed an increase of the apparent transition temperature from 47.5 to 52.5°C, whereas the  $T_{m,app}$  of the reduced mutant decreased by more than 8°C compared to the wild-type enzyme. Urea denaturation results showed a similar trend. Measurement of the kinetic stability showed that the introduction of the disulfide bond caused a decrease in activation free energy of unfolding of 0.43 kcal mol<sup>-1</sup> compared to the wild-type enzyme and also indicated that the helix–loop–helix region was involved early in the unfolding process. The results show that MD simulations are capable of identifying mobile protein domains that can successfully be used as a target for stability enhancement by the introduction of a disulfide cross-link.**

**Keywords:** disulfide bond engineering/flexible region/  
haloalkane dehalogenase/molecular dynamics simulation/  
protein stability

### Introduction

One of the main targets in protein engineering is increasing enzyme stability. The stability of proteins is determined by a number of interactions, such as hydrophobic and electrostatic interactions and hydrogen bonding. Rational design of mutants in order to increase the stability of a protein essentially aims at either improving stabilizing interactions or reducing potentially destabilizing factors.

Because of its large stabilization potential and relatively well defined conformational requirements, the introduction of disulfide bonds is an attractive method for improving stability, but this strategy has been applied with mixed success. Stability enhancement caused by the introduction of a disulfide is generally attributed to a lowered entropy of the unfolded state (Flory, 1956). This chain entropy model theoretically allows

a quantitative prediction of the effect of the introduction of a disulfide bond on the stability (Pace *et al.*, 1988). However, Doig and Williams (Doig and Williams, 1991) claim that enthalpic changes of the unfolded state also contribute to the increased stability caused by a cross-link and an increasing amount of evidence indicates that stabilizing effects on the folded state also play a role (Kuroki *et al.*, 1992; Tidor and Karplus, 1993; Clarke *et al.*, 1995a; Vogl *et al.*, 1995). In addition to the classical thermodynamic approach, Clarke and Fersht (Clarke and Fersht, 1993) showed that an enzyme can be stabilized kinetically by introducing disulfide bonds between regions involved in the early stages of unfolding. Mansfeld *et al.* (Mansfeld *et al.*, 1997) also attributed the extreme stabilization observed in a thermolysin-like protease to a disulfide bond introduced in an area involved in the partial unfolding processes that determine thermal stability.

Although the introduction of disulfide cross-links theoretically should have a stabilizing effect, the opposite is often observed. This destabilization can be attributed to negative effects on packing and flexibility, causing unfavorable enthalpic contributions to offset the entropic profit gained by the cross-link. With the disposition of high-resolution structures, substantial information has become available concerning the spatial requirements of disulfide bonds (Thornton, 1981). This resulted in the development of computer programs that can predict suitable sites for their introduction (Pabo and Suchanek, 1986; Hazes and Dijkstra, 1988; Sowdhamini *et al.*, 1989). Although these programs have facilitated the search for potentially successful substitutions to a large extent, introduction of a stabilizing disulfide bond remains a matter of trial and error if it is not possible to identify regions in the enzyme that are important for stabilization. The earliest reports on successful introductions of disulfide bonds exploited native unpaired cysteine residues, making the choice for the position of the second cysteine rather straightforward (Villafranca *et al.*, 1983; Perry and Wetzel, 1984). When stabilization by *de novo* introduction of both cysteines was shown to be feasible (Pantoliano *et al.*, 1987), the number of theoretically possible disulfide bonds again outnumbered the practical possibilities. For haloalkane dehalogenase (DhlA) the program SSBOND (Hazes and Dijkstra, 1988) found as many as 59 amino acid pairs that fulfilled the positional constraints for introduction of a disulfide bond (see Results). Comparison with related, sometimes more thermostable variants of the protein can lead to a successful choice (Takagi *et al.*, 1990; Yamaguchi *et al.*, 1996; Li *et al.*, 1998), but in most cases the selected location is based on an 'educated guess'.

Clearly, at present, a bottleneck in designing more stable proteins is the lack of accurate methods to determine the positions leading to an effective stabilization. Secondary structure elements such as helices and sheets will often complicate efficient disulfide bond formation because of their rather tight steric constraints (Thornton, 1981). Thus, in order to obtain a disulfide bond that positively contributes to the stability of a

protein, cross-linking should preferably involve a flexible region (Matsumura *et al.*, 1989). This will probably have a smaller destabilizing effect on the folded conformation since the adjacent residues are more likely to adjust in order to compensate for negative strain effects. Disulfide introduction in T4 lysozyme demonstrated that the most stabilizing disulfide cross-links were situated in flexible regions as indicated by the temperature factors of the native X-ray structure (Matsumura *et al.*, 1989). Subsequent analysis of the X-ray structure of one of these mutants (Pjura *et al.*, 1990) indeed revealed considerable shifts in main chain atoms of neighboring residues and only a slight strain on the disulfide bond itself. The cross-link did not impose rigidity on the residues that showed flexibility in the native enzyme. Weber *et al.* (Weber *et al.*, 1985) reported a similar observation after introduction of a chemical cross-link in ribonuclease A.

A powerful method to investigate dynamic properties of a protein at the atomic level is provided by molecular dynamics (MD) simulations. Modern computers allow simulations of several nanoseconds, which is long enough to determine the dominant contributions to atomic fluctuations. Therefore, MD simulations should provide an attractive means to identify flexible regions in proteins that can serve as a target for stability enhancement. Using Dhla as a model, we exploited the possibility of identifying such regions and studied the effect of cross-linking on stability.

Dhla from *Xanthobacter autotrophicus* GJ10 (Dhla) catalyzes the hydrolytic dehalogenation of a range of small halogenated alkanes (Keuning *et al.*, 1985). From the X-ray structure (Franken *et al.*, 1991) the enzyme appears to belong to the  $\alpha/\beta$ -hydrolase fold family. The members of this family share a common main domain, consisting of an eight-stranded  $\beta$ -sheet surrounded by  $\alpha$ -helices, on which the position of the catalytic residues is preserved (Ollis *et al.*, 1992). Most of the  $\alpha/\beta$ -hydrolase fold enzymes show an excursion between  $\beta$ -strands 6 and 7, which appears to be highly diverse in size and structure. In Dhla this second domain consists of five  $\alpha$ -helices and is called the cap domain, since it more or less shelters the active site that is located in a hydrophobic cavity on the interface of the two domains. Previous work on Dhla showed that the cap domain is very important in determining substrate specificity (Pries *et al.*, 1994; Schanstra *et al.*, 1996, 1997) and also indicated that flexibility of this domain is probably controlling the release of the halide ion, which is the main rate-limiting reaction step in the conversion of 1,2-dichloroethane (Schanstra and Janssen, 1996; Krooshof *et al.*, 1999).

In this report we describe the results of the MD simulation of Dhla with a chloride ion bound in the active site. The results indicate the presence of a mobile helix-loop-helix region in the cap domain, involving residues 184–211. Introduction of a disulfide bond between residues 201 and 16, anchoring this flexible region to the main domain, led to an enhanced stability towards both thermal and chemical denaturation.

## Materials and methods

### MD simulations

The MD simulation was performed using the GROMOS87 forcefield (van Gunsteren and Berendsen, 1987) with adjusted Lennard-Jones interactions between aliphatic carbon atoms and water oxygen atoms (van Buuren *et al.*, 1993) and explicit hydrogens added to aromatic residues (van der Spoel *et al.*,

1996b). Polar hydrogens were included explicitly, whereas non-polar hydrogens were implicitly included by the use of united atoms. For all but the initial simulations, a simple form of polarization was included in the force field. This was necessary because in an initial simulation the chloride ion, present in the starting structure, left the active site almost instantaneously (within 20 ps). This is due to the fact that the interaction of the chloride ion with its hydrophobic environment is too weak in the force field used.

A direct charge-induced dipole polarization term:

$$V = -\frac{1}{2} \alpha_i f^2 \frac{e^2}{r_i^4}$$

was added to the potential energy for each atom of Glu56, Trp125, Phe128, Phe172, Trp175, Phe222, Pro223 and Val226, as these amino acids all contained at least one atom within a distance of 5 Å from the chloride ion in the X-ray structure. In this equation,  $\alpha_i$  is the polarizability of atom  $i$ ,  $r_i$  the distance of atom  $i$  to the chloride ion,  $e$  the elementary charge and  $f = 1/4\pi\epsilon_0$ . Atomic polarizabilities were taken from Applequist *et al.* (Applequist *et al.*, 1972), augmented with values of 1.39 and 0.32 Å<sup>3</sup> for aromatic C and H, respectively, and 1.40 and 0.23 Å<sup>3</sup> for amide N and H, respectively. These values were derived from polarizabilities of benzene, fluorobenzene, trimethylamine and ammonia (CRC, 1986).

The polarization term adds >30% to the stabilization energy of chloride (in a typical structure from the simulation the polarization energy added –26.8 kcal mol<sup>-1</sup> to the Coulomb energy of –72.9 kcal mol<sup>-1</sup> between chloride and protein atoms) and keeps the chloride ion in place.

As a starting structure, the X-ray structure at pH 6 with a resolution of 2.1 Å (Brookhaven Data Bank entry 2DHE) was used (Verschuere *et al.*, 1993b). To compensate for the charge of the chloride ion, His298 was assumed to be protonated at both ring nitrogens. All other histidines were assumed to be electrically neutral and protonated at only one ring nitrogen. To select the protonated nitrogen, the structure was checked for the presence of possible hydrogen bonds by looking at the closest hydrogen bond acceptor. Histidine residues 37 and 54 were protonated at the N<sub>δ1</sub> position, and residues 102 and 305 at the N<sub>ε2</sub> position. Periodic boundary conditions were applied by the use of a truncated octahedron filled with equilibrated water molecules. Each water molecule of which the oxygen atom had a distance to any non-hydrogen protein atom or water molecule of <2.3 Å was removed. The minimum distance between protein atoms and the walls of the box was taken as 7.5 Å and the volume of the box was 174 000 Å<sup>3</sup>. The protein (including the chloride ion) possessed a negative charge of –17 e. To obtain an electrically neutral system, 17 sodium ions were added as counterions. This was done by calculating the electric potential at the positions of all oxygen atoms of water molecules and replacing 17 water molecules with the lowest potential by a sodium ion. In total the system contained 16 047 atoms of which 3171 were protein atoms (including the chloride ion).

An energy minimization was performed using the steepest descent method for 100 steps without position restraining. During the first 10 ps of the simulation, harmonic position restraints were applied to all protein atoms, using a force constant of 9000 kJ nm<sup>-1</sup> mol<sup>-1</sup>. After that, the position restraints were removed and the system was initially simulated for 1 ns and later continued for another 500 ps.

Initial velocities were taken from a Maxwell–Boltzmann distribution at 298 K. The temperature was kept constant by coupling solute and solvent separately to a thermal bath at 298 K with a coupling constant  $\tau_T = 0.1$  ps. Pressure was kept constant by coupling to a pressure bath at 1.0 bar, using a coupling constant  $\tau_P = 0.5$  ps (Berendsen *et al.*, 1984). Bond lengths were constrained using the SHAKE method with a relative tolerance of 0.0001 (Ryckaert *et al.*, 1977). Non-bonded interactions were calculated using a twin cut-off radius: within a short cut-off radius of 8 Å Coulomb and Lennard–Jones interactions were calculated every time step, whereas Coulomb interactions within a radius of 12 Å were calculated every 10 steps. The step size was 0.002 ps. The behavior of the secondary structure elements was calculated with the program DSSP (Kabsch and Sander, 1983). B-factors derived from the trajectory were defined as:

$$B = \frac{8}{3} \pi^2 \langle |\Delta r|^2 \rangle$$

where  $\langle |\Delta r|^2 \rangle$  is the mean square atomic displacement relative to the average position.

The simulation was performed using the GROMOS87 software package (van Gunsteren and Berendsen 1987). Non-bonded routines were rewritten from scratch for use on a Cray J90 parallel computer. The simulation took ~200 h of real time per nanosecond. Analysis of secondary structure elements, root mean square deviation (RMSD) and B-factors was performed with GROMACS (van der Spoel *et al.*, 1996a).

#### Design of the disulfide bond

To select a suitable site for the introduction of a disulfide bond connecting the helix–loop–helix region and the main domain the disulfide bond prediction computer program SSBOND (Hazes and Dijkstra, 1988) was used. As a starting structure the X-ray structure at pH 8.2 and a resolution of 1.9 Å (Brookhaven Protein Data Bank entry 1EDE) was used (Verschuere *et al.*, 1993a). The program selects potential sites based upon a series of conformational and energy constraints derived from disulfide bonds in proteins for which a three-dimensional structure has been determined. The program allows user input to adjust the energy constraints, so disulfides that do not exactly match any of the known disulfide conformations can be generated. The DhIA mutant Asp16Cys/Ala201Cys was generated after allowing the  $C_{\beta 1}$ – $S_{\gamma 1}$ – $S_{\gamma 2}$ – $C_{\beta 2}$  dihedral angle to deviate  $\pm 10^\circ$ .

#### Mutagenesis

The single mutants Asp16Cys and Ala201Cys were obtained according to Kunkel (Kunkel, 1985). *Escherichia coli* strain BW313 was used for the production of single-stranded uracil containing DNA of plasmid pGELAF(+), an expression vector based on pET-3d (Studier *et al.*, 1990) with the dehalogenase gene under control of the T7 promoter (Schanstra *et al.*, 1993). The primer used for the Ala201Cys mutation was TGACCGAATGCGAGGC, with the mutated codon in bold and the introduced *BsmI* site underlined. For the Asp16Cys mutation the following primer was used: GCAATCTCTGCC-AGTATCC. Successful mutations could be detected by the absence of a *BsaBI* site caused by this primer. After separate introduction of each mutation the double mutant was obtained by ligating the appropriate *BclI*–*PstI* DNA fragments, which were isolated from *E. coli* strain KA1271, a *dam*<sup>−</sup> strain. The oligonucleotides for mutagenesis were supplied by Euro-

sequence BV (Groningen, The Netherlands). Sequences of mutated genes were established by the Center for Biomedical Technology, University of Groningen (The Netherlands).

#### Expression and purification

For the expression of the double mutant we used *E. coli* strain BL21(DE3) (Studier *et al.*, 1990), a strain expressing the T7 RNA polymerase of bacteriophage T7. After overnight growth at 30°C, the transformed *E. coli* cells were collected from plates, resuspended in 1 ml of LB medium, and used to inoculate 100 ml of LB (containing 100 µl/ml ampicillin) to an OD<sub>600</sub> of 0.05–0.1. This preculture was grown at 30°C until it reached an OD<sub>600</sub> of ~1.0 and was added to the 1 l main culture. At an OD<sub>600</sub> of 1.0, IPTG (0.4 mM) was added and growth was continued for 15 h at 20°C. Crude extract was prepared as described earlier (Schanstra *et al.*, 1993). After passage through a 0.2 µm filter, the cell extract was applied to a Resource™Q column (Pharmacia Biotech) on a Pharmacia liquid chromatography controller LCC-500. The buffer used during purification was TEMAG [10 mM Tris–SO<sub>4</sub>, pH 7.5, 1 mM EDTA, 1 mM β-mercaptoethanol, 3 mM sodium azide and 10% (v/v) glycerol]. Bound protein was eluted using TEMAG buffer with a gradient of 0–0.5 M ammonium sulfate. Purified protein was stored at 4°C in TEMAG without ammonium sulfate.

#### Verifying the formation of the disulfide bond

Purified protein was analyzed for disulfide bond formation by comparing reduced and non-reduced samples with SDS–polyacrylamide gel electrophoresis (PAGE) using Coomassie brilliant blue staining. Prior to the gel electrophoresis the protein samples were incubated at room temperature for at least 15 min in denaturation buffer [15 mM Tris–Cl (pH 6.8) containing 5% SDS, 25% glycerol and 0.05% bromophenol blue] which was supplemented with 15% β-mercaptoethanol if the reduced form of the protein was analyzed. Free sulfhydryl groups were quantitatively determined using 5,5'-dithiobis (2-nitrobenzoic acid) (DTNB) (Ellman, 1959). This compound reacts stoichiometrically with the free sulfhydryl groups of the enzyme and subsequent release of the aromatic thiolate can be followed quantitatively by measuring the absorbance at 412 nm. The reactions with DTNB were performed by mixing 900 µl of enzyme (0.25 mg/ml in TEAG) with 100 µl of 0.13% (w/w) DTNB in 50 mM Tris buffer, pH 8.2. Incubations at room temperature were continued until the absorbance at 412 nm had stabilized. The number of reacting sulfhydryl groups was calculated from the amount of released 2-nitro-5-thiobenzoic acid using a calibration curve with cysteine.

To establish the formation of the correct disulfide bond and for the identification of the DTNB-modified sulfhydryl groups, trypsin digestions were performed. Prior to the digestion, the protein was dialyzed against 0.2 M ammonium acetate/ammonium bicarbonate buffer (pH 8). Complete digestion of 1 nM protein was obtained after incubating for 4 h at 37°C using 1% (w/w) TCPK-grade trypsin (Worthington Biochemical Corp., Lakewood, NJ). Peptide fragments were analyzed by reversed-phase HPLC using a 250×4.6 mm Nucleosil 100 C854 column (Alltech, Deerfield, IL). For identification, peptide fragments containing the DTNB-modified sulfhydryl groups, liquid chromatography–mass spectrometry (LC–MS) was used as described earlier (Pries *et al.*, 1995).

#### Stability determination

Protein stability was analyzed by determining the residual activity as a function of temperature as well as urea concentra-



tion. Thermal stability was assayed by following the loss of activity during controlled increase of temperature. Identical aliquots of 50  $\mu\text{g}$  of enzyme in 100  $\mu\text{l}$  of Tris- $\text{SO}_4$  (pH 8.2) were heated in a thermocycler (Progene, Techne; New Brunswick Scientific, Edison, NJ) at a rate of  $1^\circ\text{C}/\text{min}$ . At definite temperatures a vial was removed and immediately cooled on ice, after which the residual activity was determined.

Urea denaturation was assayed by incubating 20  $\mu\text{g}$  of protein in 200  $\mu\text{l}$  of 50 mM Tris- $\text{SO}_4$  (pH 8.2) with a certain urea concentration. Residual activity was determined after an incubation period of 1.5 h at room temperature. For stability determination of the reduced mutant the Tris- $\text{SO}_4$  buffer was supplemented with 100 mM dithiothreitol (DTT).

Dehalogenase activities were determined by measuring the rate of alcohol production in time. Thermal and urea denaturation samples were incubated at  $30^\circ\text{C}$  with 5 mM DCE as substrate. After a fixed reaction period the alcohol product was extracted with diethylether containing 0.05 mM bromohexane as the internal standard. The ether layer containing alcohol and substrate was analyzed on a Chrompack 438S gas chromatograph using a CPWax 52CB column (Chrompack, Middelburg, The Netherlands), coupled to an FID detector (Schanstra *et al.*, 1996). Residual activity was expressed as a percentage of the initial activity and the apparent denaturation transition temperature was obtained using a sigmoid function from a curve-fitting program (Sigma Plot; Jandel Scientific, San Rafael, CA).

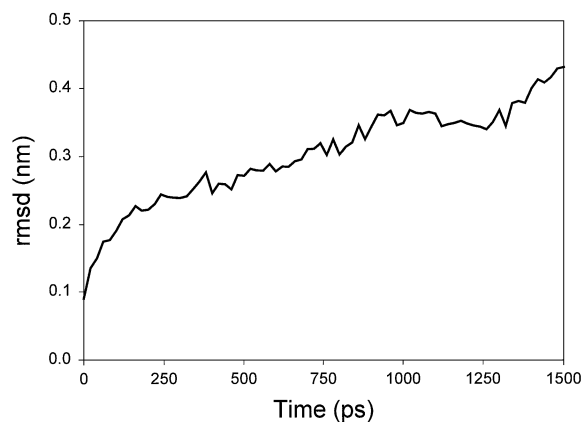
Kinetic stability was measured by following fluorescence changes after rapidly mixing the enzyme with denaturing urea concentrations. The experiments were performed on an Applied Photophysics model !SX17MV stopped-flow spectrofluorometer. The excitation wavelength was 290 nm and subsequent fluorescence emission was detected through a 320 nm cut-off filter. Reactions were performed at  $30^\circ\text{C}$  and initiated by mixing one volume of protein (1 mg/ml) with 10 volumes of a concentrated urea solution in Tris- $\text{SO}_4$  (pH 8.2). All fluorescence traces could be fitted to a single exponential from which the rate constant of unfolding  $k_u$  was obtained.

## Results

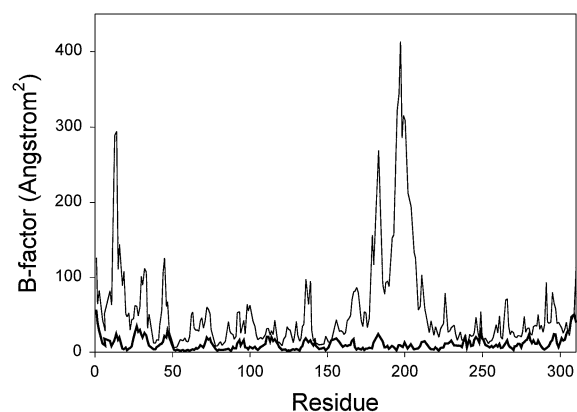
### MD simulations

To identify potentially flexible regions in Dh1A, an MD simulation of the enzyme was performed. During the first 1 ns of the simulation the RMSD from the crystal structure showed a steady increase. Since this might suggest that the structure was starting to unfold, the simulation was continued for another 500 ps. After 1 ns the RMSD stabilized around a value of  $\sim 0.35$  nm, to increase again after  $\sim 1.4$  ns (Figure 1). Close inspection of the trajectory revealed that after  $\sim 1.25$  ns some (charged) residues from the protein in the central periodic box started to interact with residues from one of the protein's images. The stabilization after 1 ns indicated that the structure remained intact, but we only considered the trajectory until 1.2 ns as being reliable and restricted further analysis to this time interval.

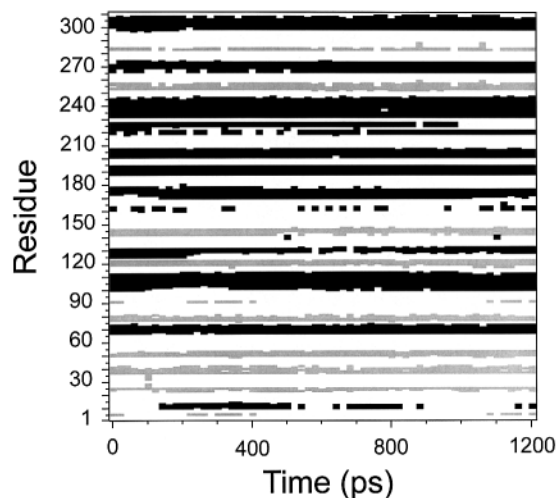
The stability of the enzyme was analyzed by inspecting two different properties: the RMSD of the molecule with respect to the X-ray structure and the stability of secondary structure elements. Figure 2 shows the B-factors derived from the simulation (for  $C_\alpha$  atoms only) compared to crystallographic B-factors. Most of the mean square displacement is concentrated in the region between residues 180 and 210 (comprising



**Fig. 1.** MD simulation of Dh1a. The RMSD from the crystal structure is plotted versus time. The trajectory beyond 1.25 ns is unreliable due to contacts between periodic images.

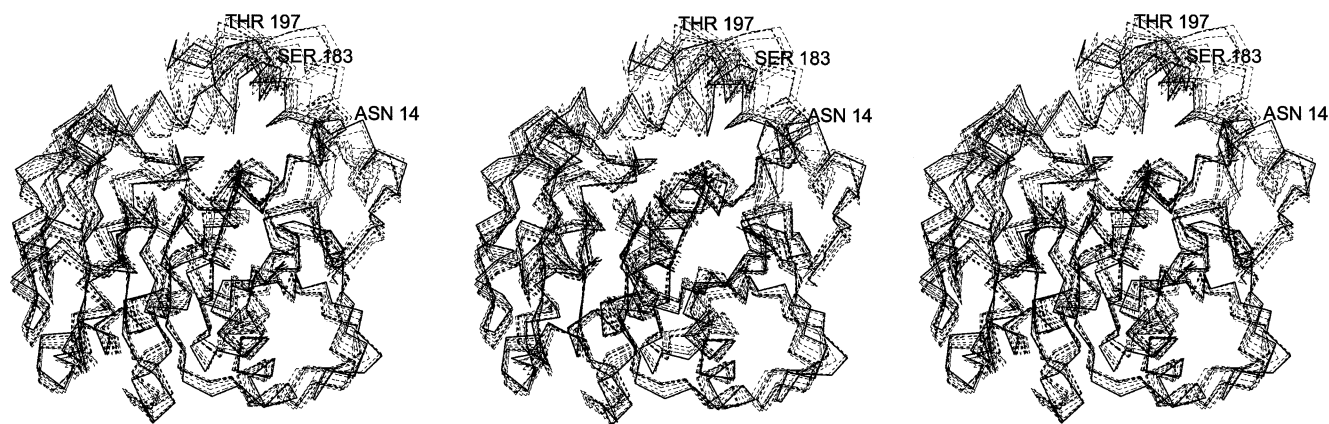


**Fig. 2.** B-factors from  $C_\alpha$  atoms: thick line, crystallographic B-factors; thin line, B-factors derived from the simulated trajectory.



**Fig. 3.** Analysis of the secondary structure elements versus time. The  $\alpha$ -helical structures are shown in black and  $\beta$ -sheets in grey.

helices 6 and 7 in the X-ray structure), with distinct peaks at Ser183 and Thr197. The crystallographic B-factors do not show any notable increase in this particular region. Analysis of secondary structure elements (Figure 3) shows that the helices in this region are highly stable, implying that the observed large RMSD values are not caused by unfolding of the structure. Therefore, it can be concluded that in the cap



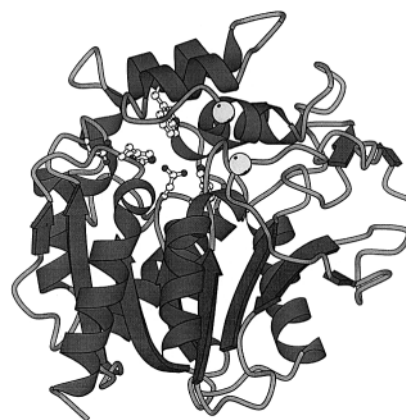
**Fig. 4.** Stereoview of the starting structure (solid line) and the same structure after 1.2 ns simulation (dashed line). Intermediate structures are projections on the first three essential degrees of freedom (see text).

domain of Dh1a relatively large motions occur and that this observed flexibility is inherent to the structure. The nature of the motions is illustrated in Figure 4, which shows a stereoview of the  $C_{\alpha}$ -backbone of the starting and final structure of the 1.2 ns simulation, as well as some intermediates. The structures shown are a projection of the trajectory onto the space of the first three eigenvectors of an essential dynamics analysis (Amadei *et al.*, 1993). The large motions appear to originate from an opening of the helix-loop-helix region between residues 185 and 211.

#### Site-directed mutagenesis

To study the effect of rigidifying the mobile part of the enzyme, a disulfide bond was introduced between this most flexible helix-loop-helix region and the main domain. Wild-type Dh1a contains four cysteine residues, located at positions 52, 150, 233 and 278, but none of them is involved in a disulfide bond. Using the computer program SSBOND (Hazes and Dijkstra, 1988) residue pairs were identified that might be capable of forming a disulfide bond when they would be replaced by a cysteine. A primary search using standard values for the structural parameters of the various torsion angles and S-S bond length revealed 59 possible candidates. Eight pairs had one or two residues within the mobile helix-loop-helix region, but none of them would result in a cap domain/main domain connection. This result reflects the small number of H bonds and electrostatic interactions between the two domains: in wild-type enzyme only 13 hydrogen bonds and four salt bridges are found (Verschuere *et al.*, 1993a). The main force to keep the domains together is probably of hydrophobic nature. A second search, in which the  $C_{\beta 1}-S_{\gamma 1}-S_{\gamma 2}-C_{\beta 2}$  dihedral angle was allowed to deviate  $10^{\circ}$  from its standard value of  $\pm 90^{\circ}$ , revealed 19 additional pairs. Two of them satisfied the desired positional criteria: the Asp16/Ala201 and the Ala195/Phe290 pair. Since Phe290 is flanking the conserved catalytic residue His289, changing this amino acid into a cysteine residue participating in a disulfide bond would probably affect the catalytic properties of the enzyme. Therefore, we chose to introduce a disulfide bond between the residue positions 16 and 201 (Figure 5).

The single mutations were introduced separately by Kunkel mutagenesis using wild-type single-stranded DNA as a template. Subsequent ligation of complementary *BclI*-*PstI* DNA fragments of pGELAF(+) yielded the double mutant Asp16Cys/Ala201Cys. The mutations did not affect the production of the enzyme. The mutant could be produced at expression



**Fig. 5.** Schematic representation of the structure of Dh1A. Residues D16 and A201, which were mutated to cysteines, are indicated with spheres. Active site residues are depicted as ball and sticks.

levels similar to that of wild-type enzyme and the mutations did not cause formation of insoluble aggregates, which was sometimes observed with other mutants when this T7 expression system was used (Schanstra *et al.*, 1993). The introduction of the two cysteine residues also did not have a negative effect on the catalytic activity of Dh1A and the reduced mutant exhibited similar specific activities for DCE and DBE as its oxidized counterpart.

#### Disulfide bond formation

After purification, the presence of a disulfide bond between the two mutated residues was confirmed by analyzing the mobility differences in non-reducing and reducing SDS-PAGE. The introduction of a disulfide bond is known to reduce the radius of gyration of the protein, causing the oxidized sample to move faster through the gel (Pollitt and Zalkin, 1983). Whereas under reducing conditions the bands of wild-type enzyme and mutant were indistinguishable, the non-reduced mutant enzyme moved faster through the gel, indicating the presence of a disulfide bond. Under these non-reducing conditions the wild-type reference sometimes showed multiple bands, indicating adventitious disulfide bond formation. Additional HPLC analysis of trypsin digests of the oxidized mutant enzyme showed the appearance of a large peak upon the loss of peak T3 and T18, which have been assigned to the fragments containing amino acids 12-35 and 194-220 (Krooshof, 2000). This indicated that the disulfide bond in the mutant was indeed

formed between Cys16 and Cys201. The identity of the adventitious bands in wild-type enzyme could not be established. SDS-PAGE analysis of mutant protein samples incubated with different reductant concentrations revealed that at least 100 mM  $\beta$ -mercaptoethanol was needed for complete reduction of the disulfide bond. Because disulfide bonds are rarely formed in *E.coli* cytosol due to the strongly reducing environment (Ziegler and Poulsen, 1977), we assume that the disulfide bond forms spontaneously during cell lysis and purification, since the  $\beta$ -mercaptoethanol concentration in the purification buffer is only 1 mM.

Detection of free thiol groups with DTNB (Ellman, 1959) revealed unexpected long-range structural consequences of the mutations. Although none of the native cysteine residues is located in the vicinity of the mutated residues, quantitative analysis showed that in the oxidized form of Asp16Cys/Ala201Cys only three out of four free sulfhydryl residues were modified with DTNB, whereas in the wild-type enzyme all four cysteine residues were reactive. DTNB reacts with the conjugate base of a thiol, so the rate of reaction varies with the pH and the  $pK_a$  of the thiol. Apart from the presence of a disulfide bond, non-stoichiometric reactivity can also be caused by inaccessibility of the sulfhydryl group, for example, caused by a buried position inside the protein or complex formation with metal ligands (Chu *et al.*, 1996; Pregel and Storer, 1997). To determine which of the cysteine residues was no longer capable of reacting with DTNB, we analyzed trypsin digests of the DTNB-modified mutant enzyme with LC-MS. Only the DTNB-modified peptide fragments T4 and T25, containing Cys52 and Cys278, respectively, could be detected. This implies that either Cys150 or Cys233 is in the oxidized form of the mutant enzyme no longer accessible for DTNB modification. This non-reacting sulfhydryl residue could very well be one of the residues involved in the adventitious disulfide bond formation observed in the wild-type enzyme, since this never occurred in the oxidized mutant.

#### Thermostability

The thermal denaturation of DhIA is not a reversible process (Krooshof, 2000). This irreversibility precludes a formal and detailed study of the thermodynamics of unfolding. Therefore, thermal denaturation of wild-type enzyme and both forms of the mutant enzyme was analyzed by measuring the residual activity of identical samples that were subjected to a gradual temperature increase. To be able to compare the oxidized and reduced form of the mutant with wild-type enzyme, the apparent denaturation transition value  $T_{m,app}$  was defined as the temperature at which the residual activity had decreased to 50% of the initial value.

The denaturation curves show that the introduction of the disulfide bond leads to a substantially increased thermostability (Figure 6). The wild-type enzyme shows an apparent transition temperature of  $47.5 \pm 0.5^\circ\text{C}$ , whereas the  $T_{m,app}$  of the oxidized mutant has increased to  $52.7 \pm 0.5^\circ\text{C}$ . These results are in good agreement with temperature-induced denaturation monitored by circular dichroism (not shown), indicating that no renaturation takes place during the assay process. The reduced mutant enzyme is obviously less stable than the wild-type enzyme and has a  $T_{m,app}$  of  $39.0 \pm 0.5^\circ\text{C}$ . The slope of the curve, which is a measure of the cooperativity of the unfolding process, is significantly lower for the reduced mutant enzyme. For a reversible process the slope at  $T_m$  is proportional to the unfolding enthalpy, but since the denaturation of DhIA is

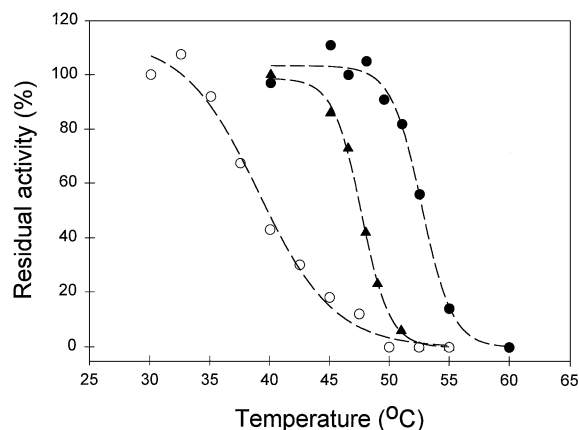


Fig. 6. Thermal denaturation profiles of wild-type DhIA (triangles), the oxidized form of the D16C/A201C mutant (closed circles) and reduced form of D16C/A201C (open circles).

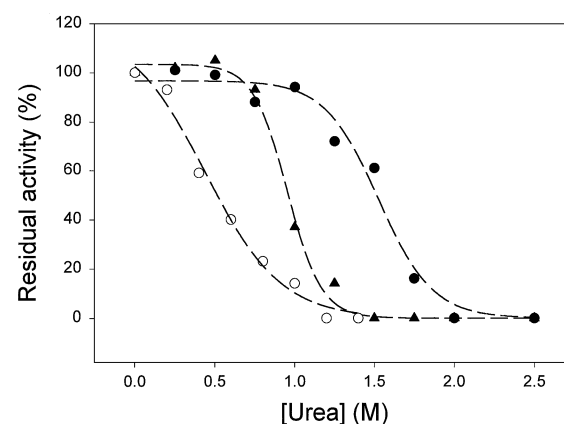


Fig. 7. Urea denaturation curves for wild-type DhIA (triangles), the oxidized form of the D16C/A201C mutant (closed circles) and reduced form of D16C/A201C (open circles).

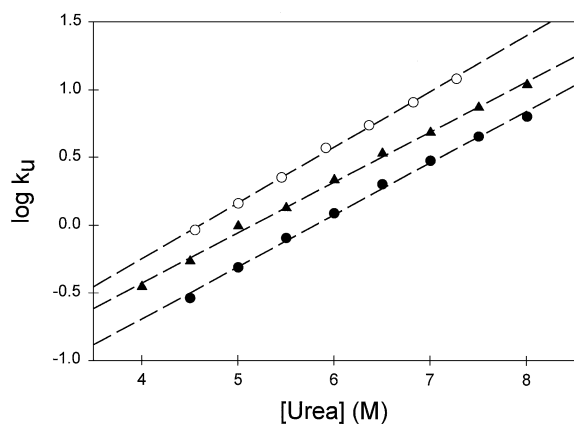
not reversible, the thermodynamic parameters of the unfolding process cannot be determined.

#### Urea unfolding

Since products of thermal denaturation are often less extensively unfolded than those of urea unfolding (Pace, 1975), the stability of DhIA was further examined by analysis of the residual activity of enzyme samples exposed to various concentrations of urea (Figure 7). Under the applied assay conditions, wild-type enzyme showed a rather sharp transition between 0.75 and 1.5 M urea, whereas the reduced form of the Asp16Cys/Ala201Cys mutant already started unfolding at very low urea concentrations. However, due to a much more gradual transition, complete unfolding occurred at only slightly lower urea concentrations than for wild-type enzyme. The oxidized form of the mutant was considerably more stable. This enzyme started unfolding at 1.25 M and for complete unfolding a concentration of 2 M urea was required.

The irreversibility of unfolding that was observed upon thermal denaturation, was also encountered using the chemical unfolding procedure. Attempts to refold urea-denatured protein by dialysis against TEMAG revealed that this irreversibility was mainly caused by incorrect disulfide bond formation between the native sulfhydryl groups. In TEMAG containing 1 mM  $\beta$ -mercaptoethanol, a reactivation of 27% was observed for wild-type enzyme, whereas for the oxidized Asp16Cys/





**Fig. 8.** Effect of urea on the rate of unfolding ( $k_u$ ) of wild-type DhIA (triangles), the oxidized form of the D16C/A201C mutant (closed circles) and reduced form of D16C/A201C (open circles).

Ala201Cys mutant only 12% of the initial activity could be restored. When unfolding and dialysis of the wild-type enzyme took place under strongly reducing conditions (500 mM DTT), as much as 80% of the enzymatic activity could be regained. This shows that the enzyme essentially is capable of refolding but that this refolding process is very vulnerable to improper disulfide bond formation.

#### Kinetic stability

The mobility that was observed in the MD simulation of DhIA indicates that helices 6 and 7 of the cap domain may be involved in the early stages of unfolding or even may represent the area where unfolding is initiated. Engineered disulfide bonds can be used as probes to follow unfolding and refolding and they are able to decrease the rate of unfolding by stabilizing early-unfolding elements (Clarke and Fersht, 1993). Therefore, we investigated the kinetic stability of wild-type enzyme and both forms of the Asp16Cys/Ala201Cys mutant.

Unfolding rates were measured using stopped-flow fluorescence. Rapid mixing of the enzyme with denaturing urea concentrations revealed single exponential fluorescence traces from which the rate constant of unfolding ( $k_u$ ) could be obtained. The dependence of the rate constant for unfolding on urea concentrations can be described as (Clarke and Fersht, 1993):

$$\log k_u = \log k_u^{\text{H}_2\text{O}} + m_{k_u} [\text{urea}] \quad (1)$$

with  $k_u^{\text{H}_2\text{O}}$  as the rate constant for unfolding in the absence of urea. The experimental plots of  $\log k_u$  versus the urea concentrations show that for the reduced Asp16Cys/Ala201Cys mutant enzyme the unfolding rates have doubled compared to the wild-type enzyme (Figure 8). However, the oxidized mutant unfolds at all urea concentrations ~2-fold slower than the wild-type enzyme, indicating an increased stability. The value of  $m_{k_u}^{\text{H}_2\text{O}}$ , which is a measure of the dependence of  $\log k_u$  on urea concentration (Equation 1), does not significantly differ between the three species. For a reversible unfolding process,  $m$  depends on the size and composition of the polypeptide chain that is freshly exposed to solvent by unfolding (Greene and Pace, 1974) and reflects the cooperativity of the transition.

Since the rate of unfolding is related to the difference in activation free energy between the folded protein and the transition state of unfolding,  $k_u$  can be used to compare the activation free energy of unfolding ( $\Delta G_{\ddagger-F}$ ) of two proteins A and B:

$$\Delta\Delta G_{\ddagger-F} = \Delta G_{\ddagger-F}^A - \Delta G_{\ddagger-F}^B = -RT \ln(k_u^A / k_u^B) \quad (2)$$

The increase in activation free energy  $\Delta\Delta G_{\ddagger-F}$  caused by the introduction of the disulfide bond is 0.43 kcal mol<sup>-1</sup> compared to the wild-type enzyme and 0.88 kcal mol<sup>-1</sup> compared to the reduced form of the mutant.

#### Discussion

We successfully enhanced the stability of DhIA by introducing a disulfide cross-link in a mobile region of the enzyme identified by MD simulations. Only a few examples of thermostability enhancements in  $\alpha/\beta$ -hydrolase fold enzymes have been reported, all of them concerning lipases. The increase in stability was achieved either by random mutagenesis (Shinkai *et al.*, 1996) or the introduction of a consensus sequence found in most lipases (Patkar *et al.*, 1998). Stability enhancement by the introduction of a disulfide bond, based on comparison with the X-ray structures of three disulfide bond containing extracellular lipases, was reported for *Penicillium camembertii* lipase (Yamaguchi *et al.*, 1996).

Although crystallographic B-factors can also provide information on flexibility, the mobility of helices 6 and 7 detected in the MD simulations could not be observed from these B-factors (Figure 2). This is probably caused by the structural arrangement of the molecules in the crystal lattice (Verschuere *et al.*, 1993a). The great majority of crystal contacts involve cap domain residues interacting with main domain amino acids of a neighboring molecule. This results in B-factors similar to those of the  $\beta$ -sheet; the average B-factors of the  $\beta$ -sheet and of the cap domain helices are, respectively, 11.5 and 10.0 Å<sup>2</sup>, whereas the helices of the main domain have a mean B-value of 15.3 Å<sup>2</sup>. Solvent accessibility data of the cap domain already gave an indication that the observed rigidity was at least partly artificial, since these values did not show the usual correlation with crystallographic B-factors (Verschuere *et al.*, 1993a). The current results clearly emphasize the limitations of X-ray crystallography in determining flexibility of proteins in solution. The need for a more dynamical approach was also nicely illustrated by the results of a hydrogen exchange study on a disulfide bond mutant of barnase (Clarke *et al.*, 1995b). Although the crystallographic B-factors in this region indicated an increased mobility in the mutated area, the increased lifetime of the hydrogen bonds showed that the local stability had actually increased due to the introduction of the cross-link.

Theoretically, the reduction in entropy resulting from a cross-link in a random-coil peptide can be calculated using Equation (3) (Pace *et al.*, 1988):

$$\Delta S = -2.1 - (3/2)R \ln n \quad (3)$$

where  $\Delta S$  is expressed in cal mol<sup>-1</sup> K<sup>-1</sup> and  $n$  is the number of residues between the side chains that are cross-linked. In terms of free energy, the unfolded state of the disulfide-containing mutant is thus expected to be over 5 kcal mol<sup>-1</sup> less stable than the unfolded state of the reduced form of the mutant. Since for DhIA, thermal unfolding is an irreversible process, it is in principle not possible to determine  $\Delta G$  of unfolding and calculate differences in stability between the different species. However, comparison of changes in  $T_m$  of reversible protein systems for which  $\Delta G$  of unfolding has been determined can give an estimate of the change in free energy. According to Pantoliano *et al.* (Pantoliano *et al.*, 1987) an increase in  $T_m$  in 1°C corresponds to an increased stability of

~0.22 kcal mol<sup>-1</sup>. For the oxidized Asp16Cys/Ala201Cys mutant of DhIA, this would suggest an increased stability of ~1.1 kcal mol<sup>-1</sup> compared to wild-type enzyme and 3.0 kcal mol<sup>-1</sup> compared to its reduced counterpart.

Our results show that introduction of the disulfide cross-link in DhIA enhanced the kinetic stability of the enzyme, suggesting that the cap domain region encompassing helices 6 and 7 is probably involved early in the unfolding process. Theoretically, the cross-linking of flexible regions could have a destabilizing effect on the native state, since it lowers the entropy of the protein. However, in the case where the regions connected by the disulfide are far apart from each other in the transition state, this state of the protein will be destabilized even more, resulting in a net stabilization of the native form. The results indicate that MD simulations may be a valuable tool for identifying flexible regions involved in the early stages of the unfolding process. The introduction of a disulfide bond in such a region can hamper or block the pathway to the transition state, slowing down the overall unfolding process. Therefore, MD simulations combined with the introduction of disulfide cross-links provide a promising tool for enhancing the kinetic stability of proteins.

## Acknowledgements

We thank Ivo Ridder for help with operating the computer program SSBOND. M.G.P. was supported by the Netherlands Foundation for Chemical Research (CW) with financial aid from the Netherlands Organization for Scientific Research (NWO).

## References

- Amadei, A., Linssen, A.B.M. and Berendsen, H.J.C. (1993) *Proteins*, **17**, 412–425.
- Applequist, J., Carl, J.R. and Fung, K. (1972) *J. Am. Chem. Soc.*, **94**, 2952–2960.
- Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R. (1984) *J. Chem. Phys.*, **81**, 3684–3690.
- Chu, Y., Lee, E.Y., Reimann, E.M., Wilson, S.E. and Schlender, K.K. (1996) *Arch. Biochem. Biophys.*, **334**, 83–88.
- Clarke, J. and Fersht, A.R. (1993) *Biochemistry*, **32**, 4322–4329.
- Clarke, J., Henrick, K. and Fersht, A.R. (1995a) *J. Mol. Biol.*, **253**, 493–504.
- Clarke, J., Hounslow, A.M. and Fersht, A.R. (1995b) *J. Mol. Biol.*, **253**, 505–513.
- CRC Handbook of Chemistry and Physics* (1986) 67th edn. Lide, D.R. ed. CRC Press Inc., Boca Raton, FL.
- Doig, A.J. and Williams, D.H. (1991) *J. Mol. Biol.*, **217**, 389–398.
- Ellman, G.L. (1959) *Arch. Biochem. Biophys.*, **82**, 70–77.
- Flory, P.J. (1956) *J. Am. Chem. Soc.*, **78**, 5222–5235.
- Franken, S.M., Rozeboom, H.J., Kalk, K.H. and Dijkstra, B.W. (1991) *EMBO J.*, **10**, 1297–1302.
- Greene, R.F., Jr and Pace, C.N. (1974) *J. Biol. Chem.*, **249**, 5388–5393.
- Hazes, B. and Dijkstra, B.W. (1988) *Protein Eng.*, **2**, 119–125.
- Kabsch, W. and Sander, C. (1983) *Biopolymers*, **22**, 2577–2637.
- Keuning, S., Janssen, D.B. and Witholt, B. (1985) *J. Bacteriol.*, **163**, 635–639.
- Krooshof, G.H., Floris, R., Tepper, A.W.J.W. and Janssen, D.B. (1999) *Protein Sci.*, **8**, 355–360.
- Krooshof, G.H. (2000) Structure–function Relationships in Haloalkane Dehalogenase. Thesis, University of Groningen, The Netherlands.
- Kunkel, T.A. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 488–492.
- Kuroki, R., Inaka, K., Taniyama, Y., Kidokoro, S., Matsushima, M., Kikuchi, M. and Yutani, K. (1992) *Biochemistry*, **31**, 8323–8328.
- Li, Y., Coutinho, P.M. and Ford, C. (1998) *Protein Eng.*, **11**, 661–667.
- Mansfeld, J., Vriend, G., Dijkstra, B.W., Veltman, O.R., Van den Burg, B., Venema, G., Ulbrich-Hofmann, R. and Eijssink, V.G. (1997) *J. Biol. Chem.*, **272**, 11152–11156.
- Matsumura, M., Becktel, W.J., Levitt, M. and Matthews, B.W. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 6562–6566.
- Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J. *et al.* (1992) *Protein Eng.*, **5**, 197–211.
- Pabo, C.O. and Suchanek, E.G. (1986) *Biochemistry*, **25**, 5987–5991.
- Pace, C.N. (1975) *CRC Crit. Rev. Biochem.*, **3**, 1–43.
- Pace, C.N., Grimsley, G.R., Thomson, J.A. and Barnett, B.J. (1988) *J. Biol. Chem.*, **263**, 11820–11825.
- Pantoliano, M.W., Ladner, R.C., Bryan, P.N., Rollence, M.L., Wood, J.F. and Poulos, T.L. (1987) *Biochemistry*, **26**, 2077–2082.
- Patkar, S., Vind, J., Kelstrup, E., Christensen, M.W., Svendsen, A., Borch, K. and Kirk, O. (1998) *Chem. Phys. Lipids*, **93**, 95–101.
- Perry, L.J. and Wetzel, R. (1984) *Science*, **226**, 555–557.
- Pjura, P.E., Matsumura, M., Wozniak, J.A. and Matthews, B.W. (1990) *Biochemistry*, **29**, 2592–2598.
- Pollitt, S. and Zalkin, H. (1983) *J. Bacteriol.*, **153**, 27–32.
- Pregel, M.J. and Storer, A.C. (1997) *J. Biol. Chem.*, **272**, 23552–23558.
- Pries, F., van den Wijngaard, A.J., Bos, R., Pentenga, M. and Janssen, D.B. (1994) *J. Biol. Chem.*, **269**, 17490–17494.
- Ryckaert, J.P., Cicotti, G. and Berendsen, H.J.C. (1977) *J. Comp. Phys.*, **23**, 327–341.
- Schanstra, J.P. and Janssen, D.B. (1996) *Biochemistry*, **35**, 5624–5632.
- Schanstra, J.P., Rink, R., Pries, F. and Janssen, D.B. (1993) *Protein Expr. Purif.*, **4**, 479–489.
- Schanstra, J.P., Ridder, I.S., Heimeriks, G.J., Rink, R., Poelarends, G.J., Kalk, K.H., Dijkstra, B.W. and Janssen, D.B. (1996) *Biochemistry*, **35**, 13186–13195.
- Schanstra, J.P., Ridder, A., Kingma, J. and Janssen, D.B. (1997) *Protein Eng.*, **10**, 53–61.
- Shinkai, A., Hirano, A. and Aisaka, K. (1996) *J. Biochem. (Tokyo)*, **120**, 915–921.
- Sowdhamini, R., Srinivasan, N., Shoichet, B., Santi, D.V., Ramakrishnan, C. and Balaram, P. (1989) *Protein Eng.*, **3**, 95–103.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
- Takagi, H., Takahashi, T., Momose, H., Inouye, M., Maeda, Y., Matsuzawa, H. and Ohta, T. (1990) *J. Biol. Chem.*, **265**, 6874–6878.
- Thornton, J.M. (1981) *J. Mol. Biol.*, **151**, 261–287.
- Tidor, B. and Karplus, M. (1993) *Proteins*, **15**, 71–79.
- van Buuren, A.R., Marrink, S.J. and Berendsen, H.J.C. (1993) *J. Phys. Chem.*, **97**, 9206–9212.
- van der Spoel, D., Berendsen, H.J.C., Apol, E., Meulenhoff, P.J., Sijbers, A.L.T.M. and van Drunen, R. (1996a) *Gromacs User Manual*. Laboratory of Physical Chemistry, University of Groningen, The Netherlands. Internet: <http://rugmd0.chem.rug.nl/~gmx>.
- van der Spoel, D., van Buuren, A.R., Tieleman, D.P. and Berendsen, H.J.C. (1996b) *J. Biomol. NMR*, **8**, 229–238.
- van Gunsteren, W.F. and Berendsen, H.J.C. (1987) *Gromos Manual. BIOMOS, Biomolecular Software*. Laboratory of Physical Chemistry, University of Groningen, The Netherlands.
- Verschuere, K.H., Franken, S.M., Rozeboom, H.J., Kalk, K.H. and Dijkstra, B.W. (1993a) *J. Mol. Biol.*, **232**, 856–872.
- Verschuere, K.H., Seljee, F., Rozeboom, H.J., Kalk, K.H. and Dijkstra, B.W. (1993b) *Nature*, **363**, 693–698.
- Villafranca, J.E., Howell, E.E., Voet, D.H., Strobel, M.S., Ogden, R.C., Abelson, J.N. and Kraut, J. (1983) *Science*, **222**, 782–788.
- Vogl, T., Brengelmann, R., Hinz, H.J., Scharf, M., Lotzbeyer, M. and Engels, J.W. (1995) *J. Mol. Biol.*, **254**, 481–496.
- Weber, P.C., Sheriff, S., Ohlendorf, D.H., Finzel, B.C. and Salemme, F.R. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 8473–8477.
- Yamaguchi, S., Takeuchi, K., Mase, T., Oikawa, K., McMullen, T., Derewenda, U., McElhane, R.N., Kay, C.M. and Derewenda, Z.S. (1996) *Protein Eng.*, **9**, 789–795.
- Ziegler, D.M. and Poulsen L.L. (1977) *TIBS*, **2**, 79–81.

Received May 15, 2001; revised November 23, 2001; accepted December 7, 2001