Structural features of halophilicity derived from the crystal structure of dihydrofolate reductase from the Dead Sea halophilic archaeon, *Haloferax volcanii*

Ursula Pieper[1], Geeta Kapadia[1], Moshe Mevarech[2] and Osnat Herzberg[1]*

Background: The proteins of halophilic archaea require high salt concentrations both for stability and for activity, whereas they denature at low ionic strength. The structural basis for this phenomenon is not yet well understood. The crystal structure of dihydrofolate reductase (DHFR) from *Haloferax volcanii* (*hv*-DHFR) reported here provides the third example of a structure of a protein from a halophilic organism. The enzyme is considered moderately halophilic, as it retains activity and secondary structure at monovalent salt concentrations as low as 0.5 M.

Results: The crystal structure of *hv*-DHFR has been determined at 2.6 Å resolution and reveals the same overall fold as that of other DHFRs. The structure is in the apo state, with an open conformation of the active-site gully different from the open conformation seen in other DHFR structures. The unique feature of *hv*-DHFR is a shift of the α helix encompassing residues 46–51 and an accompanied altered conformation of the ensuing loop relative to other DHFRs. Analysis of the charge distribution, amino acid composition, packing and hydrogen-bonding pattern in *hv*-DHFR and its non-halophilic homologs has been performed.

Conclusions: The moderately halophilic behavior of *hv*-DHFR is consistent with the lack of striking structural features expected to occur in extremely halophilic proteins. The most notable feature of halophilicity is the presence of clusters of non-interacting negatively charged residues. Such clusters are associated with unfavorable electrostatic energy at low salt concentrations, and may account for the instability of *hv*-DHFR at salt concentrations lower than 0.5 M. With respect to catalysis, the open conformation seen here is indicative of a conformational transition not reported previously. The impact of this conformation on function and/or halophilicity is unknown.

Addresses: [1]Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville MD 20850, USA and [2]Department of Molecular Microbiology and Biotechnology, George S Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, 69978, Israel.

*Corresponding author.
E-mail: osnat@carb.nist.gov

Key words: dihydrofolate reductase, *Haloferax volcanii*, halophilic archaea, X-ray crystallography

Received: **9 October 1997**
Revisions requested: **3 November 1997**
Revisions received: **11 November 1997**
Accepted: **24 November 1997**

Structure 15 January 1998, **6**:75–88
http://biomednet.com/elecref/0969212600600075

Introduction

Microorganisms that are adapted to living under extreme saline conditions have been found in many hypersaline environments, such as in the Dead Sea and in the Great Salt Lake, Utah. In order to survive at high salt concentrations these microorganisms have to maintain an osmotic balance with their external environment. The halophilic archaea cope with the external salinity by accumulating KCl intracellularly to concentrations up to 4 M [1].

Most of the enzymes of halobacteria are active and stable at high salt concentrations and lose activity at salt concentrations lower than 2 M [2–3]. Sequence comparisons show that, in general, the halophilic proteins contain an excess of negatively charged amino acids over positively charged amino acids, and the number of negatively charged amino acid residues is higher than that in their non-halophilic homologs. The additional negative charges are located mostly on the protein surface, presumably helping to stabilize the protein molecule by competing with the salt for hydration [2]. It has also been proposed that hydrophobic interactions play an important role in the ability of these proteins to cope with the salt stress in a hypersaline environment [4–6].

The crystal structures of two enzymes from the halophilic archaeon *Haloarcula marismortui* have been reported: malate dehydrogenase (MDH) and 2Fe–2S ferredoxin. Based on the analysis of the structure of halophilic MDH, an enzyme that functions at NaCl concentrations higher than 2 M, it was concluded that an increase in the number of salt bridges compared with the non-halophilic homologs enhances stability at high salt concentrations [7]. Enrichments of salt bridges were not found, however, in the crystal structure of 2Fe–2S ferredoxin from the same archaeon [8]. Instead, the entire surface of the ferredoxin protein is coated with acidic residues, with the exception of the active site. In addition, in this protein two amphipathic

helices enriched with negatively charged residues are inserted near the N terminus and may confer further stability. In contrast to MDH, the ferredoxin is considered only a moderately halophilic protein, because it is stable and functional down to salt concentrations of 0.4 M [9].

Dihydrofolate reductase from *Haloferax volcanii* (*hv*-DHFR) is a moderately halophilic enzyme, being stable and active at a concentration of NaCl or KCl higher than 0.5 M. A rapid loss of structure and activity is observed at lower salt concentrations [9]. The long term stability of the enzyme under these conditions has not been determined. Statistical analysis of the amino acid composition of 22 DHFRs showed that the increase in negative charge in the halophilic enzyme is insignificant compared with other exchanges of amino acids within the sample of 22 sequences [10].

Functionally, DHFR is an essential enzyme that catalyzes the NADPH-linked reduction of 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate. This reaction is necessary to regenerate tetrahydrofolate following the production of thymidylate, one of the DNA building units. Thus, DHFR is coupled to cell division, and both DHFR and thymidylate synthase are target enzymes in cancer chemotherapy. DHFR is inhibited by antineoplastics, such as methotrexate. Moreover, species-dependent differences in the enzyme's sensitivity to inhibitors are responsible for the efficacy of the antibacterial compound trimethoprim and the antimalarial compound pyrimethamine.

Recently, a new *Haloferax volcanii* gene was isolated that codes for a second DHFR (M Ortenberg and M Mevarech, unpublished data). The lack of cross-hybridization between the two DHFR genes, and the level of amino acid sequence identity (~40%), suggest that one of the genes may have been obtained by lateral gene transfer. The physiological role of each of the two DHFRs is currently under investigations.

We report here the crystal structure of (*hv*-DHFR), the first protein structure to be determined from *Haloferax volcanii*. DHFR is an excellent subject for comparative structural studies, because the three-dimensional structures of this enzyme have been obtained at high resolution from *Escherichia coli* [11–12], *Lactobacillus casei* [11–13], *Pneumocystis carinii* [14], chicken [15], mouse [16] and human [17]. Despite the low level of sequence identity, the three-dimensional structures of DHFR from these different organisms are similar. In this paper, we discuss the structural differences between *hv*-DHFR and its non-halophilic counterparts. The findings are compared with the information available for two other examples of structures of halophilic proteins. In addition, the *hv*-DHFR structure reveals a novel conformational transition that may accompany substrate binding.

Results and discussion

Structure determination

The automated sequence alignment procedures FASTA [18] and MPsrch [19] show that the highest level of amino acid sequence identity between *hv*-DHFR and other DHFRs of known structure is with the enzyme from *E. coli* (32% identity; Table 1). At that level of homology, the Molecular Replacement (MR) method seemed suitable for structure determination. A number of structures of DHFR from *E. coli* (*ec*-DHFR) and from other organisms were available as search models in the Protein Data Bank [20]. In addition, a model of the *hv*-DHFR, developed using comparative modeling methods, was available [21]. Nevertheless, attempts to use each of these structures in searches that employed various MR computer programs were unsuccessful. Ultimately, following the experience of Leahy *et al.* [22], the structure was determined using a composite model of seven different coordinate sets superpositioned on each other (Table 1). All sidechains were included in the model and each molecule was given equal weight (Figure 1).

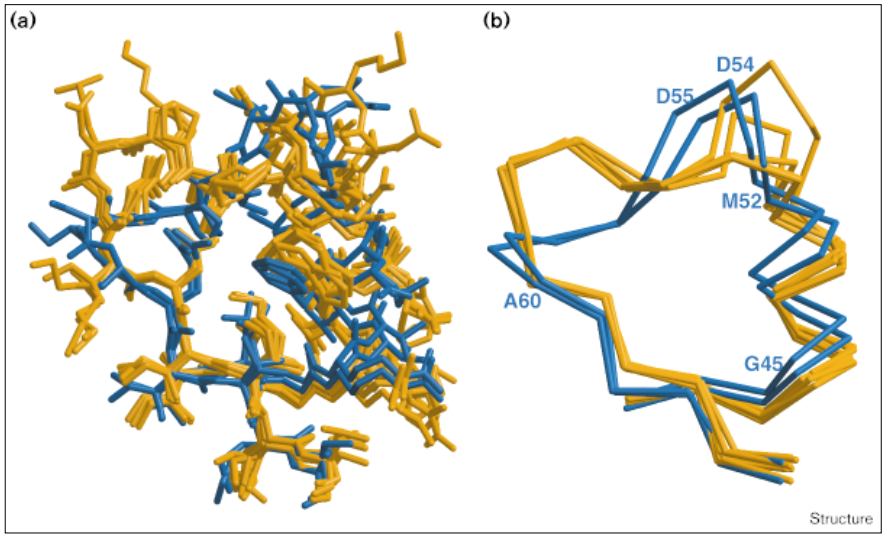
Using the MR program AMoRe [23], the self-rotation function confirmed that there are two independent molecules in the asymmetric unit, consistent with the density measurements of the crystals and calculation of cell content. The cross-rotation function did not show any significant peak higher than the background. A translation search was performed using the 25 highest peaks of the rotation function, yielding two solutions with correlation coefficients above the background. A translation search with the molecules led to a significantly higher correlation coefficient. The correct solution was confirmed by a rigid-body refinement that increased the correlation coefficient further and reduced the R value ($R = \Sigma_h ||F_o| - |F_c|| / \Sigma_h |F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor amplitudes, respectively).

Table 1

DHFR coordinates used in molecular replacement.

	PDB entry code	Sequence identity with *hv*-DHFR (%)
DHFR from *E. coli*		32
complex with methotrexate	4dfr	
holoenzyme	5dfr	
complex with NADP	6dfr	
complex with folate and NADP	7dfr	
DHFR from chicken liver complex with NAD or NADP	8dfr	25
DHFR from *L. casei* complex with NADPH and methotrexate	3dfr	23
hv-DHFR (Fidelis' model)		100

Figure 1



Superposition of the search models used in the molecular replacement (yellow) with the refined structure of *hv*-DHFR (blue). The figure shows **(a)** an all-atom representation and **(b)** a Cα trace. Residues 43–65 correspond to a region with one of the largest deviations between the search models and the actual structure; some of these residues are labeled. (The figure was generated with the program RASTER3D [47,48].)

To assess the importance of using a multimolecule search model, several tests were carried out with single-molecule search models, and with a combined model comprising only the *ec*-DHFR structures. The results are summarized in Table 2, and show that the correct solution could also be identified for the combined *ec*-DHFR model and for two of the *ec*-DHFR structures taken separately. However, the discrimination between the right solution and the next highest incorrect solution is poor, in particular for the single-molecule search models.

A structural sequence alignment (Figure 2) shows that the sequence identity between *hv*-DHFR and *ec*-DHFR is actually 28%, lower than predicted by the sequence alignment. This level of homology may be marginal for current MR methods, explaining the difficulties encountered in applying the method and later during the process of structure refinement.

Refinement

After an initial simulated-annealing cycle, using the structure of *hv*-DHFR modeled by Fidelis as the starting model, approximately 70% of the polypeptide chain was associated with well defined electron density. Building of the remaining chain involved numerous reiterations of the refinement cycles. These reiterations were combined with the omission

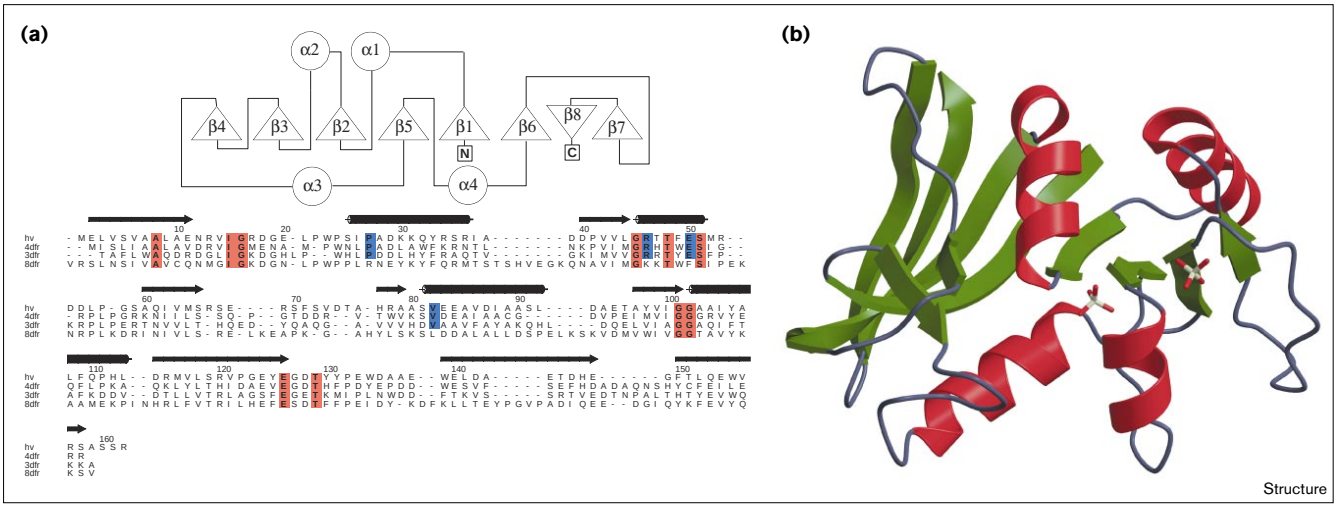
Table 2

Molecular replacement results.

	Combined search model from seven coordinate sets	Combined search model using five *E. coli* coordinate sets	3dfr	4dfr (molecule B)	7dfr	*hv*-DHFR model
Correct rotation function solution	3rd and 13th peak	6th and 9th peak	11th peak	7th and 16th peak	7th and 22nd peak	2nd peak
Correct translation function solution (one molecule)*	c = 18.9/R = 55.0 c = 18.7/R = 54.6 (c = 15.4/R = 55.9)	c = 18.1/R = 54.4 c = 17.9/R = 54.6 (c = 15.0/R = 56.3)	No correct solution	c = 15.8/R = 54.1 c = 16.9/R = 55.0 (c = 15.6/R = 55.2)	c = 16.3/R = 55.3 c = 17.0/R = 56.0 c = 14.7/R = 56.3)	No correct solution
Correct translation function solution (two molecules)*	c = 31.1/R = 48.5 (c = 21.0/R = 51.5)	c = 28.4/R = 49.2 (c = 20.6/R = 51.3)		c = 22.1/R = 51.5 (c = 17.1/R = 52.7)	c = 24.4/R = 50.3 (c = 18.7/R = 51.9)	
Rigid-body refinement*	c = 45.7/R = 43.6 (c = 35.9/R = 48.5)	c = 42.9/R = 44.2 (c = 36.5/R = 46.7)		c = 38.6/R = 45.7 (c = 36.4/R = 46.9)	c = 46.3/R = 44.4 (c = 44.3/R = 46.8)	

*c = correlation coefficient (%); R = R value (%). The value in parentheses indicates the next highest incorrect peak. No correct rotation function solution could be found for single models of 5dfr, 6dfr, 8dfr and the other molecule in the asymmetric unit of 4dfr. PDB codes are as defined in Table 1.

Figure 2



The sequence and fold of DHFR from *H. volcanii*. **(a)** Secondary structure scheme and a structural alignment of the amino acid sequences of *hv*-DHFR with DHFR from *E. coli* (4dfr), *L. casei* (3dfr), and chicken liver (8dfr). The invariant residues are highlighted in red and conserved residues in the bacterial DHFRs are highlighted in blue. The secondary structure according to Kabsch and Sander [49], as defined in *hv*-DHFR, is indicated above the sequence by arrows (β strands) and rods (α helices). Residue numbers follow the *hv*-DHFR sequence. **(b)** The overall fold of DHFR from *H. volcanii* (the phosphate ions found in the structure are shown in stick representation). Helices are shown in red and β strands in green. (The ribbon representation was generated with the programs MOLSCRIPT and RASTER3D [47,48,50].)

of segments of structure that required major adjustments because of large differences between their actual conformation and the conformations of the respective segments in any of the search models; this is illustrated in Figure 1 for the α helix and loop regions comprising residues 43 to 58. When the process was completed, the quality of the electron-density map was very good. The final R value is 0.184 for all data between 7.0–2.6 Å for which $F > 2\sigma(F)$, and 0.209 for all data without amplitude cut-off.

The model includes the two molecules in the asymmetric unit, termed A and B. Molecule A includes all but the first residue and the C-terminal 159–162 residues. These residues were associated with regions of weak electron density, indicative of disorder. For molecule B, all but residues 1, 123–125, and 160–162 are included. The model includes three phosphate ions and 80 water molecules. The root mean square (rms) deviations from ideal bond length and bond angle values of the standard geometry compiled by Engh and Huber [24] are 0.013 Å and 3.2°, respectively. The mainchain conformations of two amino acid residues (Asp18B and Ala60A, where A and B refer to the monomer of the asymmetric unit) are sterically strained [25]. As with all other known DHFR structures, the peptide bond between Gly101 and Gly102 is *cis* (except where specified otherwise, the numbering scheme is based on the *hv*-DHFR sequence; Figure 2). Superposition of the Cα atoms of the two monomers in the asymmetric unit results in an rms deviation of 0.6 Å, indicating that the structure of the two molecules is very similar. Their average crystallographic temperature factor values are 28.7 $Å^2$ and 23.1 $Å^2$, for molecules A and B, respectively.

Overall structure

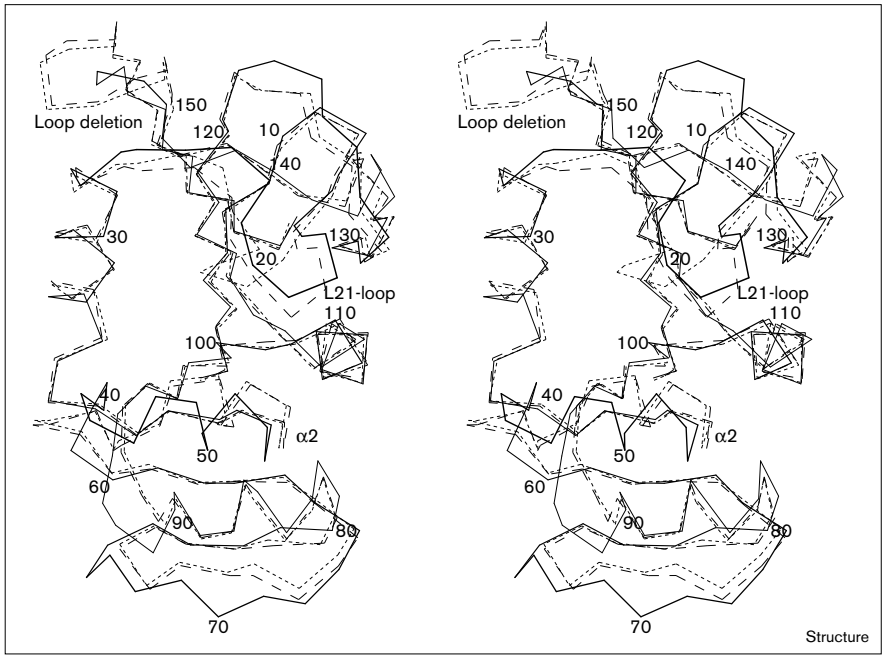
hv-DHFR folds into an eight-stranded β sheet consisting of seven parallel strands and a single C-terminal antiparallel strand (Figure 2). Four α helices pack against the β sheet, two on each side. The fourth α helix ends with three residues forming a 3_{10} helix. The fold of *hv*-DHFR is closer to that of the DHFRs of prokaryotic origin (e.g. *E. coli* and *L. casei*) than to those from eukaryotic organisms. The region of largest difference corresponds to the α2 helix and the ensuing loop region, as shown in Figure 1. Another notable difference is the deletion of six residues in an extended loop region between β strands β7 and β8 of *hv*-DHFR compared with *ec*-DHFR. Consequently, the loop forms a type II open tight turn.

The L21 loop

The various crystal structures of DHFR that have been reported have shown that ligand binding to DHFR is accompanied by loop and domain motion. Three different conformations, closed, open and occluded, have been identified in *ec*-DHFR at different states of the catalytic cycle. The conformational changes mainly affect the loop comprising amino acid residues 13–20, termed the M20 loop [26]. The structures suggest that the M20 loop is predominately closed over the reactants in the holoenzyme, Michaelis, and transition-state complexes. During the remainder of the reaction cycle, when nicotinamide is

Figure 3

Superposition of the Cα trace of *hv*-DHFR (solid line), *ec*-DHFR in the closed conformation (long dashes; PDB entry code 1rx9) and *ec*-DHFR in the occluded conformation (short dashes; PDB entry code 1rx5). Every tenth residue of *hv*-DHFR is labeled.



not bound, the loop protrudes into the nicotinamide–ribose binding pocket. Upon changing from the closed to the occluded conformation, the central portion of the loop rearranges from a β-hairpin loop type III′ to a 3_{10} helix through an intermediate open conformation, which allows the nicotinamide–ribose into or out of its binding pocket. It has been proposed that the conformational changes play a role in proton transfer and transition-state stabilization [26].

The structure of *hv*-DHFR reported here represents the apo state of the enzyme with phosphate ions bound in the nicotinamide-binding site. The conformation of the loop analogous to M20, L21, is very similar to that of the closed conformation in *ec*-DHFR, although residues 17–20 form a type I′ β-hairpin loop instead of the type III′ turn observed in the closed conformation of *ec*-DHFR. The L21 loop exhibits some of the highest crystallographic temperature factors in the structure and in molecule B the sidechain of Asp18 is disordered, indicating possible conformational flexibility. Figure 3 shows a superposition of the Cα atoms of *hv*-DHFR with those of *ec*-DHFR in both the occluded and closed states. Despite the so-called closed conformation of the L21 loop in *hv*-DHFR, access to the active site is not blocked. The α2 helix is shifted by approximately 2.3 Å compared with the respective helix in *ec*-DHFR, so that nicotinamide–ribose can diffuse into the binding pocket (Figures 3 and 4).

The nucleotide-binding site

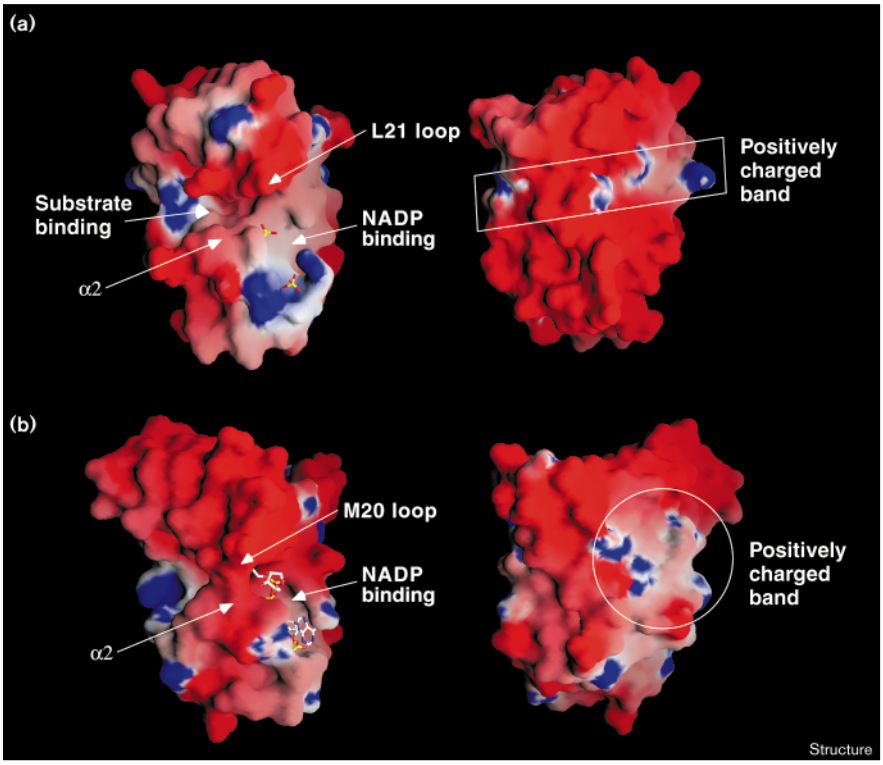
The electrostatic environment at the nucleotide-binding site is similar in all known DHFR structures. In general, the adenine moiety forms only van der Waals and hydrophobic interactions with the protein. One exception to this is in *ec*-DHFR, where an additional interaction with the sidechain amide group of Gln102 occurs. Instead of a glutamine residue, *hv*-DHFR contains a leucine in this position, Leu108, the sidechain of which cannot be involved in an electrostatic interaction.

In all DHFR structures, a positively charged pocket is formed to accommodate the phosphate groups of the nucleotide. Although the structure of *hv*-DHFR does not contain the nucleotide, two phosphate ions were identified in molecule B at positions were the phosphate groups of NADPH are expected to bind (Figure 5a). Molecule A contains only one phosphate ion, and the second site is occupied by what was interpreted in the electron density as a water molecule. The *cis* peptide bond between two invariant glycine residues, which is common to all known DHFR structures, is also present in *hv*-DHFR (between Gly101 and Gly102). A hydrogen bond is formed between the mainchain amide group of Gly102 and one of the phosphate ions in molecule B, or a water molecule in molecule A. This interaction is equivalent to the interaction seen in all the structures of the NADP-bound forms of DHFR between the amide group of the *cis* peptide bond and the phosphoryl group of the nucleotide.

The substrate-binding site

Crystallographic studies of complexes of *ec*-DHFR with either the inhibitor methotrexate [11] or the substrate folate [27] have shown that these structurally related

Figure 4



A comparison of the molecular surfaces of halophilic and non-halophilic DHFRs. **(a)** Representation of the molecular surface of *hv*-DHFR and **(b)** *ec*-DHFR in the closed conformation (PDB entry code 1rx9). The colors highlight the electrostatic potential as calculated by the program GRASP [33] with negative potential in red and positive potential in blue. The orientation is the same as in Figure 3, with the C terminus of the polypeptide chain at the top of the figure. The figure illustrates that whilst *hv*-DHFR exhibits an open conformation the L21 loop of this protein adopts the same closed conformation as that of the M20 loop in the closed conformational state of *ec*-DHFR. For each molecule, the left image corresponds to the surface where the active site is located and the right image corresponds to the opposite face.

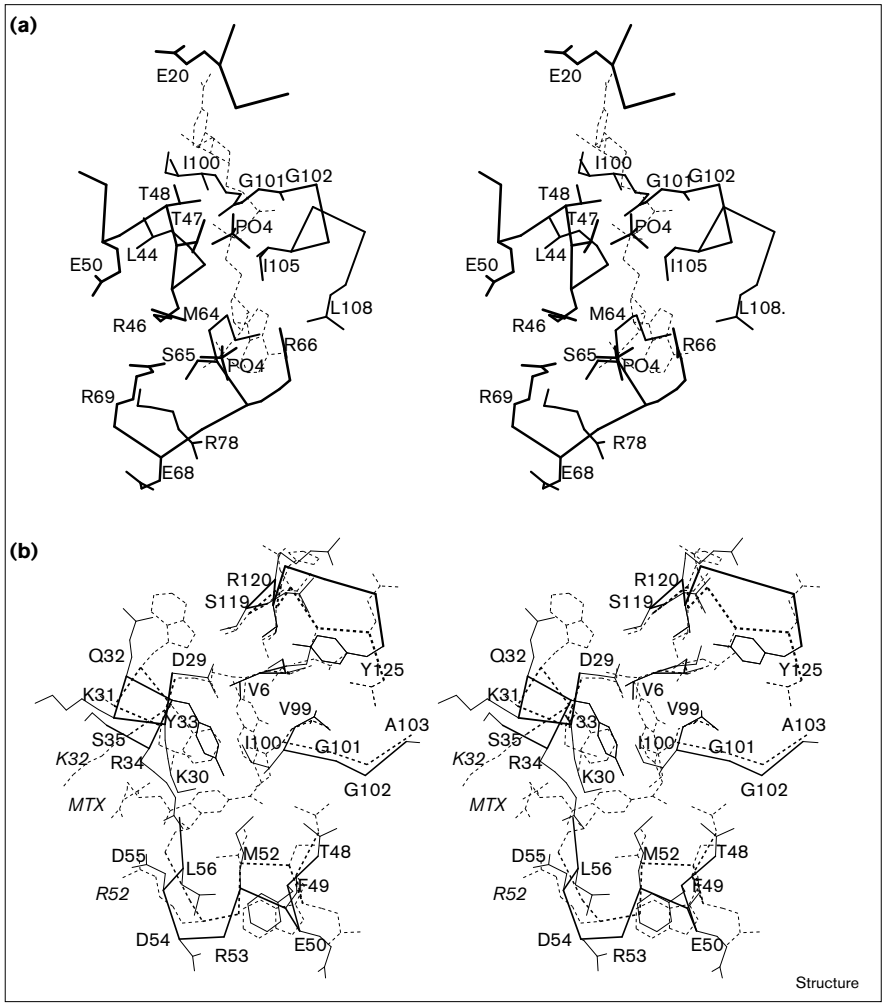
compounds bind in a deep hydrophobic cleft which bisects the DHFR molecule and contains a single polar residue, Asp27. Although *hv*-DHFR was crystallized in the presence of methotrexate, inhibitor binding could not be accounted for in the electron-density map. In fact, superposition of the *hv*-DHFR and the methotrexate-bound *ec*-DHFR (PDB entry code 4dfr) shows that some of the sidechains of residues located on the α1 helix block methotrexate access to the *hv*-DHFR cleft (Figure 5b).

Key amino acid residues involved in substrate binding are either conserved or replaced conservatively in *hv*-DHFR. Some of these residues maintain the same conformation as in other DHFRs of known structure, but there are also notable differences. In the binding site of the pteridine moiety, the catalytic residue Asp29 (Asp27 in *ec*-DHFR) exhibits the same conformation as in all other structures. Ser119 in *hv*-DHFR is equivalent to a threonine residue in the other DHFRs (Thr113 in *ec*-DHFR). The hydroxyl group of this threonine residue is hydrogen bonded to the catalytic aspartate in the methotrexate-bound structures, and is bridged to the 2-amino group of the pteridine moiety via a water molecule. This water molecule is present in all previously determined DHFR structures, independent of the state of substrate binding, but is not present in *hv*-DHFR. Furthermore, a second bridging water molecule is hydrogen bonded to Asp27 and Trp21 and to the substrate/inhibitor in the *E. coli* enzyme. This water molecule is also missing in the current *hv*-DHFR structure. The absent water molecules may just reflect the limited resolution of the data (2.6 Å), thus the functional significance of this finding is unclear.

Another key interaction of the pteridine moiety is formed between the 4-amino group of methotrexate and two mainchain carbonyl oxygen atoms in both the *E. coli* and *L. casei* DHFR (*lc*-DHFR) structures. The analogous interactions in *hv*-DHFR would involve the carbonyl groups of Val6 and Ile100. These are oriented appropriately for binding even though the substrate-binding site is unoccupied. In contrast, the sidechain of Tyr33 (analogous to a phenylalanine that forms an aromatic–aromatic interaction with the pteridine ring in the other DHFR structures) blocks access to the pocket (Figure 5b). A slight shift of the α1 helix of *hv*-DHFR compared with the respective helix in *ec*-DHFR induces the conformation that prevents binding. Assuming that the binding mode of the pteridine moiety of methotrexate in *hv*-DHFR is the same as that seen in *ec*-DHFR, the α1 helix and the sidechain of Tyr33 must undergo a conformational transition to accommodate the methotrexate.

The binding site for the benzoylcarbonyl-L-glutamate-α-carboxylate moiety of methotrexate is blocked in the *hv*-DHFR structure by the sidechain of Arg34 (Figure 5b). Another protein segment of *hv*-DHFR that should contact

Figure 5



The binding sites of *hv*-DHFR. **(a)** Stereo representation of the nucleotide-binding site in *hv*-DHFR. Only the Cα trace is shown, except for residues Gly101–Gly102 where all mainchain atoms are shown to highlight the nonproline *cis* peptide bond. An NADP molecule is modeled into the binding site (dashed line), in the conformation found in *ec*-DHFR; its phosphate groups were superimposed onto the phosphate ions identified in the crystal structure of *hv*-DHFR. **(b)** Stereo representation of the substrate-binding site in *hv*-DHFR superimposed with the substrate-binding site of *ec*-DHFR including the bound inhibitor methotrexate (dashed line; PDB entry code 4dfr). Virtual bonds between Cα atoms and an all-atom sidechain representation are shown. Residues of *ec*-DHFR are labeled in italics.

this part of methotrexate comprises residues 50–59. This segment adopts a drastically different conformation to that observed in either the methotrexate-bound or unbound *ec*-DHFR structures and in the methotrexate-bound *lc*-DHFR. Although the 50–59 segment of *hv*-DHFR does not block access to methotrexate, it carries Asp55 which is oriented towards the position where the glutamate moiety of methotrexate should bind, by analogy to *ec*-DHFR and *lc*-DHFR. To avoid repulsive charge interactions, Asp55 must change conformation, and this may be accompanied by global changes in the 50–59 segment.

It is interesting to note that the arginine residue that interacts with the α-carboxylate group of methotrexate in the known structures (Arg57 in *ec*-DHFR) is replaced by a serine in *hv*-DHFR. Only Arg34 in *hv*-DHFR could form favorable electrostatic interaction with the α-carboxylate group of methotrexate (upon conformational adjustments). However, Arg34 is structurally equivalent to Arg32/Lys32 in *ec*-/*lc*-DHFRs, residues that do not form close interactions with methotrexate.

The carboxylate group of Asp55 in *hv*-DHFR is oriented towards the expected position of the negatively charged groups of methotrexate, an energetically unfavorable orientation for binding. Asp55 and Asp54 constitute a unique insertion in the sequence of *hv*-DHFR compared with other DHFRs, and may be responsible in part for the very different conformation of the 50–59 peptide. On the opposite side of the substrate cavity (Figure 5b), two lysine residues of *hv*-DHFR, Lys30 and Lys31, may counter the electrostatic effect of the two aspartate residues on binding. Single- and double-mutant enzymes have been produced, in which each of the lysine residues was replaced by a hydrophobic residue, resulting in a comparable increase in k_{cat} and K_m [28]. The increase in the K_m values may be attributed to the elimination of the positive charges which compensate for the unfavorable electrostatic contribution of Asp54 and Asp55 to binding.

The above observations indicate that *hv*-DHFR may utilize different loop and sidechain conformations for binding the benzoylcarbonyl-L-glutamate-α-carboxylate moiety, compared with those observed for *ec*-DHFR and *lc*-DHFR. This proposal awaits experimental confirmation, by determination of the structure of the methotrexate- or folate-bound enzyme. Nevertheless, the blocked access to the substrate-binding site implies that the enzyme must undergo some conformational transition. In *ec*-DHFR, the M20 loop alternates between conformations that open or close the active-site depression depending on the state of substrate or cofactor binding [26]. Although the conformation of the equivalent loop in the structure of *hv*-DHFR (the L21 loop) is similar to the closed conformation seen in *ec*-DHFR, the L21 loop does not prevent access to the active site (Figure 4) because the α2 helix and the ensuing loop on the opposite side of the cleft are shifted compared with the other DHFR structures (Figure 3). Conformational adjustments within the α2 helix and loop provide an alternative mode of recognition not realized in previously determined DHFR structures. The α1 helix must also undergo conformational transition, although not as extensive, in order to remove the steric hinderance of the sidechains of Tyr33 and Arg34.

hv-DHFR does not exhibit striking halophilic features

hv-DHFR is a moderately halophilic enzyme as it is active and stable at monovalent salt concentrations as low as 0.5 M. Amino acid sequences of halophilic proteins typically show an enlarged number of negatively charged residues compared with their non-halophilic homologs, with glutamic acid residues being the most abundant of these [2]. As carboxylate groups are capable of binding more water molecules than other protein groups [29], it has been proposed that these negatively charged residues help to form a hydration shell and prevent the protein from aggregation. At low salt concentrations, the high concentration of negatively charged residues is repulsive, resulting in protein instability.

Consistent with the moderate halophilicity of *hv*-DHFR, the polar, nonpolar, and charged amino acid contents are similar to those of the DHFRs from *E. coli* and *L. casei* [10] (Table 3). Although *hv*-DHFR contains a higher percentage of negatively charged amino acid residues (18.5%) than its non-halophilic counterparts (*ec*-DHFR, 15.7%; *lc*-DHFR, 13.6%), the difference is not statistically significant [10]. Moreover, the number of negatively charged residues is not uncommon in non-halophilic enzymes. Examination of the charge distribution in the context of the three-dimensional structure is required to gain better understanding of the halophilic properties of the enzyme.

Another notable aspect of the amino acid sequence is that while the percentage of positively charged residues is very similar to that of the other two bacterial DHFRs of known structure (9–10%), *hv*-DHFR exhibits an enrichment in arginine residues compared with lysine residues. The only two lysine residues in *hv*-DHFR are associated with the substrate-binding site, as discussed above. The positive charge is more localized in lysine than in arginine. An analysis of the distributions of water molecules around well ordered amino acid residues in high resolution crystal structures [30] shows that arginine sidechains tend to bind more water molecules than lysine sidechains. The increased arginine to lysine ratio has also been observed in other halophilic proteins [5]. A comparison of MDH amino acid sequences [31] reveals an arginine to lysine content of 8Arg:25Lys and 8Arg:21Lys for the non-halophilic enzymes from pig mitochondria and *E. coli*, respectively, whereas the Arg:Lys content for the halophilic enzyme is 15Arg:8Lys. A similar tendency is observed in 2Fe–2S ferredoxin: the Arg:Lys ratios for the non-halophilic *Anabaena* and halophilic enzymes are 0Arg:4Lys and 3Arg:3Lys, respectively [8].

Table 3

Amino acid composition (%) of bacterial DHFRs with known structure.

Amino acid residues	*H. volcanii*	*E. coli*	*L. casei*
Nonpolar	44.4	45.9	43.8
Polar	27.8	28.9	32.1
Charged	27.7	25.1	24.1
Arg	13(8.0)	9(5.6)	8(4.9)
Lys	2(1.2)	6(3.8)	9(5.6)
Glu	16(9.9)	12(7.5)	7(4.3)
Asp	14(8.6)	13(8.2)	15(9.3)

Previous comparative modeling led to the proposal that the structure of *hv*-DHFR shows an asymmetrical charge distribution over the protein surface, with positively charged amino acids centered around the active site and negative charges on the opposite side of the enzyme [9]. Indeed the crystal structure reveals that the highest concentration of positively charged residues are located in the vicinity of the nucleotide- and substrate-binding sites, whereas most of the remaining surface of the protein is negatively charged. However, because the substrate and cofactor are negatively charged compounds, the asymmetrical charge distribution is of functional importance and is also seen in the non-halophilic DHFRs, including *ec*-DHFR [32]. Figure 4 illustrates the electrostatic potential on the surfaces of *hv*-DHFR and *ec*-DHFR calculated with the program GRASP [33]. The figure shows that both proteins exhibit a highly negatively charged surface. In addition to the accumulation of positive charges in the active-site area, both proteins show positively charged regions on the opposite side of the protein. In *ec*-DHFR, a positively charged patch is centered around

Table 4

Statistics of the hydrogen-bonding pattern and salt bridges in *hv*-DHFR, *ec*-DHFR and *lc*-DHFR.

Data set (resolution)	Crystallization conditions	Intermolecular salt bridges*	Intermolecular hydrogen bonds*	Intermolecular salt bridges*	Intermolecular hydrogen bonds*
hv-DHFR (2.6 Å)	2.4 M phosphate pH 8.0 0.5% PEG 1K 2 mM methotrexate	2/12†	15/19†	8/6†	194/207†
ec-DHFR 4dfr (*E. coli*) (1.7 Å)	20% ethanol 5 mM Ca acetate His-HCl pH 6.8	8/7†	12/29†	6/6†	228/236†
ec-DHFR 7dfr (*E. coli*) (1.7 Å)	12% PEG 6K 3% ethanol phthalate buffer pH unknown	0	36	7	207
lc-DHFR 3dfr (*L. casei*) (2.5 Å)	55% NH_4SO_4 10 mM Tris pH 7.4	2	14	9	239

*The cut-off distance for salt bridges and hydrogen bonds was 3.5 Å. †The two numbers indicate the values for two independent molecules in the asymmetric unit. The PDB codes are as defined in Table 1.

residues Arg33, Lys38, Arg57 and Lys109. A positively charged region, comprising Lys30, Lys31, Arg34, Arg36, Arg115 and Arg157, is arranged differently in the halophilic enzyme. The region extends from the active site and forms a band between two negatively charged regions. One of these regions, comprising residues 121–156, contains a total of 11 negatively charged residues at the bottom of the β sheet. The second negative potential surface includes residues 2, 39–40, 83–95 and 114. Whether these different modes of forming positively charged regions have any implications for halophilic behavior is presently unknown.

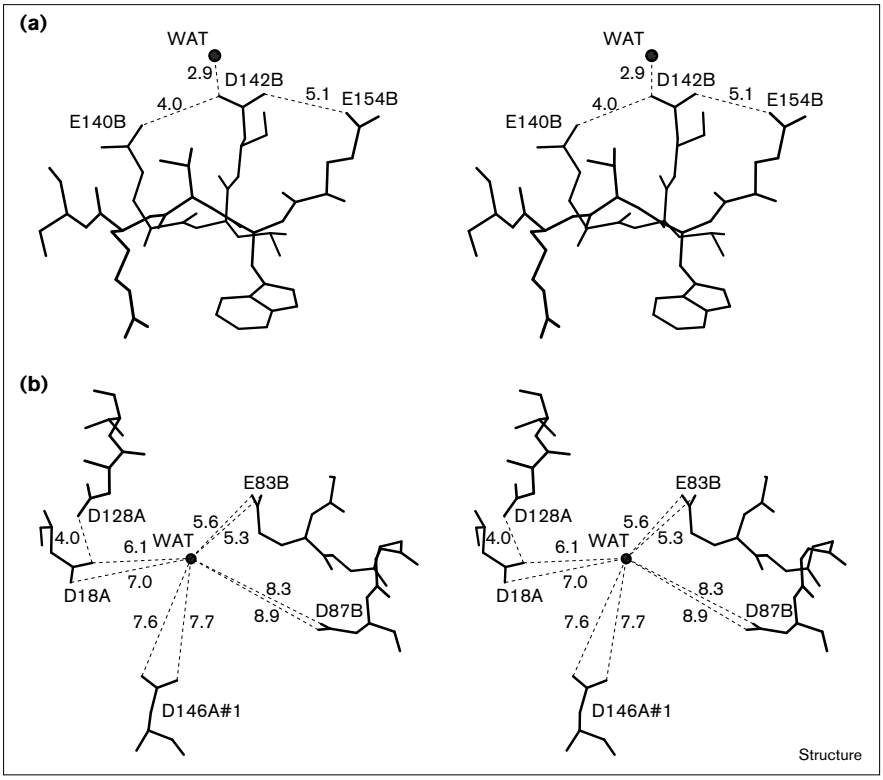
The dominance of negatively charged residues within the C-terminal segment of *hv*-DHFR is reminiscent of the first 62 residues of halophilic ferredoxin, which are highly enriched with carboxylate groups. The example of *hv*-DHFR is not as dramatic, however, and this same feature is also observed in the *ec*-DHFR sequence, where only negatively charged residues are located between residues 110 and 157 (numbering according to *ec*-DHFR).

The structure of MDH from *H. marismortui*, an extremely halophilic protein [7], contains more salt bridges than its non-halophilic homologs. It has been proposed that these salt bridges, together with the excess of acidic over basic residues, promote the stability of the protein at high salt concentrations. The structure of the 2Fe–2S ferredoxin from the same organism, that is only a moderately halophilic enzyme [34], lacks this feature [8]. Similarly, there is no significant increase in the number of salt bridge interactions in *hv*-DHFR (Table 4).

The crystal structure of halophilic 2Fe–2S ferredoxin includes a similar number of water molecules to that of a lysozyme structure determined at the same resolution using data collected at the same temperature. The water molecules in 2Fe–2S ferredoxin, however, form 40% more hydrogen bonds than in the lysozyme structure. A reliable analysis of solvent structure could not be made for *hv*-DHFR, because the resolution of the diffraction data was not sufficiently high.

The potential to form negatively charged clusters is expected to contribute most to the instability of halophilic proteins at low salt concentrations. The previous prediction of clusters made by Böhm and Jaenicke [9], based on comparative modeling, can be now assessed critically in the context of the experimentally determined structure of *hv*-DHFR. Two clusters of three carboxylate groups were highlighted by the authors. One proposed cluster, comprising Asp146, Glu148, and Glu154, is at error because Glu154 is in fact quite remote from the other two residues, with its sidechain oriented towards the opposite face of the β sheet than Asp146 and Glu148. The error arose from sequence misalignment compared with the structural alignment, a common problem of comparative modeling. Furthermore, Glu154 forms a salt bridge with Arg36, countering the carboxylate charge. The second proposed cluster, comprising Asp39, Asp40 and Glu95, is also inaccurate because the sidechains of Asp39 and Asp40 are oriented in the opposite direction to that of Glu95. The predicted structure is based on sequence alignment which considers Glu95 and a previous residue as an insertion compared with most other DHFRs, including the *E. coli*

Figure 6



Stereo representation of two clusters of negatively charged residues in *hv*-DHFR. Water molecules are depicted as black spheres and hydrogen bonds as dashed lines. **(a)** Intramolecular cluster; **(b)** intermolecular cluster at the interface of three molecules.

enzyme (see sequence alignment [9]). In fact, the structural alignment shows that there is no insertion in that region, and the polypeptide chain follows closely the trace of the *ec*-DHFR structure. In general, the *hv*-DHFR crystal structure shows that the sidechains of clustered negatively charged residues tend to orient as far away as possible from each other.

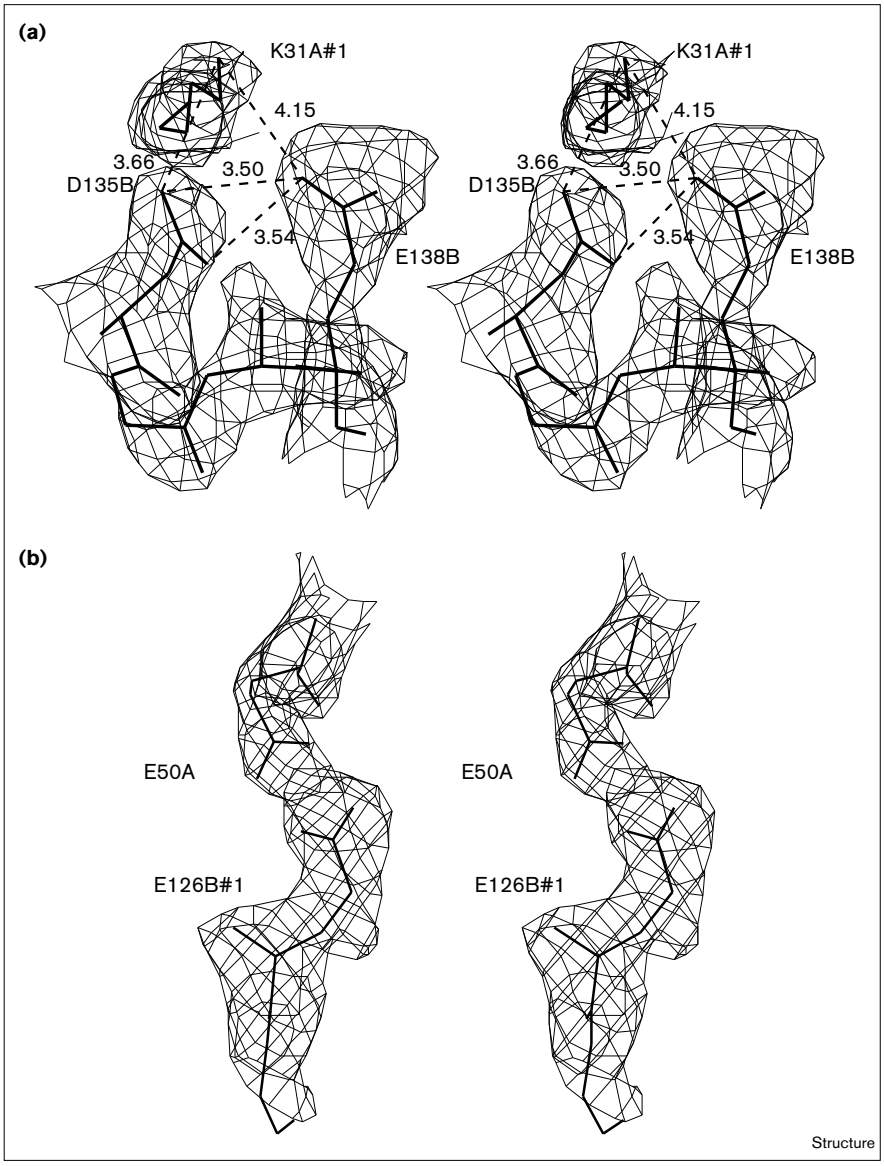
Examples of regions of the *hv*-DHFR structure where several negatively charge amino acid residues are located in close proximity to each other but avoid direct electrostatic interactions are shown in Figure 6. Glu140, Asp142 and Glu154 of molecule B are located on two adjacent β strands with a distance of 4–5 Å between their carboxylate groups (Figure 6a). A water molecule is coordinated to the central aspartate residue and the carboxylate group of Glu154 forms salt bridges to two arginines (not shown), reducing the repulsion between the negative charges.

To counteract the excess negative charges, binding of cations to the protein is expected [35]. Indeed, a number of potassium ions have been identified in the crystal structure of the halophilic ferredoxin [8]. The limited resolution of the diffraction data of the *hv*-DHFR crystals prevents identification of cations. Nevertheless, a cluster of five negatively charged residues at the interface of three neighboring molecules in the crystal has been identified (Figure 6b). A fairly strong positive peak in the difference electron-density map is located at the center of the solvent region surrounded by the cluster; the peak is located rather a long distance away from any of the carboxylate groups. A water molecule was assigned and refined at this position, although well ordered water molecules are not expected at such a location. It is quite possible that this peak represents a cation whose position may be partially restricted because of the surrounding negative charges. Long range restriction of ions around proteins has been observed in accurate molecular dynamic simulations performed in the context of the crystal of *Streptomyces griseus* protease A [36].

Only two direct intramolecular carboxyl–carboxylate interactions at a 3.5 Å cut-off distance occur in the structure of *hv*-DHFR (Figure 7). Figure 7a shows the interaction between Asp135 and Glu138 in molecule B, which also involves a lysine residue of a neighboring molecule. The lysine residue may mimic the role of a positively charged ion *in vivo* in neutralizing the unfavorable interaction. However, in molecule A, where no crystal contacts are observed, the sidechain of Glu138 is disordered.

The second unfavorable interaction occurs between two glutamic acid residues of neighboring molecules (Figure 7b). There is no positively charged residue located in the

Figure 7



Stereo representation of the carboxy–carboxylate interactions in *hv*-DHFR. The difference Fourier electron-density map is shown together with the model. The coefficients $2F_o–F_c$ and calculated phases were used; the map is contoured at 1σ level. **(a)** Intramolecular interaction between an aspartate and a glutamate residue. **(b)** Intermolecular glutamate–glutamate interaction.

vicinity. Carboxy–carboxylate interactions are associated with an abnormally high pK_a value of approximately 6, and have been observed in other crystal structures (see for example [37]). However, the DHFR crystals were obtained at a much higher pH (pH 8). Although the electron density associated with these residues is clear, the model of this region should be regarded with low confidence. The crystallographic temperature factors of the residues are high and amino acid residues 123–125, preceding Glu126 of molecule B, are disordered.

The two adjacent aspartate residues, Asp54 and Asp55, may be relevant to function. Although their sidechains point in opposite directions in the apo state of the enzyme, they may change orientation as the loop carrying them undergoes conformational transitions during the catalytic cycle. These two negatively charged residues are a unique feature of halophilic DHFR, thus they may play a role in linking the halophilic properties of the enzyme to function. This hypothesis could be tested by site-directed mutagenesis.

It has been suggested, that at high salt concentration hydrophobic interactions predominate and the protein molecules assume more tightly packed structures [6]. Therefore, the hydrogen bonding between amino acid residues has been analyzed to obtain information about the packing. The number of hydrogen bonds found in *hv*-DHFR, two *ec*-DHFR structures and *lc*-DHFR is similar, with no significant difference in the pattern (Table 4).

Table 5

Surface area and volume in various DHFR structures.

	Surface area (Å^2)*	Volume (Å^3)*
hv-DHFR	44.1/44.2†	123.2/122.1†
3dfr (*L. casei*)	46.5	130.0
4dfr (*E. coli*)	44.1/44.7†	127.9/130.0†
5dfr (*E. coli*)	43.9	123.7
6dfr (*E. coli*)	42.8	123.5
7dfr (*E. coli*)	43.7	123.6
8dfr (chicken)	46.2	131.7
1dyr (*P. carinii*)	46.5	136.4
1dhf (human)	48.0/45.8†	132.3/136.4†

*The surface area and volume are normalized to the number of amino acid residues. †The two values correspond to the two molecules in the asymmetric unit. The PDB codes are provided.

The surface areas and the volumes of various DHFR structures, calculated with the program GRASP and normalized to the number of amino acid residues, are summarized in Table 5. No significant trend is indicated from these data that would confirm the above proposal.

Biological implications

The delicate balance between stabilizing and destabilizing interactions in proteins renders them marginally stable. Proteins often respond to extreme changes in temperature, hydrostatic pressure, pH and salinity by unfolding, activity loss, and aggregation. Adaptation to extreme environments involves changes in the amino acid sequence that ensure the appropriate fold, substrate binding and solubility of the protein molecule in aqueous solution. The structural basis for adaptation to extreme halophilic conditions is addressed in this study.

Many halophilic proteins contain excess negatively charged amino acid residues in comparison with their non-halophilic counterparts. The presence of these negatively charged residues probably helps to solubilize these proteins when the availability of water molecules is restricted by high salt concentration. The amino acid sequence of dihydrofolate reductase from the Dead Sea archaeon *Haloferax volcanii* (*hv*-DHFR) is moderately enriched with negatively charged residues relative to the non-halophilic DHFRs. Nevertheless, evidence from the study presented here raises the possibility that this may not be an essential requirement for halophilicity. *hv*-DHFR functions at monovalent ion concentrations as low as 0.5 M, but its optimal activity is at 3–4 M. Recognizing the fact that many non-halophilic proteins are active at high salt concentrations, the question that is addressed here is why halophilic proteins unfold and loose activity at low salt concentrations. The structure of *hv*-DHFR lacks the striking features previously identified as characterizing halophilicity in ferredoxin and malate dehydrogenase. Neither the hypernegatively charged domain of ferredoxin nor the internal salt bridges of malate dehydrogenase occur in *hv*-DHFR. There are, however, clusters of negative charges on the surface of the molecule, which do not form close interactions but may be relevant to the moderately halophilic properties of the protein. These residues are implicated in halophilicity because they are energetically repulsive at physiological pH unless they are counteracted by cation binding.

The *hv*-DHFR structure provides new information about the apo state of the enzyme, showing considerable shifts of some loops and helical units compared with other DHFRs of known structure. These shifts indicate that the enzyme may undergo conformational transitions upon substrate binding that are different from those identified previously.

Materials and methods

Crystallization, data collection and data processing

DHFR from *H. volcanii* was over-produced by gene amplification in response to the addition of trimethoprin to the bacterial growth media [38]. The protein was purified as previously reported [39]. Crystals were obtained using the hanging drop vapor diffusion method at room temperature. The reservoir solution contained 2.4 M phosphate buffer (pH 7.8–8.0), 0.25–0.5% polyethylene glycol (PEG) 1000 and 2 mM methotrexate. Equal volumes of reservoir solution and 10 mg/ml protein solution in 0.1 M phosphate buffer at pH 7.0 were mixed. Crystals achieved their final size within two weeks. The crystals were monoclinic, belonging to the space group C2 ($a = 70.9$ Å, $b = 59.5$ Å, $c = 78.2$ Å, $\beta = 95.0°$). Density measurements using the water saturated xylene-bromobenzene gradient tube method [40] and the self-rotation function using AMoRe [23] showed that there were two crystallographically independent molecules in the asymmetric unit.

X-ray diffraction data to a resolution of 2.6 Å were collected at room temperature on a Siemens 3-axis goniostat using a single crystal of approximate dimensions of $0.2 \times 0.2 \times 0.3$ mm^3. Graphite monochromated CuKα radiation was generated with a Rigaku RU200BH rotating anode. The data were processed using the program suite XENGEN [41]. A total of 10 496 unique reflections (98% complete set) were collected at 2.55 Å resolution, with an R_{merge} value of 0.073 ($R_{merge} = \Sigma_h\Sigma_i|<I(h)> - I(h)_i|/\Sigma_h\Sigma_i I(h)_i$ for equivalent reflections). The data redundancy is 2.8, and the mean value of $I/\sigma(I)$ is 18.2. At the highest resolution shell between 2.71–2.55, the data is 88% complete and the mean $I/\sigma(I)$ is 2.4. The structure was solved by MR with the program AMoRe [23] using a multimolecule search model as described in the Results section.

Refinement

Two molecules of the modeled structure of *hv*-DHFR [21] were positioned in the cell according to the MR solutions and were initially subjected to X-PLOR simulated-annealing refinement in the temperature range 3000K to 300K [42]. Data in the range 7.0–3.0 Å for which $F \geq 2\sigma(F)$ were included. The R factor was 0.280 ($R_{free} = 0.380$; [43]). Approximately 70% of the initially refined model was consistent with the electron-density map. The computer graphics program TURBO-FRODO was used for model building [44]. Subsequent positional refinement using noncrystallographic symmetry restraints as implemented in X-PLOR and including only the 70% of the molecule

evaluated as correct did not improve the quality of the map in regions of missing amino acid residues. Gradual tracing of the remaining residues involved an additional seven cycles of simulated annealing and temperature-factor refinement at 2.6 Å resolution, followed by noncrystallographic symmetry averaging using RAVE [45,46]. At this stage the R factor was reduced to 0.220 ($R_{free} = 0.320$). The quality of subsequent electron-density maps enabled the addition of three phosphate ions and 80 water molecules. Next, noncrystallographic restraints were released and after further refinement the R factor was reduced to 0.180 ($R_{free} = 0.300$). Finally, a cycle of positional and temperature-factor refinement was performed, adding the randomly selected data previously used for the R_{free} calculation and resulting in an R factor value of 0.184 for data between 7.0–2.6 Å resolution for which $F \geq 2\sigma(F)$.

The crystals have been obtained from solutions containing methotrexate, an inhibitor of DHFR. In addition, trimethoprim, another inhibitor of DHFR, and NADPH were present during the expression of the protein. Thus, the site of the cofactor and substrate binding were scrutinized for binding of any of these compounds. Initially, residual electron density was observed in both sites, but it was not extensive enough to account for either methotrexate, trimethoprim, or NADPH. Moreover, both binding sites are partially blocked as discussed above, supporting the interpretation that this crystal structure represents the apo state of the enzyme. Phosphate ions and water molecules fully account for the density in the binding sites.

Accession numbers

The atomic coordinates have been deposited in the Brookhaven Protein Data Bank (entry code 1dvr).

Acknowledgements

We thank John Moult for useful discussions and Krzysztof Fidelis for providing the model of halophilic DHFR. This work was supported in part by NIH grant RO1-AI27175 (OH) and by US Israel Binational Science Foundation (MM). Ursula Pieper was a recipient of a research fellowship from the Deutsche Forschunggemeinschaft.

References

1. Christian, J.H.B. & Waltho, J.A. (1962). Solute concentrations within celles of halophilic and non-halophilic bacteria. *Biochem. Biophys. Acta* **65**, 506–508.
2. Lanyi, J.K. (1974). Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriol. Rev.* **38**, 272–290.
3. Jaenicke, R. (1981). Enzymes under extremes of physical conditions. *Annu. Rev. Biophys. Bioeng.* **10**, 1–67.
4. Eisenberg, H. & Wachtel, E.J. (1987). Structural studies of halophilic proteins, ribosomes, and organelles of bacteria adapted to extreme salt concentrations. *Annu. Rev. Biophys. Biophys. Chem.* **16**, 69–92.
5. Jaenicke, R. (1991). Protein stability and molecular adaptation to extreme conditions. *Eur. J. Biochem.* **202**, 715–728.
6. Eisenberg, H., Mevarech, M. & Zacchai, G. (1992). Biochemical, structural, and molecular genetic aspects of halophilism. *Adv. Protein Chem.* **43**, 1–62.
7. Dym, O., Mevarech, M. & Sussman, J.L. (1995). Structural features that stabilize halophilic malate dehydrogenase from an archaebacterium. *Science* **167**, 1344–1346.
8. Frolow, F., Harel, M., Sussman, J.L., Mevarech, M. & Shoham, M. (1996). Insights into protein adaptation to a saturated salt environment from the crystal structure of a halophilic 2Fe–2S ferredoxin. *Nat. Struct. Biol.* **3**, 452–458.
9. Böhm, G. & Jaenicke, R. (1994). A structure-based model for the halophilic adaptation of dihydrofolate reductase from *Halobacterium volcanii*. *Protein Eng.* **7**, 213–220.
10. Böhm, G. & Jaenicke, R. (1994). Relevance of sequence statistics for the properties of extremophilic proteins. *Int. J. Peptide Protein Res.* **43**, 97–106.
11. Bolin, J.T., Filman, D.J., Matthews, D.A., Hamlin, R.C. & Kraut, J. (1982). Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. I. General features and binding of methotrexate. *J. Biol. Chem.* **257**, 13650–13662.
12. Filman, D.J., Bolin, J.T., Matthews, D.A. & Kraut, J. (1982). Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. II. Environment of bound NADPH and implications for catalysis. *J. Biol. Chem.* **257**, 13663–13672.
13. Matthews, D.A., Alden, R.A., Freer, S.T., Zuong, N.-h. & Kraut, J. (1979). Dihydrofolate reductase from *Lactobacillus casei*, stereochemistry of NADPH binding. *J. Biol. Chem.* **254**, 4144–4151.
14. Champness, J.N., Achari, A., Ballantine, S.P., Bryant, P.K., Delves, C.J. & Stammers, D.K. (1994). The structure of *Pneumocystis carinii* dihydrofolate reductase to 1.9 Å resolution. *Structure* **2**, 915–924.
15. Matthews, D.A., *et al.*, & Kraut, J. (1985). Refined crystal structures of *Escherichia coli* and chicken liver dihydrofolate reductase containing bound trimethoprim. *J .Biol. Chem.* **260**, 381–391.
16. Stammers, D.K., *et al.*, & North, A.C. (1987). The structure of mouse L1210 dihydrofolate reductase–drug complexes and the construction of a model of human enzyme. *FEBS Lett.* **218**, 178–184.
17. Oefner, C., D'Arcy, A. & Winkler, F.K. (1988). Crystal structure of human dihydrofolate reductase complexed with folate. *Eur. J. Biochem.* **174**, 377–385.
18. Pearson, W.R. & Lipman, D.J. (1988). Improved tools for biological sequence analysis. *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
19. Sturrock, S.S. & Collins, J.F. (1993). *Mpsrch Version 1.3.* Biocomputing Research Unit, University of Edinburgh, UK.
20. Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F. & Weng, J. (1987). *Protein Data Bank, in Crystallographic Databases – Information Content, Software Systems, Scientific Applications. (*Allen, F.H., Bergerhoff, G. & Sievers, R., eds.), pp. 107–132, Data Commission of the International Union of Crystallography, Cambridge, UK.
21. Fidelis, K. (1996). Second Meeting on the critical assessment of techniques for protein structure prediction, Target No. T0001CM120_1. http://PredictionCenter.llnl.gov.
22. Leahy, D.J., Axel, R. & Hendrickson, W.A. (1992). Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6 Å resolution. *Cell* **68**, 1145–1162.
23. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst. A* **50**, 157–163.
24. Engh, R.A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst. A* **47**, 392–400.
25. Ramakrishnan, C. & Ramachandran, G.N. (1965). Stereochemical criteria for polypeptide and protein chain conformation. *Biophys. J.* **5**, 909–933.
26. Sawaya, M.R. & Kraut, J. (1997). Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: crystallographic evidence. *Biochem.* **36**, 586–603.
27. Bystroff, C., Oatley, S.J. & Kraut, J. (1990). Crystal structures of *Escherichia coli* dihydrofolate reductase: the $NADP^+$ holoenzyme and the folate·$NADP^+$ ternary complex. Substrate binding and a model for the transition state. *Biochem.* **29**, 3263–3277.
28. Blecher, O., Goldman, S. & Mevarech, M. (1993). High expression in *Escherichia coli* of the gene coding for dihydrofolate reductase of the extremely halophilic archaebacterium *Haloferax volcanii*. Reconstitution of the active enzyme and mutation studies. *Eur. J. Biochem.* **216**, 199–203.
29. Kuntz, I.D. (1971). Hydration of macromolecules. IV. Polypeptide conformation in frozen solutions. *J. Am. Chem. Soc.* **93**, 514–516.
30. Thanki, N., Thornton, J.M. & Goodfellow, J.M. (1988). Distributions of water around amino acid residues in proteins. *J. Mol. Biol.* **202**, 637–657.
31. Cendrin, F., Chroboczek, J., Zaccai, G., Eisenberg, H. & Mevarech, M. (1993). Cloning, sequencing, and expression in *Escherichia coli* of the gene coding for malate dehydrogenase of the extremely halophilic archaebacterium *Haloarcula marismortui*. *Biochemistry* **32**, 4308–4313.
32. Bajorath, J., Kitson, D., Kraut, J. & Hagler, A. (1991). The electrostatic potential of *Escherichia coli* dihydrofolate reductase. *Proteins* **11**, 1–12.
33. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamical properties of hydrocarbons. *Proteins* **11**, 281–296.
34. Werber, M.M. & Mevarech, M. (1978). Purification and characterization of a highly acidic 2Fe–2S ferredoxin from *Halobacterium* of the Dead Sea. *Arch. Biochem. Biophys.* **187**, 447–456.
35. Zaccai, G., Cendrin, F., Haik, Y., Borochov, N. & Eisenberg, H. (1989). Stabilization of halophilic malate dehydrogenase. *J. Mol. Biol.* **208**, 491–500.

36. Avbelj, F., Moult, J., Kitson, D.H., James, M.N. & Hagler, A.T. (1990). Molecular dynamics study of the structure and dynamics of a protein molecule in a crystalline ionic environment, *Streptomyces griseus* protease A. *Biochem.* **29**, 8658–8676.
37. Herzberg, O. & James, M.N.G. (1988). Refined crystal structure of troponin C from turkey skeletal muscle at 2.0 Å resolution. *J. Mol. Biol.* **203**, 761–779.
38. Rosenshine, I. & Mevarech, M. (1989). Isolation and partial characterization of plasmids found in three *Halobacterium volcanii* isolates. *Can. J. Microbiol.* **35,** 92–95.
39. Zusman, T., Rosenshine, I., Böhm, G., Jaenicke, R., Leskiw, B. & Mevarech, M. (1989). Dihydrofolate reductase of the extremely halophilic archaebacterium *Halobacterium volcanii.* The enzyme and its coding gene. *J. Biol. Chem.* **264**, 18878–18883.
40. Low, B.W. & Richards, F.M. (1952). The use of the gradient tube for the determination of crystal densities. *J. Am. Chem. Soc.* **74**, 1660–1666.
41. Howard, A.J., Gilliland, G.L., Finzel, B.C., Poulos, T., Ohlendorf, D.O. & Salemme, F.R. (1987). The use of an imaging proportional counter in macromolecular crystallography. *J. Appl. Cryst.* **20**, 383–387.
42. Brünger, A.T. (1992). *X-PLOR Version 3.1: a System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
43. Brünger, A.T. (1992). Free R-value: novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475.
44. Roussel, A. & Cambillau, C. (1989). TURBO_FRODO. In *Silicon Graphics Geometry Partners Directory*. pp. 77–78, Silicon Graphics, Mountain View, CA.
45. Jones, T.A. (1992). a, yaap, asap, @#*? a set of averaging programs. In *Molecular Replacement, Proceedings of the CCP4 Study Weekend.* (Dodson, E.J., Glover, S. & Wolf, W., eds), pp. 92–105, SERC Daresbury Laboratory, Warrington, UK.
46. Kleywegt, G.J. & Jones, T.A. (1994). Halloween masks and bones. In *From first Map to Final Model, Proceedings of the CCP4 Study Weekend.* (Bailey, S., Hubbard, R. & Waller, D., eds), pp. 59–66, SERC Daresbury Laboratory, Warrington, UK.
47. Bacon, D.J. & Anderson, W.F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219–220
48. Merritt, E.A. & Murphy, M.E.P. (1994). Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Cryst. D* **50**, 869–873.
49. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.
50. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950.