Halophilic Adaptation: Novel Solvent Protein Interactions Observed in the 2.9 and 2.6 Å Resolution Structures of the Wild Type and a Mutant of Malate Dehydrogenase from *Haloarcula marismortui*‡

Stéphane B. Richard,§,|| Dominique Madern,§ Elsa Garcin,§,⊥ and Giuseppe Zaccai*,§,@

Institut de Biologie Structurale, CEA-CNRS, 41 Avenue des Martyrs, F-38027 Grenoble Cedex 1, France, and Institut Laue Langevin, BP 156X, 38042 Grenoble, France

Received May 3, 1999; Revised Manuscript Received October 22, 1999

ABSTRACT: Previous biophysical studies of tetrameric malate dehydrogenase from the halophilic archaeon *Haloarcula marismortui* (*Hm* MalDH) have revealed the importance of protein−solvent interactions for its adaptation to molar salt conditions that strongly affect protein solubility, stability, and activity, in general. The structures of the E267R stability mutant of apo (−NADH) *Hm* MalDH determined to 2.6 Å resolution and of apo (−NADH) wild type *Hm* MalDH determined to 2.9 Å resolution, presented here, highlight a variety of novel protein−solvent features involved in halophilic adaptation. The tetramer appears to be stabilized by ordered water molecule networks and intersubunit complex salt bridges "locked" in by bound solvent chloride and sodium ions. The E267R mutation points into a central ordered water cavity, disrupting protein−solvent interactions. The analysis of the crystal structures showed that halophilic adaptation is not aimed uniquely at "protecting" the enzyme from the extreme salt conditions, as may have been expected, but, on the contrary, consists of mechanisms that harness the high ionic concentration in the environment.

To understand molecular adaptation to extreme salinity, it is important to address fundamental problems involving protein stabilization and solubility (*1*, *2*). This understanding is also essential for biotechnological applications such as the engineering of enzymes to function in special solvents for industrial or environmental use. In particular, what is the nature of protein−solvent and protein−protein interactions under very high salt conditions? Under these conditions, the concept of ionic strength is no longer sufficent and specific salt effects must be taken into account (*3*).

Halophilic Archaea grow optimally in molar NaCl, in marine salterns, and in hypersaline lakes such as the Great Salt Lake or the Dead Sea. They balance the external high salt concentration by accumulating intracellular KCl close to saturation (*4*−*6*). Stable, soluble, and active in high salt concentrations, proteins purified from these organisms generally unfold when the salt concentration decreases, e.g., below 2 M KCl (*2*). It has been shown from experiments with proteins in various salt solutions that, depending on the particular protein−ion−water interactions, certain salts favor the folded state whereas others favor unfolding (*7*, *8*).

Protein−solvent interactions are the focus of our studies on *Hm*[1] MalDH, a 132 kDa homotetramer with 303 residues per monomer (*9*, *10*). Through the availability of extensive biochemical and biophysical data, this enzyme has become a paradigm for halophilic proteins. Its gene has been cloned and the protein expressed in *Escherichia coli*, and the recombinant enzyme and stable mutants have been successfully renatured by dialysis against high salt (*11*, *12*). The protein is unstable under an unusually wide range of salt conditions and temperatures, making it a particularly interesting model for the study of the thermodynamics of solvent and salt effects (*2*, *13*−*16*), and it has been found that each salt type and concentration is associated with a particular mechanism dominating stability (*13*). Enthalpy terms dominate the activation free energy of dissociation and unfolding in KCl, in NaCl, and in an intermediate concentration range of $MgCl_2$, whereas entropy-driven mechanisms related to the hydrophobic effect and "salting-out" dominate in potassium phosphate or $(NH_4)_2SO_4$ (*13*, *17*). In high concentrations of NaCl or KCl, *Hm* MalDH and two other halophilic proteins, a polypeptide elongation factor and a Gra-PDH, have been shown to bind significantly large amounts of salt and water that may be characteristic of structural adaptation to halophilic conditions (*10*, *18*, *19*). Sequence comparisons between homologous proteins have shown that halophilic proteins have a significant excess of acidic residues over basic residues (*12*). Following experiments on the correlation

‡ The coordinates of the E266R mutant at 2.6 Å resolution and of the wild type at 2.95 Å resolution of MalDH from *H. marismortui* were deposited in the Protein Data Bank with accession codes 2hlp and 1d3a, respectively.

* To whom correspondence should be addressed: Institut de Biologie Structurale, 41 rue Jules Horowitz, F-38027 Grenoble Cedex 1, France. Phone: 33-4-76-88-95-73. Fax: 33-4-76-88-54-94. E-mail: zaccai@ibs.fr.

§ CEA-CNRS.

|| Present address: Structural Biology Laboratory, The Salk Institute for Biological Studies, P.O. Box 85800, San Diego, CA 92186-5800.

⊥ Present address: The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037.

@ Institut Laue Langevin.

[1] Abbreviations: apo, enzyme free of cofactor or substrate; *B*, crystallographic Debye−Waller factor; F_c, calculated structure factor amplitudes; F_o, observed structure factor amplitudes; *Hm*, *Haloarcula marismortui*; MalDH, L-malate dehydrogenase; holo, enzyme complexed with its cofactor; Gra-PDH, glyceraldehyde-3-phosphate dehydrogenase; LDH, L-lactate dehydrogenase; LURE, Laboratoire pour l'Utilisation du Rayonnement Eléctromagnétique.

10.1021/bi991001a CCC: $19.00
Published on Web 01/14/2000

between stability and salt binding, a stabilization−solvation model has been proposed in which the acidic side chains located at the protein surface cooperatively bind a network of hydrated salt ions that contribute to stabilization (*20*, *21*).

So far, the crystallographic structures of only three proteins from extreme halophiles have been published. That of dihydrofolate reductase from *Haloferax volcanii* appears to be similar to homologous proteins from nonhalophilic organisms. The gene product that has been analyzed, however, may reflect the result of lateral gene transfer rather than the endogenous halophilic enzyme (*22*). The structure of [2Fe-2S] ferredoxin from *Haloarcula marismortui* at 1.9 Å resolution revealed an acidic surface as well as extensive hydration networks (*23*). The crystal structure of tetrameric *Hm* MalDH, complexed with its NADH cofactor, determined to 3.2 Å resolution (PDB entry 1HLP), also revealed an acidic protein surface (*24*). Furthermore, an increase in the number of salt bridge clusters with respect to homologous structures suggested similarities with thermophilic proteins (*24*). The resolution limit, however, did not allow visualization of water molecules or bound solvent ions, so important questions remained unanswered. For example, how are intersubunit salt bridges stabilized in the very high salt concentrations of the halophilic environment? Are protein−solvent interactions involved? Electrostatic calculations on the crystal structures and homology-based modeling have also highlighted the role of protein−solvent interactions in halophilic adaptation (*25*−*27*).

The role of acidic residues in halophilic adaptation has been investigated by site-directed mutagenesis before the structure of *Hm* MalDH had been determined. Note that because we have now adopted the standard residue numbering for LDH according to ref *28*, the E267 residue corresponds to E243 in ref *12*. The E267R mutant is more sensitive to temperature than the wild type protein and requires significantly higher concentrations of NaCl or KCl for the equivalent stability. Its salt-dependent catalytic activity and K_M values, however, are not affected compared to those of the wild type protein (*12*). In the stabilization−solvation model, the decreased stability level of the mutant would be due to the perturbation of the solvation shell (*20*, *21*). We therefore collected crystallographic data on E266R *Hm* MalDH and the wild type (wt) protein with the goal of obtaining information about protein−solvent interactions. The resolution attained for the wt structure was limited to 2.9 Å. On the other hand, 2.6 Å resolution was achieved for the E267R mutant, providing significant new information about solvent−protein interactions.

The main results are as follows. Except for the mutated site, the structures of the E267R mutant and wt proteins are identical. The position 267 side chain points into a central cavity in the tetramer containing a network of 88 ordered water molecules. A solvent binding site that stabilizes important catalytic site residues was identified. A novel family of large complex salt bridges was observed, binding solvent chloride and sodium ions at the dimer−dimer and monomer−monomer interfaces, respectively. These data were analyzed to provide new insights into halophilic molecular adaptation.

MATERIALS AND METHODS

Crystallization. The wt and E267R protein were overexpressed in *E. coli*, renatured, and purified, and the cofactor was removed as described previously (*11*, *12*). The proteins were stored as ∼50 mg/mL stock solutions in 4 M NaCl and 50 mM Tris (pH 7.6), with 10 mM CsCl included in the wt solution only. Good diffracting crystals of wt *Hm* MalDH complexed with NADH in a concentrated NaCl solution had been obtained previously by using MPD (2-methyl-2,4-pentanediol) as a precipitating agent (*24*). The phase diagram of MPD/NaCl/H_2O and crystallization of *Hm* MalDH have been studied in detail (*29*). In fact, protein solubility decreases at low salt concentrations, and crystals appear upon dilution of the initial drop, in a "reverse" vapor diffusion method (*30*). This approach was used for both the wt and mutant enzymes, by using "sitting" drops. A 7 μL drop of 50 mM Tris (pH 7.6) containing 1.8 M NaCl, 58% MPD (v/v) (with 4.3 mM CsCl for the wt only), and 14 mg/mL protein was equilibrated against a 1 mL reservoir containing only MPD (60−70% v/v) in water. Fragile crystals up to 480 μm × 240 μm × 80 μm grew in 4−6 weeks. A constant temperature of 6 °C, required for the stability of the mutant protein, was maintained throughout, for purification, crystallization, and data collection. The crystals of both proteins belong to the orthorhombic space group $C222_1$ with the cell parameters given in Table 1. The asymmetric unit consists of one dimer. The solvent content of the crystals was about 70%, calculated according to the method described in ref *31*.

X-ray Data Collection. A full X-ray data set was obtained from three different crystals in each case, to a final resolution of 2.59 Å for the mutant and 2.94 Å for the wt protein, by using synchrotron radiation on the DW-32 wiggler beamline (*32*) at LURE (Orsay, France). The X-ray wavelength was 0.975 Å. The data were processed with XDS (*33*), scaled, and merged with the CCP4 suite (*34*) (Table 1).

Model Building and Refinement. The 3.2 Å structure of wt *Hm* MalDH (*24*), from which the NADH cofactor was removed, was used as starting model to determine the 2.6 Å resolution structure of the E267R mutant; the final model for the mutant was then used as a starting model for the apo wt structure to 2.94 Å resolution. Crystallographic refinement and rebuilding at the graphics interface were carried out in a cyclic process of gradual improvement of the model (*35*). Refinement was performed with the programs X-PLOR and CNS (*36*, *37*), by using specific parameters (*38*) and diffraction data between 8 and 2.59 Å. The statistical cross-validation method was applied by following the evolution of both the *R* factor, *R*, and the "free" *R* value, R_{free}, in all refinement and rebuilding cycles (*39*). As the use of constrained noncrystallographic symmetry (NCS) always produced an increase in R_{free}, NCS constraints were progressively relaxed during the cycles for the mutant structure and in the later stages during the refinement of the wt structure (*40*, *41*). "Solvent flattening" techniques were not used since we were interested in the solvent structure. The $2mF_o - DF_c$ and $mF_o - DF_c$ maps were calculated with the programs SIGMAA, EXTEND, and FFT from the CCP4 suite (*34*), X-PLOR, and the Groningen BIOMOL package (B. W. Dijkstra, unpublished results) for Bhat-omit maps $3mF_o - 2DF_c$ (*42*, *43*); all the data were included in the map calculations, as the low-resolution terms are important for placing the solvent (*44*). Assessment of the quality of the structure was carried out, after each refinement step, with the program PROCHECK (*45*). The visualization was performed on an 02/R500SC computer with IRIX 6.3 (Silicon

Table 1: Crystallographic Parameters and Data Processing and Refinement Statistics

	apo E267R	apo wt
space group	$C222_1$	$C222_1$
cell constants (Å)		
a	114.16	115.78
b	129.98	131.33
c	123.45	125.45
content of the asymmetric unit	2 monomers	2 monomers
solvent content (%)	69	69
X-ray source (LURE)	DW-32	DW-32
wavelength (Å)	0.975	0.975
temperature (°C)	4−6	6
no. of crystals	3	3
resolution range (Å)	28.63−2.59	30.0−2.94
final shell resolution range (Å)	2.75−2.59	3.07−2.94
no. of total reflections	159628	99423
no. of unique reflections	26445 (941)	19960 (3237)
no. of reflections in the working set	23573	15920
no. of reflections in the test set	1195	1003
R_{merge} (overall)[a] (%)	8.6	10.0 (33.2)
multiplicity	3.8	4.9 (5.2)
completeness (%)	88.5 (71.3)	94.2 (94.5)
model content of the asymmetric unit		
no. of amino acid residues	606	606
no. of protein non-H atoms	4606	4606
no. of solvent molecules (water)	250	45
no. of heterogen atoms		
chloride	2	2
sodium	1	1
R factor[b] (20.0−2.59 Å) (%)	18.8	18.8
R_{free}[c] (20.0−2.59 Å) (%)	23.2	24.4
B values		
from Wilson plot (Å^2)	32.7	
average *B* factor (Å^2)	42.4	61.5
rms deviations in *B* factors		
bonded main chain atoms (Å^2)	4.7	
bonded side chain atoms (Å^2)	8.1	
rms standard deviation from ideal values		
bond lengths (Å)	0.009	0.024
bond angles (deg)	1.5	2.09
dihedrals (deg)	21.7	23.34

[a] $R_{merge} = \{\Sigma_h\Sigma_i[I(h,i) - \langle I(h)\rangle]\}/\Sigma_h\Sigma_i[I(h,i)] \times 100$, where *R* is the merging *R* factor, $I(h,i)$ is the intensity value of the *i*th measurement of reflection *h*, and $\langle I(h)\rangle$ is the corresponding mean value of *h* for all *i* measurements of *h*. The summation is over all measurements. In parentheses are listed the values of R_{merge}, completeness, and multiplicity of the highest-resolution shell. [b] $R = \Sigma_{hkl}(|F_o - F_c|)/\Sigma|F_o|$. [c] Cross-validation was used through all refinement steps, by using 5% of the reflections.

Graphics) using the program O (*46*); building was carried out following the indications of the OOPS macro in O (*47*).

Assignment of Solvent Peaks. Solvent molecules were progressively added to the model during refinement and rebuilding cycles. Water molecules were at first assigned to positive extra-protein electron density peaks that appeared in both the $2F_o - F_c$ ($+1.0\sigma$ level) and $F_o - F_c$ ($+3.0\sigma$ level) maps. After visual inspection, a water molecule was conserved only if it could make at least one H-bond with the protein or another water molecule. After each cycle, the solvent peaks were critically evaluated with respect to electron density in the $2F_o - F_c$ and $F_o - F_c$ maps, the temperature factor (which should be similar to that of neighboring groups), and the structure environment. It was then decided whether to reject the assignment, maintain it as a water molecule, or explore the possibility of it being an ion (according to its electron density, its coordination, and the stereochemistry of the environment).

Quality and Analysis of the Model. Final refinement rounds were performed by using all diffraction data and employing

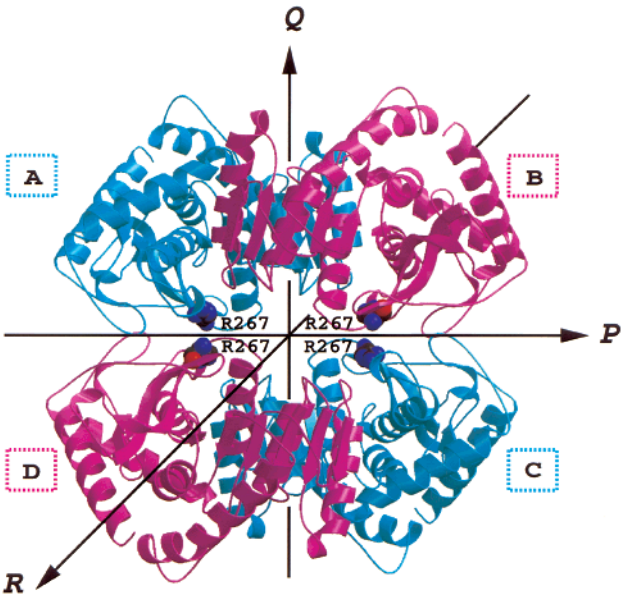

FIGURE 1: The 2.59 Å resolution structure of the apo (E267R) *Hm* MalDH tetramer. The mutated position 267 side chains, on either side of the *P* axis, are drawn in CPK mode. A−D identify the four protein subunits, and *P*−*R* identify the three 2-fold axes in the tetramer.

only energy minimization and temperature-factor refinement to produce a final assessment of the quality of the structure. As NCS was not constrained in the final refinement, differences between NCS-related monomers were analyzed "skeptically" (*41*). All the main chain dihedral angles (Φ and Ψ) of non-glycine and non-proline residues of the final E267R and wt models are within the most favored and additional allowed regions of the Ramachandran plot (data not shown). The error in the coordinates of the model is estimated to be ∼0.30 and 0.33 Å (and 0.40 and 0.44 Å for the cross-validated coordinate error, respectively) by the Luzzati plot for the E267R mutant and wt structures, respectively (*48*). Calculations of intra- and intermolecular distances, and of accessible surfaces, for sorting and statistical characterization of bound water molecules, were performed with the programs VOIDOO (*49*), and LSQKAB and CONTACT from the CCP4 package (*34*). The figures were drawn with MOLSCRIPT, BOBSCRIPT (*49*, *50*), Raster3D (*51*), and the program O (*46*).

Residue Numbering and Structure Description. The structure of *Hm* MalDH (Figure 1) is described according to the nomenclature proposed for dehydrogenases (*52*). The crystallographic dimer−dimer interface (between the A−B and C−D dimers) is along the *P* molecular axis. The two molecular 2-fold axes *Q* and *R* are between the monomers (A−B and C−D) and along the central cavity, respectively. Due to its high degree of similarity to LDH (*11*, *24*, *53*), the numbering of residues and abbreviations for secondary structure elements of *Hm* MalDH were given according to those used for *Squalus acanthias* LDH (*54*). This system was adopted at a time when the exact sequence of this LDH was not known. Consequently, some residue numbers are not used (21, 82, 104, and 300), and in cases where the same number refers to more than one amino acid residue, these are distinguished by letters (e.g., 29A, 29B, 54A, 54B, 54C, 132A, 132B, 210A, and 210B). The common numbering allows one to recognize readily important active site residues,

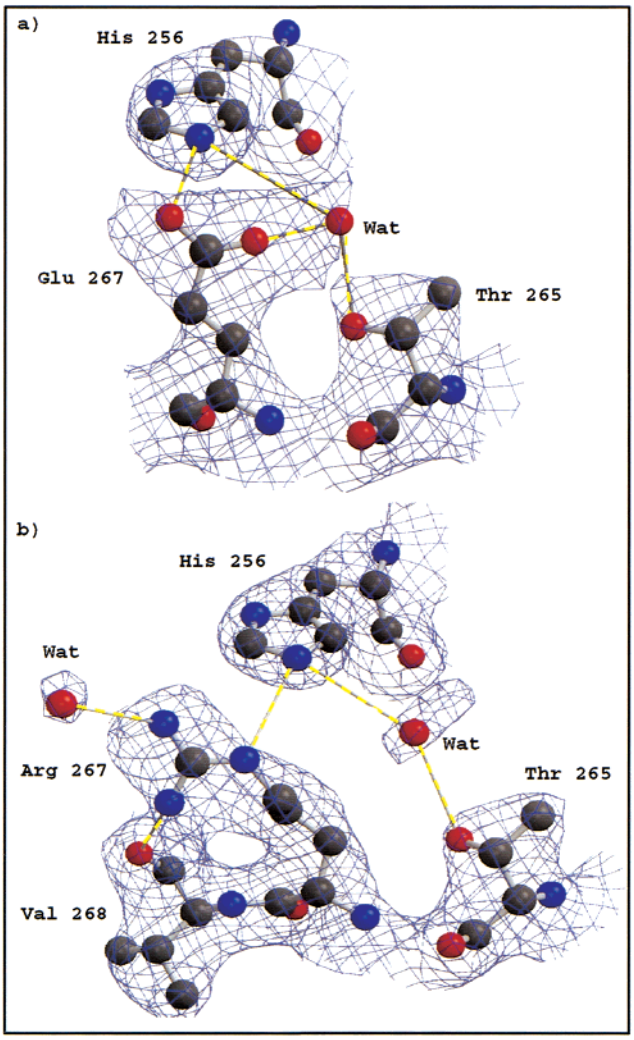

FIGURE 2: Ball-and-stick representation of the side chain at position 267. (a) E267 in the apo wt *Hm* MalDH at 2.94 Å resolution. (b) R267 in the mutant at 2.59 Å resolution. The carbon, nitrogen, and oxygen atoms are represented as gray, blue, and red spheres, respectively. The yellow thick and thin lines represent hydrogen bonds.

such as R102, R109, D168, R171, and H195, which are strictly conserved in all MalDHs and LDHs (*55*).

RESULTS

The 2.6 Å Resolution Structure of Apo E267R Hm MalDH Compared to the Apo wt at 2.9 Å Resolution. Compared to the previous 3.2 Å resolution structure (*24*), and as observed in other MalDHs and LDHs, the $\alpha_{1G/2G}$-helix was found to be more extended and a *cis*-proline (position 141) was clearly defined in the mobile catalytic loop. There is a root-mean-square deviation of 0.27 Å between the mutant and apo wt backbone C_α structures, showing that the mutation did not modify the structure of the main chain. Due to the lack of a structure with NAD at equivalent resolution, putative structural modifications induced by cofactor binding were not analyzed.

Mutation of the E267 Acidic Solvent-Exposed Side Chain. The electron density maps around the side chain of residue 267 (Figure 2) illustrate the quality of the data. In the apo wt structure, Glu 267 points into the central cavity and interacts with the protein via one intramolecular H-bond with His 256. In the mutant, Arg 267 forms a hydrogen bond with Wat 108 of the central cavity and is within hydrogen bond distance of His 256 and Val 268. The water molecule that bridges His 256 and Thr 265 is conserved in the two structures.

Water Structure. The solvent accessible surfaces in the tetramer are shown in Figure 3a. At first, about 250 nonprotein electron density peaks per asymmetric unit (dimer) were assigned as water molecules. Most are located in the first hydration shell, and make hydrogen bonds with the protein. None of these indicated the presence of ordered sulfate or phosphate ions left behind from the purification procedure. A few, however, were found to correspond to salt ions during the crystallographic analysis (see below). Many of the water molecules are associated in networks. The most extensive (88 hydrogen-bonded water molecules in a quasi-linear arrangement) fills the central cavity (along the *P* molecular axis), linking the four subunits (Figure 3b); smaller networks involving no more than six water molecules were found in the cofactor binding site, at the top of the monomer−monomer interface (*Q* molecular axis), and at the outer surface of the protein.

Complex Salt Bridges. Salt bridge clusters were identified at the monomer−monomer (*Q*) and dimer−dimer (*P*) interfaces (Figure 4). The cluster at the monomer−monomer interface (cluster 1) had not been analyzed in the 3.2 Å resolution structure of the wild type, and the clusters at the dimer−dimer interface (clusters 2 and 3) had been only partially constructed (*24*).

Cluster 1 is a small complex salt bridge involving six charged residues connected by a solvent molecule that was tentatively identified as a sodium ion (see below); it involves Glu 247(A), Lys 243(A), Asp 65(B), and Gln 62(B), interacting with the four symmetry-related side chains across the *Q* axis (the letter in parentheses refers to the monomeric subunit, Figure 1). There are two copies of cluster 1 in the structure linking monomers A to B, and C to D.

Cluster 2 is made up of Asp 209(A), Arg 292(D), and Glu 301(D) interacting with Asn 186(A) and the main chain oxygen atom of Gly 266(D). There are four copies of this complex salt bridge in the structure, located at the dimer−dimer interface, and linking the *P*-related subunits.

Cluster 3 is an extended complex salt bridge located around the *P* molecular axis. Its global composition is two Glu, two Arg, and two Asp residues and two Lys side chains connecting the two homodimers, A−B and C−D, respectively. Two similar groups of four residues (Lys 205, Asp 211, Arg 207, and Glu 188) are connected by an acid−acid pair (Glu−Glu with 2.7 and 3.0 Å distances between facing oxygen atoms). The two arginine residues (Arg 207) located on either side of the acidic pair are within hydrogen bonding distance, and may contribute to charge neutralization. This is not obvious in the two-dimensional representation of Figure 4. In three dimensions, however, it is clear that the two Glu 188 and the two Arg 207 residues are arranged in a cluster of four residues (two acidic and two basic) at a "connecting position" of the complex salt bridge, around the *P* molecular and 2-fold crystallographic axis. In addition, a small network of water molecules is hydrogen bonded to one side of the Glu−Glu pair in the cluster.

Chloride Ion Binding. The difference Fourier maps ($F_o - F_c$) showed high positive values ($+7.0\sigma$) at the extremities

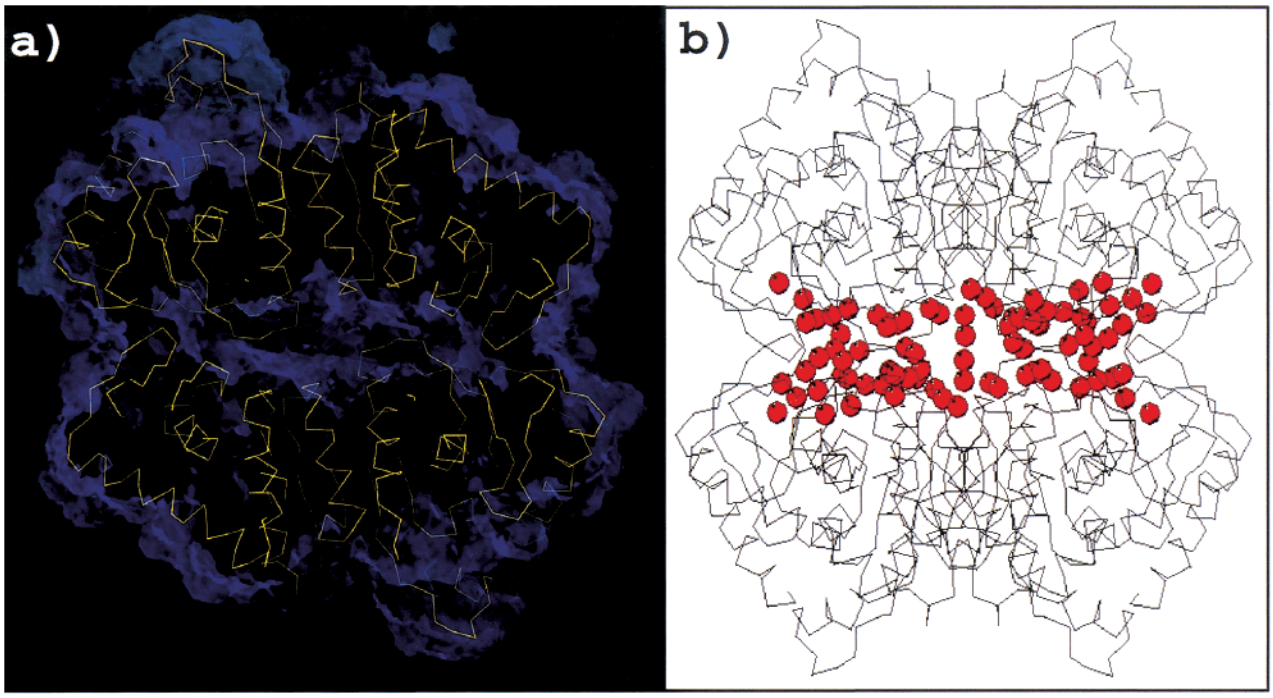

FIGURE 3: (a) Inter dimer cavity calculation performed with the program VOIDOO (*67*) on the final model of the mutant using a 1.4 Å probe. (b) Representation of the main water molecule network. The C_α backbone of the structure is shown in black lines, and the water molecules are shown in CPK mode. About 88 water molecules are hydrogen bonded in a large network filling the central cavity in the *Hm* (E267R) MalDH tetramer. This network is in direct contact with the position 267 side chain, and links the four subunits to each other.

of cluster 3, suggesting that a solvent molecule with high electron density is bound to the Lys 205 side chain (Figure 5). This peak was first modeled as a water molecule, but the associated temperature factor reached the minimum value allowed by the refinement program, characteristic of an electron gap between the model and the data. When the peak was modeled as a chloride ion, the residual positive electron density peak in the $F_o - F_c$ map disappeared. Cluster 3, therefore, has an anion binding site at each extremity. The chloride ion contacts are given in Figure 4 and Table 2. They include the positively charged Nz atom of Lys 205; Lys residues, known to be more common at the edges of complex salt bridges (*61*), are perfect candidates for a strong unidirectional interaction with the chloride ion. The anion is also close to the NH direction of two peptide groups, in a geometry similar to what has been described for sulfate and phosphate ions (*62*). Here, the chloride is bound via two weak bonds with the main chain nitrogen atom of Asp 306 from the same monomer as Lys 205 and to the main chain nitrogen from Asp 211 from the adjacent monomer. We cannot fully confirm that the chloride ion has a correct coordination because only three of its ligands are observed; the anion, however, is located at the surface of the protein, and the remaining ligands could well be solvent molecules that are not crystallographically ordered.

Sodium Ion Binding. At the top of the monomer−monomer interface, we found an acid−acid pair formed by the two symmetrically related Glu 247 side chains, which is not observed in other LDHs or MalDH (cluster 1, Figure 4). Just above the plane formed by these two side chains is located a solvent peak that was first modeled as water. Unlike the chloride ion (36 electrons), a sodium ion (10 electrons) cannot be distinguished from H_2O (10 electrons) by its electron density. There are various clues, however, which suggest that this solvent peak corresponds to a sodium ion. First, the coordination of the atom involves not less than nine bonds: four with the oxygen atoms of the acid−acid pair and five with other solvent atoms. A less well-defined water molecule, located 4 Å away from this solvent peak, could make up a tenth neighbor. The coordination geometry of the peak is distorted octahedral. This can be seen in the approximately planar square with the Na^+ in the center, formed by Wat 13, Wat 46, Wat 101, and Wat 116, and the Glu residues on either side (Figure 6). This coordination makes it very unlikely that the solvent peak represents a water molecule; in solution or in ice, a water molecule makes four to five hydrogen bonds with its neighbors. Sodium ions, on the other hand, have six or nine nearest neighbors (*62*) at distances of about 2.4 Å (*63*). Second, the acid−acid pair and its solvent environment suggest a metal binding site (*58*), which could easily coordinate a sodium ion in this case, as there is a relatively high concentration of sodium ions in the crystallization medium. Third, it has been shown that *Hm* MalDH binds a large number of salt ions (*13*). We propose, therefore, that a sodium ion is present at this position, even though the distances to its nearest neighbors may be slightly large (2.4−3.5 Å, Table 2). The valence calculation, according to the formula used for Na^+ screening (*64*), gives a value smaller than expected for a sodium ion at this site, but higher by three standard deviations than the values usually found for a water molecule.

The Catalytic Site Is Stabilized by a Bound Solvent Water Molecule or Ion. A buried solvent molecule (first modeled as a water molecule, Wat 2) could also correspond to a sodium ion by criteria based on coordination and bond lengths. It has eight ligands between 2.2 and 3.9 Å, of which six are

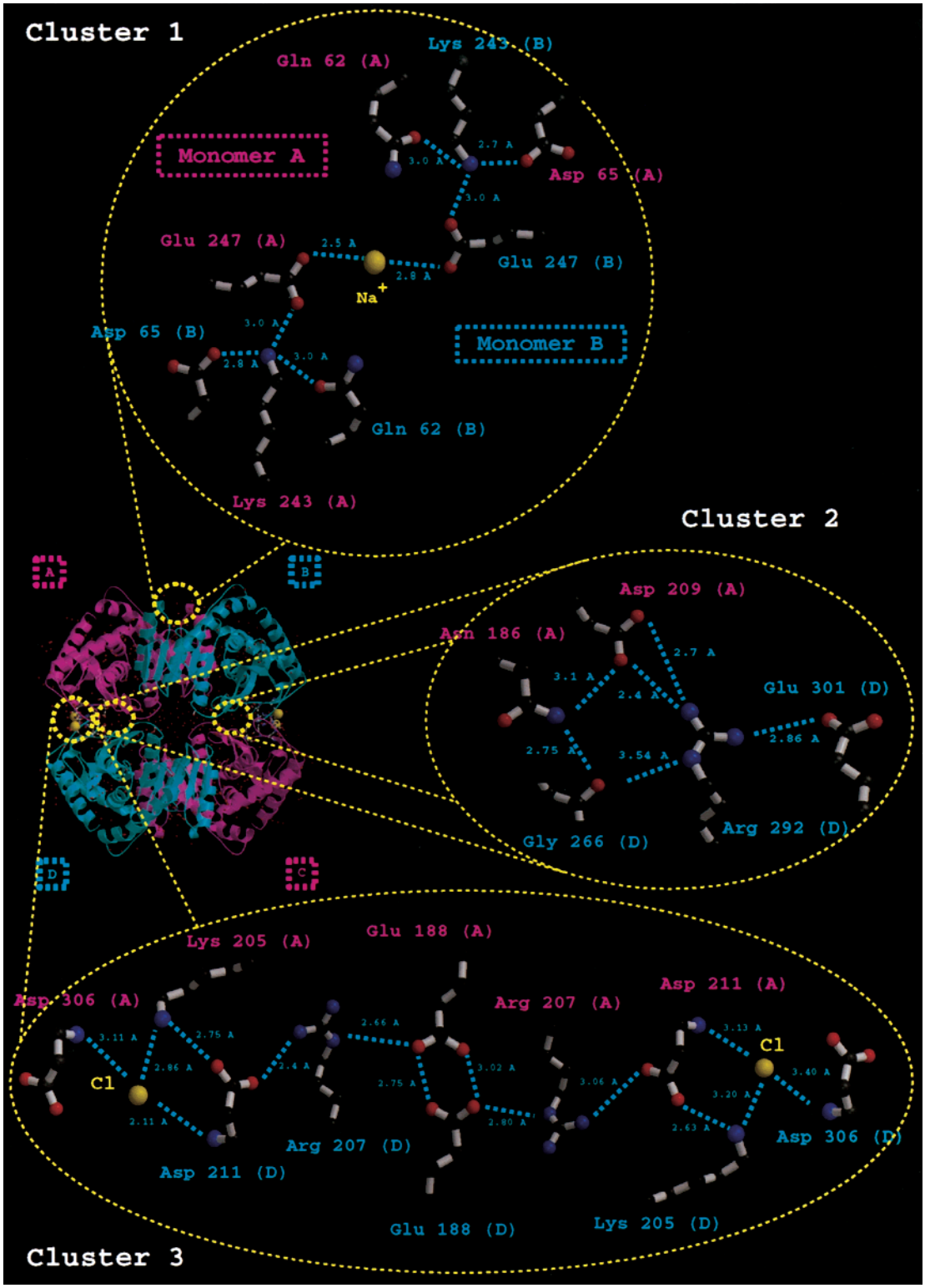

FIGURE 4: Schematic two-dimensional representation of the salt bridge and hydrogen bonds for the three complex salt bridge clusters found in the 2.6 Å resolution structure of *Hm* (E267R) MalDH (colors used are red for oxygen, blue for nitrogen, and gray for carbon). In cluster 1, at the *Q* axis monomer−momnomer interface, a central sodium ion links two triad salt bridges, involving Gln 62, Asp 65, Glu 247, and Lys 243 from two *Q*-related monomers. Cluster 2 is at the dimer−dimer interface and represents a small complex salt bridge involving three charged residues across the *P* molecular axis. Cluster 3 is a large complex salt bridge located around the *P* molecular axis, involving eight side chains (Glu 188, Lys 205, Arg 207, and Asp 211 from each monomer) and two chloride ions. There are also bonds involving main chain atoms in Arg 207, but they are not shown in the "exploded view".

oxygen atoms; it interacts with the side chain of Asp 168 and with main chain atoms in Ile 191, Gly 193, and Glu 194. These contacts stabilize a β-turn allowing His 195 to be correctly located with respect to Asp 168. In fact, a striking conserved feature in the sequences and structures of MalDH and LDH is the interacting His 195−Asp 168 pair, which acts as a proton relay in the active site. In *S. acanthias* LDH, the role of Wat 2 is filled by a cysteine side chain.

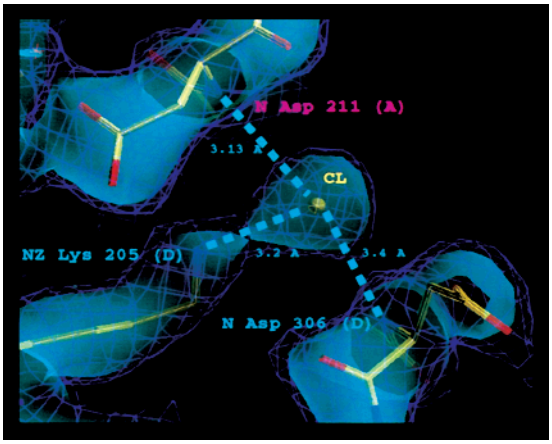

FIGURE 5: Plot of the chloride ion environment. The Nz of Lys 205, the N atom of Asp 306, and the N atom of the *P*-symmetrically related Asp 211 are involved in the interactions with the chloride ion. The $2mF_o - DF_c$ map is drawn at two different contour levels ($+1.0\sigma$ in broken lines and $+1.2\sigma$ in transparency).

Table 2: Coordination of Bound Cl^- and Na^+

site	interface	ligand	*d* (Å)	*B*
Cl^-_1	A−D	Nz Lys 205(A)	2.86	27.3
		N Asp 211(D)	3.40	31.7
		N Asp 306(A)	3.10	42.6
Cl^-_2	A−D	Nz Lys 205(B)	3.20	40.6
		N Asp 211(C)	3.12	30.9
		N Asp 306(B)	3.40	44.8
Na^+	A−B	OE1 Glu 247 (A)	3.58	37.9
		OE2 Glu 247 (A)	2.76	46.8
		OE1 Glu 247 (B)	3.72	28.3
		OE2 Glu 247 (B)	2.81	37.7
		Wat 8	3.98	36.6
		Wat 13	3.71	54.5
		Wat 46	3.03	39.2
		Wat 101	3.10	49.8
		Wat 116	3.38	45.8
		Wat 161	3.75	48.8

DISCUSSION

Proteins in High Salt. What are the problems for protein folding, stabilization, and solubility that have to be solved by halophilic adaptation? Folding and stabilization of monodomain proteins should in fact be favored by the mildly "salting-out" conditions provided by high concentrations of NaCl or KCl (*8*, *16*). Processes involving domain surfaces such as intersubunit interactions and macromolecule solubility, on the other hand, are in competition with the high-salt environment (*15*). Intersubunit complex salt bridges, for example, often found in thermophilic proteins, were also observed in *Hm* MalDH, raising the question of their stability at very high salt concentrations.

Effect of the Mutation. The E267R *Hm* MalDH mutant was constructed before the structure had been determined with the purpose of perturbing a hypothetical solvation layer around the enzyme (*12*). If it were the number of negative charges on the protein surface which determines low salt instability, then one would expect the mutant to remain stable to lower salt concentrations. This is not the case. E267R *Hm* MalDH requires higher concentrations of the solvent salt for stability similar to that of the wild type. It is not directly clear from the local structure around the position 267 side chain why the mutation reduces protein stability, unless it is due to a perturbation of solvent organization in the cavity necessary for tetramer stability. The destabilization may also be due to the electrostatic consequence of replacing an acidic group with a basic one; however, the similarity of the mutant and wt structures is more in favor of an effect on the solvent. Higher-resolution data of the wt protein are required to explore these hypotheses further.

Ordered and Disordered Hydration Layers in Hm MalDH. About 500 nonprotein electron density peaks per tetramer were assigned as water molecules in the crystal structure of the enzyme. However, about 200 molecules of salt and almost 3000 molecules of water were found to be weakly bound to the *Hm* MalDH tetramer by densimetry, analytical centrifugation, and forward scattering of neutrons (*10*). Significantly fewer solvent peaks (water molecules or ions) than this are observed, suggesting that the water−salt shell is not as crystallographically ordered as the 88-molecule network in the central cavity, for example. Does this network include salt ions? Clearly, it does not contain chloride ions. But sodium ions and water cannot be distinguished by their electron number alone. However, Na^+ ions have a higher coordination number than water molecules. By this criterion, it is likely that the central cavity electron peaks correspond mostly to water. This may be due to the fact that the protein surface within the cavity contains both acidic and basic side chains and is, on the whole, significantly less electronegative than the outer surface. It has been shown by small angle scattering experiments that hydration shells around proteins, which are not fully observed by crystallography, have a higher density than the bulk solvent (*65*). The thermodynamics and crystallography experiments taken together suggested, therefore, that there are two types of hydration layers in *Hm* MalDH, one in the central cavity composed essentially of ordered water molecules and excluding salt and a high-density disordered layer made up of water molecules and salt ions on the outer protein surface.

Halophilic Adaptation at Interfaces. A complex salt bridge is seen at the top of the monomer A−monomer B interface involving symmetry-related residues forming cluster 1 and ending in a Glu−Glu pair apparently locked in by a Na^+ site (cluster 1, Figure 4). Structural features where two acidic residues face each other within hydrogen-bonding distances have been described previously (*57*); they commonly correspond either to metal ion binding sites or to purely protein−protein intersubunit interactions as seen in cluster 3 (Figure 4) (*58*). In a modeling study of dihydrolipoamide dehydrogenase from *H. volcanii*, an intersubunit acidic cluster has been proposed as a potential K^+ binding site but it was not determined whether the site belonged to a complex salt bridge (*66*). At the dimer−dimer interface, four copies of a small salt bridge (cluster 2) and two copies of an extended complex salt bridge network locked in at the extremities by chloride ions (cluster 3) surround the central solvent cavity. We note that Glu−Glu pairs were found in both clusters 1 and 3. Their charge neutralization, however, appears to be coped with differently in each case. Whereas in cluster 1 the Glu−Glu pair forms a metal binding site for a hydrated sodium ion, at the dimer−dimer interface it forms an intersubunit group with two Arg side chains and water molecules. It has been shown by statistical analysis that acid pairs other than metal binding sites are frequently the Glu−Glu pair interacting with nearby positively charged Arg residues (*58*). As described above, an organized network of 88 water molecules in the

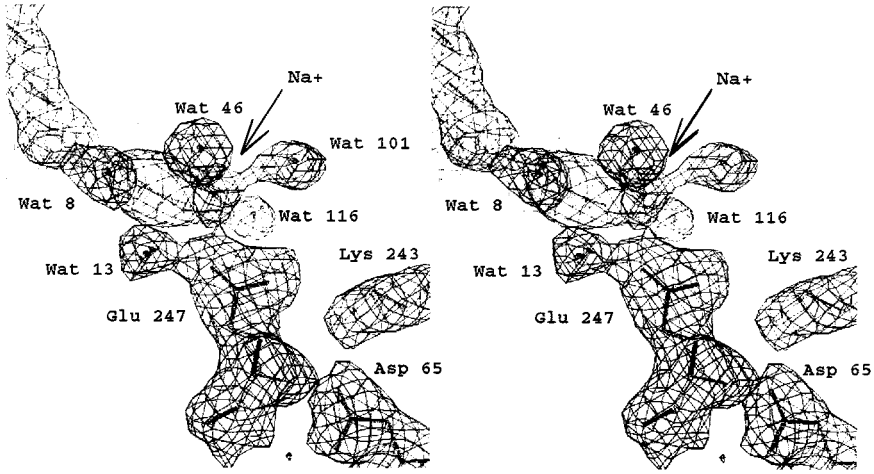

FIGURE 6: Stereoplot of the sodium ion environment. The figure shows a $2mF_o - DF_c$ map of one of the two salt bidge triads linked to the sodium ion. The map was calculated at 2.6 Å, and contoured at $+1.0\sigma$.

central cavity bridges the gap between the four subunits at the intersection of the *P*−*R* axes contributing to the stabilization of the tetramer. The complex salt bridges at the dimer−dimer interface are accessible to the water network in the central cavity. Complex salt bridges have been investigated in 94 protein structures (*59*), in which arginine most often serves as a key connector and/or branching unit because its geometry allows three possible interaction directions. A similar central cluster of four charged side chains has been described in the structure of a hyperthermophilic glutamate dehydrogenase (*60*).

In ref *53*, it is shown that *Hm* MalDH is a representative of an LDH-like group of MalDHs, in which some structural reorganization with respect to dimeric MalDH drives the association of two active dimers (A−B and C−D) into a tetrameric native enzyme. Consistent with these functional constraints, specific salt bridges clusters are more extended at dimer−dimer interfaces than at monomer−monomer interfaces within an active unit.

Is Salt Ion Binding a Halophilic Adaptation Feature That Stabilizes Complex Salt Bridges in the Extreme Salt Environment? A novel feature revealed by the structure presented here is complex intersubunit salt bridges that appear to be locked in by solvent ion binding sites: sodium ions at the top of the *Q* monomer−monomer interfaces and chloride ions at the dimer−dimer interface. Complex intersubunit salt bridges are often seen in thermophilic proteins (*67*−*70*). They do not appear, however, in the structure of the hyperthermophilic formylmethanofuran:tetrahydromethanopterin formyltransferase from *Methanopyrus kandleri*, which requires high concentrations of lyotropic salts [K_2HPO_4 or $(NH_4)_2SO_4$] to be stable (*71*). In *Hm* MalDH, solvent ions appear to be recruited by the complex salt bridges to contribute via their binding energy to stabilization and offer protection against charge neutralization in the high-salt environment. At high salt concentrations, the negative free energy of ion binding to the sites formed by the salt bridge networks will counterbalance screening of protein-charged groups by solvent ions. This might well be an important feature of halophilic adaptation.

Role of Ion and NADH Binding in Hm MalDH Folding. NADH binding stabilizes the wild type protein at low salt concentrations (*15*). In ref *53*, the roles of the salt bridge clusters 2 and 3 in the stabilization of the tetramer were explored through the systematic mutagenesis of the arginine residues that are involved (*53*). An inactive monomer and an active dimer of the mutant were observed, depending on the solvent conditions. The monomer could be reactivated by either NADH or high NaCl concentrations. Cofactor binding alone was not sufficient, however, to induce tetramer formation from the dimeric species, whereas a high NaCl concentration was very effective. The observation of specific ion sites supports the hypothesis that either ion binding at high salt concentrations or NADH binding induces a monomer conformation that leads to enzyme oligomerization. NADH binding may well be equivalent to Na^+ binding in favoring dimerization, but it cannot replace high-salt solvent interactions involved in the association of two dimers to form the tetramer, indicating an important role for Cl^- binding at this interface.

Extensive biochemical and biophysical studies on *Hm* MalDH provided evidence that specific protein−solvent interactions should be considered as an integral part of the active protein structure. This view is confirmed by the 2.6 Å resolution structure of E266R *Hm* MalDH, which now represents the best characterized description of the enzyme. The results emphasize that halophilic adaptation is not aimed uniquely at "protecting" the enzyme from the extreme salt, as may have been expected, but, on the contrary, consists of mechanisms that harness the high ionic concentration in the environment.

ACKNOWLEDGMENT

We are very grateful to Professors Joel Sussman and Moshe Mevarech for their advice and scientific insight concerning this work and to Drs. Fréderic Vellieux and Christine Ebel for critical reading of the manuscript and many exciting discussions. Thanks are also due to Dr. Thierry Prangé and the staff of the DW-32 beamline at LURE, to Dr. Michel Roth and the staff of the D2AM beam line at ESRF (Grenoble, France), and to Dr. Otto Dideberg for providing access to his laboratory facilities.

REFERENCES

1. Jaenicke, R. (1991) *Eur. J. Biochem. 202*, 715−728.
2. Eisenberg, H., Mevarech, M., and Zaccai, G. (1992) *Adv. Protein Chem. 43*, 1−61.

3. Glasstone, S. (1938) *Recent Advances in Physical Chemistry*, 3rd ed., J. and A. Churchill Ltd., London.
4. Christian, J. H. B., and Waltho, J. A. (1962) *Biochim. Biophys. Acta 65*, 506−508.
5. Ginzburg, M., Sachs, L., and Ginzburg, B. Z. (1970) *J. Gen. Physiol. 55*, 187−207.
6. Oren, A. (1993) *FEMS Microbiol. Rev. 13*, 415−440.
7. Arakawa, T., Bhat, R., and Timasheff, S. N. (1990) *Biochemistry 29*, 1924−1931.
8. Von Hippel, P., and Schleich, T. (1969) in *Structure of Biological Macromolecules* (Timasheff, S. N., and Fasman, G. D., Eds.) pp 417−575, Marcel Dekker, New York.
9. Mevarech, M., Eisenberg, H., and Neumann, E. (1977) *Biochemistry 16*, 3781.
10. Bonneté, F., Ebel, C., Zaccai, G., and Eisenberg, H. (1993) *J. Chem. Soc., Faraday Trans. 89* (15), 2659−2666.
11. Cendrin, F., Chroboczek, J., Zaccai, G., Eisenberg, H., and Mevarech, M. (1993) *Biochemistry 32*, 4308−4313.
12. Madern, D., Pfister, C., and Zaccai, G. (1995) *Eur. J. Biochem. 230*, 1088−1095.
13. Bonneté, F., Madern, D., and Zaccaï, G. (1994) *J. Mol. Biol. 244*, 436−447.
14. Ebel, C., Faou, P., Franzetti, B., Kernel, B., Madern, D., Pascu, M., Pfister, C., Richard, S., and Zaccai, G. (1998) In *Microbiology and Biogeochemistry of Hypersaline Environments* (Oren, A., Ed.) CRC Press, Boca Raton, FL (in press).
15. Ebel, C., Faou, P., and Zaccai, G.(1999) *J. Cryst. Growth 196*, 395−402.
16. Ebel, C., Faou, P., Kernel, B., and Zaccai, G. (1999) *Biochemistry* (in press).
17. Madern, D., and Zaccai, G. (1997) *Eur. J. Biochem. 249*, 607−611.
18. Ebel, C., Guinet, F., Langowski, J., Urbanke, C., Gagnon, J., and Zaccai, G. (1992) *J. Mol. Biol. 223*, 361−371.
19. Ebel, C., Altekar, W., Langowski, J., Urbanke, C., Forest, E., and Zaccai, G. (1995) *Biophys. Chem. 54*, 219−227.
20. Zaccai, G., Cendrin, F., Yeheskiel, H., Borochov, N., and Eisenberg, H. (1989) *J. Mol. Biol. 208*, 491−500.
21. Zaccai, G., and Eisenberg, H. (1990) *Trends Biochem. Sci. 9*, 333−337.
22. Pieper, U., Kapadia, G., Mevarech, M., and Herzberg, O. (1998) *Structure 6* (1), 75−88.
23. Frolow, F., Harel, M., Sussman, J. L., Mevarech, M., and Shoham, M. (1996) *Nat. Struct. Biol. 3*, 452−458.
24. Dym, O., Mevarech, M., and Sussman, J. L. (1995) *Science 267*, 1344−1346.
25. Böhm, G., and Jaenicke, R. (1994) *Protein Eng. 7*, 213−220.
26. Britton, K. L., Stillman, T. J., Yip, K. S. P., Forterre, P., Engel, P. C., and Rice, D. W. (1998) *J. Biol. Chem. 273*, 9023−9030.
27. Elcock, A. H., and McCammon, J. A. (1998) *J. Mol. Biol. 280* (4), 731−748.
28. Eventoff, W., Rossmann, M. G., Taylor, S. S., Torff, H.-J., Meyer, H., Keil, W., and Kiltz, H.-H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74* (7), 2677−2681.
29. Richard, S., Bonneté, F., Dym, O., and Zaccai, G. (1995) In *Archaea: A laboratory manual* (Dassarma, S., Ed.) pp 149−153, Cold Spring Harbor Laboratory Press, Plainview, NY.
30. Jeruzalmi, D., and Steitz, T. A. (1997) *J. Mol. Biol. 274*, 748−756.
31. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
32. Fourme, R., Dhez, P., Benoit, J.-P., Kahn, R., Dubuisson, J.-M., Besson, P., and Frouin, J. (1992) *Rev. Sci. Instrum. 63*, 982−987.
33. Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 67−71.
34. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr., Sect. D 50*, 760−763.
35. Kleywegt, G. J., and Jones, T. A. (1993) *ESF/CCP4 Newsletter 29*, 26−28.
36. Brünger, A. T. (1992) *X-PLOR*, version 3.1, Yale University Press, New Haven, CT.
37. Brünger, A. T., Adams, P. D., Clore, G. M., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nigles, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr., Sect. D. 54*, 905−921.
38. Engh, R. A., and Huber, R. (1991) *Acta Crystallogr., Sect. A 47*, 392−400.
39. Brünger, A. T. (1992) *Nature 355*, 472−475.
40. Kleywegt, G. J., and Brünger, A. T. (1996) *Structure 4*, 897−904.
41. Kleywegt, G. J. (1996) *Acta Crystallogr., Sect. D 52*, 842−857.
42. Baht, N. T. (1988) *J. Appl. Crystallogr. 21*, 279−281.
43. Vellieux, F. M. D., and Dijkstra, B. W. (1997) *J. Appl. Crystallogr. 30*, 396−399.
44. Podjarny, A. D., Howard, E. I., Urzhumtsev, A., and Grigera, J. R. (1997) *Proteins 28*, 302−312.
45. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
46. Jones, T. A., and Kjeldgaard, M. O. (1993) *O-The manual*, Uppsala University Press, Uppsala, Sweden.
47. Kleywegt, G. J., and Jones, T. A. (1993) *ESF/CCP4 Newsletter 29*, 20−24.
48. Luzzati, P. V. (1952) *Acta Crystallogr. 5*, 802−810.
49. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
50. Esnouf, R. M (1997) *J. Mol. Graphics 15*, 132−134.
51. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
52. Rossmann, M. G., Adams, M. J., Buehner, M., Ford, G. C., Hackert, M. L., Liljas, A., Rao, S. T., Banaszak, L. J., Hill, E., Tsernoglou, D., and Webb, L. (1973) *J. Mol. Biol. 76*, 533−537.
53. Madern, D., Ebel, C., Mevarech, M., Richard, S., Pfister, C., and Zaccai, G. (2000) *Biochemistry 39*, 1001−1010.
54. Abad-Zapareto, C., Griffith, J. P., Sussman, J. L., and Rossmann, M. G. (1987) *J. Mol. Biol. 198*, 445−467.
55. Goward, R. C., and Nicholls, D. J. (1994) *Protein Sci. 3*, 1883−1888.
56. Hendrickson, W. A. (1976) *Acta Crystallogr., Sect. A 32*, 922−923.
57. Zou, J., Flocco, M. M., and Mowbray, S. L. (1993) *J. Mol. Biol. 233*, 739−752.
58. Flocco, M. M., and Mowbray, S. L. (1995) *J. Mol. Biol. 254*, 96−105.
59. Musafia, B., Buchner, V., and Arad, D. (1995) *J. Mol. Biol. 254*, 761−770.
60. Yip, K. S. P., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, M. W., Robb, F. T., Vetriani, C., Maeder, D., and Rice, D. W. (1998) *Eur. J. Biochem. 255*, 336−346.
61. Chakrabarti, P. (1993) *J. Mol. Biol. 234*, 463−482.
62. Shannon, R. D. (1976) *Acta Crystallogr., Sect. A 32*, 751−767.
63. Vaney, M. C., Maignan, S., Ries-Kraut, M., and Ducruix, A. (1996) *Acta Crystallogr., Sect. D 52*, 505−517.
64. Nayal, M., and Di Cera, E. (1996) *J. Mol. Biol. 256*, 228−234.
65. Svergun, D. I., Richards, S., Koch, M. H. J., Sayers, Z., Kuprin, S., and Zaccai, G. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 2267−2272.
66. Jolley, K. A., Russell, R. J. M., Hough, D. W., and Danson, M. J. (1997) *Eur. J. Biochem. 248*, 362−368.
67. Yip, K. S. P., Stillman, T. J., Britton, K. L., Artymiuk, P. J., Baker, P. J., Sedelnikova, S. E., Engel, P. C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R., and Rice, D. W. (1995) *Structure 3*, 1147−1158.
68. Tanner, J., Hecht, R., and Krause, K. (1996) *Biochemistry 35* (2), 597−609.
69. Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M., and Pearl, L. H. (1997) *J. Mol. Biol. 271*, 789−802.
70. Villeret, V., Clantin, B., Tricot, C., Legrain, C., Roovers, M., Stalon, V., Glansdorff, N., and Van Beeumen, J. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 2801−2806.
71. Ermler, U., Merckel, M. C., Thauer, R. K., and Shima, S. (1997) *Structure 5*, 635−646.

BI991001A