Insights into the Molecular Basis of Salt Tolerance from the Study of Glutamate Dehydrogenase from *Halobacterium salinarum*

(Received for publication, October 16, 1997, and in revised form, December 3, 1997)

K. Linda Britton‡, Timothy J. Stillman‡, Kitty S. P. Yip‡, Patrick Forterre‡, Paul C. Engel¶, and David W. Rice¶‡

From the ‡The Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, United Kingdom, §Institute Genetique et Microbiologie, Batiment 409, Universite Paris-Sud, 91405, Orsay Cedex, France, and ¶Department of Biochemistry, University College Dublin, Belfield, Dublin 4, Ireland.

A homology-based modeling study on the extremely halophilic glutamate dehydrogenase from *Halobacterium salinarum* has been used to provide insights into the molecular basis of salt tolerance. The modeling reveals two significant differences in the characteristics of the surface of the halophilic enzyme that may contribute to its stability in high salt. The first of these is that the surface is decorated with acidic residues, a feature previously seen in structures of halophilic enzymes. The second is that the surface displays a significant reduction in exposed hydrophobic character. The latter arises not from a loss of surface-exposed hydrophobic residues, as has previously been proposed, but from a reduction in surface-exposed lysine residues. This is the first report of such an observation.

In highly saline environments, for example in salt lakes or in desiccating salt marshes, where salt concentrations can exceed 3 M, the dominant microorganisms are extremely halophilic Archaea (1). These halophilic organisms accumulate inorganic ions within the cell at concentrations equivalent to or greater than that of the environment (2), and their proteins are therefore specialized to function under high salt conditions. The ease with which these organisms are grown and the absence of a necessity for aseptic conditions makes them very attractive for commercial applications including, among others, production of bio-degradable plastics (3) and cosmetics (4). Furthermore, the structures of this enzyme from *Cs* and *Pf* are also very similar (10), strongly suggesting that the core of the three-dimensional structures of GluDH are highly conserved, and therefore these structures can serve as models for all other hexameric GluDHs. The sequences of the GluDHs from *Cs* (11) and *Pf* (13, 14) were aligned against one another using their three-dimensional structures as a guide. The alignment of the sequence of the GluDH from *Cs* (6) against those of the *Cs* and *Pf* enzymes was greatly simplified by its similarity to the latter (47% identity) (Fig. 1). Of the 68 residues strongly conserved across the family of GluDHs, 57 are conserved in the *Hs* enzyme. This suggests that the key residues concerned with the maintenance of the catalytic properties and structural framework of the enzyme are not modified by the necessity of the halophilic enzyme to operate in high salt conditions. Throughout this paper, unless specified, the *Hs* GluDH numbering is used to identify equivalent residues in the other GluDH sequences.

**Construction of the Model of *Hs* GluDH**—To date the structures of six glutamate dehydrogenases have been determined; *C. symbiosum* (8) *Escherichia coli*,*3 Neospora crassa*,*2 Pyrococcus furiosus* (10), *Thermotoga maritima* (15), and *Thermococcus litoralis*. The r.m.s. fit based on 436 residues of the *Cs* structure to that of *E. coli* is approximately 0.8 Å despite a sequence identity of only 52%. Even between the *Cs* and *Pf* enzymes (identity 32%), the r.m.s. fit based on 210 out of 419 equivalent residues is only 1.0 Å. Thus, given the close sequence relationship between the *Hs* and *Pf* enzymes (identity 47%), we have chosen to model the *Hs* enzyme onto the *Pf* structure using the other GluDH structures as a guide where small insertions or deletions make these more appropriate.

The sequence information from the alignment together with the program FRODO (16) was used to produce an atomic model for the halobacterial GluDH based primarily on the structure of the more closely related *Pf* enzyme. The major differences between this model and the structure of the *Pf* enzyme relate to differences in the path of the main chain caused by the occurrence of small insertions and deletions. Such differences are found in six regions and involve only 35 residues out of the 428 residues present in the model. First, at the N terminus, the level of sequence homology between *Pf* and *Hs* GluDH is particularly poor, and an additional 21 residues are found in the latter. In comparison, the clostridial enzyme has an additional 15 residues with respect to pyrococcal GluDH, with these residues folding to form a "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

To whom correspondence should be addressed. Tel.: 44 114 2224242; Fax: 44 114 2272907; E-mail: D.Rice@Sheffield.ac.uk.

The abbreviations used are: GluDH, glutamate dehydrogenase; *Hs*, *Halobacterium salinarum*; *Cs*, *Clostridium symbiosum*; *Pf*, *Pyrococcus furiosus*; r.m.s., root mean square; e, electron.
Fig. 1. Alignment of the sequences of the GluDHs (GDH) from Cs, Hs, and Pf. The amino acids for each sequence are identified by the single letter code, with those in lower case representing the N-terminal amino acids that are disordered in the x-ray structures of Cs and Pf GluDH and where, in the Hs enzyme, there is no equivalent in either of the other two GluDH sequences. These residues have not been included in the analysis. Positions of insertions and deletions between the sequences are indicated by –. The secondary structural elements of the three-dimensional structure of the Cs GluDH are shown above, with helices represented by cylinders and strands shown as arrows. A similar representation of the three-dimensional structure of the Pf GluDH is shown below. The numbering displayed beneath the alignment corresponds to the halophilic enzyme sequence. Residues that are located within 6 Å of any atom of the bound glutamate substrate or the dinucleotide cofactor are indicated by ○ and ●, respectively. Where residues lie within this limit for both bound substrate and cofactor, the residues are highlighted by ●. The solvent-accessible surface area of the side chains in the hexamer of the three GluDHs were calculated by the method of Lee and Richards.
an extra helix. We have therefore chosen to ignore the first six of the additional residues in the halopholic enzyme and to model the remainder as a helix, as found in clostridial GluDH. Inspection of the model appears to support this, as all six acidic residues in this region fall on an exposed face of the helix, and the only other charged residue (His-11) is partially exposed to solvent.

Elsewhere, the Cs GluDH structure was only used to model those loop regions where differences in length in the alignment occurred between the two archaeal enzymes but where similarities in length with the clostridial enzyme were noted. This affects the structure around residues 186–189, 226–228, and 250–253 inclusively in the Hs GluDH. The construction of this hybrid model therefore involved some local rebuilding and geometry regularization at the “annealing” points between the two structures but also at one other location where there was a deletion of one residue in the Hs sequence with respect to both the Cs and Pf GluDH structures (between residues 292 and 293 in Hs GluDH). Finally, the Hs sequence is one residue shorter at the C terminus.

This model was then used to provide the structural backbone onto which the relevant side-chain residues were substituted to produce a model for the Hs GluDH. If side chains were in common between the GluDH structure being used (primarily that of the Pf enzyme) and the Hs GluDH sequence, their conformations were retained. Where side chains required substitution, the position of the replaced residue was maintained as far as possible unless this introduced unacceptable steric clashes. Analysis of the final model using PROCHECK (17) suggested that the torsion angle distribution is typical for that seen in a high resolution protein structure, and there are no residues with disallowed Ramachandran angles.

**Analysis of the Model**—The solvent-accessible surface areas for each atom of both a monomer and the hexamer of the three enzymes were calculated using the algorithm of Lee and Richards (18), excluding the solvent molecules of the models. The resulting solvent-accessible areas for each residue were expressed as a fraction of the total solvent-accessible surface area for each type of amino acid (19). Atoms that recorded different solvent-accessible areas between the monomer and the hexamer were defined as the buried surface area on hexamer assembly. The definitions of Miller et al. (20) for nonpolar, polar, and charged constituents of proteins were used to tabulate the chemical composition of the surface.

**RESULTS**

**Differences in Amino Acid Composition**

The comparison of the amino acid compositions between halophilic proteins and their mesophilic counterparts has highlighted the emergence of three general trends in halophilic proteins: an excess of acidic over basic residues, an increase in the “borderline” hydrophobic residues serine and threonine, and a collective decrease in the strongly hydrophobic residues valine, isoleucine, leucine, and phenylalanine (21). Analysis of the amino acid compositions for the mesophilic, halophilic, and hyperthermophilic GluDHs (Table I) shows that the halophilic GluDH contains significantly more acidic residues, with 64% of the total number of charged residues being either aspartate or glutamate compared with 51 and 53% of such residues in the Pf and Cs enzymes, respectively. In total, the halophilic enzyme contains 77 acidic and 44 basic residues in each subunit, which gives rise to an overall negative charge of 198 for the hexamer. This compares with the significantly lower values for the net charge of −18 and −42 for the hexamers of Pf and Cs GluDH, respectively.

Although the number of glutamate, histidine, and arginine residues is similar for all three enzymes, the number of aspartate residues increases from 24 to 37 between the hyperthermophilic and halophilic GluDHs. At the same time, there is a dramatic reduction in the number of lysines from 32 to 13 in the same two enzymes. Similar increases in acidic residues and lowered lysine content have been found in a comparison between the sequences of elongation factor EF-Tu from *Halobacterium marismortui* and its counterpart from the mesophile *Methanococcus vannielli* (22). Consideration of the other potentially significant differences in composition shows there to be an increase in the proportion of threonine and a decrease in the proportion of phenylalanine in the halophilic GluDH compared with both enzymes and a reduction in isoleucine and increase in serine with respect to the Pf enzyme, all of which are fully consistent with the trends noted by Lanyi (21).

**Location of the Sequence Differences on the Three-dimensional Structure of GluDH**

**Analysis of Changes in the Buried Core**—The subunit structure of the Pf GluDH is shown schematically in Fig. 2. Each subunit in this hexameric enzyme is organized into two domains separated by a deep cleft, which forms the active site (Fig. 2). Sequence substitutions of totally buried residues almost exclusively involves conservative replacements from within the set of hydrophobic amino acids (Fig. 1). Substitutions within the subset of largely buried strongly hydrophobic residues (defined as having between 0 and 20% surface area accessible to solvent) commonly involves exchanges from within the set of hydrophobic amino acids, although the replacement by borderline hydrophobic residues such as threonine is also observed (Fig. 1). The replacement of hydrophobic residues that are partially exposed to solvent (defined as having at least 20% of the residue surface accessible to solvent) frequently involves modification to a polar or charged amino acid.

**Analysis of the Intersubunit Interface**—Analysis of the nature of the subunit surface that is buried on assembly of the hexamer (Table II) shows that the smaller proportion of nonpolar constituents observed for the solvent-accessible surface of the halophilic enzyme is repeated at this interface. Not surprisingly and in contrast to the solvent-accessible surface, the

| Amino Acid | Cs | Hs | Pf |
|-----------|----|----|----|
| Asp | 24 | 37 | 24 |
| Asn | 20 | 13 | 15 |
| Ala | 43 | 49 | 39 |
| Arg | 18 | 24 | 20 |
| Val | 45 | 39 | 35 |
| Pro | 20 | 24 | 19 |
| Gly | 48 | 33 | 34 |
| Glu | 34 | 40 | 36 |
| Gln | 16 | 12 | 11 |
| Cys | 2 | 3 | 1 |
| Met | 17 | 10 | 12 |
| Ile | 20 | 22 | 35 |
| Leu | 27 | 27 | 24 |
| Lys | 27 | 13 | 32 |
| His | 6 | 7 | 5 |
| Tyr | 21 | 15 | 20 |
| Phe | 19 | 4 | 10 |
| Trp | 5 | 9 | 10 |
| Ser | 20 | 22 | 14 |
| Thr | 17 | 31 | 23 |
| Total | 449 | 434 | 419 |

*References for the sequences given are Cs (11), Hs (6), and Pf (13, 14).*
number of charged residues and the charge balance at the intersubunit interface for these three GluDHs is very similar (Table II). This similarity is not surprising but strongly suggests that there are no gross errors associated with the modeling of the Hs enzyme.

The Nature of the Solvent-accessible Surface Area—Characteristics of the surface accessible to the solvent in the hexamers of the hyperthermophilic and mesophilic GluDHs from Pf and Cs have been analyzed previously (10). This study showed that although the proportion of the nonpolar surface was constant between these two enzymes, there was an increase in the charged nature of the hexamer surface in the hyperthermophilic Pf enzyme (24 and 28% of the solvent-accessible area in the Cs and Pf enzymes, respectively) and a corresponding decrease in the occurrence of polar groups relative to the Cs GluDH. Extending the study to include the Hs GluDH model (Table II) revealed a smaller proportion of nonpolar components and an even larger proportion of solvent-accessible charged groups relative to the Cs and Pf enzymes. Further examination of the area of the molecular surface that carries a formal charge has revealed that in each of the three GluDHs, the solvent-accessible surface area contains a higher proportion of negatively charged groups compared with the proportion of positively charged groups. However, this ratio is dramatically increased for the halophilic enzyme model, resulting in a surface predominantly covered in negatively charged residues. This is consistent with results obtained from the structure determinations of the malate dehydrogenase and the ferredoxin from the halophile Haloarcula marismortui (23, 24). The net charge density for the hexamer of the halophilic GluDH is $-2.6 \times 10^3$ eÅ$^{-2}$, far greater than the values of $-0.9 \times 10^3$ eÅ$^{-2}$ and $-0.5 \times 10^3$ eÅ$^{-2}$ for its mesophilic and hyperthermophilic counterparts, respectively, and comparable to the reported net charge densities of other halophilic proteins (24). These significant differences in the acidic nature of the hexamer surface can be seen both as a function of the electrostatic potential and in terms of the distribution of the charged residues on the protein surface (Fig. 3).

If we consider only those residues that are either partially or wholly accessible to the solvent, it is evident from the comparison of the three GluDHs that the larger total of exposed acidic

![A MOLSCRIPT representation of the fold of the polypeptide chain of a single subunit of the Pf GluDH is shown schematically (27).](image)

In this view, the 3-fold axis of the GluDH hexamer runs vertically. The enzyme is organized into two domains separated by a deep cleft. Domain I (upper) is responsible for subunit assembly, whereas domain II (lower) contains the nucleotide binding site. The elements of secondary structure are labeled with numbers representing helices (1–14) and letters for strands (a–m).

### Table II

| Analysis of the overall solvent-accessible surface areas for the GluDHs from Cs, Hs, and Pf calculated by the method of Lee and Richards (18) for both the monomers and hexamers Analysis of the chemical composition of the solvent-accessible surface area and the surface areas that are buried at the subunit interfaces in the hexameric GluDHs from the Cs, Hs, and Pf, respectively, in terms of nonpolar, polar, and charged constituent atoms. These calculations are based on the method of Miller et al. (20). |
|---|---|---|
| **Cs** | **Hs** | **Pf** |
| Solvent-accessible surface areas (Å$^2$) | | |
| Monomer | 18,500 | 18,500 | 17,000 |
| Hexamer | 88,000 | 93,000 | 82,500 |
| Characteristics of the solvent-accessible region (%) | | |
| Nonpolar | 47 | 43 | 48 |
| Polar | 29 | 25 | 24 |
| Negatively charged | 13 | 24 | 17 |
| Positively charged | 11 | 8 | 11 |
| Characteristics of the subunit interface region (%) | | |
| Nonpolar | 51 | 44 | 48 |
| Polar | 26 | 28 | 25 |
| Negatively charged | 7 | 11 | 10 |
| Positively charged | 16 | 17 | 17 |
Fig. 3. Comparative views of the surfaces of the halophilic, mesophilic, and thermophilic GluDHs to illustrate the different surface properties.  

a–c, views down the 3-fold axis of the molecular surfaces of the Hs, Cs, and Pf enzymes, respectively, to show the electrostatic potential at 0 M salt concentration, prepared using the program GRASP (28) (red corresponds to a surface potential less than −20 kcal(mol–electron)$^{-1}$; blue corresponds to a potential greater than +20 kcal(mol–electron)$^{-1}$).

d, the same view of the molecular surface of the hexameric GluDH from Hs prepared using the MIDASPLUS program (29, 30), highlighting the distribution of charged residues on the enzyme surface. All guanidinium, imidazole, and amino groups present in arginine, histidine, and lysine residues are shown in blue, and carboxyl groups associated with aspartate and glutamate residues are shown in red.

e–g, views parallel to the 2-fold axis (drawn in the same manner as d) of the Cs, Pf, and Hs enzymes, respectively.

h, a close-up of the molecular surface in the active-site region of the Hs enzyme. The location of the bound glutamate (green) and coenzyme (magenta) are shown, revealing the lack of negatively charged residues in the vicinity of the two binding sites, in sharp contrast with the remainder of the protein surface.
residues in the halophilic enzyme is dominated by the increase in aspartate. In fact, 36 aspartate residues per subunit fall into this category in the Hs enzyme compared with 22 and 23 aspartates in the mesophilic and hyperthermophilic counterparts. An equally dramatic reduction of solvent-accessible lysines is also observed with only 11 lysines of this type in the Hs GluDH model compared with 27 and 32 in the structure of the Cs and Pf enzymes, respectively.

We have also analyzed the solvent-accessible areas of the side-chain components of those residues for which significant differences in composition have been noted. The solvent-accessible area for aspartate in the Hs enzyme is more than double that found in the mesophilic and hyperthermophilic counterparts (Table III), reflecting the significant increase in composition associated with this residue type. Similarly, the solvent-accessible area associated with serine and threonine side chains is also greater in the model of Hs GluDH. Thus, 29 of the 31 threonines have some part of the residue solvent-accessible in the Hs GluDH model compared with 15 of the 17 and 18 of the 23 threonines of the Cs and Pf GluDH structures, respectively, implying that the increase in threonine residues is concentrated at the protein surface of this enzyme. Likewise, consideration of the serine residues present in each of the three GluDHs shows there to be 20 of the 22, 17 of the 20 and 12 of the 14 accessible to the solvent in the Hs, Cs, and Pf enzymes, respectively. Analysis of the Hs GluDH model places many of these surface-accessible serine and threonine residues adjacent to exposed acidic groups. Inspection of the model of the Hs enzyme suggests that many of these residues occupy positions where they could potentially hydrogen bond to nearby carboxyl groups.

Examination of the characteristics of the solvent-accessible area shows that there is a somewhat lower proportion of exposed hydrophobic groups in the model of the halophilic GluDH, with 43, 47, and 48% of the surface being hydrophobic in the Hs, Cs, and Pf enzymes, respectively (Table II). Analysis of the composition for a number of different residue types of the hydrophobic-accessible surface area is presented in Table III. This shows that in the solvent-accessible hydrophobic surface area of the Hs model, a reduction of some 2000 Å² is due to changes associated with the subset of strongly hydrophobic amino acid residues (Phe, Ile, Leu, and Val). Moreover, this reduction is offset by the increase in hydrophobic surface brought about by the additional aspartic acid residues. Of all the other changes, one stands out as being predominant. This involves a reduction of 3600 Å² and 6100 Å² in the exposed hydrophobic surface of the halophilic enzyme compared with the Cs and Pf enzymes, respectively, due to the reduction in the number of exposed lysine residues and the loss of the associated exposed alkyl component of these side chains. Moreover, of the lysines in the Hs model, the average solvent-accessible hydrophobic surface per residue is somewhat lower than for the other two GluDHs (175, 210, and 260 Å² for the Hs, Cs, and Pf enzymes, respectively). The explanation for this lies in the fact that those lysines that are retained in the Hs enzyme form the subset of such residues that are involved in the functional properties of GluDH and are in fact strongly buried in the structure (defined as a residue with less than 20% solvent-accessible surface). Thus, although there are 18 and 25 exposed lysines in the Cs and Pf enzymes, there are only 5 in the halophilic counterpart. In contrast, the number of strongly buried lysines is more constant across the three enzymes (8, 9, and 7 for the Hs, Cs, and Pf GluDHs, respectively).

To our knowledge the reduction in hydrophobic surface associated with the depletion of lysine in halophilic enzyme has not been reported even for those comparisons between structures of equivalent halophilic and mesophilic enzymes. Therefore to extend this analysis we have also compared the recently determined structure of the halophilic malate dehydrogenase (23) with the counterpart lactate dehydrogenase from dogfish (Protein Data Bank depositions 1HLP and 6LDH, respectively). Comparison of the amino acid compositions of these two enzymes also shows a lower number of lysines in the halophile (29 and 8 lysine residues in the dogfish and halophilic enzymes, respectively). Furthermore, solvent-accessible surface area calculations revealed a lower surface contribution of nonpolar components in the tetrameric halophilic malate dehydrogenase (42% compared with 50% in the mesophilic counterpart). Analysis of the contributions to this surface from the subset of strongly hydrophobic residues showed a fall of 900 Å² in the halophilic enzyme, insufficient to account for the overall reduction in hydrophobic solvent-accessible surface (5200 Å²). However, the lower lysine content again results in a loss of 5100 Å² in hydrophobic surface, which accounts for the difference between the two enzymes. Furthermore, the reduction in lysine content reported for H. marismortui elongation factor EF-Tu compared with the counterpart from the mesophile M. vannielii (22) is also consistent with our observations.

d) Analysis of the Active Site—The active site of this enzyme family has been located following analysis of the binary complexes of Cs GluDH with NAD⁺ and glutamate (8, 9) and of the Pf enzyme with NAP⁺. Analysis of the region around the glutamate binding pocket has shown that there are 20 amino acid residues that have at least 1 atom lying within 6 Å of any atom of the glutamate substrate (Fig. 1). Of these 20 residues, 16 are completely conserved across all three enzymes. At three of the remaining four positions (110, 161, and 164) the residues of the hyperthermophilic and halophilic enzymes are identical, and the final difference involves a substitution at residue 114 to cysteine in Hs GluDH, again a substitution that can be found in other members of the GluDH family (11). Overall, therefore, it would appear that the halophilic enzyme is remarkably similar to its mesophilic and hyperthermophilic counterparts in the region of the active site. Interestingly, we note that the assignment of the region of the enzyme surface that forms the active site in Hs GluDH would have been a potentially simple procedure even in the absence of direct structural information,

TABLE III

| Side chain | Total solvent-accessible area | Total hydrophobic solvent-accessible area |
|------------|-------------------------------|----------------------------------------|
|            | Cs Hs Pf                       | Cs Hs Pf                               |
|            | Å²  Å²                         | Å²  Å²                                  |
| Asp        | 5900 13,700 6500               | 1800 3500 1800                          |
| Glu        | 12,700 16,600 13,500           | 4200 4900 4500                          |
| Arg        | 6600 5100 4700                | 2500 2400 1700                          |
| Lys        | 10,000 3800 14,300             | 5700 2100 8200                          |
| His        | 1300 1000 1300                | 900 600 900                             |
| Ser        | 2300 4500 1400                | 1500 2100 1100                          |
| Thr        | 2700 4500 3500                | 2000 3000 2300                          |
| Phe, Ile, Leu and Val | 7200 5300 7500 | 7200 5300 7500 |

9 K. L. Britton, T. J. Stillman, K. S. P. Yip, and D. W. Rice, manuscript in preparation.
since it is the only region on the protein surface not to be dominated by the almost uniform coverage by acidic residues (Fig. 3). This may prove to be a general feature of many halophilic enzymes.

Analysis of Ion Pair Networks

The recently reported structure of the halophilic malate dehydrogenase (23) highlighted the formation of clusters of ion pairs, a feature shared by, although more prominent in, the thermophilic malate dehydrogenase from *Thermus flavus*. This feature is absent from the mesophilic dogfish lactate dehydrogenase and is thought to be related to the superior thermal properties of the *T. flavus* and the halophilic enzymes. The recent structure determination of the GluDH from the hyperthermophile *Pf* and its comparison with its counterpart from the mesophile *Cs* has also highlighted a potential role for ion pairs in the determinants of the thermal stability of this enzyme (10). Examination of the halophilic GluDH model suggests that the dramatic 18-residue ion pair cluster in the *Pf* enzyme, which is located across a region of the interface between pairs of dimers, is only partially retained in the *Hs* model, creating two symmetry-related ion-pair networks comprised of four residues (Fig. 4). These findings are consistent with the work on malate dehydrogenase (23) and may explain the apparently greater thermal stability of the halophilic GluDH compared with its mesophilic counterpart (25).

**Halophilic Addition**

The analysis of the structure of a halophilic 2Fe-2S ferredoxin (24) introduced the concept of halophilic addition. This describes the additional contribution made to the solvent-accessible surface of acidic residues arising from an insertion of an extra small domain of 33 residues that is rich in carboxylates (14 such residues in total), found near the N terminus of this protein. Compared with the mesophilic *Cs* enzyme, the halophilic GluDH is only six residues longer at the N terminus, with none of these being acidic. However, when compared with the hyperthermophilic enzyme, an N-terminal extension of 21 residues, including 6 acidic residues, can be seen. Although the proportion of acidic residues in the N-terminal region of the *Hs* enzyme (29%) is higher than in the protein as a whole (18%), taken together with the comparison with the mesophilic enzyme, the data from the analysis of GluDH are not strongly supportive of the presence of an additional carboxylate-rich domain.

**DISCUSSION**

Two distinct properties are commonly associated with enzymes from halophiles. The first of these is their ability to catalyze reactions under conditions of extremely high salt. For example, the GluDH from *Hs* is still active in 4 M KCl, whereas that from the mesophilic *Cs* is potently inhibited at such high concentrations. The second common property of these halophilic enzymes is that from the mesophile *Cs* concentrations. The second common property of these halo-ionic strength. For example, the GluDH from *Hs* is markedly unstable below 1 M KCl unless the salt is replaced by a compatible solute such as betaine (26). One immediate question that arises therefore is whether these two properties are related in molecular terms or whether they represent different aspects of the structure/function relationships.

Three features emerge as potentially significant in the comparison of the model of the halophilic enzyme with its non-halophilic counterparts. The first of these is that the surface of the model of the halophilic GluDH has shown that it is deco-

<FIG. 4. A schematic representation of the residues involved in the most extensive ion pair cluster of the Pf/GluDH involving 18 residues in total across part of the interface region between subunits of the assembled hexamer. Each amino acid is identified by its Pf residue number. The residues highlighted by boxes represent those that appear to be maintained in the Hs GluDH from the sequence alignment. The Pf residues concerned and their Hs equivalents are Glu-120 (E120) and Glu-141 (E141), Arg-124 (R124) and Arg-145 (R145), Asp-157 (D157) and Asp-178 (D178), Arg-414 (R414) and Arg-430 (R430), respectively.>

6 M. Kalinowski and P. C. Engel, unpublished results.
between two glutamate residues, where the presence of a third carboxylate would be unfavorable (24). Finally, the marked reduction in the number of surface lysine residues is a further dominant feature of the halophilic GluDH and, while helping to increase the overall negative charge on the protein, also serves to decrease the hydrophobic fraction of the solvent-accessible surface. To our knowledge, this is the first report of such an observation. Moreover, our analysis of the structure of the halophilic malate dehydrogenase compared with a structure for a mesophilic counterpart strongly supports this finding, showing both a decrease in the number of lysine residues and the consequent marked reduction in the contribution of solvent-accessible hydrophobic surface. At present, the significance of this observation is unclear, but it may be that the presence of significant numbers of alkyl groups on the enzyme surface may well serve to disrupt the production of a well connected hydration shell required in such saline environments, and therefore the long alkyl tails associated with lysines are particularly unfavorable.

For the future, this comparative analysis provides a hypothesis on the molecular basis of salt tolerance that is clearly testable by site-directed mutagenesis. One challenge therefore will be to rationally engineer such properties into mesophilic enzymes to exploit them in an industrial context. To date, this has not been accomplished, and although the structural data on halophilic enzymes may now point the way forward, we should be cautious in assuming that we now understand the structural basis of this phenomena and can manipulate it at will.

Acknowledgments—We thank the United Kingdom Biotechnology and Biological Sciences Research Council, the Wellcome Trust, the European Union Biotechnology Program, and the New Energy and Industrial Development Organization for financial support. The Krebs Institute is a designated Biotechnology and Biological Sciences Research Council Biomolecular Sciences Center.

REFERENCES

1. Woese, C. R., and Fox, G. E. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5088–5090
2. Eisenberg, H., Mevarech, M., and Zaccaci, G. (1992) Adv. Protein Chem. 43, 1–62
3. Lillo, G. J., and Ridgway-Valera, F. (1990) Appl. Environ. Microbiol. 56, 2517–2521
4. Galinski, E. A., and Tindall, B. J. (1992) in Molecular Biology and Biotechnology of Extremophiles (Herbert, R. D., and Sharp, R. J., eds) pp. 76–114, Blackie and Sons, London
5. Flam, F. (1994) Science 265, 471–472
6. Benachenhou, N., and Baldacci, G. (1991) Mol. Gen. Genet. 230, 345–352
7. Benachenhou-Lahlin, N., Labedan, B., and Forterre, P. (1994) Gene 140, 17–24
8. Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W., and Stillman, T. J. (1992) Proteins Struct. Funct. Genet. 12, 75–86
9. Stillman, T. J., Baker, P. J., Britton, K. L., and Rice, D. W. (1993) J. Mol. Biol. 234, 1131–1139
10. 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
11. Teller, J. K., Smith, R. M., McPherson, M. J., Engel, P. C., and Guest, J. R. (1992) Eur. J. Biochem. 206, 151–159
12. Britton, K. L., Baker, P. J., Rice, D. W., and Stillman, T. J. (1992) Eur. J. Biochem. 209, 851–859
13. Maras, B., Valiante, S., Chiaraluce, R., Consalvi, V., Politi, L., De Rosa, M., Bossa, F., Scandurra, R., and Barra, D. (1994) J. Protein Chem. 13, 253–259
14. Eggen, R. I. L., Geering, A. C. M., Waldkotter, K., Antranikan, G., and de Vos, W. M. (1993) Gene 132, 143–148
15. Knapp, S., de Vos, W. M., Rice, D., and Ladenstein, R. (1997) J. Mol. Biol. 267, 916–932
16. Jones, T. A. (1985) Methods Enzymol. 115, 157–171
17. Laskovski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallog. 26, 283–281
18. Lee, B., and Richards, F. M. (1971) J. Mol. Biol. 64, 379–400
19. Singh, J., and Thornton, J. M. (1992) Atlas of Protein Side Chain Interactions, IRL Press at Oxford University Press, Oxford
20. Miller, S., Janin, J., Lester, A. M., and Chothia, C. (1987) J. Mol. Biol. 196, 641–656
21. Lanyi, J. K. (1974) Bacteriol. Rev. 38, 272–290
22. Baldacci, G., Guinet, F., Tillit, J., Zaccaci, G., and de Recondo, A.-M. (1990) Nucleic Acids Res. 18, 507–511
23. Dym, O., Mevarech, M., and Sussman, J. L. (1995) Science 267, 1344–1346
24. Prohl, F., Harel, M., Sussman, J. L., Mevarech, M., and Shoham, M. (1996) Nat. Struct. Biol. 3, 452–458
25. Bonete, M. J., Camacho, M. L., and Cadenas, E. (1987) Int. J. Biochem. 19, 1149–1155
26. Cadenas, Q., and Engel, P. C. (1994) Biochem. Mol. Biol. Int. 33, 785–792
27. Kraulis, P. J. (1991) J. Appl. Crystallog. 24, 946–950
28. Nicholls, A., and Honig, B. (1991) J. Comput. Chem. 12, 435–445
29. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) J. Mol. Graphics 6, 13–27
30. Huang, C. C., Pettersen, E. F., Klein, T. E., Ferrin, T. E., and Langridge, R. (1991) J. Mol. Graphics 9, 230–236