Protein Engineering vol.15 no.3 pp.215–223, 2002

A low-resolution 3D model of the tetrameric alcohol dehydrogenase from *Sulfolobus solfataricus*

Rita Casadio[1], Pier Luigi Martelli[1], Antonietta Giordano[2], Mosè Rossi[2,3] and Carlo A.Raia[2,4]

[1]Laboratory of Biocomputing, Centro Interdipartimentale per le Ricerche Biotecnologiche (CIRB), Bologna and Laboratory of Biophysics, Department of Biology, University of Bologna, Via Irnerio 42, 40126 Bologna, [2]IBPE, Consiglio Nazionale delle Ricerche, Via G. Marconi 10, 80125 Naples and [3]Università di Napoli 'Federico II', Via Mezzocannone 16, 80134 Naples, Italy

[4]To whom correspondence should be addressed.
E-mail: raia@dafne.ibpe.na.cnr.it

We describe the computation of a model of the thermophilic NAD-dependent homotetrameric alcohol dehydrogenase from the archaeon *Sulfolobus solfataricus* (SsADH). Modeling is based on the knowledge that each monomer contains two Zn ions with catalytic and structural function, respectively. In the database of known structures, proteins with similar functions are either dimers containing two zinc ions per monomer or tetramers with one zinc ion per monomer. In any case, the sequence identity of the target to the possible templates is low. A threading procedure is therefore developed which includes constraints taking into account residue conservation both at the zinc ion binding and at the monomer–monomer interaction sites in the tetrameric unit. The model is consistent with previously reported data. Furthermore, cross-linking experiments are described which support the computed tetrameric model.
Keywords: cross-linking/multiple sequence alignment/protein modeling/thermostable alcohol dehydrogenase/threading

Introduction

Alcohol dehydrogenase (ADH) (EC 1.1.1.1) from the acidothermophilus archaeon *Sulfolobus solfataricus* (SsADH) is an NAD-dependent metalloenzyme that catalyzes the reversible oxidation reaction of alcohols to their corresponding aldehydes or ketones (Ammendola *et al.*, 1992). SsADH is an oligomeric metalloprotein endowed with remarkable thermostability and thermophilicity (Giordano *et al.*, 1999). In this context, the enzyme offers an excellent opportunity to investigate the extent to which thermostability is based on intra- and inter-subunit interactions, and also on the role of structural zinc and coenzyme. ADHs are widely distributed in Eucarya, Bacteria and Archaea (Sun and Plapp, 1992; Raia *et al.*, 2001); therefore, a careful comparison of the structure–stability–function relationship in the different species will also shed light on general rules of adaptability to different temperatures.

The sequence of SsADH (347 residues) was determined by gene and peptide analysis (Ammendola *et al.*, 1992). The tetrameric structure was defined by Cannio *et al.* (1996), who also reported the expression of the SsADH gene in *Escherichia coli.* Although SsADH is fairly resistant to denaturation at elevated temperatures, it shows a lower specific activity when compared with ADH from mesophilic or moderately thermophilic sources (Raia *et al.*, 2001). However, selective carboxymethylation of Cys38, a ligand of the catalytic zinc in SsADH, results in a more active although less stable enzyme (Raia *et al.*, 1996). On the other hand, the substitution of Asn249 with a tyrosine residue increases both SsADH activity and thermal stability (Giordano *et al.*, 1999). Preliminary crystallographic studies have been performed with apo and holo SsADH (Pearl *et al.*, 1993). However, the unsuccessful molecular replacement with crystals of apo-form and the twinning phenomenon which occurs on crystallization of the holo-form (Esposito *et al.*, 1998) hamper X-ray analysis.

This prompted us to develop a model of the tetrameric SsADH enzyme, which is described in this paper. Given the low sequence similarity of SsADH to the ADH structures currently known at atomic resolution, a procedure based on comparison (homology modeling) is not applicable. Our model is generated by adopting a threading procedure which includes constraints derived from both functional and structural characterization of the enzyme. In particular, we took advantage of the notion that SsADH contains two zinc atoms per subunit, with a catalytic and a structural function, respectively. In this respect SsADH is similar to the mesophilic dimeric two-zinc-containing horse liver enzyme (HLADH), although only 24–25% identical in sequence (Ammendola *et al.*, 1992). HLADH is known in two forms, apo and holo, depending on the absence or presence of bound NAD, which promotes a well documented rigid body inter-domain rotation of the catalytic domain with respect to the binding domain and a consequent closing of the inter-domain cleft (Colonna-Cesari *et al.*, 1986). However, although HLADH is the best homology model, it exists as a dimer. It was therefore necessary to use the structures of the two tetrameric ADHs in the PDB, from *Thermoanaerobacter brockii* (holo form only) (TbADH) and *Clostridium beijerinckii* (apo and holo forms) (CbADH), to model the quaternary organization, even though these enzymes share 24–25% sequence identity with SsADH and contain only one zinc per subunit. Experimental results consistent with the model are described.

Materials and methods

Chemicals

Dimethyl adipimate (DMA), dimethyl suberimidate (DMS) and ethylene glycol bis(succinimidylsuccinate) (EGS) were purchased from Pierce Chemical (Rockford, IL). Electrophoresis reagents and apparatus were obtainedfrom Bio-Rad (Hercules, CA). Other chemicals were A grade substances from Sigma Chemical (St. Louis, MO) or Applichem (Darmstadt, Germany).

Proteins

Preparation of coenzyme-free recombinant wild-type SsADH was performed as described previously (Giordano *et al.*, 1999). For the cross-linking experiments, the SsADH apo form was dialyzed against 0.1 M TEA–HCl, pH 8.5 and the holo form was prepared by adding to the latter 0.8 mM NAD.

Cross-linking

Cross-linking of SsADH with DMA, DMS and EGS was performed in 0.2 M TEA–HCl, pH 8.5, for 1 h at 25°C with a final protein concentration of 0.2 mg/ml and a protein to bifunctional reagent ratio of 1:6.2 and 4.4 (μg/μg). Reactions were stopped by denaturing the samples for 15 min at 90–100°C in the presence of 1–2% SDS and 2-mercaptoethanol. Gel electrophoretic separation was performed essentially as described (Davies and Stark, 1970) using 5% polyacrylamide rod gels. Amounts of 2–10 μg of protein were loaded on to each disk gel. Protein bands were visualized by staining with Coomassie Brilliant Blue G-250 and quantification of the band intensity was achieved using a Bio-Rad Model GS-710 imaging densitometer.

The spacer lengths of the bifunctional reagents DMA, DMS and EGS are 8.6, 11.0 and 16.1 Å, respectively. Owing to the flexibility of the polypeptide chain and the limited accuracy of the model, it is assumed that each reagent can span two neighboring lysine residues provided that their separation in the model is within ±2 Å, the length of the spacer between two reacting groups. Moreover, it is assumed that all Lys residues involved in cross-linking have the same reactivity.

Multiple sequence alignment

All the sequences of two zinc-containing ADHs were searched in the SwissProt database (Release 37, December 1998). The resulting 87 sequences were aligned with CLUSTALW (Thompson *et al.*, 1994). The residues involved in the binding of the two zinc atoms were deduced from this alignment. Similarly, the sequences of 14 tetrameric ADHs were searched and aligned to extract the sequence features that contribute to the oligomerization.

Structural alignments

The 3D crystal structures of ADHs were extracted from the PDB files and were superimposed with the BRAGI program (ftp.gbf.de/pub/Bragi) to obtain pairwise structural alignments.

Secondary structure prediction

The secondary structure of SsADH was predicted with a program based on neural networks (Jacoboni *et al.*, 2000) and with a consensus based procedure (Cuff and Barton, 1999); the secondary structural elements of the known ADHs structures were defined using the DSSP program (Kabsch and Sander, 1983). When necessary, solvent exposure of residues was evaluated with the DSSP program.

Model building and evaluation

The comparative model building was performed with the MODELLER program (Šali and Blundell, 1993). For a given alignment, 10 model structures were built and were evaluated with the PROCHECK suite of programs (Laskowski *et al.*, 1993). Only the best evaluated model was retained after the analysis. Tetrameric assemblages were obtained by superimposing with the BRAGI program the modeled monomer with each subunit of the tetrameric templates. Close contacts were removed with the DEBUMP program of the WHATIF package (Vriend, 1990).

Docking of the coenzyme

Docking simulation of NADH coenzyme into the monomer was performed with the AutoDock 3 program, which allows flexible docking by means of Lamarckian Genetic algorithms (Morris *et al.*, 1998). Fifty independent runs were performed, each one processing a population of 100 conformations for 27 000 generations with rates of mutation and crossover set to 0.02 and 0.8, respectively. The elitism parameter was set to 2. The docked conformation with the lowest value of estimated free binding energy was retained for further analysis. In order to compute the parameters to be used in docking the coenzyme to the SsADH holo model, the binding of NADH in the site of HLADH was performed. The result agrees with the conformation of co-crystallized coenzyme (PDB file 2OHX): the root mean square deviation (r.m.s.d.) value between the real and the docked conformations is 0.95 Å; the energy of interaction is –20.76 kcal/mol (–21.93 kcal/mol from intermolecular interactions, 1.18 kcal/mol from intramolecular interactions) and the free energy of binding is –16.64 kcal/mol, as estimated by the program AutoDock 3.

Results and discussion

Basis of the modeling procedure

A key feature of our modeling procedure stems from the following considerations: SsADH is a two-zinc-containing protein with low sequence identity (⩽25%) to all the other ADH sequences of known 3D structure; the functional unit of the target is a homotetramer; and the enzyme can bind NADH. In this respect, one should consider that the PDB database contains 31 atomic resolved structures of ADHs, corresponding to only eight different sequences from different species; 28 of these structures (six sequences) are two-zinc-containing dimeric ADHs, while the remaining three are one-zinc-containing tetrameric ADHs. There is no two-zinc tetrameric ADH structure available. Furthermore, 26 of these structures are holo and two-zinc dimers; two are apo and two-zinc dimers; two are holo and one-zinc tetramers; and one is apo and one-zinc tetramer. An interesting observation useful to our procedure is that the functional set of ADHs has monomers (containing either two or one zinc ions) whose structure is within 1.8 Å of the r.m.s.d. (data not shown). This confirms that ADH monomers have a rather well conserved structure irrespective of the level of sequence identity of the different species, number of zinc ions and aggregation state of the monomeric unit. This observation makes it highly probable that the monomer of our target will adopt a similar backbone conformation. Furthermore, if we require that the template structure has two zinc ions and is endowed with both the holo and apo forms, we are forced to adopt only HLADH as a template. However, the aggregation state of HLADH is dimeric whereas we know from experimental observations that our target is tetrameric and presumably with inter-monomer interactions different from those of the dimer. Therefore, for monomer aggregation we should adopt different targets, with a tetrameric aggregation state. A solution to this problem can be to search the ADH database of sequence and structures to find residues in the target which are conserved both in the two-zinc-containing and in the tetrameric enzymes. This would constrain threading of the target to the templates.

Multiple sequence alignment

In order to cope with our requirements, from the 90 ADH sequences contained in the SwissProt database, 87 were selected as containing two zinc ions. Multiple sequence alignment highlights that 38 residues are conserved in more than 90% of the 87 sequences (in Figure 1, only 20 out of the 87 aligned sequences are shown for practical reasons). Cys38, Gly67, His68, Glu69, Gly72, Gly87, Cys101, Cys104, Cys112, Cys154, Gly184 and Gly216 (numbering according to SsADH)

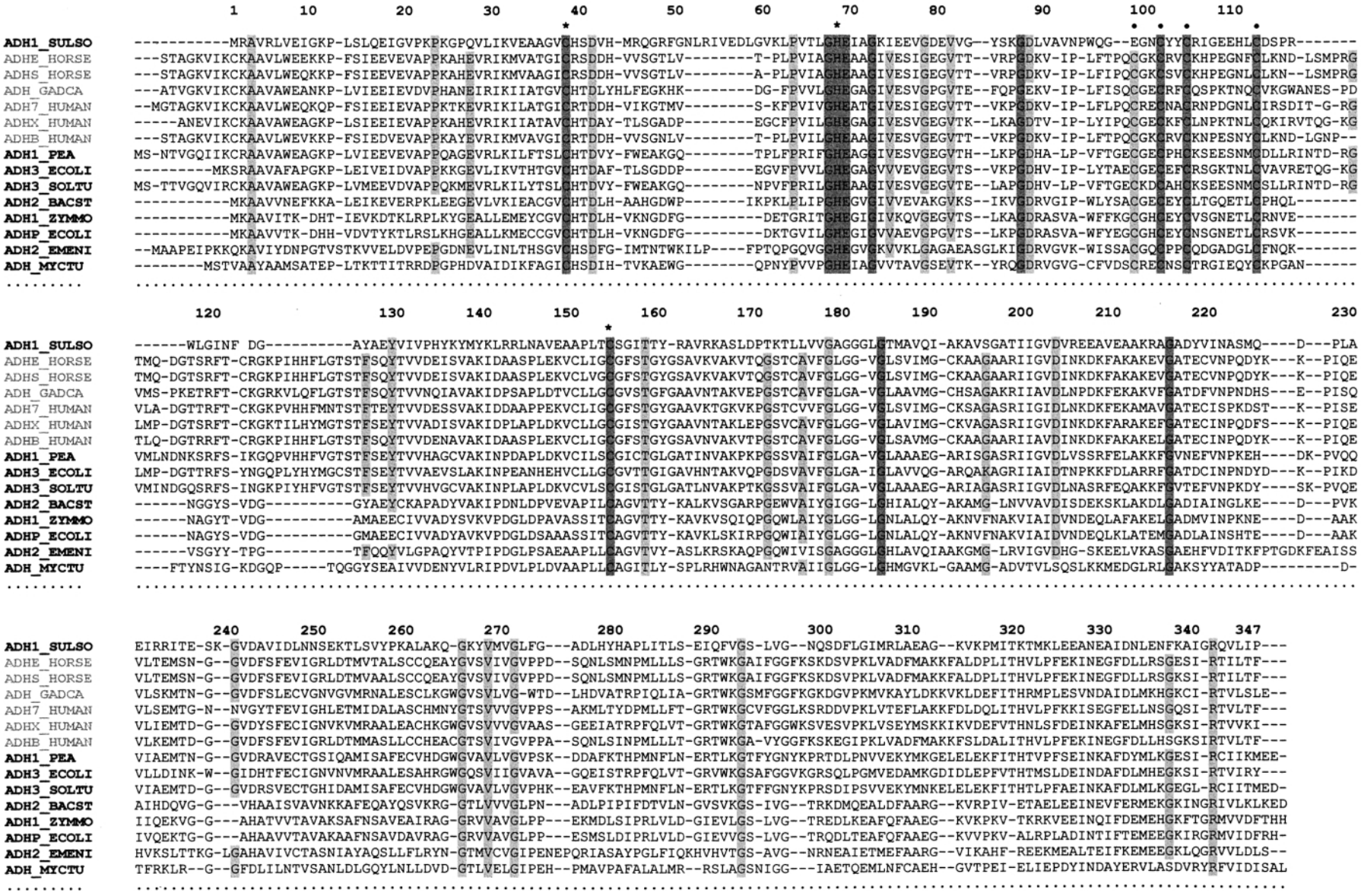

Fig. 1. Multiple sequence alignment of the 87 two-zinc-containing alcohol dehydrogenases. Only 15 out of 87 aligned sequences are shown; the names of the six sequences with known structures are highlighted in gray. Residues in dark gray columns are conserved in all the sequences and residues in light gray columns are conserved in at least 90% of the sequences. The numbers of residues refer to SsADH. Asterisks above columns indicate the residues that coordinate the catalytic zinc atom. Dots above columns indicate the residues that coordinate the structural zinc atom.

are conserved in all the sequences. Cys38, His67 and Cys154 constitute the binding site of the catalytic zinc. Glu98, Cys101, Cys104 and Cys112 are the residues involved in the binding of the structural zinc atom. Glu98 is substituted in all the other sequences with Cys (Vallee and Auld, 1990). Theoretical and experimental approaches have shown that this mutation does not affect the binding of a zinc atom and contributes to the thermostability of the enzyme, possibly by influencing the electrostatic interactions of the binding site (Ammendola *et al.*, 1995). Overall, 28 of the 38 highly conserved residues are found in the SsADH sequence; 11 of them are glycines that can be markers of the presence of a loop in the structure. This first step constrains the residues involved in binding both the structural and catalytic zinc ions.

A second alignment is in order to constrain inter-monomer interactions. If the selection among the ADHs in the SwissProt database is made considering those sequences that are found to aggregate as tetramers, 24 chains are selected, apart from the number of zinc ions present in the monomer. The multiple sequence alignment of the 24 tetrameric ADHs indicates 27 highly conserved residues (Figure 2). It should be noted that 22 sequences (including our target) are, however, two-zinc-containing ADHs and that in this second alignment the metal binding regions are also well conserved. This second alignment highlights those residues which are conserved in the tetrameric chains: Pro22, Cys38, Asp41, Gly67, Hys68, Glu69, Gly72, Gly78, Val81, Gly87, Asp88, Asp123, Gly124, Thr158, Gly178, Gly184, Ala188, Ala191, Gly216, Gly240, Gly265 and Gly293 (numbering according to SsADH). These markers can be used to constrain the alignment of our target with the one-zinc-containing monomers of the tetrameric structures.

Selection of templates and structural alignment

The results described above prompted us to adopt two sets of template structures, since we are interested in modeling both the holo and apo forms of SsADH. We considered as prototypes for the holo form the 1.8 Å resolved structure of the dimeric holo-HLADH (PDB code: 2OHX) and the structure of the tetrameric TbADH solved in the holo form at 2.5 Å (PDB code: 1YKF). The structure of the apo HLADH (PDB code: 8ADH, 2.4 Å resolution) and the structure of the apo tetrameric CbADH (PDB code: 1PED, 2.15 Å resolution) were used as templates for the apo form. The sequence of CbADH is, however, 75% homologous with that of TbADH (Korkhin *et al.*, 1999). Structural alignment between the two sets of templates indicates that the main difference between HLADH and the two bacterial enzyme structures is the conformation of the loop coordinating the structural zinc ion, present only in HLADH. Moreover, loop 114–143 in HLADH differs considerably from the corresponding loop 105–117 in TbADH and CbADH. The first one is longer and comprises two strands, whereas the others are unstructured (Figure 3).

Alignment of the target with the templates

In order to align the target structure with the templates, we first predicted the secondary structure of the target with well consolidated methods of protein secondary structure prediction [neural network-based and implemented in house (Jacoboni *et al.*, 2000); the results were also compared with a method based on a consensus procedure (Cuff and Barton, 1999)]. The target sequence is then aligned with the structural alignment of the templates taking into consideration the information derived from the multiple sequence alignments described above (Figure 3) and from the secondary structure prediction. This procedure is essentially an expertise-driven one, satisfying the constraints of the two-zinc binding sites and those of the tetrameric form. The loop coordinating the structural zinc atom in SsADH (residues 98–110) is modeled only on the structure of HLADH. The loop, which is similar in the two tetrameric bacterial structures and differs in HLADH, is aligned to exclude the corresponding portion of the HLADH chain. This is based on the finding that multiple sequence alignment and secondary structure prediction indicate that the corresponding segment on SsADH (residues 115–123) is more similar to the tetrameric structures.

The segment comprising residues 52–60 of SsADH is not aligned with any template sequence. This is an insertion in an exposed loop of the templates and its structure is automatically built and optimized by the modeling algorithm.

The procedure outlined above is similar for modeling both the apo and holo forms of SsADH.

The monomer model

The low-resolution model that we obtain after energy minimization is the monomer unit, which comprises two domains, separated by a deep cleft where the catalytic zinc atom is bound. The quality of the model was checked with Procheck (Laskowski *et al.*, 1993).

Characteristic features derived from the templates are as follows: the coenzyme binding domain (residues 156–295) is a Rossmann fold; the catalytic domain (residues 1–155 and 296–347) is an all-β GroES-like fold; and 27% of the residues are in helical structure and 20% are in a β-strand conformation. The structural zinc atom is bound in an external loop (Figure 4).

The holo and the apo forms of the SsADH models differ by a r.m.s.d. of 1 Å; this value is to be compared with an r.m.s.d. of 1 and 0.8 Å, respectively, for the holo and apo forms of HLADH and CbADH. Structural alignment of the forms of SsADH reveals that the apo is more open than the holo form, a characteristic common to all the apo and their holo counterparts of the structural database. The rigid body rotation of two domains is, however, of lesser extent with respect to that of HLADH, possibly due to the modeling procedure. The holo model has an r.m.s.d. to its templates of 0.7 and 1.1 Å to HLADH and CbADH, respectively; the apo model has an r.m.s.d. to its templates of 0.9 and 0.9 Å to HLADH and CbADH, respectively.

A salient feature of the model is that the two tryptophan residues contained in the monomer are buried (Trp95 and Trp117), in both the apo and holo forms, as detected by measuring the relative solvent accessibility. This is consistent with SsADH intrinsic fluorescence studies, showing that the fluorescence spectrum of enzyme is a relatively blue one with λ_{max} = 321 and 319 nm for the apo and holo form, respectively (Giordano *et al.*, 1999).

The tetramer model

The tetrameric organization is built automatically by BRAGI, by performing structural alignment of the apo and holo forms with the monomer of the tetramers selected as targets in our modeling procedure: CbADH for the apo form and TbADH for the holo form, respectively. The tetrameric organization consists of two interacting dimers in which interactions between β-strands in positions 289–294 and between stretches 273 and 282 stabilize monomer–monomer interaction (Figure 5).

In the assembled tetramer, the four catalytic sites are at the vertexes of a tetrahedron and ~45 Å distant. The distance

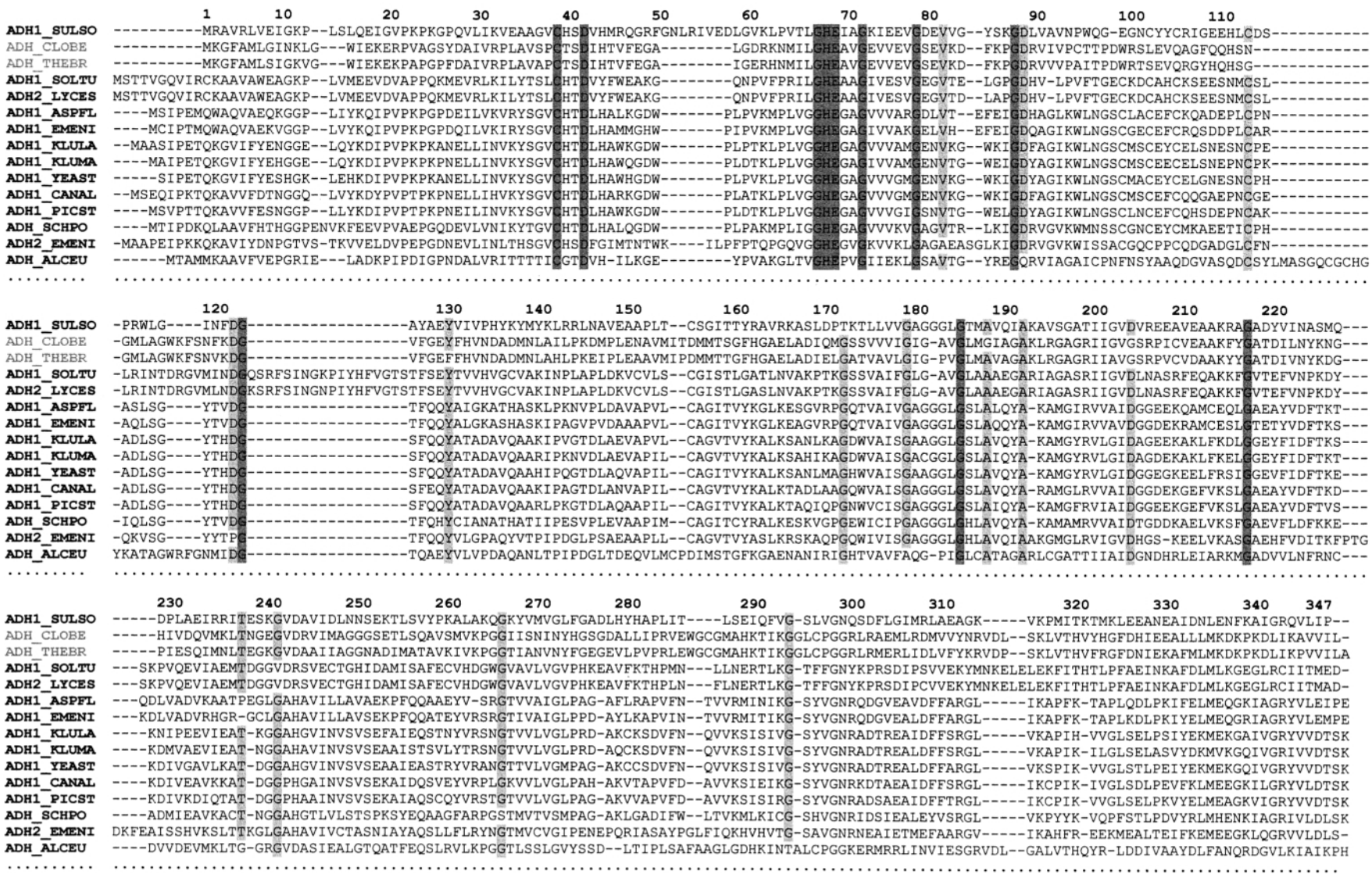

Fig. 2. Multiple sequence alignment of the 25 tetrameric alcohol dehydrogenases. Only 15 sequences are shown; the names of the two sequences with known structures are highlighted in gray. See also the legend of Figure 1.

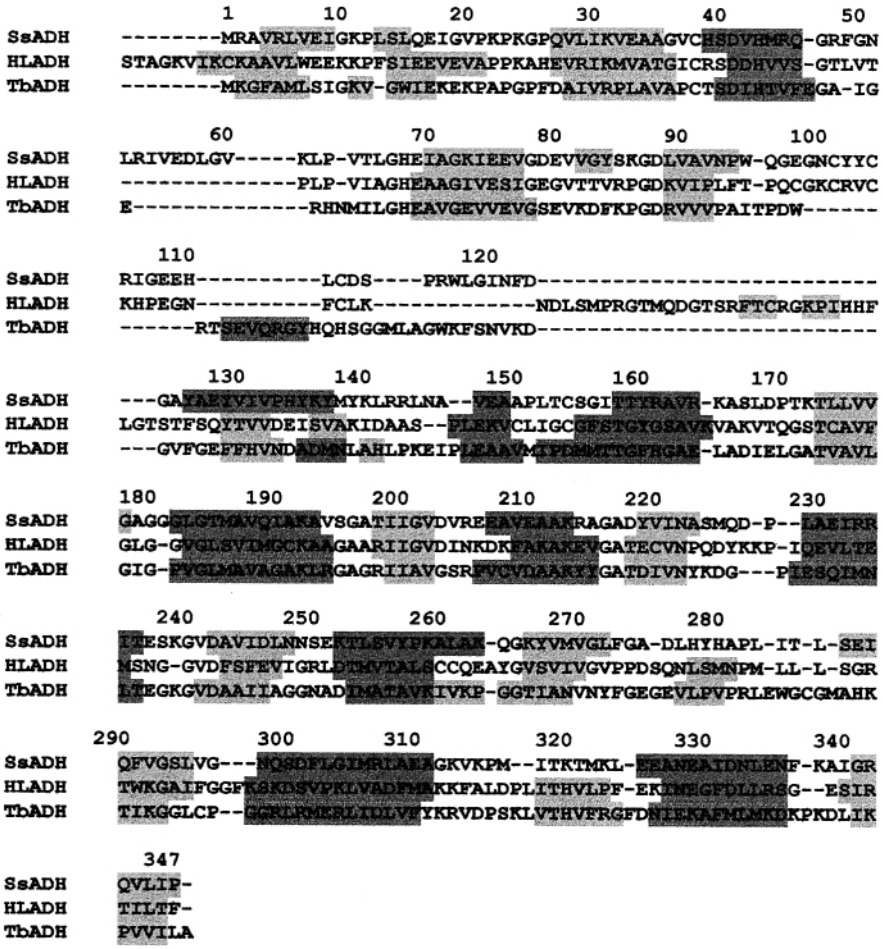

Fig. 3. Alignment of SsADH with the templates. α-Helices are highlighted in dark gray and β-strands in light gray.

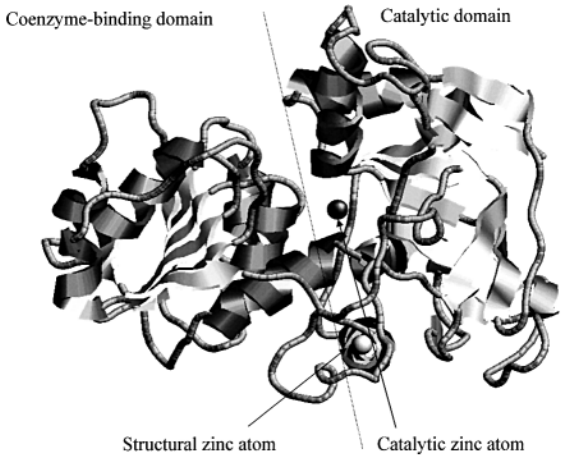

Fig. 4. Cartoon representation of the model of the SsADH monomer in holo form. The helices are depicted in dark gray and the strands in light gray. The two zinc atoms are drawn as spheres: the dark gray one represents the catalytic zinc and the light gray one represents the structural zinc. The RASMOL program was used for visualization (Sayle and Milner-White, 1995)

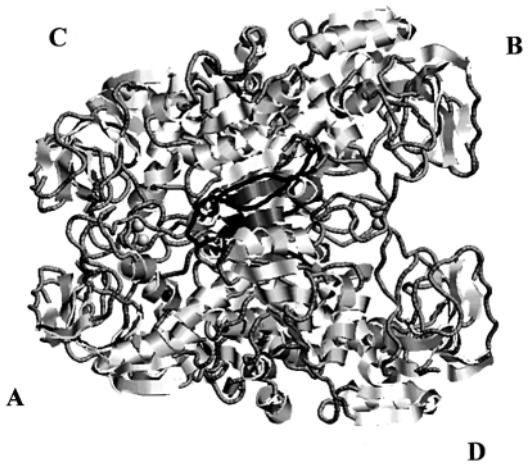

Fig. 5. Cartoon representation of the model of the SsADH tetramer in holo form. The dark gray spheres represent the four catalytic zinc atoms and the light gray spheres represent the four structural zinc atoms. The sequence stretches interacting between A and B subunits are coloured in dark gray.

between the structural zinc atoms of subunits A and C in the holo and apo forms is 11 and 12 Å, respectively.

Interestingly, the loop containing the structural zinc ion is less solvent-exposed in the holo than the apo form. The solvent accessibility values (in $Å^2$) of the side chain of the four structural zinc ligands are the following: Glu98 70 and 24, Cys101 10 and 9, Cys104 0 and 0 and Cys112 0 and 0 for the apo and holo form, respectively. Glu98 was replaced by a cysteine residue by site-directed mutagenesis (Ammendola *et al.*, 1995), thus restoring the structural zinc binding site of mesophilic ADHs, which is characterized by four cysteine residues (Vallee and Auld, 1990). The Glu98CysSsADH mutant proved equally active but less thermostable than native SsADH,

Table I. Features of the SsADH model

	Holo form	Apo form
Intrachain H-bonds (per subunit)	214	211
Interchain H-bonds	60	64
Intrachain salt bridges (per subunit)	16	15
Interchain salt bridges	28	24
Solvent-accessible surface of the monomer (Å^2)	16960	17530
Solvent-accessible surface of the tetramer (Å^2)	53550	55320
No. of lysine pairs in the dimer at a distance <13 Å	5	9
No. of pairs in the tetramer at a distance of 13–16 Å	14	16

suggesting that at least part of the SsADH thermostability is due to the presence of the glutamate in its structural metal binding site. The higher accessibility of Glu98 residue with respect to the other three ligands supports the hypothesis that the replacement of Cys by Glu represents a significant achievement in the evolutionary adaptation of alcohol dehydrogenase to thermophilic conditions (Ammendola *et al.*, 1992).

Glu108, Glu109, Lys263 and Lys266 of the monomers of each dimer are ion-bridged in both the apo and holo forms. The two dimers interact with four different surface patches: the first includes residues 301–309, 170–173 and 193–198 of subunit A and interacts with the same residues of subunit D; the second includes the same segments between the B and C subunits; the third patch is composed of residues 96–110, 135–137 and 301–299 of subunit A which interacts with the corresponding surface of subunit C; the fourth includes the same segments between the B and D subunits.

Twenty more ion-bridges stabilize the tetramer of the holo form: Glu98 and Lys136 of subunits A and C interact and similarly those of subunits B and D; the couples Asp218–Lys313, Asp218–Arg307, Asp242–Arg307 and Asp169–Arg164 can be detected between the subunits A and D and the subunits B and C.

In the apo form, 16 more interdimeric ion-bridges are detectable: Glu98–Lys136 which stabilize A with C and B with D; Asp169–Arg164, Asp169–Arg161 and Asp218–Arg307 are detected between subunits A and D and subunits B and C.

The overall electrostatic and chemico-physical properties of the model are listed in Table I. The elevated number of ion pairs suggests that SsADH retains its enzymatic activity at higher temperatures because of greater stability of subunit interaction in the tetramer. Moreover, the additional ion pairs in the holo enzyme strongly agree with the experimental observation that holo- is less sensitive than apo-SsADH to the disaggregating action of protein denaturants (Giordano *et al.*, 1999; Raia *et al.*, 2001).

Therefore, the stabilizing role of coenzyme in archaeal tetrameric ADH seems to involve both subunit interaction and the solvent accessibility of the structural zinc loop.

Binding of NADH

We estimated NADH binding to the holo form with a docking procedure. The result indicates that NADH binds to its pocket with an interaction rather similar to that detected in the holo form of HLADH (2OHX). The r.m.s.d. value to the ligand in HLADH is 1.4 Å (Figure 6). The interaction energy of the docked conformation of NADH is –22.30 kcal/mol (–23.74 kcal/mol intermolecular, 1.45 kcal/mol intramolecular) and the estimated free energy of binding is –18.44 kcal/mol. Comparison of this value with that estimated for HLADH, –16.64 kcal/mol, suggests a tighter binding of the coenzyme to the archaeal than mammalian ADH. Interestingly, the K_d values are 0.02 and 1.1 μM for the SsADH–NADH and HLAD–NADH complex, respectively (Iweibo and Weiner, 1972; Raia *et al.*, 2001). Asp203 binds the adenine ribose and so it determines the specificity for NADH with respect to NADPH (similarly to Asp223 in HLADH and yeast ADH; Fan *et al.*, 1991). Nicotinamide ribose interacts with Asn248 and Asn249 (Figure 6). It was previously suggested that Asn249Tyr substitution increases SsADH activity, possibly by decreasing ligand-binding affinity (Giordano *et al.*, 1999). This finding in the model supports the early hypothesis that Tyr249 sterically and electrostatically contrasts the interaction of the coenzyme with the binding site. Trp95 is found at a 6 Å distance from the nicotinamide moiety of the reduced coenzyme. This is consistent with the very intense energy transfer band centered at 422 nm occurring upon addition of only a stoichiometric amount of NADH (Giordano *et al.*, 1999) and the hypothesized location of the Trp95 residue at the catalytic site (Raia *et al.*, 1996).

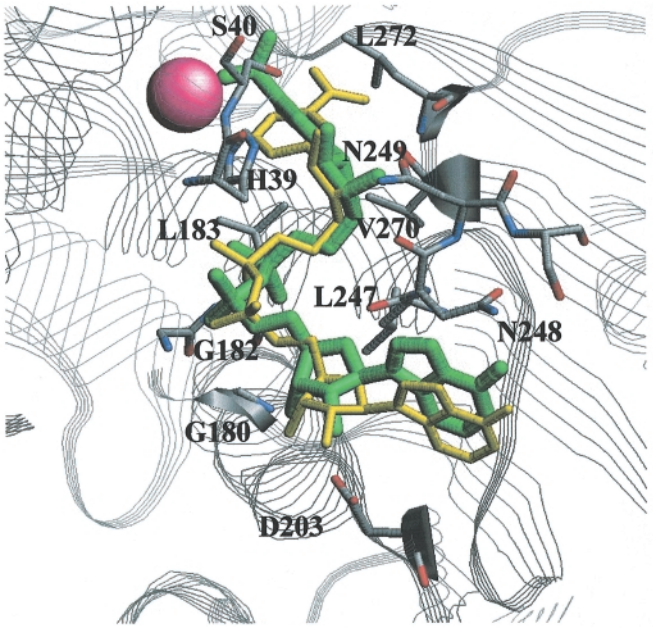

Fig. 6. Docking of NADH coenzyme into the active site of SsADH model in holo form. The conformation coloured in green represents the best energy docked structure. Residues of the model interacting within 0.30 nm with the coenzyme are represented with a stick representation. The magenta sphere represents the catalytic zinc atom. For comparison, the conformation of NADH bound to the HLADH (as found in the 2OHX pdb file) has been superimposed and coloured in yellow.

Through-space distance between lysine residues

The SsADH molecule contains 23 lysine residues per subunit, most of which change significantly their solvent exposure and reciprocal intra- and intersubunit distances upon the rearrangement of the apo to the holo form, as indicated by the different solvent accessibility values (data not shown). This feature can be tested using a cross-link methodology with the aim of evaluating the differences in through-space distance between lysine residues and indirectly to assess the consistency of the tetramer assembly.

A measure of the lysine pair distance in the model indicates that below a cut-off distance of 14 Å only AD and BC dimers can be cross-linked, provided that a cross-linker of a suitable

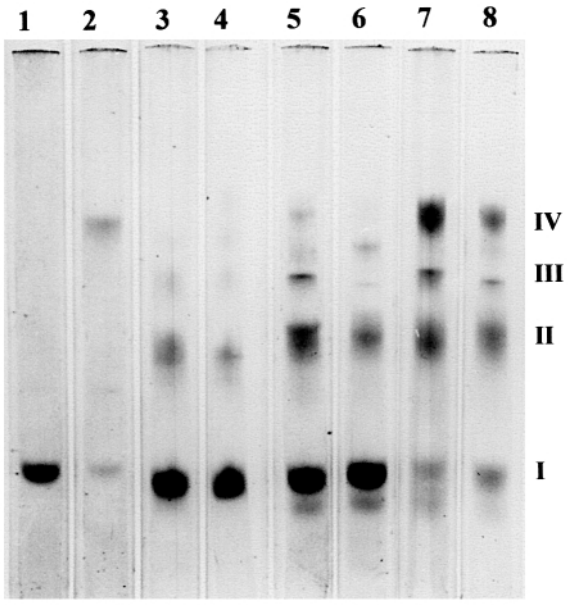

Fig. 7. SDS disk gel electrophoresis of wild-type SsADH untreated (gels 1, 2) and treated with DMA (gels 3, 4), DMS (gels 5, 6) and EGS (gels 7, 8). Apo (gels 1, 3, 5, 7) and holo (gels 2, 4, 6, 8) enzymes (7 μg each) were treated with the bifunctional reagents as described in Materials and methods and were then denatured under reducing conditions for 15 min at 90°C. Untreated samples (2 μg each) were denatured under the same conditions, except for holo SsADH which was incubated for 20 min at 50°C (gel 2) in order to obtain a partial denaturation. Apo and holo SsADH completely denature at 90°C, excluding that the band of the cross-linked tetramer contains also monofunctionally reacted and unreacted protein. I, II, III and IV correspond to the monomeric, dimeric, trimeric and tetrameric species, respectively.

length is used. At greater distances (>14 Å), cross-linking of lysine pairs in the AB and DC dimers makes the formation of tetramers more likely.

The apo and holo forms have six (K136, -140, -192, -213, -239 and -313) and three Lys residues (K192, -213 and -313), respectively, which are solvent exposed and suitable to be conjugated by bifunctional reagents spanning up to 14 Å. Only AD and BC dimers and no tetramer can be formed through these residues pairs. This suggests that apo SsADH will yield more cross-links than the holo enzyme and that the production of tetramer will be due only to intermolecular cross-linking.

According to our model, four Lys residues of the apo (K266, -172, -24 and -136) and only two Lys residues of the holo form (K136 and -24) are found at a distance of about 16 Å and solvent exposed. Their cross-linking can probably give tetramers.

Samples of SsADH were treated with bifunctional reagents spanning a distance from about 8 to 16 Å in the presence and absence of NAD and the resulting reaction mixtures were then studied by SDS electrophoresis (Figure 7). Quantification of the band intensity (Table II) shows that the extent of cross-linking achieved with the three reagents is significantly higher for the apo than the holo form of SsADH and that the cross-linked tetramer is the predominant species only with the longer reagent EGS, in agreement with the theoretical prediction gathered from our SsADH model.

Conclusion

To our knoweledge, this is the first report of a modeling procedure of a tetramer of a chain with low sequence identity with the structures of the database and the first model of a two-zinc-containing tetrameric ADH. Our threading procedure takes into consideration the following key observations: (1) all the monomers of the ADHs present in the database can be structurally aligned with an r.m.s.d. of 1.8 Å, indicating that in spite of different sources and number of bound metal ions (one or two zinc ions), the monomeric unit is structurally well conserved independently of the aggregation state (dimers or tetramers); (2) residues coordinating both the catalytic and structural zinc ions are well conserved in the multiple sequence alignment; (3) residues at the monomer interface in sequences annotated as tetramers in the sequence database are conserved in the multiple sequence alignment.

The computed model is a low-resolution model whose Cα backbone is within 2 Å from the ADH monomers of the structure database. However, docking of NADH to the binding site reveals that this portion of the protein is computed with an accuracy sufficient to dock the ligand as in HLADH. Static fluorescence data and energy transfer measurements, previously performed on the same protein, are consistent with the model. Cross-linking experiments are in agreement with predictions of dimer and tetramer formation as evaluated by measuring the lysine pair distances in the model.

Although the experimental data only support and do not fully validate the SsADH model, our computational approach is strengthened by the critical assessment that residues are evolutionarily conserved at both the catalytic and structural zinc ion binding sites and at the monomer interface.

Table II. Percentage of the cross-linked species and the monomer after modification of SsADH (apo and holo forms) with the bifunctional reagents DMA, DMS and EGS

Band	Intensity (%)[a]					
	DMA		DMS		EGS	
	apo	holo	apo	holo	apo	holo
Tetramer	8	6	11	8	39	29
Trimer	14	10	11	4	14	11
Dimer	22	18	31	24	31	36
Monomer	55	66	47	64	16	24

[a]Relative intensity within each pattern of bands of gels 3–8 in Figure 7.

Acknowledgements

R.C. and P.L.M. were supported partially by a grant from the Ministero della Università e della Ricerca Scientifica e Tecnologica (MURST) for the project 'Structural, Functional and Applicative Prospects of Proteins from Thermophiles' and by a grant for a target project in Biotechnology from the Italian Centro Nazionale delle Ricerche (CNR), both to R.C. We are grateful to Domenico Carlomagno for his invaluable contribution in performing the cross-linking experiments.

References

Ammendola,S., Raia,C.A., Caruso,C., Camardella,L., D'Auria,S., De Rosa,M. and Rossi M. (1992) *Biochemistry*, **31**, 12514–12523.

Ammendola,S., Raucci,G., Incani,O., Mele,A., Tramontano,A. and Wallace A. (1995) *Protein Eng.*, **8**, 31–37.

Cannio,R., Fiorentino,G., Carpinelli,P., Rossi,M. and Bartolucci,S. (1996) *J. Bacteriol.*, **178**, 301–305.

Colonna-Cesari,F., Perahia,D., Karplus,M., Eklund,H., Brändén,C.-I. and Tapia,O. (1986) *J. Biol. Chem.*, **261**, 15273–15280.

Cuff,J.A. and Barton,G.J. (1999) *Proteins*, **34**, 508–519.

Davies,G.E. and Stark,G.R. (1970) *Proc. Natl Acad. Sci. USA*, **66**, 651–656.

Esposito,L. *et al.* (1998) *Acta Crystallogr.*, **D54**, 386–390.

Fan,F., Lorenzen,J.A. and Plapp,B.V. (1991) *Biochemistry*, **30**, 6397–6401.

Giordano,A., Cannio,R., La Cara,F., Bartolucci,S., Rossi,M. and Raia,C.A. (1999) *Biochemistry*, **38**, 3043–3054.
Iweibo,I. and Weiner,H. (1972) *Biochemistry*, **11**, 1003–1018.
Jacoboni,I., Martelli,P.L., Fariselli,P., Compiani,M. and Casadio,R. (2000) *Proteins*, **41**, 535–544.
Kabsch,W. and Sander,C. (1983) *Biopolymers*, **22**, 2577–2637.
Korkhin,Y., Kalb (Gilboa),A.J., Peretz,M., Bogin,O., Burstein,Y. and Frolow,F. (1999) *Protein Sci.*, **8**, 1241–1249.
Laskowski,R.A., MacArthur M.W., Moss,D.S. and Thornton,J.M. (1993) *J. Appl. Crystallogr.*, **26**, 283–291.
Morris,G.M., Goodsell,D.S., Halliday,R.S., Huey,R., Hart,W.E., Belew,R.K. and Olson,A.J. (1998) *J. Comput. Chem.*, **19**, 1639–1662.
Pearl. L.H., Damasi,D., Hemmings,A.M., Sica,F., Mazzarella,L., Raia,C.A., D'Auria,S. and Rossi,M. (1993) *J. Mol. Biol.*, **229**, 782–784.
Raia,C.A., Caruso,C., Marino,M., Vespa,N. and Rossi,M. (1996) *Biochemistry*, **35**, 638–647.
Raia,C.A., Giordano,A. and Rossi,M. (2001) *Methods Enzymol.*, **331**, 176–195.
Šali,A. and Blundell,T.L. (1993) *J. Mol. Biol.*, **234**, 779–815.
Sayle,R.A. and Milner-White,E.J. (1995) *Trends Biochem. Sci*., **20**, 374–376.
Sun,H.-W. and Plapp,B.V. (1992) *J. Mol. Evol.*, **34**, 522–535.
Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) *Nucleic Acids Res.*, **22**, 4673–4680.
Vallee,B.L. and Auld,D.S. (1990) *Biochemistry*, **29**, 5647–5659.
Vriend,G. (1990) *J. Mol. Graphics*, **8**, 52–56.

Received April 17, 2001; revised November 21, 2001;
accepted December 7, 2001