Structure of a Hyperthermophilic Tungstopterin Enzyme, Aldehyde Ferredoxin Oxidoreductase

Michael K. Chan, Swarnalatha Mukund, Arnulf Kletzin, Michael W. W. Adams, Douglas C. Rees*

The crystal structure of the tungsten-containing aldehyde ferredoxin oxidoreductase (AOR) from *Pyrococcus furiosus*, a hyperthermophilic archaeon (formerly archaebacterium) that grows optimally at 100°C, has been determined at 2.3 angstrom resolution by means of multiple isomorphous replacement and multiple crystal form averaging. AOR consists of two identical subunits, each containing an Fe_4S_4 cluster and a molybdopterin-based tungsten cofactor that is analogous to the molybdenum cofactor found in a large class of oxotransferases. Whereas the general features of the tungsten coordination in this cofactor were consistent with a previously proposed structure, each AOR subunit unexpectedly contained two molybdopterin molecules that coordinate a tungsten by a total of four sulfur ligands, and the pterin system was modified by an intramolecular cyclization that generated a three-ringed structure. In comparison to other proteins, the hyperthermophilic enzyme AOR has a relatively small solvent-exposed surface area, and a relatively large number of both ion pairs and buried atoms. These properties may contribute to the extreme thermostability of this enzyme.

The recent discovery of life at temperatures of 100°C or higher is having increasing ramifications in biology and biotechnology (*1–4*). One of the fundamental questions arising from the study of these hyperthermophilic organisms concerns the structural and energetic basis of protein stability at extremely high temperatures, under conditions where many proteins from mesophilic organisms would be denatured. As part of a coordinated effort to establish the structural and functional properties of these intriguing macromolecules, we have initiated the structure determination of several proteins from hyperthermophilic organisms (*5*). We now describe the structure of one of the more unusual of these proteins, the aldehyde ferredoxin oxidoreductase (AOR) from the hyperthermophilic archaeon *Pyrococcus furiosus* (T_{opt} = 100°C); this protein catalyzes the reversible oxidation of aldehydes to their corresponding carboxylic acids with the accompanying reduction of the redox protein ferredoxin.

AOR was discovered because of the unusual growth requirement of *P. furiosus* for the third-row transition element tungsten (*6, 7*). Many enzymes utilize first-row transition elements at their active sites, but only two second- or third-row transition metals, tungsten and molybdenum, have such a biological role (*8*). Hence, one of the critical questions in studying proteins containing tungsten or molybdenum focuses on understanding the intrinsic properties that contribute to the incorporation of these metals compared to more traditional first row transition elements. The versatile redox chemistry of molybdenum and tungsten enables enzymes containing these elements to catalyze fundamental steps in the biological utilization of carbon, nitrogen, and sulfur. All tungsten and molybdenum enzymes, with the notable exception of the FeMo-cofactor of nitrogenase, belong to a family known as oxotransferases, which catalyze the transfer of an oxo group to or from substrate in a two-electron redox reaction (*9–11*). Although most members of this family are molybdoenzymes, including xanthine oxidases that convert xanthine to uric acid, sulfite oxidases involved in the metabolism of sulfur, and nitrate reductases that form nitrite from nitrate, several tungsten-containing enzymes have been isolated (*12*). To date, these tungstoenzymes have been found mainly in thermophilic anaerobes, and function either in CO_2 activation or, as in the case of *P. furiosus* AOR, in the conversion of aldehydes to carboxylic acids.

The catalytic site in these oxotransferase enzymes is a labile, oxygen-sensitive, metal cofactor consisting of a tungsten or molybdenum ion bound to an organic component. This ligand, termed molybdopterin (MPT), was first identified by Rajagopalan and coworkers (*13*), who isolated and characterized its di(carboxamidomethyl) derivative by fluorescence and mass spectroscopy. While additional studies have demonstrated that all eukaryotic molybdoenzymes contain this unmodified form of MPT, prokaryotic forms usually contain an alternative version with either a GMP (*14–17*), CMP (*18, 19*), AMP (*17*), or IMP (*17*) nucleotide appended to the molybdopterin via a pyrophosphate linkage. The role of the additional nucleotide in the chemistry of this system is unknown. As with molybdenum enzymes, all tungsten oxotransferases contain either the MPT or modified MPT ligand (*20*). The model for the molybdenum cofactor proposed by Rajagopalan *et al.* (*11, 21*) based on biochemical and spectroscopic analyses, consists of a single MPT ligand bound to molybdenum via two dithiolene sulfurs, as shown in Fig. 1A. Direct confirmation of this model, however, has been hindered by the lack of a three-dimensional structure for any molybdenum or tungsten oxotransferase. To establish the structural features of tungsten and MPT cofactors, and to gain insight into the structural origins of thermal stability in hyperthermophilic proteins, we determined the structure of *P. furiosus* AOR with x-ray diffraction methods.

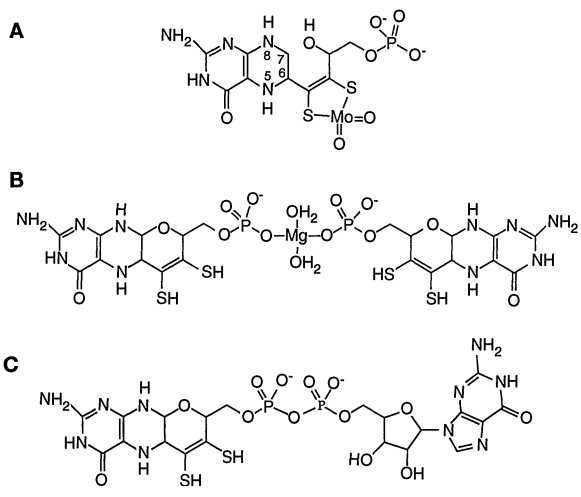

Fig. 1. Comparison of molybdopterin structures. (**A**) Rajagopalan's model of the molybdenum cofactor consisting of a single molybdopterin (MPT) ligand bound to molybdenum (*11, 21*); (**B**) the structure of the bis(MPT) ligand determined from x-ray analysis of the *P. furiosus* AOR protein; (**C**) proposed structure of molybdopterin guanine dinucleotide (MGD) based on the revised MPT structure. A single molecule of MGD and related ligands could possibly coordinate metals in a similar fashion to the bis(MPT) ligand by utilizing the sugar hydroxyl oxygens in place of the second dithiolene group.

M. K. Chan and D. C. Rees are in the Division of Chemistry and Chemical Engineering 147-75CH, California Institute of Technology, Pasadena, CA 91125, USA. S. Mukund, A. Kletzin, and M. W. W. Adams are in the Department of Biochemistry and Molecular Biology, and Center for Metalloenzyme Studies, University of Georgia, Athens, GA 30602, USA.

*To whom correspondence should be addressed.

The growth of *P. furiosus* and purification of AOR under anaerobic reducing conditions have been described (*7*). Details of the crystallization and structure determination are provided in Table 1. The AOR model contains 9888 total atoms (123 atoms in five cofactor sites, 9386 protein atoms and 379 water molecules) and has been refined to an *R* factor of 0.155 for 55520 observed reflections in the resolution range between 10 and 2.3 Å (no σ cutoff, 89 percent complete), with root-mean-square deviations from ideal bond distances and angles of 0.009 Å and 2.58°, respectively. The AOR molecule (Fig. 2, A and B), consists of two identical 605 residue (66 kD) subunits of known sequence (*22*). Three different types of metal sites are found in the AOR protein dimer; an Fe_4S_4 cluster and a tungsten cofactor that are contained within each subunit, and a single tetrahedral metal atom located at the dimer interface. The Fe_4S_4 cluster and tungsten cofactor within one subunit are in close proximity, but the mononuclear center is positioned ~25 Å from these groups. The Fe_4S_4 clusters and tungsten cofactors in different subunits of the dimer are separated by ~50 Å.

Structure of the metal sites. The mononuclear site is symmetrically coordinated by Glu^{332} and His^{383} from each subunit in a tetrahedral fashion. The metal site is tentatively identified as iron on the basis of anomalous scattering of the atom and a metal analysis of the AOR protein although direct confirmation of this assignment by Mössbauer spectroscopy is needed. Because of the ~25 Å distance between this site and the other metal centers, this mononuclear metal is not likely to play a redox role; rather it may be involved in stabilizing the protein dimer at high temperatures.

The tungsten cofactor fits into a pocket located at the center of each subunit (Fig. 2). The structure of this site (Fig. 3) reveals that the molybdopterin does bind to the metal via the dithiolene sulfurs, as proposed by Rajagopalan *et al.* (*21*). Surprisingly, however, two molybdopterin molecules, instead of the expected one molecule, are coordinated to the tungsten. The arrangement of the tungsten and the two pairs of dithiolene sulfurs may be described as a distorted square pyramid, with an angle between the planes of the molybdopterin ligands of ~97°. No protein ligands are coordinated to the tungsten atom, although our analysis of the electron density around the tungsten suggests that two additional coordination sites may be occupied by glycerol or oxo ligands (or both) to yield an overall distorted trigonal prismatic arrangement at the tungsten site. The glycerol presumably comes from the protein storage buffer and may represent a substrate analog. Unless substantial rearrangements in the tungsten coordination sphere occur during substrate oxidation, the presence of two MPT ligands in AOR enforces cis coordination for the substrate and oxo groups,

Table 1. Summary of AOR structure determination. Native and derivative data sets for each of the three crystal forms (*42*) were collected from single crystals at room temperature on either a RAXIS imaging plate system or a Siemens area detector mounted on rotating anode generators producing CuKα (λ = 1.54 Å) radiation. Diffraction data sets were processed with either RAXIS (provided by Rigaku), XENGEN (*43*), or XDS (*44*) software, as appropriate, and then locally scaled to the native data set for that crystal form with the use of the program package ROCKS (*45*). An additional native data set for crystal form 1 was collected at λ = 1.08 Å at the Stanford Synchrotron Radiation Laboratory, with a MAR imaging plate detector, and processed with MOSFLM (*46*). Data collection statistics from native and derivative crystals are summarized below. The positions of five anomalous scattering sites in crystal form 1 were determined from inspection of the anomalous difference Patterson maps, calculated from a native data set collected on the RAXIS system. Because of changes in the scattering factors for Fe and W between 1.54 Å and 1.08 Å, the Fe_4S_4 clusters and W sites were initially differentiated on the basis of anomalous Fourier maps calculated from data sets collected at these two wavelengths. In addition to anomalous scattering phasing information for crystal form 1, different Pt derivatives could be obtained by varying the time (12 hours to 2 weeks) of soaking crystals of each form in 10 mM K_2PtCl_4. The locations of the heavy atom sites were determined by difference Fourier analysis and refined in PHARE (*47*) to generate initial experimental phases for each crystal form. The figure of merits for crystal forms 1, 2, and 3 to 3.0 Å resolution were 0.60, 0.50, and 0.41, respectively. The multiple isomorphous replacement phases were refined by noncrystallographic symmetry averaging within and between the different crystal forms (*48*). The initial polyalanine model was built with TOM/FRODO (*50*), and the final model was constructed with the program O (*51*) as the sequence became available. The TNT program (*52*) was used for the initial stages of the model refinement, and positional and B factor refinements at the latter stages were calculated with X-PLOR (*53*). Model phases calculated at each stage were combined with the experimental phases to generate improved maps. All heavy atom derivative Pt sites were located near histidine or methionine side chains. The model quality was also assessed with the programs PROCHECK (*54*), VERIFY-3D (*55*), and ERRAT (*56*), which were consistent with a correct structure. More than 90 percent of the (ϕ, ψ) angles of the AOR model are in the most favored regions of the Ramachandran plot (*54*), and the secondary structure regions are listed in (*57*).

Crystal	Space group	Crystal parameters						
		a (Å)	*b* (Å)	*c* (Å)	α (°)	β (°)	γ (°)	*Z*
Native 1	$P2_12_12_1$	80.9	108.3	159.8	90.0	90.0	90.0	8
Native 2	P1	85.5	54.1	85.4	105.3	93.5	93.4	2
Native 3	P1	55.7	67.0	85.2	95.4	86.6	92.9	2

Data collection and processing statistics

Data set	Resolution (Å)	Reflections		Complete (%)	R_{sym}*	Isomorphous *R* factor†	Sites (No.)	Phasing power‡
		Measured	Unique					
Native 1	2.32	237,105	56,373	92.4	0.049		5§	1.33
SSRL	3.0	86,751	22,541	79.0	0.086			
$PtCl_4$ 1	2.37	114,240	35,556	62.1	0.095	0.096	2	0.68
$PtCl_4$ 2	2.67	242,839	37,676	92.5	0.120	0.149	8	1.37
Native 2	2.95	56,398	26,114	74.8	0.090		4§	0.67
$PtCl_4$ 3	2.05	227,138	81,661	84.5	0.124	0.154	8	1.43
Native 3	2.30	106,070	48,579	82.7	0.074		4§	0.97
$PtCl_4$ 4	2.31	103,475	48,981	82.8	0.078	0.206	10	1.4

*R_{sym} describes the agreement between multiple observations of symmetry related reflections, based on the absolute value of the intensity differences. †Mean fractional change between native and derivative structure factor amplitudes. ‡Ratio of the heavy atom scattering amplitude to estimated phasing model error, calculated to 3 Å resolution. §Native anomalous scattering sites.

which would have stereochemical consequences for the mechanism of the oxotransfer reaction. The side chains of residues Glu^{313} and His^{448} are in the vicinity of the substrate binding site on the tungsten and could participate in proton transfers associated with oxidation-reduction reactions.

In addition to interactions between the dithiolene sulfurs and tungsten, the two MPT ligands are also linked through their phosphate groups, which coordinate axial sites of the same magnesium ion. The two MPT ligands are approximately related by a twofold rotation about an axis that passes through both the tungsten and magnesium sites. The magnesium ion exhibits octahedral geometry. In addition to the two coordination sites filled by phosphate oxygens, two waters bind at cis sites on the side facing the tungsten atom, and the two backbone carbonyl oxygens from residues Asn^{93} and Ala^{183} bind at cis sites on the Mg directed away from the tungsten. The two waters form part of an intricate, asymmetric hydrogen bonding network within the tungsten cofactor (Fig. 3B). Each water is hydrogen-bonded to at least one phosphate oxygen and an N-5 nitrogen of the pterin ring, with one of the two waters also within hydrogen bond distance of the ether oxygen of one of the MPT ligands. The origin of this slight asymmetry is due to the hydrogen bonding of the side chain of Arg^{76} to the ether oxygen of the other MPT group.

The refined AOR structure also indicates that the MPT ligand itself is slightly different from that previously postulated (Figs. 1 and 3). The crystallographic analysis, as illustrated by the electron density map shown in Fig. 3A, reveals that the MPT groups in AOR contain a third ring formed by closure of the side-chain hydroxyl with the pterin ring at C-7. Such a cyclization would not be unprecedented in that related compounds have been synthesized (*23*). A plausible mechanism for formation of the third ring would be by attack of the side-chain hydroxyl on the C-7 carbon of a 5,6 dihydropterin (Fig. 1). Provided that this ring closure is reversible, there is no inconsistency between the three-ringed structure seen in the protein, and the two-ringed form observed for the isolated pterin (*11*, *21*). The nonplanar structure of the pterin ring system seen in AOR is consistent with a reduced form for the protein-bound pterin, which would likely be equivalent to a dihydropterin (possibly the 5,6-dihydropterin) in the ring-opened form. This revised MPT structure may be a common feature in all molybdopterin-containing enzymes.

The revised MPT structure may also be relevant to the structure of the various dinucleotide derivatives found in other molybdenum-cofactor systems (Fig. 1). Comparison of the dimolybdopterin to the modified molybdopterin dinucleotides reveals an overall similarity. In each case, two aromatic base systems attached to potential bidentate liganding groups are connected via a phosphate backbone. It is possible that molybdenum or tungsten cofactors containing these modified MPT dinucleotides do have a structure similar to that observed in AOR, with the only difference being replacement of the second pterin in AOR by a nucleotide, and consequent metal ligation by the sugar hydroxyl oxygens. There is precedent for metal coordination by sugar hydroxyl groups in this fashion (*24*). Such binding may help to explain why the sugar ring is the only common feature among the modified molybdopterin dinucleotides, and why differences appear to exist in the ligation of the Mo sites in different molybdoenzymes. An alternative possibility is that the nucleotide-containing end of the modified MPT cofactor could adopt a somewhat different conformation, that would permit metal coordination by groups on the nucleotide ring or, perhaps by protein side chains.

Electron paramagnetic resonance, magnetic circular dichroism, and resonance Raman studies indicate that AOR contains an Fe_4S_4 cluster located near a paramagnetic center (*7*). Indeed, the x-ray structure reveals that the Fe_4S_4 cluster is positioned approximately 10 Å from the tungsten atom. The Fe_4S_4 cluster is buried ~6 Å below the van der Waals surface of the protein. This arrangement is consistent with the postulated role of the Fe_4S_4 cluster as an intermediary for electron transfer between the tungsten cofactor and ferredoxin, the physiological electron acceptor of AOR. Four cysteine ligands, provided by the Sγ of Cys^{288}, Cys^{291}, Cys^{295}, and Cys^{494}, coordinate the Fe_4S_4 cluster. Whereas the first three Cys residues are part of a characteristic iron-sulfur cluster binding sequence (*22*), the polypeptide conformation of this region is distinct from that adopted by the Fe_4S_4 binding site found in ferredoxins. The Fe_4S_4 cluster is linked to one of the two molybdopterins of the tungsten cofactor by two distinct sets of interactions. The side

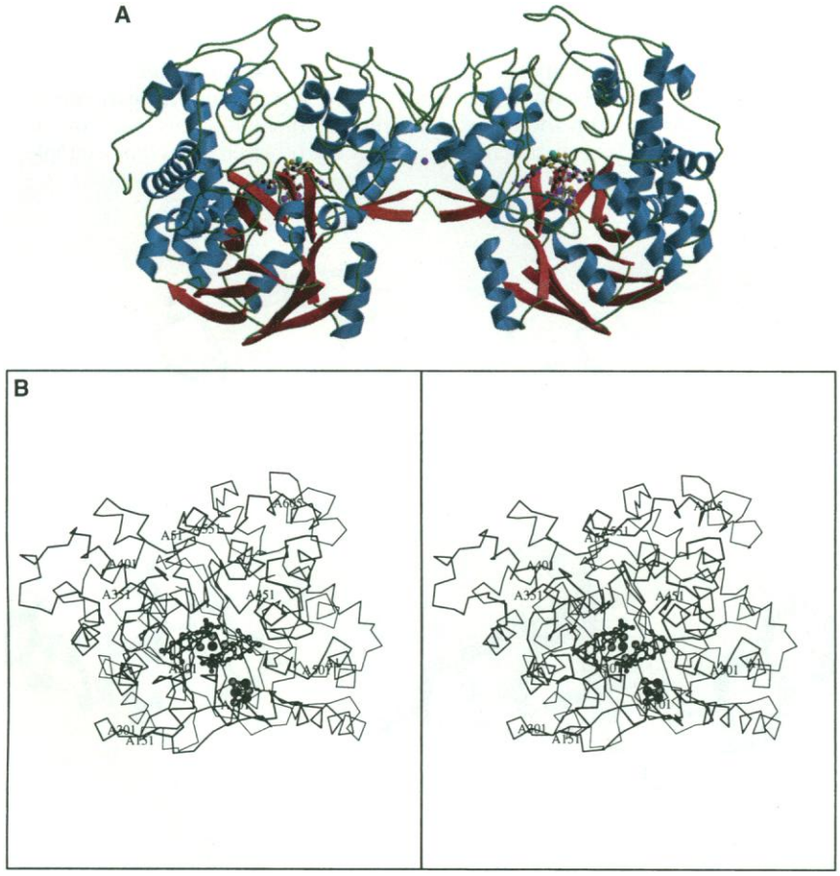

Fig. 2. Structure of the AOR protein. (**A**) Ribbons diagram of the AOR dimer viewed perpendicular to the dimer twofold axis. The secondary structures are color-coded with α helical regions as light blue, β sheet regions as red, and the remainder as green. Atoms in the cofactors are colored by elements, with carbon as black, nitrogen as blue, oxygen as red, sulfur as yellow, phosphate as magenta, magnesium as orange, iron as purple, and tungsten as cyan. (**B**) Stereodiagram of the polypeptide chain of a single AOR subunit. This figure was prepared with the programs MOLSCRIPT and Raster 3D (*58*).

chain of Arg^{76} bridges these two groups by forming hydrogen bonds to an inorganic sulfur of the Fe_4S_4 cluster, and to two sites on the molybdopterin, the ether oxygen and a phosphate oxygen. In addition, the Sγ of Cys^{494}, an Fe_4S_4 cluster ligand, is positioned to accept a hydrogen bond from the pterin ring nitrogen N-8, which is nearest the ether linkage of the molybdopterin. These interactions could provide electron transfer pathways between the two centers. Furthermore, this arrangement suggests that the pterin ligand does not merely play a passive structural role, but may be an active participant in the redox chemistry of AOR.

Protein structure. The binding sites for the tungsten cofactor and Fe_4S_4 cluster are located at the interfaces of the three domains that are formed by each AOR subunit. The location of redox centers at domain or subunit interfaces appears to be a general property of multicentered redox proteins (*25*). Domain 1 (residues 1 to 210) forms a base on which the saddle-like tungsten pterin cofactor sits, while domains 2 (residues 210 to 417) and 3 (residues 418 to 605) surround the opposite surface of the tungsten cofactor, and provide residues that form specific polar and ionic interactions with the different metal centers. Each subunit exhibits a pseudo twofold axis that coincides approximately with the twofold axis of the tungsten cofactor. This rotation axes passes through the center of domain 1, and approximately relates domains 2 and 3. The twofold symmetric arrangement of the polypeptide chain in domain 1 is particularly striking (Fig. 4A). The similarity of domains 2 and 3 is less apparent, which may result from the involvement of domain 2 in both generating the dimer interface, and in providing most of the Fe_4S_4 cluster binding site.

Domain 1 consists primarily of 12 β strands arranged in two six-stranded, predominantly antiparallel β sheets that are related approximately by the twofold rotation axis described above (Fig. 4A). The two sheets are joined by antiparallel β sheet interactions between the β strands containing residues 89 to 95 and 182 to 186. Residues Asn^{93} and Ala^{183}, which coordinate the magnesium ion through their carbonyl oxygens, are located within these two β strands. Within domain 1 is residue Arg^{182}, which forms a salt bridge to the phosphate group of the molybdopterin nearer the Fe_4S_4 cluster, whereas additional hydrogen bonding interactions to the phosphates are formed by the hydrogens on the main chain amide nitrogens of residues 93, 95, 183, 185, and 186. Residue Arg^{76}, which hydrogen bonds both a molybdopterin ether oxygen and an Fe_4S_4 inorganic sulfur, is also located within this domain.

The regular secondary structure of domains 2 and 3 consists primarily of α helices, with few β sheet regions. Although these two domains are related approximately by the twofold axis described above, this structural relationship is less apparent than for domain 1 and the tungsten cofactor. Consistent with an underlying structural similarity between the two domains, residues 338 to 344 and 489 to 495 in domains 2 and 3, respectively, contain Asp-X-X-Gly-Leu-(Cys or Asp)-X sequences, where the Asp carboxylate group and Leu main chain carbonyl oxygen are arranged to bind the primary amine group of their respective MPT ligand (Fig. 4, B and C). Asp^{343} and Cys^{494} interact with the secondary ring nitrogen, N-8, near the ether linkage, while the amide nitrogens from the carboxy terminal end of each motif, Thr^{344} and Leu^{495}, form a hydrogen bond with a secondary ring nitrogen on the pterin. This sequence motif occurs in both domains at the end of α helices (Fig. 4C), with the Gly residue adopting the left-handed helical (Lα) conformation characteristic of a helical capping interaction (*26*). An additional polar interaction to one of the two MPT rings is generated by a hydrogen bond between the side chain of Lys^{450} and a pterin carbonyl oxygen.

In addition to maintaining specific binding interactions with the pterin, another role for domains 2 and 3 may be to regulate substrate access to AOR. With the use of a 1.4 Å radius probe, a long hydrophobic channel leading from the tungsten site to the surface of the protein may be defined (*27*) between these two domains. The nonpterin coordination sites on the tungsten atom occupy one end of this cavity, consistent with the role of these coordination sites in substrate binding and catalysis. A 3_{10} helix formed by residues 495 to 500 runs parallel to the length of the cavity. Within this helix, the side chains of Phe^{496} and Phe^{499} are directed in a parallel fashion toward the central axis of the cavity. The channel seems large enough to accommodate a variety of substrates, which may help to explain the ability of AOR to oxidize both aliphatic and aromatic aldehydes. In contrast to AOR, the tungsten-containing enzyme, formaldehyde ferredoxin oxidoreductase (FOR), which is also found in hyperthermophiles such as *P. furiosus* and *Thermococcus litoralis* (*22*, *28*), is specific for small (one to three carbon) aldehydes. *Pyrococcus furiosus* AOR and *T. litoralis* FOR appear to be structurally related as their sequences are very similar [59 percent with 38 percent identity (*22*)]. Hence, different channel sizes in these two enzymes might account for the differences in their substrate specificities.

Although all molybdenum-tungsten oxotransferase enzymes that have been characterized contain similar pterin cofactors, it is not yet known whether any of the protein structural motifs found in AOR exist in the other enzymes. For example, with

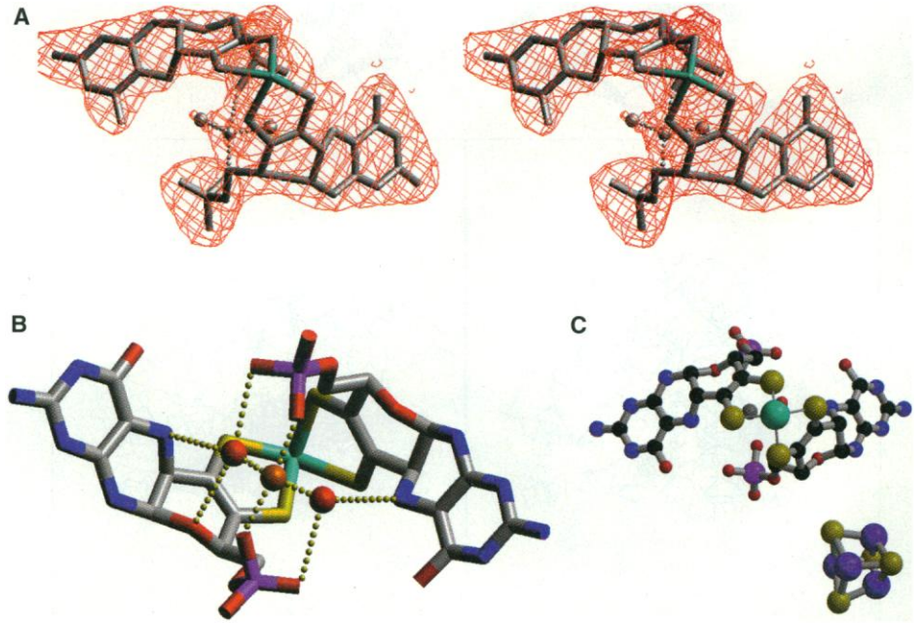

Fig. 3. Structure of the tungsten cofactor. (**A**) $F_o - F_c$ omit-refined electron density map of the tungsten-dimolybdopterin cofactor as viewed along the pseudo twofold axis. The contour level is three times the standard deviation of the electron density map. (**B**) Stick diagram of the internal hydrogen bonding network within the tungsten cofactor itself. The atom coloring scheme is the same as in Fig. 2, except that carbon atoms are gray. (**C**) Ball-and-stick figure of the tungsten cofactor and Fe_4S_4 cluster of AOR. (A) and (B) were prepared with the program SETOR (*59*) and (C) was prepared with the programs MOLSCRIPT and Raster 3D (*58*).

the exception of *T. litoralis* FOR, only NH_2-terminal sequences are available for the three other types of tungstoenzymes known: carboxylic acid reductases (CAR) from acetogens, formylmethanofuran dehydrogenases from methanogens, and glyceraldehyde-3-phosphate ferredoxin oxidoreductase from *P. furiosus* (*12*, *22*, *28*). Of these, only CAR shows similarity to the NH_2-terminal sequences of AOR and FOR (*22*). In contrast, the complete sequences are available for several molybdenum-containing oxotransferases, including various xanthine oxidases and nitrate reductases. However, these show no similarity to the AOR or FOR sequences (*22*). Thus, it is important to establish the structures of other oxotransferases to see whether these have truly different structures or whether the underlying structural similarity to AOR is obscured by sequence divergence.

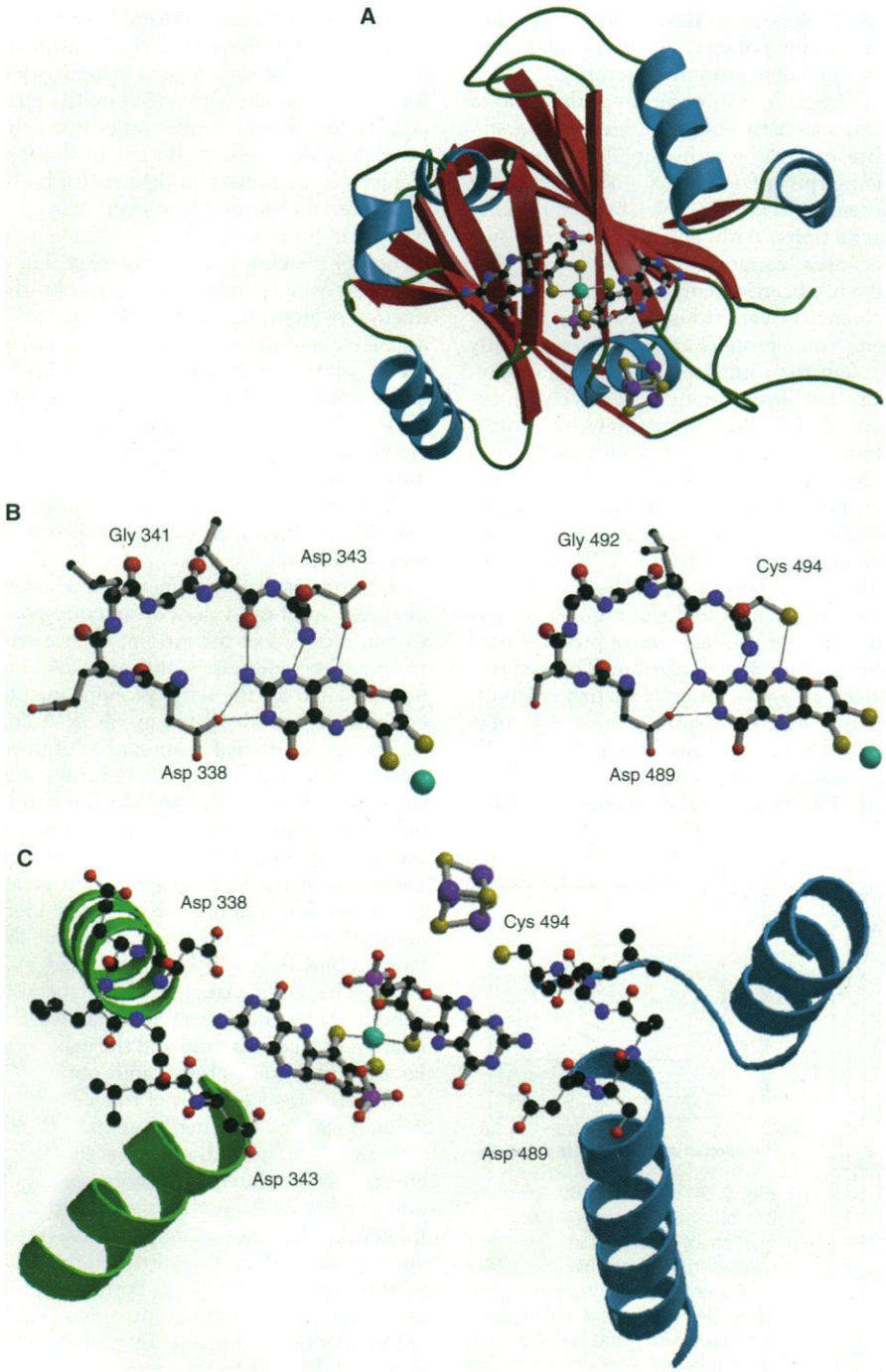

Fig. 4. Protein-molybdopterin interactions. (**A**) Ribbons diagram of the interaction of the tungsten cofactor and domain 1, illustrating the β sheet arrangement and the overall twofold symmetry of both the domain and the pterin cofactor. (**B**) Hydrogen bonding of the tungsten cofactor to domains 2 and 3. The DXXGL(D,C)X motifs in both domains 2 and 3, which are approximately related by the twofold axis passing through the cofactor, form multiple hydrogen bonds to each pterin ligand. (**C**) Ribbons diagram illustrating the interaction of the pterin ligands of the tungsten cofactor with α-helical regions of domains 2 and 3 that contain the sequence motif detailed in (B). This figure was prepared with the programs MOLSCRIPT and Raster 3D (*58*).

Protein thermostability. One of the striking properties of AOR is the extreme thermal stability of the protein structure. Although calorimetric studies of AOR stability have not yet been performed, the temperature for optimum aldehyde oxidation activity is more than 95°C, which sets a lower limit on the temperature range for AOR stability. Since AOR is the first hyperthermophilic enzyme to be structurally characterized, and in the absence of a mesophilic counterpart for direct comparison, an assessment of the origins of the thermal stability of AOR must necessarily be based on analysis of rather general features of protein structure and stability established primarily from mesophilic structures. The primary, secondary and tertiary structures of AOR are unexceptional for water-soluble, globular proteins. The amino acid composition of AOR (*22*) is close to the average observed for prokaryotic proteins, with the exception of glutamine, which occurs at a reduced frequency. The secondary structure composition of AOR (45 percent helix and 14 percent β sheet) is also within the range seen for water-soluble, globular proteins, and the tertiary packing interactions between the α helices and β strands are also standard. No disulfides exist in AOR and although metal complexation will contribute to protein stability, the number of metal binding sites is not exceptionally high.

The importance of ion pairing to stability has been suggested from both structural and genetic comparisons of mesophilic and thermophilic proteins (*29*–*31*), although some recent studies have questioned the significance of the contributions of these interactions to protein stability (*32*). To evaluate the possible significance of electrostatic interactions for the thermostability of AOR, we tabulated the numbers of ionic interactions per residue for a sampling of 30 monomeric and oligomeric proteins [from the Brookhaven Protein Data Base (*33*)] that contained ~300 to ~2400 residues (about half the size of the AOR monomer to twice the size of the AOR dimer), for comparison to the value determined for AOR (Fig. 5A). The number of ionic interactions per residue was defined as the difference between the number of interactions between oppositely charged side chains and the number of interactions between side chains of like charge, divided by the total number of residues in the protein. Interac-

tions were defined as occurring between polar (oxygen or nitrogen) atoms of potentially ionizable side chains (Arg, Asp, Glu, His, and Lys) that are positioned ≤4 Å apart (*34*). The average number of ion-pairs per residue indicated in Fig. 5A, ~0.04, is close to that observed in Barlow and Thornton's earlier study (*34*). The data in this comparison indicate that AOR does contain a relatively high number of salt bridges per residue, although other proteins also have a similar percentage of stabilizing ionic interactions. Thus, while ion pairs may contribute to the stability of AOR, they cannot represent the sole determinant of thermostability.

Richards pioneered the application of the concepts of molecular surface area and volume to the analysis of macromolecular structures (*35*). From this work, a general picture of water-soluble globular proteins has emerged emphasizing (i) the relatively smooth, polar surface that minimizes the surface energy in an aqueous environment, and (ii) the efficiently packed, relatively nonpolar interior. The solvent accessible surface of AOR can be defined by means of the ACCESS algorithm of Richmond and Richards (*36*), the group radii of Chothia (*37*), and a probe radius of 1.4 Å. The surface area for the AOR dimer and one subunit are 35,740 Å^2 and 19,050 Å^2, respectively, which indicates that about 1200 Å^2 per subunit are buried in the dimer interface. A dimeric protein the size of AOR (134 kD) is expected to have about a 17 percent larger accessible surface area (41,800 Å^2) on the basis of a survey of water-soluble oligomeric proteins (*38*). The accessible surface area of AOR was found to be 56 percent apolar, 27 percent polar, and 17 percent charged, values which are comparable to the values 57 percent, 24 percent, and 19 percent, respectively, observed in a survey of water-soluble proteins (*39*). Hence, the surface of AOR appears to have reduced area, but comparable polarity, when compared to other globular protein structures.

In general, minimization of the ratio of surface area to volume increases the stability of an object by simultaneously reducing the unfavorable surface energy and increasing the attractive interior packing interactions. Although the effects of surface area variation on protein stability have not been systematically explored, the influence of variations in packing interactions on protein stability, especially through the formation or removal of cavities, has been more extensively documented (*40*). Not surprisingly, the introduction of cavities into a protein is often associated with decreased protein stability. The possibility that minimization of the surface area to volume ratio might contribute to the stability of AOR was computationally assessed by evaluating the relative surface area and the number of buried atoms for the same set of proteins used in the ion-pair calculation. The relative surface area was obtained by first calculating the accessible surface area, A_0, for a particular protein, and then dividing by the expected surface area for a protein of that size. From a least squares fit of log (A_0) as a function of log (N), where N is the total number of atoms in the protein model, the expected area, A_c, of a generic protein as a function of N can be fit to the equation:

$$A_c = 15.0\ N^{0.866}$$

which reproduces the observed areas to within ±15 percent, comparable to the results of other studies (*38*, *39*). An estimate of the extent of surface area minimization for a given protein is provided by the ratio A_0/A_c. Variations in the ratio from the average value 1 reflect changes in the molecular shape or surface roughness (or both) that alter the accessible surface area of a particular protein. Evaluation of the efficiency of packing is more complex, but a relatively simple indicator is given by the fraction of atoms in a protein with no (zero) accessible surface area, n_B, because atoms are only buried if completely surrounded by other atoms. The larger the value of n_B, the greater the fraction of interior atoms that are present in the structure, which provides an estimate of the extent of atom sequestration from solvent exposure. The value n_B may be obtained directly from the surface area calculation.

In terms of these parameters, low values of A_0/A_c and high values of n_B correspond to minimization of the ratio of surface area to volume of a protein. Values of A_0/A_c and n_B calculated for the set of proteins used to establish A_c are illustrated in Fig. 5, A and B. The more detailed examination of these factors presented in Fig. 5B indicates that the relative surface area and the fraction of buried atoms in proteins are negatively correlated. Qualitatively, this observation seems reasonable, since in general the greater the surface area of a protein, the more atoms that will be on the surface and the fewer atoms that will be buried, and vice versa. The AOR structure simultaneously exhibits both the minimum value for the relative surface area ratio and the maximum fraction of buried atoms, indicating that AOR has a reduced ratio of surface area to volume relative to other proteins. A plausible inference from this result is that these effects may be related to the enhanced thermal stability of AOR. Consistent with this inference, the same general trends (especially in the surface area) are seen for three pairs of structures for which both mesophilic and thermophilic structures are available (PDB identifiers for the mesophilic and thermophilic structures, respectively, are indicated in parentheses): glyceraldehyde phosphate dehydrogenase (1GPD and 1GD1), phosphofructokinase (1PFK and 3PFK), and phosphoglycerate kinase (PGK and 1PHP), although this pattern is not exhibited by lactate dehydrogenase (9LDT and 1LDB). A role for improved packing in

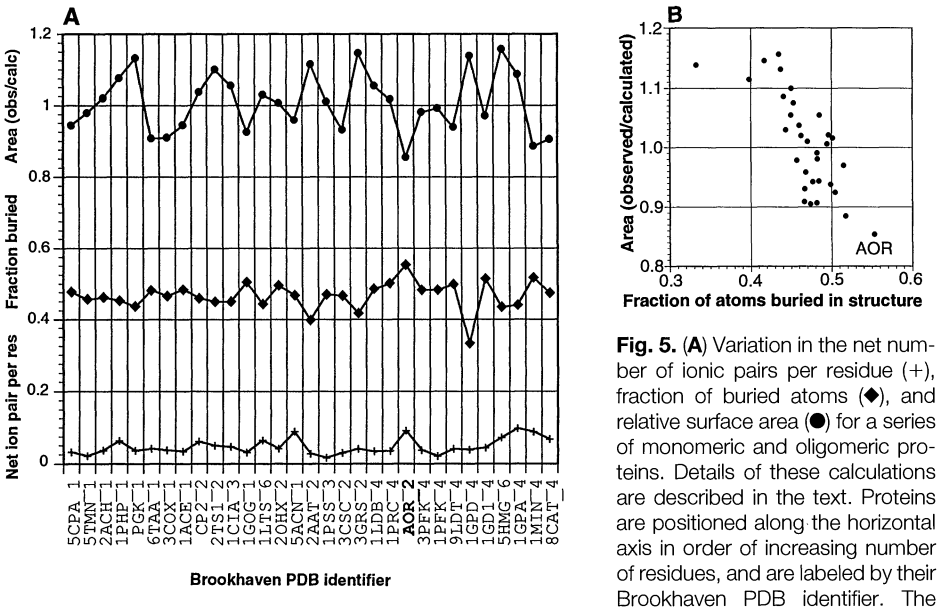

Fig. 5. (**A**) Variation in the net number of ionic pairs per residue (+), fraction of buried atoms (◆), and relative surface area (●) for a series of monomeric and oligomeric proteins. Details of these calculations are described in the text. Proteins are positioned along the horizontal axis in order of increasing number of residues, and are labeled by their Brookhaven PDB identifier. The number of subunits in each protein is indicated by the suffix to the PDB identifier. Two proteins, yeast phosphoglycerate kinase (PGK) and the *Clostridium pasteurianum* nitrogenase iron protein (CP2), have been refined but not yet deposited (*60*). (**B**) Relationship between the fraction of buried atoms and the relative surface area, calculated for the proteins used in (A), demonstrating the anti-correlation between these two parameters. The position of AOR is indicated.

the enhanced stability of hyperthermophilic enzymes has been previously discussed on the basis of biochemical and biophysical comparisons of homologous enzymes from mesophilic and thermophilic sources (*41*). The structural basis by which the surface area is reduced and the fraction of buried atoms is increased in AOR relative to other proteins remains an open question, as is the energetic basis for the possible correlation of these parameters with thermal stability.

The AOR structure demonstrates that hyperthermostability can be achieved without requiring any new types of interactions to stabilize the folded conformation. Rather than being the consequence of one dominant type of interaction, it appears that the stability of AOR may reflect a number of subtle interactions that both minimize the surface energy while maximizing packing and electrostatic interactions. It remains to be determined, however, if this is a general feature of hyperthermophilic enzymes. Similarly, to what extent the AOR structure is representative of other molybdenum or tungsten containing enzymes is also not clear. Surprisingly, the family of proteins represented by AOR shows no sequence similarity to any other molybdoenzyme, yet, the oxotransferase function of all of these proteins is achieved with similar cofactors, namely molybdopterin and typically one or more iron-sulfur clusters. The structure of the pterin site and an adjacent iron-sulfur cluster determined in AOR may therefore reflect a common motif in an apparently diverse group of enzymes, whose unifying feature is the use of a second or third row transition metal in catalysis.

REFERENCES AND NOTES

1. K. O. Stetter, G. Fiala, G. Huber, R. Huber, G. Segerer, *FEMS Microbiol. Rev.* **75**, 117 (1990).
2. M. W. W. Adams and R. M. Kelley, Eds., *Biocatalysis at Extreme Temperatures: Enzyme Systems Near and Above 100°C, ACS Sym. Ser.* **498**, pp. 1–215 (1992).
3. M. W. W. Adams, *Annu. Rev. Microbiol.* **47**, 627 (1993).
4. _____, *FEMS Microbiol. Rev.* **15**, 261 (1994).
5. M. W. Day *et al.*, *Protein Sci.* **1**, 1494 (1992).
6. S. Mukund and M. W. W. Adams, *J. Biol. Chem.* **265**, 11508 (1990).
7. _____, *ibid.* **266**, 14208 (1991).
8. M. J. Kendrick, M. T. May, M. J. Plishka, K. D. Robinson, *Metals in Biological Systems* (Horwood, New York, NY, 1992).
9. R. C. Bray, *Q. Rev. Biophys.* **21**, 299 (1988); J. C. Wootton *et al.*, *Biochim. Biophys. Acta* **1057**, 157 (1991); R. S. Pilato and E. I. Stiefel, in *Bioinorganic Catalysis*, J. Reedijk, Ed. (Dekker, New York, 1993), pp. 131–188; E. I. Steifel, D. Coucouvanis, W. E. Newton, Eds., *Molybdenum Enzymes, Cofactors, and Model Systems, ACS Symp. Ser.* **535**, pp. 1–387 (1993).
10. R. H. Holm, *Coord. Chem. Rev.* **100**, 183 (1990); J. H. Enemark, C. G. Young, *Adv. Inorg. Chem.* **40**, 1 (1994).
11. K. V. Rajagopalan, *Adv. Enzymol. Rel. Areas Mol. Biol.* **64**, 215 (1991); _____ and J. L. Johnson, *J. Biol. Chem.* **267**, 10199 (1992).
12. M. W. W. Adams, in *Encyclopedia of Inorganic Chemistry*, R. B. King, Ed. (Wiley, New York, 1994), pp. 4284–4291.
13. J. L. Johnson and K. V. Rajagopalan, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6856 (1982).
14. J. L. Johnson, N. R. Bastian, K. V. Rajagopalan, *ibid.* **87**, 3190 (1990).
15. M. Karrasch, G. Boerner, R. K. Thauer, *FEBS Lett.* **274**, 48 (1990).
16. J. L. Johnson, N. R. Bastian, N. L. Schauer, J. G. Ferry, K. V. Rajagopalan, *FEMS Microbiol. Lett.* **77**, 213 (1991).
17. G. Börner, M. Karrasch, R. K. Thauer, *FEBS Lett.* **290**, 31 (1991).
18. J. L. Johnson, K. V. Rajagopalan, O. Meyer, *Arch. Biochem. Biophys.* **283**, 542 (1990).
19. D. Hettrick, B. Peschke, B. Tschisuaka, F. Lingens, *Biol. Chem. Hoppe-Seyler* **372**, 513 (1991).
20. J. L. Johnson, K. V. Rajagopalan, S. Mukund, M. W. W. Adams, *J. Biol. Chem.* **268**, 4848 (1993).
21. S. P. Kramer, J. L. Johnson, A. A. Ribeiro, D. S. Millington, K. V. Rajagopalan, *J. Biol. Chem.* **262**, 16370 (1987).
22. A. Kletzin *et al.*, in preparation.
23. R. Soyka and W. Pfleiderer, *Helv. Chim. Acta* **73**, 808 (1990).
24. J. F. Conn, J. J. Kim, F. L. Suddath, P. Blattmann, A. Rich, *J. Am. Chem. Soc.* **96**, 7152 (1974); N. S. Begum and H. Manohar, *Tetrahedron* **11**, 2823 (1992).
25. J. Kim and D. C. Rees, *Nature* **360**, 553 (1992).
26. J. S. Richardson and D. C. Richardson, *Science* **240**, 1648 (1988); R. Aurora, R. Srinivasan, G. D. Rose, *ibid.* **264**, 1126 (1994).
27. G. J. Kleywegt and T. A. Jones, *Acta Crystallogr.* **D50**, 178 (1994).
28. S. Mukund and M. W. W. Adams, *J. Biol. Chem.* **268**, 13592 (1993); in preparation.
29. M. F. Perutz and H. Raidt, *Nature* **255**, 256 (1975); M. F. Perutz, *Science* **201**, 1187 (1978).
30. J. E. Walker, A. J. Wonacott, J. I. Harris, *Eur. J. Biochem.* **108**, 581 (1980).
31. G. J. Davies *et al.*, *Proteins* **15**, 283 (1993).
32. S. Dao-pin, U. Sauer, H. Nicholson, B. W. Matthews, *Biochemistry* **30**, 7142 (1991); Z. S. Hendsch and B. Tidor, *Protein Sci.* **3**, 211 (1994).
33. F. C. Bernstein *et al.*, *J. Mol. Biol.* **112**, 535 (1977).
34. D. J. Barlow and J. M. Thornton, *ibid.* **168**, 867 (1983).
35. F. M. Richards, *Annu. Rev. Biophys. Bioeng.* **6**, 151 (1977).
36. T. J. Richmond and F. M. Richards, *J. Mol. Biol.* **119**, 537 (1978).
37. C. Chothia, *Nature* **254**, 304 (1975).
38. S. Miller, A. M. Lesk, J. Janin, C. Chothia, *ibid.* **328**, 834 (1987).
39. S. Miller, J. Janin, A. M. Lesk, C. Chothia, *J. Mol. Biol.* **196**, 641 (1987).
40. W. A. Lim and R. T. Sauer, *Nature* **339**, 31 (1989); A. E. Eriksson *et al.*, *Science* **255**, 178 (1992); W. A. Lim, A. Hodel, R. T. Sauer, F. M. Richards, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 423 (1994); F. M. Richards and W. A. Lim, *Q. Rev. Biophys.* **26**, 423 (1994); Y. Harpaz, M. Gerstein, C. Chothia, *Structure* **2**, 641 (1994).
41. A. Wrba, A. Schweiger, V. Schultes, R. Jaenicke, P. Závodszky, *Biochemistry* **29**, 7584 (1990).
42. Because of the oxygen-sensitivity of AOR, crystals were grown by the capillary method [M. M. Georgiadis *et al.*, *Science* **257**, 1653 (1992)] with 20 μl of protein (30 mg/ml in 50 mM tris, pH 8.0, containing 10 percent glycerol, 2 mM dithionite, 2 mM dithiothreitol) and 30 μl of precipitant, in an inert atmosphere glove box. Three different conditions yielded diffraction quality crystals: (native 1) 0.18 M Na_2MoO_4, 0.1 M tris, pH 8.0, and 30 percent 4K PEG; (native 2) 0.2 M sodium acetate, 0.1 M sodium cacodylate, pH 6.5, 30 percent 8K PEG; and (native 3) 0.2 M $MgCl_2$, 0.1 M tris, pH 8.0, 30 percent 4K PEG, and 10 percent glycerol. Cell constants and space groups for these three forms are summarized in Table 1. Each crystal form contains one AOR dimer in the asymmetric unit.
43. A. J. Howard *et al.*, *J. Appl. Crystallogr.* **20**, 383 (1987).
44. W. Kabsch, *ibid.* **21**, 916 (1988).
45. G. N. Reeke Jr. *ibid.* **17**, 125 (1984); P. Bethge, *ibid.*, p. 215.
46. A. G. W. Leslie, in *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, no. 26, (Daresbury Laboratory, Warrington, United Kingdom, 1992).
47. PHARE was originally distributed by G. Bricogne (MRC, Cambridge, England) and subsequently modified for the VAX by J. L. Crawford and J. E. Ladner.
48. The location of the non-crystallagraphic NCS twofold between the AOR subunits in the $P2_12_12_1$ crystal form was established from an analysis of the heavy atom and anomalous scattering sites. The NCS twofold within each *P1* crystal form was located using the Crowther fast rotation function [R. A. Crowther, in *The Molecular Replacement Method*, M. G. Rossman, Ed. (Gordon and Breach, New York, 1972), pp. 173–178], as were the cross rotation angles between each crystal form. After determination of the NCS relationships, the initial experimental maps were averaged within and between the three crystal forms, which is effectively equivalent to sixfold averaging. For the averaging, a dimer mask was generated with a modification of the Wang algorithm (*49*). Several iterations of averaging produced electron density maps of sufficient quality to trace the chain and to build a polyalanine model. At the end of averaging, the final *R* factors between the observed structure factors and structures factors calculated from the averaged maps were 23.5, 20.2, and 22.0 percent to 3 Å resolution, for crystal forms 1, 2, and 3, respectively.
49. B. C. Wang, *Methods Enzymol.* **115**, 90 (1985).
50. T. A. Jones, *ibid.*, p. 151.
51. _____, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr.* **A47**, 110 (1991).
52. D. E. Tronrud, L. F. Ten Eyck, B. W. Matthews, *ibid.* **A43**, 489 (1987).
53. A. T. Brünger, *J. Mol. Biol.* **203**, 803 (1968).
54. R. A. Laskowski, R. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **26**, 283 (1993).
55. R. Lüthy, J. U. Bowie, D. Eisenberg, *Nature* **356**, 83 (1992).
56. C. Colovos and T. O. Yeates, *Protein Sci.* **2**, 1511 (1993).
57. The α helical and β sheet regions have been identified as (*54*): α-helical residues: 25-31, 34-45, 98-105, 135-137, 142-153, 165-169, 189-196, 214-230, 232-241, 237-241, 246-252, 269-272, 274-280, 312-317, 327-340, 344-360, 365-368, 380-390, 396-400, 403-410, 413-416, 433-441, 454-457, 472-491, 495-497, 503-514, 520-541, 545-548, 553-557, 573-583, 593-599; β sheet residues: 7-13, 17-24, 54-67, 76-82, 89-95, 109-113, 119-125, 128-133, 158-164, 174-178, 182-186, 198-206, 282-286, 296-301, 305-308.
58. P. J. Kraulis, *J. Appl. Cryst.* **24**, 946 (1991); E. A. Merritt and M. E. P. Murphy, *Acta. Crystallogr.* **D50**, 869 (1994).
59. S. V. Evans, *J. Mol. Graphics* **11**, 134 (1993).
60. D. C. Rees, in preparation.

61. We thank B. T. Hsu, J. Schlessman, J. Kim, S. Faham, M. Stowell, and T. McPhillips for advice and assistance, and A. G. Myers, M. K. Johnson, P. J. Bjorkman, and J. E. Bercaw for discussions. Supported by USPHS grant GM50775 (D.C.R.); ONR grant N00014-90-J-1894 (M.W.W.A.), DOE grant FG09-88ER13901 (M.W.W.A.) and NSF grant BCS-9320069 (M.W.W.A.); the Beckman Institute, the Joseph Irvine Equipment Fund, and the Howard Hughes Medical Institute (to P. J. Bjorkman) for part of the diffraction instrumentation; NIH fellowship 1F32 GM15006 (M.K.C.); and by the rotation camera facility at the Stanford Synchrotron Radiation Laboratory, which in turn is supported by the Department of Energy, Office of Basic Energy Sciences, and the National Institutes of Health Biomedical Resource Technology Program, Division of Research Resources. Coordinates have been deposited in the Brookhaven Protein Data base, not yet available.

30 September 1994; accepted 20 January 1995