Conformational variability in structures of the nitrogenase iron proteins from Azotobacter vinelandii and Clostridium pasteurianum

ARTICLE *in* JOURNAL OF MOLECULAR BIOLOGY · AUGUST 1998

Impact Factor: 4.33 · DOI: 10.1006/jmbi.1998.1898 · Source: PubMed

5 AUTHORS, INCLUDING:

Leemor Joshua-Tor
Cold Spring Harbor Laboratory
67 PUBLICATIONS **6,298** CITATIONS

SEE PROFILE

Article No. mb981898 *J. Mol. Biol.* (1998) **280**, 669–685

JMB

Conformational Variability in Structures of the Nitrogenase Iron Proteins from *Azotobacter vinelandii* and *Clostridium pasteurianum*

Jamie L. Schlessman[1], Deborah Woo[1], Leemor Joshua-Tor[1] James B. Howard[2] and Douglas C. Rees[1,3]*

[1]*Division of Chemistry and Chemical Engineering California Institute of Technology, 147-75CH Pasadena, CA 91125, USA*

[2]*Department of Biochemistry University of Minnesota School of Medicine, Minneapolis MN 55455, USA*

[3]*Howard Hughes Medical Institute, California Institute of Technology, 147-75CH Pasadena, CA 91125, USA*

**Corresponding author*

The nitrogenase iron (Fe) protein performs multiple functions during biological nitrogen fixation, including mediating the mechanistically essential coupling between ATP hydrolysis and electron transfer to the nitrogenase molybdenum iron (MoFe) protein during substrate reduction, and participating in the biosynthesis and insertion of the FeMo-cofactor into the MoFe-protein. To establish a structural framework for addressing the diverse functions of Fe-protein, crystal structures of the Fe-proteins from *Azotobacter vinelandii* and *Clostridium pasteurianum* have been determined at resolutions of 2.2 Å and 1.93 Å, respectively. These two Fe-proteins are among the more diverse in terms of amino acid sequence and biochemical properties. As described initially for the *A. vinelandii* Fe-protein in a different crystal form at 2.9 Å resolution, each subunit of the dimeric Fe-protein adopts a polypeptide fold related to other mononucleotide-binding proteins such as G-proteins, with the two subunits bridged by a 4Fe:4S cluster. The overall similarities in the subunit fold and dimer arrangement observed in the structures of the *A. vinelandii* and *C. pasteurianum* Fe-proteins indicate that they are representative of the conformation of free Fe-protein that is not in complex with nucleotide or the MoFe-protein. Residues in the cluster and nucleotide-binding sites are linked by a network of conserved hydrogen bonds, salt-bridges and water molecules that may conformationally couple these regions. Significant variability is observed in localized regions, especially near the 4Fe:4S cluster and the MoFe-protein binding surface, that change conformation upon formation of the ADP·AlF_4^- stabilized complex with the MoFe-protein. A core of 140 conserved residues is identified in an alignment of 59 Fe-protein sequences that may be useful for the identification of homologous proteins with functions comparable to that of Fe-protein in non-nitrogen fixing systems.

Keywords: nitrogen fixation; nucleotide-binding proteins; iron-sulfur proteins; metalloproteins; macromolecular structure

Introduction

Fixed nitrogen, in the form of ammonia (NH_3), is a critical starting material for fundamental processes in both agriculture and industry. Biological nitrogen fixation, effected primarily by diazotrophic bacteria, supplies more than 60% of the world's annual resources of new ammonia (Schlesinger, 1991). Juxtaposed against the corresponding commercial Haber-Bosch reaction, with its high temperature and pressure requirements, the biological process of producing ammonia from dinitrogen presents a striking

Present address: J. L. Schlessman, Department of Molecular and Medical Pharmacology, University of California, Los Angeles School of Medicine, Los Angeles, CA 90095, USA; D. Woo, NeXstar Pharmaceuticals, 650 Cliffside Drive, San Dimas, CA 91773, USA; L. Joshua-Tor, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA.

E-mail address of the corresponding author: rees@citray.caltech.edu

0022–2836/98/290669–17 $30.00/0

contrast. Such elegance, however, is countered by a steep energetic cost: while thermodynamically favorable (Alberty, 1994), biological nitrogen fixation by the metalloenzyme nitrogenase has a substantial ATP requirement that can consume up to ~40% of the ATP synthesized by a bacterium that is actively fixing nitrogen (Daesch & Mortenson, 1967).

The two-component nitrogenase system consists of the iron protein (Fe-protein) and molybdenum-iron protein (MoFe-protein) working in concert to effect nitrogen reduction (for recent reviews, see Howard & Rees, 1994, 1996; Peters *et al.*, 1995; Burgess & Lowe, 1996; Muchmore *et al.*, 1996; Seefeldt & Dean, 1997). Under conditions of molybdenum depletion, alternative nitrogenase systems containing either vanadium or iron that are homologous to the molybdenum-containing ''conventional'' nitrogenase system may be induced (reviewed by Eady, 1996). Both metalloproteins are extremely oxygen-sensitive. Fe-protein is a 64 kDa γ_2 dimer with a single 4Fe:4S cluster covalently bridging each subunit; the isolated protein can bind up to two MgADP and/or MgATP molecules per dimer, and is essential for coupling nucleotide hydrolysis to electron transfer. The structure of the *Azotobacter vinelandii* Fe-protein (denoted Av2) was solved initially at 2.9 Å resolution using a monoclinic crystal form (Georgiadis *et al.*, 1992). The MoFe-protein is a 240 kDa $\alpha_2\beta_2$ tetramer containing two types of metallocenters: the 8Fe:7S P-clusters located at the interface between the α and β-subunits of each αβ dimer, and the 1Mo:7Fe:9S:1homocitrate FeMo-cofactor contained within each α subunit. The FeMo-cofactor contains the expected site of substrate reduction, while the P-clusters are believed to participate in the transfer of electrons from the Fe-protein to the FeMo-cofactor. Extensive crystallographic studies have yielded structures of the MoFe-protein from *A. vinelandii* (Kim & Rees, 1992a,b; Muchmore, 1995; Peters *et al.*, 1997), *Clostridium pasteurianum* (Bolin *et al.*, 1993; Kim *et al.*, 1993) and *Klebsiella pneumoniae* (Smith *et al.*, 1998). These three MoFe-proteins are denoted Av1, Cp1 and Kp1, respectively.

Although the detailed mechanism of substrate reduction is still poorly defined, the general sequence of electron-transfer reactions through the nitrogenase system has been established. Three basic types of electron transfer reactions are involved: (i) the reduction of Fe-protein by electron carriers such as ferredoxins and flavodoxins *in vivo* or dithionite *in vitro*; this is conventionally envisioned as a one-electron transfer, although Fe-protein can exist in a two-electron ''super-reduced'' form that may be important for substrate reduction (Watt & Reddy, 1994; Angove *et al.*, 1997); (ii) transfer of electrons from Fe-protein to MoFe-protein in a MgATP-dependent process, with a minimal stoichiometry of two MgATP molecules hydrolyzed per electron transferred; and (iii) electron and proton transfer to the substrate at the active site within the MoFe-protein. The overall stoichiometry of dinitrogen reduction under optimal conditions has been established (Simpson & Burris, 1984):

$$N_2 + 8H^+ + 8e^- + 16MgATP \rightarrow 2NH_3 + H_2 + 16MgADP + 16P_i$$

Including obligatory hydrogen evolution coupled to dinitrogen reduction, the net process consumes eight electrons. Nitrogenase is a relatively slow-acting enzyme, with a turnover time per electron of ~5 s^{-1}. Each electron transfer step between Fe-protein and MoFe-protein appears to involve a cycle of association and dissociation of the protein complex, with dissociation the likely rate-determining step for the overall reaction (Hageman & Burris, 1978; Lowe & Thorneley, 1983, 1984). The complex of the two proteins plays a key role in the nitrogenase mechanism, since it is in this species that the coupling of ATP hydrolysis to electron transfer occurs. Recently, the structure of an $(Av2)_2Av1$ complex stabilized by AlF_4^-·ADP has been described (Schindelin *et al.*, 1997). The complex structure indicates that while relatively minor changes occur in the MoFe-protein structure relative to the uncomplexed protein, large-scale conformational changes are evident in Fe-protein, both within and between the subunits. The role of ATP hydrolysis in this process is not well understood, but there is suggestive evidence that ATP hydrolysis is coupled to a series of conformational changes, especially in the Fe-protein, that function as a molecular clock to properly sequence different events in the reaction (Howard & Rees, 1994).

Although Fe-protein is perhaps best known for its role in substrate reduction, this protein participates in a variety of other aspects of nitrogenase function, including FeMo-cofactor biosynthesis and maturation of MoFe-protein (reviewed by Allen *et al.*, 1994; Burgess & Lowe, 1996). Reflecting the multiplicity of functional roles, Fe-protein exhibits a high degree of amino acid sequence conservation (Figure 1). Two of the more diverse Fe-proteins, from *A. vinelandii* (Av2) and *C. pasteurianum* (Cp2), share 69% amino acid sequence identity (Tanaka *et al.*, 1977; Hausinger & Howard, 1982). The differences between Av2 and Cp2 occur in regions of greatest variability between Fe-protein sequences: the so-called ''60's loop'' formed by residues ~60 to 70 (Av2 sequence numbering), and the carboxy terminus. Specifically, the major features distinguishing Cp2 from Av2 include a two residue deletion at residue 65 and a 13 residue truncation of the C terminus. The functional differences between Av2 and Cp2 are most dramatically illustrated by the ability of these proteins to replace Fe-proteins from other species in substrate reduction. While Av2 can typically function with MoFe-proteins from other species, Cp2 generally does not support substrate reduction with other MoFe-proteins (Emerich *et al.*, 1981). In particular, Cp2

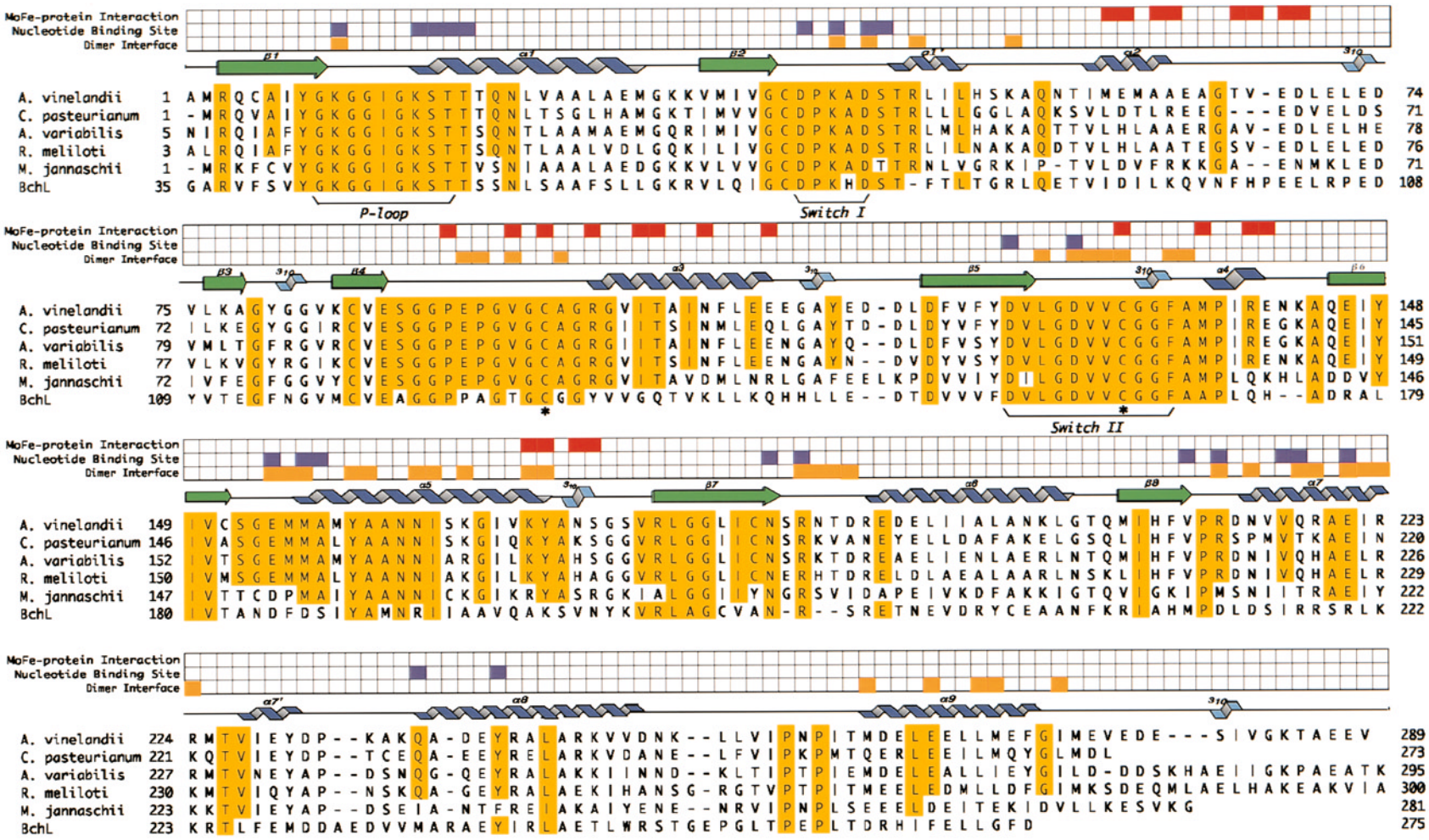

Figure 1. Alignment of selected nitrogenase Fe-proteins and related sequences. Amino acid sequences for the nitrogenase iron proteins from *A. vinelandii*, *C. pasteurianum*, *Anabaena variabilis*, and *Rhizobium meliloti*, as well as the NifH-like sequence from *Methanococcus jannaschii* and the BchL sequence from *Rhodobacter capsulatus* are included. The secondary structure for AV2 is shown as in Table 1 and Figure 2. Residues implicated in binding to the MoFe-protein and to nucleotide as identified in the AlF_4-stabilized nitrogenase complex are indicated in red and purple, respectively. Residues involved in dimer interactions within Fe-protein are indicated in orange. The two Cys residues from each subunit that coordinate the 4Fe:4S cluster are denoted with an asterisk (*). The 140 residues identical in at least 57 of 59 aligned NifH sequences are highlighted in yellow.

and Av1 form a tight-binding, inactive heterocomplex (Emerich *et al*., 1978). The contribution of the varied residues to these functional differences has been probed by mutagenesis experiments (Jacobson *et al*., 1990; Peters *et al*., 1994), that implicated residues in the 60's loop as participating in the formation of a stable Cp2-Av1 heterocomplex. Differences in the relative inhibitory efficiencies of various nucleoside diphosphate species have been reported for the *C. pasteurianum* and *A. vinelandii* nitrogenases (Weston *et al*., 1983), which presumably reflect the affinity of Cp2 and Av2 for these ligands. Given this diversity in biochemical and sequence properties, crystallographic analyses of both Av2 and Cp2 should permit an assessment of the range of structural variability present in the Fe-protein family.

A clear mechanistic picture of nitrogenase, including Fe-protein, will provide clues to the complex coupling of nucleotide hydrolysis and electron transfer, and will have implications for signal transduction in multiprotein assemblies (Georgiadis *et al*., 1992; Thorneley, 1992; Howard, 1993; Howard & Rees, 1994). Fe-protein displays structural similarity to other nucleotide-binding proteins, including signal transduction molecules such as G-proteins and *ras*, and energy transduction systems such as myosin. Consequently, detailed comparisons of the structures may yield greater insight into nucleotide-dependent transduction mechanisms. Nucleotide-binding proteins in this category, including Fe-protein, are characterized by common core elements (Milner-White *et al*., 1991; Schulz, 1992; Sprang, 1997) consisting of: (1) predominantly parallel β-sheet flanked by α-helices; (2) a phosphate-binding loop (P-loop), or Walker A motif (Walker *et al*., 1982), containing the G-X-X-X-X-G-K-S/T consensus sequence (residues 9 to 16 in Av2); and (3) two switch regions, Switch I and Switch II, that interact with the γ-phosphate group of the bound nucleoside triphosphate. Fe-protein residues corresponding to Switch I and Switch II include ~38 to 43 and 125 to 135, respectively. The conserved D-X-X-G sequence motif corresponds to residues 125 to 128 in Av2. The switch regions undergo conformational changes upon hydrolysis to the nucleoside diphosphate, with consequent loss of the interaction with the terminal phosphate group. As a result of these structural alterations, the switch regions play critical roles in transmitting information concerning the nucleotide state to other effector molecules that bind to these regions.

Essential to the detailed characterization of the structural properties of Fe-protein is the availability of high-resolution refined atomic coordinates. We present the crystal structures of the nitrogenase Fe-protein in the absence of nucleotide from *A. vinelandii* and *C. pasteurianum* at 2.2 Å and 1.93 Å resolution, respectively, to provide greater insight into the structural basis of Fe-protein function. The present Av2 structure represents a novel orthorhombic crystal form for Fe-protein, and the Cp2 structure provides the highest-resolution structure of Fe-protein to date.

Results

Fe-protein topology

As described by Georgiadis *et al*. (1992), each subunit in the Fe-protein dimer exhibits a mixed α-helix/β-sheet polypeptide fold, with a consensus topology of an eight-stranded β-sheet flanked by nine α-helices (Figure 2). A non-crystallographic 2-fold axis passing through the solvent-exposed 4Fe:4S cluster relates the two monomers in each Fe-protein dimer. The twisted β-sheet core of each monomer contains seven parallel and one short antiparallel β-strands (Table 1), with strand order 3-4-2-5-1-6-7-8. The single antiparallel strand, β3, is located at one edge of the β-sheet. Topological switch points (Brändén, 1980) formed by loops at the carboxy-terminal ends of several β strands create binding sites critical to Fe-protein function; the loops following β1 and β2 are involved with

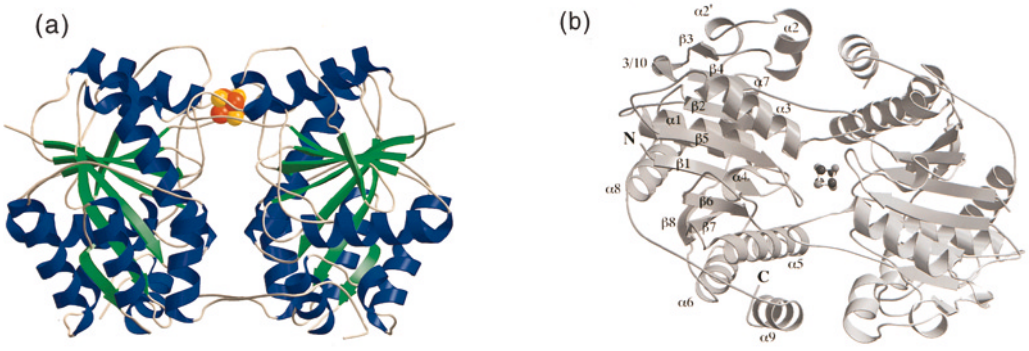

Figure 2. (a) A ribbon diagram of the Av2 dimer. Each Fe-protein monomer adopts a mixed α-helix/β-sheet topology, with the eight-strand β-sheet core (shown in green) surrounded by nine consensus α-helices (shown in blue). The 4Fe:4S cluster lies at one edge of the dimer interface; iron atoms are indicated in orange, and sulfur atoms are in yellow. The non-crystallographic 2-fold axis runs vertically between the subunits in the plane of the page. (b) A view down the 2-fold axis of the Cp2 dimer, with secondary structure elements (Table 1) indicated for one subunit.

Table 1. Secondary structure of nitrogenase Fe-protein from *Azotobacter vinelandii* and *Clostridium pasteurianum*, calculated using the program STRIDE (Frishman & Argos, 1995)

Secondary structure	AV2	CP2
β1	3-8	2-8
α1	15-29	14-27
β2	33-37	32-37
α2	57-64	56-61
3_{10}	A73-A75	
(α2′)	(B72-B75)	69-72
β3	77-79	73-75
3_{10}	A80-A82	77-79
β4	84-87	81-84
α3	98-112	95-109
β5	121-125	118-124
3_{10}	133-135	B130-B133
α4	136-141	133-138
β6	146-151(B149)	143-148
α5	155-171(B173)	152-168
β7	178-185(B182)	175-182
α6	192-203	189-200
β8	206-211(B207)	204-208
α7	215-222	212-220
(α7′)	(B227-B230)	A224-A227
α8	235-248	232-246
α9	261-271	258-268

Residues are numbered according to the Av2 or Cp2 sequence, as appropriate.

nucleotide binding, while loops following β5 and β6 contain cysteine ligands to the 4Fe:4S cluster.

Computerized comparisons between the Fe-protein and other proteins of known structure, using either the DALI (Holm & Sander, 1998) or SCOP (Murzin *et al.*, 1995) servers, reveal that the Fe-protein exhibits particularly close similarities to the GTPase domains of Ffh and FtsY (the signal recognition particle (Freymann *et al.*, 1997) and signal recognition particle receptor (Montoya *et al.*, 1997), respectively), and to two ATP-utilizing enzymes, dethiobiotin synthase (Huang *et al.*, 1994) and adenylosuccinate synthase (Poland *et al.*, 1993). The β-sheet at the core of each of these proteins contains strands corresponding to the parallel β-strands 2-5-1-6-7-8 of Fe-protein, along with the intervening α-helices. The GTPase domains of Ffh, FtsY and adenylosuccinate synthase contain an additional strand corresponding to strand β4 of Fe-protein, such that the β-sheet topology of these proteins differs from Fe-protein only by the presence of the short anti-parallel β3 in the latter. The parallel β-strands 5-1-6-7 (Fe-protein numbering) are found at the core of numerous mononucleotide-binding proteins, such as *ras* and the G-protein family (Milner-White *et al.*, 1991). The most significant variation in β-sheet topology in this family occurs at the edge of the sheet adjacent to β5 (Schulz, 1992); for example, in *ras* and related proteins, an antiparallel β-strand hydrogen bonds to the equivalent to β5, as compared to a parallel β-strand at this position in Fe-protein. It is intriguing that the GTPase proteins that most closely resemble Fe-protein are among the few G-proteins to stably exist in the nucleotide-free form (as does Fe-protein); in general, G-proteins are found exclusively with some form of bound nucleotide (Sprang, 1997).

The Av2 and Cp2 structures, refined at 2.2 Å and 1.93 Å resolution, respectively, generally resemble the structure described previously for Av2 in a monoclinic crystal form at 2.9 Å resolution. For ease of discussion, the presently determined high-resolution Av2 and Cp2 structures, the initial Av2 structure (Georgiadis *et al.*, 1992), and the Fe-protein in the $ADP{\cdot}AlF_4^-$ stabilized Av1-Av2 complex (Schindelin *et al.*, 1997) will be designated AV2, CP2, 1NIP and ALF, respectively. Unless otherwise indicated, residues are numbered according to the Av2 protein sequence (Hausinger & Howard, 1982). Beyond the overall similarities, however, there are discernible differences in these structures that reflect contributions from both sequence and conformational variations. Although tight non-crystallographic symmetry restraints were imposed on the two subunits in each dimer during the initial stages of the refinement, monitoring of R_{free} during later stages of the refinement to avoid overfitting (Brünger, 1992a; Kleywegt & Jones, 1995) indicated that both Av2 and Cp2 structures exhibited deviations from this ideal symmetry. These differences can be identified by comparison of the two monomers within a given dimer, and by comparison of subunits between different Av2 and Cp2 dimers (Figures 3 and 4). For these calculations, the various structures were superimposed using a rigid-body, least-squares algorithm developed by Kabsch (1976). The magnitudes of changes are comparable for each of the pairwise comparisons; for example, the rms deviations between the C^α positions in the two subunits of AV2, 1NIP and CP2 are 0.62 Å, 0.59 Å and 0.74 Å, respectively. When the dimers are compared by superimposing the A and B subunits of one structure onto the A and B subunits of a second structure, rms deviations of 1.13 Å, 1.33 Å and 1.13 Å are observed in the C^α positions between the AV2-1NIP, AV2-CP2 and CP2-1NIP dimers, respectively. For comparison, the rms deviations in C^α position between the AV2 and ALF structures are 1.7 Å and 4.1 Å when the superposition in based on the monomer and dimer, respectively.

Examination of Figures 3 and 4 illustrates that the rms deviations in C^α positions are not uniform throughout the polypeptide chain, but rather exhibit pronounced variations with structurally conserved regions interspersed with more variable regions. The more variable regions are not always correlated with regions of higher average temperature factor (Figure 3); for example, the region around residue 110 that has the highest average *B*-factor in Av2 exhibits few structural differences between the two subunits in the dimer. Within each Fe-protein subunit, the β-sheet core exhibits strong structural conservation; the rms deviations between superimposed C^α atoms of β-sheet resi-

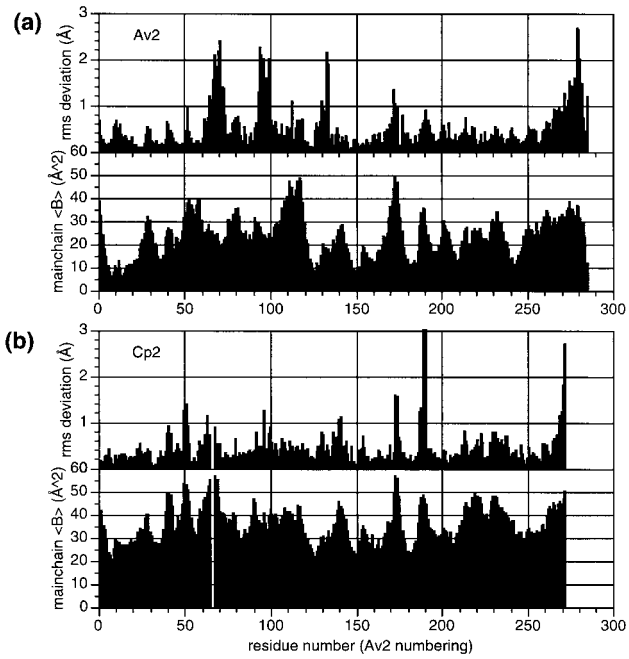

Figure 3. Plots of the structural variability of AV2 and CP2. (a) rms deviation in C^{α} position following superposition of each subunit onto the other in the AV2 dimer, plotted as a function of residue number, with the variation in average temperature factor for the main-chain atoms illustrated below. (b) The same as for (a), except calculated for the CP2 structure. The rms deviation of residue 189 (Av2 number) was truncated from the actual value of 6.3 Å in this plot. Comparison of the rms deviation and temperature factor plots reveals that regions of high structural variability are not universally correlated with regions of high average temperature factor.

dues in the two subunits of the dimer average 0.2 Å for both AV2 and CP2. When the β-sheet C^{α} atoms of CP2 are superimposed onto those of AV2, the rms deviation is 0.3 Å for each subunit and 0.6 Å for the dimer. Even between the AV2 and ALF structures, the β-sheet remains relatively invariant, as the rms deviation in C^{α} position for β-sheet residues superimposed between the individual AV2 and ALF subunits is ~0.5 Å. The α-helices surrounding the β-sheet are less structurally conserved between monomers within each Fe-protein structure, and between AV2 and CP2. Superimposing the α-helical C^{α} atoms of each subunit within AV2 or CP2 yields an rms deviation of 0.4 Å. Between CP2 and AV2, the value rises to 0.8 Å for each subunit and to 1.1 Å for the dimer. For two proteins that share 69% sequence identity, these rms values correspond well with the value (0.71 Å) expected from an empirical relationship obtained by Chothia & Lesk (1986).

In comparing AV2 and the 1NIP structures of Av2, the largest deviations occur in the 40's, 90's and C-terminal regions, where different conformations of the polypeptide chain are adopted. The differences in the 40's and 90's regions appear to reflect specific conformational differences, since these regions are similar within a given dimer. In contrast, the C-terminal regions appear generally flexible and sensitive to lattice interactions. For example, the C terminus of subunit B of one AV2 dimer packs tightly against the surface of Fe-protein near the 4Fe:4S cluster of a neighboring molecule, with the result that all 289 residues of the B subunit can be identified in the electron density maps. In the 1NIP structure, molybdate ions from the Na_2MoO_4 used for crystallization were found to bind to the P-loop; in addition, a partially occupied, endogenous ADP molecule was detected bound between the Av2 subunits in this structure. Despite inclusion of Na_2MoO_4 in the AV2 crystallization conditions, no evidence of bound molybdate was found in the P-loop of the orthorhombic crys-

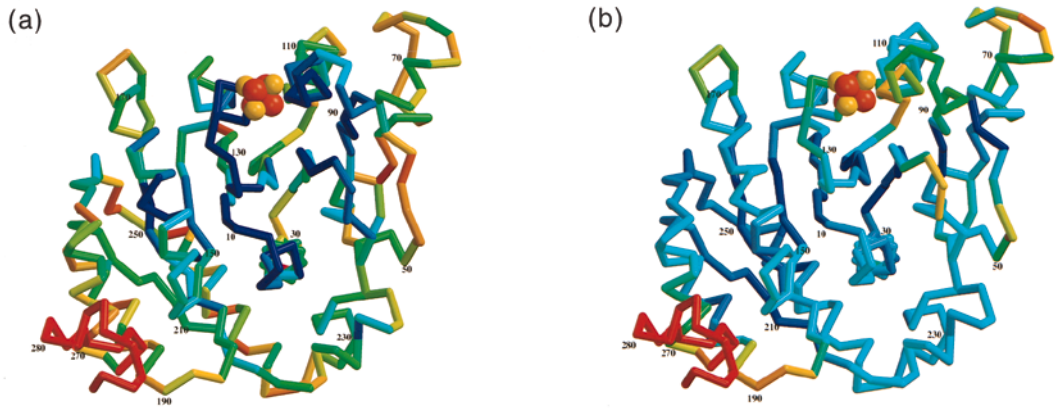

Figure 4. (a) A representation of one Fe-protein monomer colored according to rms deviation in C^{α} position, following superposition of the four AV2 and CP2 subunits. Residues with high rms deviation, such as the C terminus, are indicated in red; regions of strong structural conservation, such as the β-sheet, are indicated in dark blue. (b) A representation of one Fe-protein monomer colored according to amino acid residue conservation, following alignment of 59 Fe-protein sequences. Residues in red, such as the C terminus, indicate the regions of greatest sequence variability; those in dark blue, such as the 4Fe:4S cluster ligands, P-loop, Switch I and Switch II residues, indicate regions of strongest sequence conservation. In both (a) and (b), the view is from the dimer interface, and Av2 residue numbers are included for reference.

tals; furthermore, no ordered nucleotide was observed in this structure. It is possible that the variations between the 1NIP and AV2 structures reflect these differences in bound anion and nucleotide.

While the Av2 and Cp2 structures exhibit strong overall similarity, there are regions that show more substantial variation in their conformation. These residues, identified as having rms differences in C^{α} positions exceeding 1.5 Å, are primarily located in loop regions and include residues 28, 40 to 42, 51-52, 61 to 70, 91 to 93, 96, 98 to 101, 112-113, 116 to 118, 133-134, 173-174, 189 to 191 and 269 to the end. The deviations in the 60's loop and carboxy terminus undoubtedly reflect the deletions and more substantial sequence variations in these regions.

The six subunits present in the AV2, CP2 and 1NIP structures may be superimposed with pairwise rms deviations in C^{α} positions that average 0.98 Å. Residues with rms differences exceeding 1.5 Å are again concentrated primarily in loop regions, and include many of the same residues that were present in the AV2–CP2 comparison: 41-42, 51-52, 64-65, 68-69, 91, 96, 98 to 100, 113, 117, 173 to 75, 189, 190-191 and 269 to the end. A representation of the extent of structural variation within a single Fe-protein subunit is presented in Figure 4(a), which indicates both the structural conservation in the β-sheet core, as well as the regions of greater conformational variability. These changes reflect contributions of both sequence variations (Figure 4(b)) and conformational mobility due to intrinsic flexibility and effects of lattice interactions.

Conserved water structures

A total of 372 and 118 water molecules were incorporated into the AV2 and CP2 models, respectively. Of these, 108 pairs of water molecules in the AV2 structure and 44 pairs of water molecules in the CP2 structure were found to occupy similar positions in each of the two subunits within each respective Fe-protein dimer. The program LSQMAN was used for this analysis (Kleywegt & Jones, 1994); following superposition of the relevant subunits, two water molecules were considered equivalent if they were positioned within 2.0 Å. Of this group, 11 sets of equivalent water molecules were identified that are similarly located in the four crystallographically independent subunits of the AV2 and CP2 structures. An additional eight sets of equivalent water molecules were identified that are similarly located in three of the four subunits of AV2 and CP2. These water molecules are listed in Table 2, along with the potential hydrogen bonding partners in the protein. Most of these water molecules form multiple hydrogen bonding interactions, as expected from the conservation of these sites in the different structures. As detailed more fully below, many of these water molecules are associated with the P-loop, Switch I and Switch II regions, and are likely functionally relevant. For the purposes of discussion, these water molecules will be denoted by the residue number assigned to the water molecule interacting with subunit A of AV2.

The 4Fe:4S cluster of Fe-protein

The thiol ligands of Cys97 and Cys132 from each Fe-protein subunit coordinate the cubane 4Fe:4S cluster. The cluster is buttressed on three sides by the main-chain atoms of residues 96 to 100 and 132 to 134, as well as by the hydrophobic side-chains of Ala98, Val130 and Phe135; these invariant residues are likely to be critical for maintaining the appropriate cluster environment. Each cluster ligand is at the amino-terminal end of helical struc-

Table 2. Conserved water molecules in AV2 and CP2, identified as described in the text

Av2		Cp2		
A	B	A	B	Potential hydrogen bonding partners
23	11	7	22	Ser16 N, Wat13
57	12	2	12	Gly14 O, Ser16 O^{γ}, Thr17 N, $O^{\gamma 1}$
13	16	3	1	Lys10 O, Gly12 N, Ile13 N, Lys15 N, Wat23
34	15	31	11	Lys10 N, Val15 O, Ala157 O
21	121	20	65	Ile255 N
24	275	27	14	Asp39 N, Ser44 O^{γ}, Glu87 O^{ε}
35	61	25	37	Leu170 O, Tyr124 Oη, Lys143 O
189	44	26	30	Gln4 Nε2, Val130 O
203	133	102	9	Thr18 $O^{\gamma 1}$, Tyr240 O^{η}
107	153	8	–	Gln129 N, Leu127 O, Val130 O
18	17	–	95	Ser186 O
–	179	13	111	Val131 N
39	–	123	24	Glu110 O, Ala114 N, Tyr115 N
89	–	34	55	Lys57 O, Val75 O, Glu87 N
–	66	57	109	Ala198 O, Glu202 O^{ε}
–	353	122	59	Leu253 O
58	–	87	66	Glu202 O^{ε}, Thr260 N

For each structure, the residue number is listed for the water molecule binding to the A or B subunit of the specified protein. The symbol ''–''denotes that no water molecule was found in this site. Amino acid residues serving as potential hydrogen bonding partners are numbered according to the Av2 sequence.

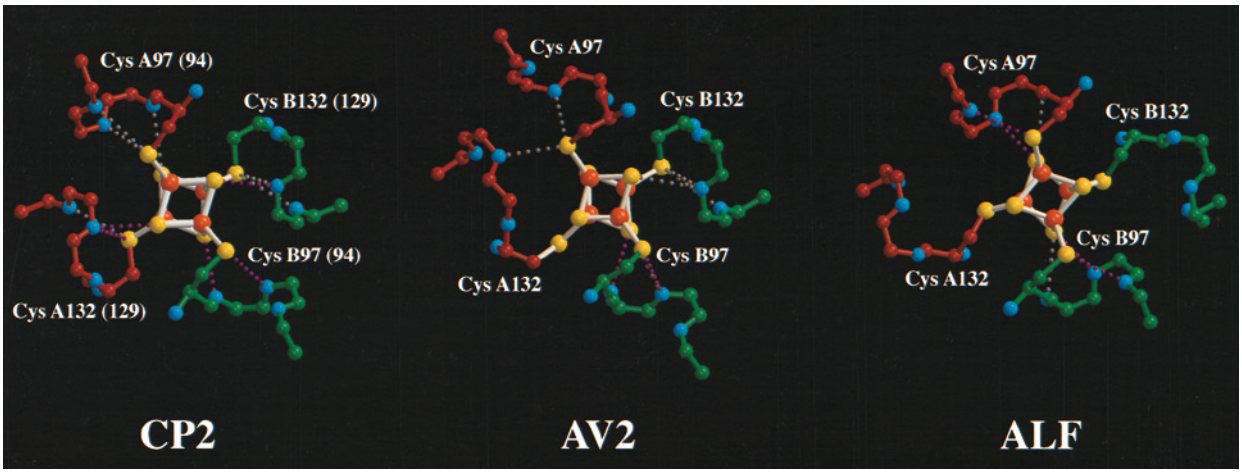

Figure 5. Ball-and-stick representations of the 4Fe:4S cluster environment in CP2, AV2 and ALF structures, viewed down the dimer 2-fold axis from the MoFe-protein binding direction. Subunit A of each protein is shown in red, and subunit B in green; iron atoms are colored orange, and sulfur atoms are yellow. The cysteine ligands to the cluster are indicated in each cluster. Potential stabilizing NH-S hydrogen bonds are indicated in purple, and ''incipient'' hydrogen bonds are shown in grey (as listed in Table 3). In general, residues neighboring the cluster in ALF show a more extended conformation than in CP2. Note that subunit A of AV2 most closely resembles that of ALF, while subunit B of AV2 more closely matches that of CP2. This variability is due largely to lattice contacts imposed by packing in the orthorhombic AV2 crystal. Residues are numbered according to the Av2 sequence, with Cp2 residue numbers included in parentheses.

tures, such that the amide NH groups are oriented towards both the cluster and ligand. NH-S hydrogen bonds (Adman *et al.*, 1975) between main-chain amide groups and sulfur atoms of the cluster and thiol groups of cysteine ligands are present in all available Fe-protein structures, although the detailed interactions vary considerably between structures (Figure 5, Table 3). In general, NH-S bonds reflect two types of structural features: (1) interaction between the S^{γ} ligand of residue i and the NH groups of either residue $i+2$ or $i+3$ in the helical structure; and (2) interactions between an inorganic sulfur atom of the cluster and the NH group positioned at either $i+1$ or $i+2$, where residue i is a cluster ligand. For the purposes of this analysis, an NH-S hydrogen bond is identified if the nitrogen and sulfur atoms are separated by less than 3.5 Å, and if the N-H-S angle (using generated hydrogen positions) is at least 120°. If the N-S distance is between 3.5 and 3.7 Å or if the N-H-S angle is less than 120°, then the NH-S bond is termed ''incipient''.

The most extensive arrangement of NH-S interactions occurs in CP2, where a crown of 11 NH-S bonds encircles the 4Fe:4S cluster symmetrically about the 2-fold non-crystallographic axis passing through the cluster center. Five of these 11 interactions satisfy the criteria for an NH-S hydrogen bond, while the remaining six are incipient. In the AV2 structure, the C terminus of a neighboring molecule spirals into hydrogen bonding distance of the cluster (Figure 6), and it is likely that these lattice interactions contribute to differences in the NH-S bond arrangement between the two subunits. There are appreciable differences between

Table 3. NH-S hydrogen bond interactions in the AV2, CP2 and ALF structures

Fe-protein	Donor	Acceptor	N···S distance (Å)	N-H···S angle (deg.)
CP2	A96 N	A94 Sγ	3.62	108.2[a]
	A97 N	A94 Sγ	3.67[a]	152.0
	A131 N	A129 Sγ	3.47	122.0
	A132 N	A129 Sγ	3.55[a]	158.1
	B96 N	B94 Sγ	3.40	147.7
	B131 N	B129 Sγ	3.65	111.3[a]
	B132 N	B129 Sγ	3.62[a]	155.2
	A95 N	S4	3.70[a]	135.4
	A131 N	S2	3.36	126.1
	B95 N	S1	3.44	143.5
	B131 N	S3	3.29	141.4
AV2	A99 N	A97 Sγ	3.70[a]	153.7
	A134 N	A97 Sγ	3.58[a]	171.4
	B99 N	B97 Sγ	3.38	161.8
	B134 N	B132 Sγ	3.65[a]	122.4
	B135 N	B132 Sγ	3.67[a]	153.1
	B98 N	S1	3.43	129.1
	B99 N	S1	3.69[a]	126.5
	B134 N	S3	3.65[a]	122.4
ALF	E100 N	E97 Sγ	3.46	162.5
	E98 N	S4	3.50	108.2[a]
	E99 N	S4	3.14	153.7
	F100 N	F97 Sγ	3.46	166.4
	F98 N	S1	3.48	113.5[a]
	F99 N	S1	3.30	157.7
	G100 N	G97 Sγ	3.47	164.4
	G98 N	S4	3.70[a]	109.0[a]
	G99 N	S4	3.35	152.9
	H100 N	H97 Sγ	3.50	166.1
	H98 N	S1	3.41	108.8[a]
	H99 N	S1	3.13	157.7

[a] Considered to be an incipient hydrogen bond (3.5 Å < N-H···S bond length < 3.7 Å or N-H···S bond angle < 120 °).

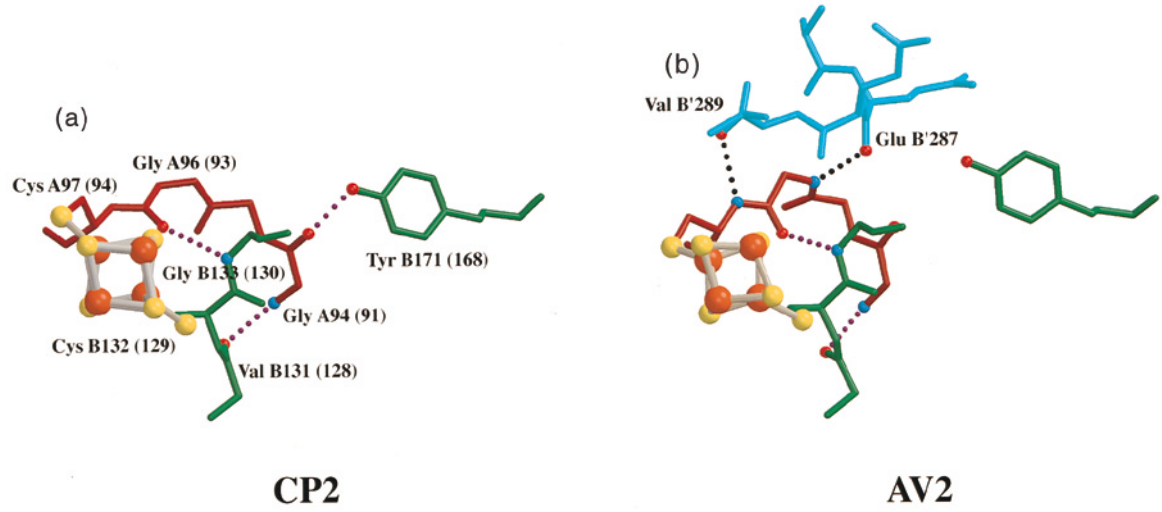

Figure 6. Structural variability in the subunit-subunit interactions near the 4Fe:4S clusters of CP2 and AV2. Subunit A of each protein is shown in red, and subunit B in green. The 4Fe:4S cluster atoms are colored as in Figure 2. Complementary residues on the opposing subunits have been omitted for ease of viewing. (a) Three intersubunit hydrogen bonds are located near the 4Fe:4S cluster in CP2. (b) Crystal packing forces distort this region in AV2, with the loss of several of these intersubunit interactions. The C terminus of a neighboring molecule (shown in cyan) forms two hydrogen bonds with the 4Fe:4S region of subunit A in AV2 at residues Gly96 and Cys97. Despite this conformational alteration near the 4Fe:4S cluster in subunit A, the corresponding region in subunit B is undisturbed.

the CP2 and AV2 structures around the 4Fe:4S cluster, especially in the region around the Cys132 ligand of the A subunit, which participates in lattice contacts in the AV2 crystal. In the ALF structure, extensive interactions between the Fe-protein and MoFe-protein occur in the region surrounding the 4Fe:4S cluster, leading to changes in the pattern of NH-S hydrogen bonds. Perhaps as a result of the lattice contacts near the cluster, the A subunit of AV2 shows stronger similarity to the ALF structure than it does to the B subunit of AV2, or to either of the CP2 and 1NIP subunits. The variation in NH-S interactions with changes in the environment of the 4Fe:4S cluster is a reflection of the conformational sensitivity of this region of Fe-protein.

Various properties of Fe-protein, such as the rate of chelation of iron atoms from the 4Fe:4S cluster or the solvent-dependence of various spectroscopic parameters (Hagen *et al.*, 1985; Lindahl *et al.*, 1985; Morgan *et al.*, 1990), have been interpreted as reflecting the solvent accessibility of the 4Fe:4S cluster. Calculations of the accessible surface area (Richards, 1977) of the 4Fe:4S clusters in AV2 and CP2 reveal relatively few differences in solvent exposure. The inorganic sulfur atoms S1 and S3 on the top surface of the 4Fe:4S cluster have 14 Å^2 and 18 Å^2 exposed surface in AV2 and CP2, respectively, while the S^{γ} atoms of the Cys97 ligand have 9 Å^2 and 19 Å^2 exposed surface, respectively, in these two structures. If the Fe-protein is computationally separated from the MoFe-protein in the ADP·AlF_4^- stabilized complex, the accessible surface areas for the inorganic sulfur atoms and the Cys97 thiol groups are 43 Å^2 and 28 Å^2, respectively, in the ALF structure. For comparison, the values obtained from an ''isolated'' AV2 cluster generated by removing all protein atoms except for the liganding cysteine ligands are 28 Å^2 and 58 Å^2 for the accessible surface area of the inorganic sulfur atoms and ligand Cys97 thiol groups, respectively. Corresponding values for the isolated ALF cluster are 50 Å^2 and 66 Å^2 for the inorganic sulfur atoms and ligand Cys97 thiol groups, respectively. Although cluster accessibility is reduced in Fe-protein relative to the isolated cluster, it is important to recognize that the cluster is solvent accessible even in the free Fe-protein structures.

Dimer Interface

Extraction of the 4Fe:4S cluster does not result in immediate loss of Fe-protein dimerization (Anderson & Howard, 1984), presumably due to the extensive intersubunit contacts. The total buried surface areas for each monomer resulting from dimer formation are calculated to be 1788 Å^2 and 985 Å^2 for AV2 and CP2, respectively; if the AV2 C terminus is truncated beyond residue 272 to omit residues interacting with the adjacent subunit, then the buried surface area drops to 957 Å^2. This is less extensive than the ALF dimer interface with a total buried surface area of 2586 Å^2, with the C-terminal residues neglected. In the AV2 dimer, residues with >30 Å^2 of buried surface area include Lys41, Glu92, Pro93, Val95, Ala98, Asp129, Val130, Val131, Cys132, Met156, Tyr159, Lys166 and Lys170. As detailed below, these residues, along with other residues adjacent in the sequence, mediate the subunit-subunit interactions through a series of primarily polar (hydrogen bond and salt-bridge) interactions.

The most extensive set of polar interactions occurs in the immediate vicinity of the 4Fe:4S cluster, and involves residues near the Cys ligands from both subunits (Figure 6). In particular, residues 92 to 97 on one subunit form a loop that abuts two helical regions on the adjacent subunit, one including residues 131 to 133 in a 3_{10} helix that precedes helix α_4, and the other involving residues 163 to 171 of helix α_5. As a result of these interactions, the C^{α} residues of Cys97 and Cys132 on opposing subunits are in relatively close proximity, with an average separation in the AV2 structure of ~5.5 Å, while the corresponding distance between the C^{α} positions for these residues on the same subunit is ~9.5 Å. Hydrogen bond interactions mediated primarily by main-chain groups are present that can stabilize the dimer interface. In the CP2 structure, three pairs of cross-subunit hydrogen bonds form between Gly96 O and Gly133 N, Gly94 N and Val131 O, and Gly94 O and Tyr171 OH (Av2 sequence numbering), reflecting the symmetrical environment of the 4Fe:4S cluster in this structure (Figure 6(a)). Distortions apparently due to lattice interactions between the carboxyl-terminal residues of a neighboring molecule and the 4Fe:4S region of AV2 prohibit formation of cross-subunit hydrogen bonds between Tyr171 OH and Gly94 O, and between GlyB96 O and GlyA133 N (Figure 6(b)). However, an additional hydrogen bond does form between GlyB94 O and LysA170 N^{ζ} in this structure.

Cross-subunit salt-bridges are evident in Fe-protein, but they are poorly conserved between the various structures. For example, Lys170 N^{ζ} forms a salt-bridge with Glu92 O^{ε} atoms from the opposing subunit in the 1NIP structure. This Glu92-Lys170 interaction had been considered significant because the residues are conserved in almost all known Fe-protein sequences. In CP2, the distances between the corresponding residues are too long to confidently assign as salt-bridges (4.2 Å, 4.4 Å), although the side-chains do seem poised to interact. In contrast, the Glu92 side-chain in AV2 seems more likely to interact with Lys166 N^{ζ} of the opposing subunit, although the distances are still slightly long (3.5 Å, 3.8 Å). Residue 166 is a positively charged residue (Arg or Lys) in all known Fe-protein sequences. Another set of salt-bridges observed in the 1NIP structure, involving the cyclic arrangement of salt-bridges between LysA41-AspB129-LysB41-AspA129, was incomplete in the CP2 structure (with only the salt bridge present between the Asp on subunit A and the Lys on subunit B), and is not observed in the AV2 structure. Indeed, side-chain atoms of Asp129 on the two subunits are separated by <5 Å in AV2, which, if these residues are ionized, would represent an energetically unfavorable interaction.

Interactions across the dimer interface occur predominantly between polar residues, although there are some contacts involving apolar residues. Side-chain atoms of Val130 on the two subunits are approaching van der Waals distance (~5 Å), suggesting that they are in contact. Residues Met156 and Tyr159 in helix α_5 approach residues in the opposing P-loop and Switch I regions, but the interactions do not appear to be particularly extensive. In neither AV2 nor CP2 is there substantial interaction between apolar surfaces at the dimer interface.

The Fe-protein C terminus provides an additional degree of intersubunit interaction in Av2, as the C-terminal residues of each monomer wrap around the body of the opposing subunit. Cross-subunit salt-bridges form between Lys224 and Glu277, and Lys233 and Glu287, and hydrogen bonds connect Arg223 of the A subunit with Lys284 O and Val289 OT of the B subunit in the AV2 model. Although the C-terminal residues are not essential for Fe-protein function (Jacobson *et al.*, such interactions may contribute to the overall stabilization of the Av2 dimer. The shorter C terminus of Cp2 does not extend far enough onto the opposing subunit to create substantial interactions.

Nucleotide binding region

Key elements of the nucleotide binding region of Fe-protein include residues in the P-loop, Switch I and Switch II regions. Although bound nucleotide is not present in the AV2 and CP2 structures, these regions are linked by a number of polar interactions, involving both protein residues and water molecules (Figure 7) and are well ordered (Figure 8). These interactions may contribute to the stability of nucleotide free Fe-protein, which contrasts sharply with the behavior of many of the GTPase proteins that are stable only with bound nucleotide.

Residues in the P-loop are positioned near the dimer interface, in the loop between β1 and α1, and are approximately 19 Å from the 4Fe:4S cluster. The P-loop conformations are highly conserved between the AV2, CP2 and 1NIP structures; the average pairwise rms deviation in C^{α} positions is 0.53 Å for the eight residues. A network of well-ordered and conserved water molecules is present in AV2 and CP2 that links together the P-loop residues. The carbonyl oxygen atom of Lys10 and the O^{γ} of Ser16 are bridged by two hydrogen bonded water molecules, Wat13 and Wat23. For reference, Wat23 is located near the position of the β-phosphate group of ADP in the ALF structure. The side-chain O^{γ} atoms of Ser16 and Thr17 hydrogen bond to Wat57, which additionally interacts with the amide N of the latter residue. Inter-residue interactions in this region include hydrogen bonds between the Lys10 side-chain N^{ζ} and the carbonyl oxygen atom of Gly11, and between the Lys15 side-chain N^{ζ} and the carbonyl oxygen atom of Ile7. In addition to these intra P-loop interactions, a conserved water molecule (Wat34) connects the NH group of Lys10 with the carbonyl oxygen atom of residues of Val150 and Ala157 on helix α_5 of the same subunit.

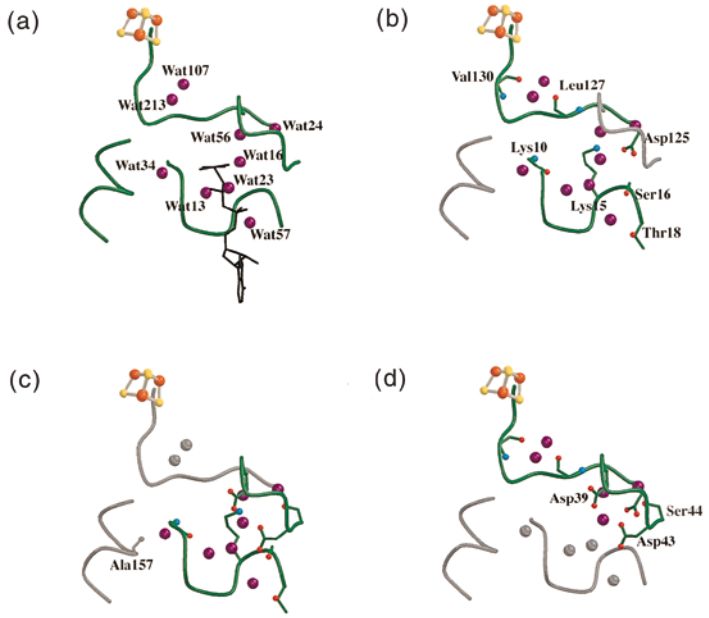

Figure 7. Conserved water molecules in AV2 and CP2, numbered according to the AV2 scheme. (a) An overview of the P-loop, Switch I and Switch II regions (shown in green), and the nine conserved water molecules found in these areas (shown in purple). The $ADP \cdot AlF_4^-$ molecule from the stabilized complex structure has been superimposed and is shown in black for reference. (b) P-loop and Switch II regions and related water molecules are highlighted, and residues involved in potential hydrogen bonds to conserved water molecules are indicated. (c) P-loop and Switch I regions and related water molecules are highlighted. (d) Switch I and Switch II regions and related water molecules are highlighted.

The structure of the Switch II region is also well conserved in the AV2 and CP2 structures, with an average pairwise rms deviation in C^{α} positions of 0.58 Å for residues 125 to 128 forming the DXXG motif. Two water molecules present in both AV2 subunits, Wat107 and Wat213, bridge the carbonyl oxygen atom of Leu127 with the amide and carbonyl oxygen groups of Val130, respectively.

In contrast to the P-loop and Switch II regions, the Switch I region (residues 38 to 44) forms a loop marked by conformational flexibility. The AV2 structure shows a significant change (rms deviation 4.0 Å) compared with the 1NIP model, with a potential salt-bridge connecting Lys41 N^{ζ} and Asp43 $O^{\delta 1}$ in subunit A. The CP2 model favors the 1NIP model; although an alternative conformation similar to that of the AV2 structure can be modeled, albeit with a less satisfactory fit to the electron density. The ALF model corresponds well with the AV2 model in this region. The observed flexibility is likely to be mechanistically relevant; in *ras*, the Switch I region undergoes significant rearrangements between the GTP and GDP-bound states and a similar mechanism may be employed in Fe-protein. In both the AV2 and CP2 structures, a main-chain to main-chain hydrogen bond is formed between Ala42 N and Asp39 O. A conserved water molecule, Wat24, links the peptide

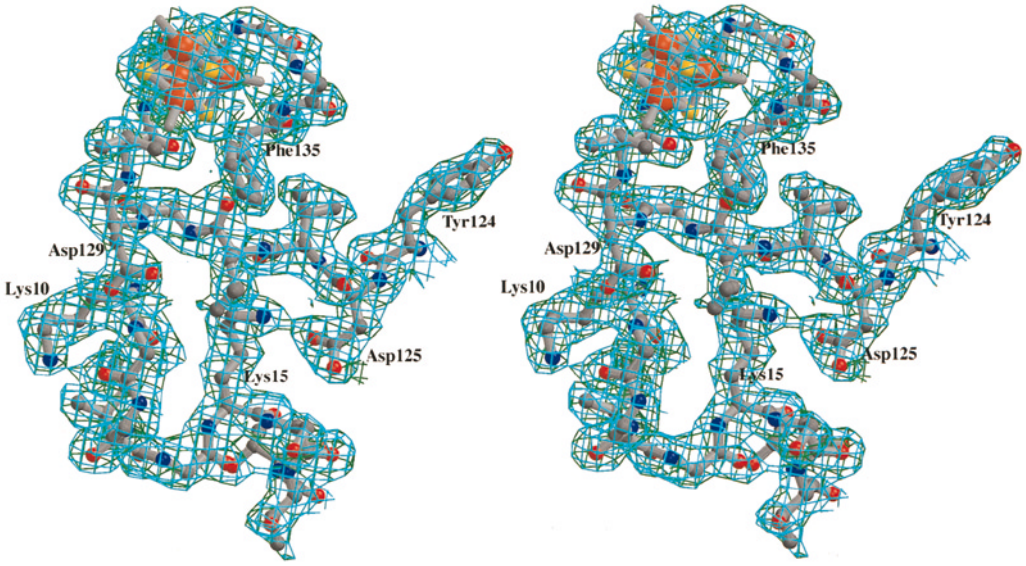

Figure 8. Stereo view of a representative $2F_o - F_c$ electron density map for AV2, calculated at 2.2 Å resolution, in the vicinity of the P-loop, Switch II and 4Fe:4S cluster. The critical salt-bridge between the Lys15 and Asp125 side-chains is illustrated. The map is contoured at 1.0σ.

group of Asp39 and the side-chain Oγ of Ser44; this water molecule additionally interacts with the side-chain group of Glu87.

An extensive network of hydrogen bonding and salt-bridge interactions link the P-loop, Switch I and Switch II regions of AV2 and CP2 (Figure 7). For convenience, these interactions are described between pairs of regions:

P-loop and Switch II regions

One of the most significant interactions is the intrasubunit salt-bridge connecting the side-chains of Lys15 and Asp125, which is one of the very few intrasubunit salt-bridges conserved in all free Fe-protein structures. Lys15 N^{ζ} also interacts with Asp125 through a water-mediated hydrogen bond (Wat31) to the carbonyl oxygen atom of that residue. In AV2, Lys15 N^{ζ} forms an additional hydrogen bond with the carbonyl oxygen atom of Val126. Several sets of protein-protein hydrogen bonds link these regions: Gly9 O to Gly128 N; Gly11 N to the side-chain of Asp129; and Thr19 $O^{\gamma 1}$ to Asp125 $O^{\delta 1}$.

P-loop–Switch I

Ser16 O^{γ} is positioned within hydrogen bonding distance of Asp43 $O^{\delta 2}$ and O. A conserved set of interactions involves Wat56 hydrogen bonding to the side-chains of Asp39 and Asp43, as well as to Wat23 involved in the previously described P-loop hydrogen bonding network. In CP2, Lys15 N^{ζ} forms a salt-bridge with Asp39 $O^{\delta 2}$.

Switch I–Switch II

The primary polar interactions linking these two regions are the hydrogen bonds between Cys38 N and Asp125 O, and between the side-chains of Ser44 and Asp125.

Discussion

A comparison of the structures of ''free'' nitrogenase Fe-protein (i.e. Fe-protein without fully occupied nucleotide binding and in the absence of MoFe-protein) from *A. vinelandii* and *C. pasteurianum* reveals similarities in the polypeptide conformations for both monomers and in the relationship between the two monomers in each dimer. The general insensitivity of the Fe-protein structure to the sequence differences between Av2 and Cp2, and to differences in crystal lattice contacts, indicates that these structures are representative of the uncomplexed, nucleotide-free form of Fe-protein. Not surprisingly, however, differences are evident between these Fe-protein structures, with the closest superpositions observed for the central β-sheet, followed by the α-helices, and the greatest structural diversity observed in the connecting loops between secondary structure elements. Although the residues that exhibit the greatest variability, both between subunits in the same dimer and between the AV2 and CP2 dimers, are distributed throughout the protein sequence, in terms of the three-dimensional structure they are concentrated in two major regions: at the surface of the Fe-protein in the vicinity of the 4Fe:4S cluster and near the C terminus (Figure 4).

The functional significance of these observations may be appreciated within the context of the ADP-AlF_4^- stabilized complex of Av1 and Av2 (Schindelin *et al.*, 1997). This structure demonstrated that substantial changes in Av2 occurred relative to the uncomplexed Av2 structure at both the tertiary and quaternary structural levels. The Av2 subunits in the complex are each rotated towards the dimer interface by ~13° about an axis perpendicular to the dimer 2-fold axis, resulting in a more extensive interface interaction. Significant tertiary structure changes further occurred in five regions associated with the nucleotide binding site, the 4Fe:4S cluster, and the MoFe-protein-binding interface: (1) the P-loop region; (2) residues 51 to 75 containing part of the MoFe-protein-binding interface; (3) residues 88 to 118 including cluster ligand Cys97 and residues involved in binding to the MoFe-protein; (4) residues 127 to 143 including cluster ligand Cys132, the Switch II region and other residues involved in binding to the MoFe-protein; and (5) residues 151 to 176, including helix α_5 found at the Fe-protein dimer interface and the MoFe-protein-binding interface. The C-terminal residues were not visible in the ALF electron density map and are likely to be conformationally disordered.

Comparison of free and complexed Fe-protein structures demonstrates that the regions of greatest structural variability in the free Fe-protein correspond to a subset of the residues that undergoes the most significant changes upon complex formation. Residues at the MoFe-protein-binding surface exhibit significant variability in Fe-protein structures, as do the C-terminal residues and adjacent regions, such as residue 186 in the adenine ring binding site (Figure 4). The flexibility seen in the surface loops that form the MoFe-protein binding region contrast with other loop regions that show little variation between the different structures of free Fe-protein. This suggests that Fe-protein residues forming the MoFe-protein-binding interface have the intrinsic ability to adopt the appropriate conformations needed to interact with the MoFe-protein during the process of substrate reduction.

Notably absent from the set of residues that exhibit conformational flexibility in the free Fe-protein are the residues that interact with the nucleotide phosphate groups, especially the P-loop and Switch II regions. The high-resolution AV2 and CP2 structures demonstrate that these regions are interconnected by an extensive series of conserved salt-bridges and hydrogen bonds, including those mediated by a set of conserved water molecules. The existence of this network suggests that conformational changes in the nucleotide-binding regions

are highly cooperative. Furthermore, it is likely that the changes observed in these regions in the ADP·AlF_4^- stabilized Av1-Av2 complex are not an intrinsic property of Fe-protein, but rather require the actual binding of nucleotide and/or MoFe-protein to trigger the required structural transitions.

Although Av2 and Cp2 represent two of the more diverse pairs of nitrogenase Fe-proteins, they still share 69% sequence identity. The high overall level of sequence conservation undoubtedly reflects the multiple functions of Fe-proteins, including interactions with a variety of protein partners. In an attempt to identify which residues and regions of Fe-protein may be functionally significant in nitrogen-fixing organisms, the amino acid sequences of 63 Fe-proteins have been aligned using the CLUSTAL-W program (Thompson *et al.*, 1994). From this alignment, a set of 85 residues was identified with the identical amino acid present in at least 61 of 63 sequences. However, if four sequences are removed from this list, a set of 140 residues is found with the identical amino acid present in at least 57 of 59 sequences (Figure 1). These 140 residues define an essentially invariant set of core Fe-protein residues that are indicated by the yellow shading in Figure 1. A spatial representation of the pattern of residue conservation along the Fe-protein backbone is shown in Figure 4(b). Not surprisingly, many of these core residues reside in the cluster, dimer interface and nucleotide-binding regions. Site-directed mutagenesis studies are addressing the roles of some of these residues, particularly those mediating the nucleotide binding properties of Fe-protein (Howard *et al.*, 1989; Gavini & Burgess, 1992; Seefeldt *et al.*, 1992; Wolle *et al.*, 1992; Seefledt & Mortenson, 1993; Lanzilotta *et al.*, 1995; Ryle & Seefeldt, 1996). However, other invariant residues, such as Pro256 and Pro258, are located in unanticipated sites. It is possible that some of these conserved residues may function in one or more of Fe-protein's additional roles, including FeMo-cofactor assembly and insertion. Future mutational studies may be required to establish the roles that these particular residues play in Fe-protein function.

Lists of conserved residues are useful in assessing functional implications of new sequences that exhibit strong similarity to the nitrogenase Fe-protein. For example, the four ''outlier'' Fe-protein sequences identified in the preceding analysis share strong sequence conservation in the cluster and nucleotide-binding regions, but they do exhibit notable differences, particularly in residues forming the dimer interface and the adenine ring binding site. Three of these sequences were identified in a previous survey (Normand & Bosquet, 1989) as forming the most distant group of Fe-protein sequences, while the fourth sequences were recently identified in the *Methanococcus jannaschii* genome (Bult *et al.*, 1996). Despite the high degree of sequence identity with Av2 and Cp2 (Figure 1), it is not clear that the *M. jannaschii* protein functions in nitrogen fixation, since no sequences were recognized in the genome that would code for either the MoFe-protein or most of the proteins required for assembly of the FeMo-cofactor (Haselkorn & Buikema, 1996). Two of the ''missing'' residues in the *M. jannaschii* Fe-protein homolog are Glu154 and Arg213 that form an intersubunit salt-bridge in the nitrogenase complex with the arginine group stacked over the adenine ring of the nucleotide (Schindelin *et al.*, 1997); a substitution of the latter residue has been identified in a mutant Fe-protein from a *Klebsiella pneumoniae* strain that is unable to fix nitrogen (Chang *et al.*, 1988). Another example is provided by the products of the plant *frxC* gene and homologous *bchL* and *bchX* bacterial genes catalyzing protochlorophyllide reduction that share significant identity (30 to 37%) with *nifH*. Again, while sequence comparisons show invariant residues in regions associated with cluster and nucleotide-binding, there are also substantial differences (Figure 1). While these proteins quite likely couple hydrolysis of nucleoside triphosphates to redox reactions from an iron-sulfur cluster, it is probable that the global functions of these Fe-protein homologs have varied, perhaps with either a change in nucleotide specificity, the incorporation of new functions, or the loss of others. The 140 residue conserved core derived from *nifH* amino acid sequence alignments may be useful as a benchmark to authenticate *nifH*-like sequences as true Fe-protein sequences.

The present view of Fe-protein reveals a molecule of multiple functions and conformations. Fe-protein is involved in nucleotide hydrolysis and electron transfer, and in FeMo cofactor biosynthesis and assembly. Additional functions and conformational changes undoubtedly await discovery, as details emerge about the interactions between Fe-protein and its various protein partners. The available structures of Fe-protein, alone and in complex with the MoFe-protein, reveal a variety of different Fe-protein conformations. Comparisons between the structures of the ADP·AlF_4^- stabilized nitrogenase complex and the recently solved (J. Peters *et al.*, unpublished results) complex between Av1 and the Leu127 deletion mutant of Av2 (Ryle & Seefeldt, 1996) show discernible differences even within the complexed Fe-proteins. Similarly, the structures of free Fe-proteins from *A. vinelandii* and *C. pasteurianum* described here are not identical. Taken together, the various structures suggest that a continuum of Fe-protein structure exists, reinforcing our perception of nitrogenase as a dynamic enzyme system.

Methods

Av2 was anaerobically purified from *A. vinelandii* OP cells (ATCC 13705) (grown on modified Burk's N_2-free medium) with a Pharmacia FPLC using a modification of the method of Burgess *et al.* (1980). All subsequent manipulations were carried out in an anaerobic chamber (Coy Products, Ann Arbor, MI). Crystals were grown by the capillary microbatch method (Georgiadis *et al.*, 1992)

with 30 μl of precipitant solution containing 25 to 30% PEG 4000, 140 to 200 mM Na_2MoO_4, 50 mM Tris-Cl (pH 8.0), and 15 μl of protein solution containing Av2 in 20% glycerol, 450 mM NaCl, 2 mM $Na_2S_2O_4$ and 50 mM Tris-HCl (pH 8.0) in 1.5 mm melting point capillaries. The crystals reached a maximum size of 0.2 mm × 0.2 mm × 0.5 mm within several months and belonged to space group $P2_12_12_1$ with cell constants $a = 73.12$ Å, $b = 90.53$ Å and $c = 91.21$ Å. This represents a new crystal form compared to the initially solved $P2_1$ form (Georgiadis *et al.*, 1992). Data were collected from a single crystal at −160°C at beamline 7-1 of the Stanford Synchrotron Research Laboratory, and processed and scaled using the programs DENZO/SCALEPACK (Otwinowski & Minor, 1997; Table 4).

The Av2 structure was solved by molecular replacement with AMORE (Navaza, 1994), using the 1NIP structure as the search model. Solvent flattening and 2-fold non-crystallographic averaging with SOLOMON (Abrahams & Leslie, 1996) were employed to improve map quality. Iterative model building and refinement were performed using the programs O (Jones *et al.*, 1991) and X-PLOR (Brünger, 1992b), but the *R*-factor and R_{free} could not be reduced below 28% and 34%, respectively, to 2.8 Å resolution. At this point, the molecular replacement was repeated using superimposed 1NIP and CP2 models. Solvent flattening and 2-fold averaging with SOLOMON at 2.2 Å resolution produced superior maps that facilitated rebuilding. Release of the 2-fold non-crystallographic symmetry restraints, following careful monitoring of R_{free}, was required to complete the refinement. The final orthorhombic Av2 model, with *R*-factor 22.3% and R_{free} 28.9% using all data to 2.2 Å resolution, consists of 575 of 578 possible protein residues, one 4Fe:4S cluster, and 372 water molecules. This model obeys reasonable geometry, with rms deviations from ideality for bond lengths and angles of 0.019 Å and 1.93°, respectively, and 85.9% of all protein residues lying in the most favored Ramachandran regions as determined with the program PROCHECK (Laskowski *et al.*, 1993). The average temperature factor is 31.7 Å^2 for all non-hydrogen atoms. Atoms for which no interpretable ($2|F_{obs}| - |F_{calc}|$) electron density was observed were conservatively modeled and assigned occupancies of 0.0. The coordinates have been deposited in the Brookhaven Protein Data Bank (PDB entry 2NIP). A view of the electron density around the P-loop, Switch II and 4Fe:4S cluster regions is illustrated by Figure 8.

Table 4. Summary of crystallographic statistics

Data sets	Av2	Cp2
Number of crystals	1	1
Space group	$P2_12_12_1$	$P2_1$
a, *b*, *c* (Å)	73.12, 90.53, 91.21	67.60, 75.87, 53.55 (β=114.17°)
Wavelength (Å)	1.08	1.5418
Resolution (Å)	2.13	1.93
Completeness (%)	92.9 (94.6)	98.8 (91.0)
Reflections	260,578	250,737
Unique reflections	29,060	36,808
$I/\sigma(I)$	6.9 (2.4)	5.8
R_{sym} (%)[a]	8.6 (30.9)	7.3
Refinement statistics		
Resolution range (Å)	50-2.2	5.0-1.93
R_{cryst} (%)[b]	22.3	20.9
R_{free} (%)[b]	29.0	29.7
rmsΔ bond lengths (Å)	0.017	0.018
rmsΔ bond angles (deg.)	1.93	2.04
⟨*B*⟩, non-hydrogen atoms (Å^2)	31.7	36.3

Numbers in parentheses indicate values for the highest resolution bin.

[a] $R_{sym} = \Sigma|I_i - \langle I\rangle|/\Sigma(\langle I\rangle)$, where *i* is the *i*th measurement and $\langle I\rangle$ is the weighted mean of *I*.

[b] $R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$. R_{free} is the same as R_{cryst}, but calculated for *n*% of the data omitted from refinement, where $n = 6$ for Av2 and $n = 10$ for Cp2.

C. pasteurianum W5 cells (ATCC 6013) were grown anaerobically using a modification of the method of R. Bare (Exxon, personal communication). Cp2 was purified using a modification of the methods of Nelson *et al.* (1983), Kornuc (1988), and R. Bare (personal communication). Crystals were grown under anaerobic conditions using the capillary microbatch method, with 40 μl of precipitant solution containing 5 to 15% glycerol, 18.75 to 22.5% PEG 4000, 230 to 270 mM $CaCl_2$, 50 mM Hepes (pH 7.5), layered with 10 μl of protein solution containing Cp2 in 20% glycerol, 2 mM $Na_2S_2O_4$, 50 mM Hepes, (pH 7.5). Crystals emerged within several days to weeks, and belonged to space group $P2_1$ with cell constants $a = 67.60$ Å, $b = 75.87$ Å, $c = 53.55$ Å and $\beta = 114.17°$. Data were collected from a single crystal maintained at −160°C on an in-house Rigaku R-AXIS IIC image plate detector and processed using software supplied by Rigaku. The data set contained 36,808 unique reflections and was 98.8% complete to 1.93 Å, with R_{merge} 7.25%.

The Cp2 structure was solved by molecular replacement using the 1NIP structure as the search model (Crowther, 1972). The 4Fe:4S cluster position corresponded with that determined from the native anomalous difference Patterson map. Iterative cycles of refinement with X-PLOR and model building in TOM/FRODO (Jones, 1985) yielded the final model, with *R*-factor and R_{free} of 22.9% and 29.7%, respectively, using all data from 5.0 to 2.0 Å resolution. The final Cp2 model includes 538 of 544 possible amino acid residues, one 4Fe:4S cluster and 118 water molecules. The model obeys standard geometry, with rms deviations from ideality of 0.018 Å and 2.04° for bond lengths and angles, respectively. The average temperature factor is 35.6 Å^2 for all non-hydrogen atoms, and 87.0% of all protein residues lie within the most favored Ramachandran regions as defined by PROCHECK (Laskowski *et al.*, 1993). As in the case of the Av2 model, atoms for which no interpretable ($2|F_{obs}| - |F_{calc}|$) electron density was observed were conservatively modeled and assigned occupancies of 0.0. The coordinates have been deposited in the Brookhaven Protein Data Bank (PDB entry 1CP2). Table 4 provides a summary of data collection and model statistics for Av2 and Cp2.

The amino acid sequences employed in the Fe-protein alignment were retrieved from the National Biomedical Research Foundation Protein Identification Resource (PIR), GenBank Genetics Sequence Data Bank (GB), and SWISS_PROT Protein sequence Database (SW) *via* the National Center for Biotechnology Information. The accession numbers for these sequences are as follows: *Alcaligenes faecalis* (SW Q44044); *Anabaena azollae* (GB L34879); *Anabaena variabilis* (GB U89346); *Anabaena* sp. L31 (SW P33178); *Anabaena* PCC7120 (SW P00457); *Azospirillum brasilense*-1 (SW P17303); *Azospirillum brasilense*-2A (PIR S15745); *Azospirillum brasilense*-2B (PIR S15747); *Azotobacter chroococcum*-1 (SW P26248); *Azotobacter chroococcum*-2 (SW P06118); *Azotobacter vinelandii*-1 (SW P00459); *Azotobacter vinelandii*-2A (SW P15335); *Azotobacter vinelandii*-2B (GB M32371); *Azotobacter vinelandii*-3

(SW P16269); *Clostridium pasteurianum*-1 (SW P00456); *Clostridium pasteurianum*-2 (SW P09552); *Clostridium pasteurianum*-3 (SW P09553); *Clostridium pasteurianum*-4 (SW P22548); *Clostridium pasteurianum*-5 (SW P09554); *Clostridium pasteurianum*-6 (SW P09555); *Cyanothece* ATCC51142 (GB AF003336); *Desulfovibrio gigas* (PIR A61030); *Fischerella* sp.-1 (GB U49514); *Fischerella* sp.-2 (GB U49515); *Frankia alni*-1 (SW P08925); *Frankia* sp.-3 (PIR S26190); *Frankia* sp.-4 (GB U53362); *Frankia* sp. FaC1 (SW P46034); *Frankia alni*-2 (SW P08925); *Herbaspirillum seropedicae* (GB Z54207); *Klebsiella pneumoniae*-1 (GB J01740); *Klebsiella pneumoniae*-2 (SW P00458); *Methanobacterium ivanovii*-1 (SW P08624); *Methanobacterium ivanovii*-2 (SW P51602); *Methanococcus maripaludis* (GB U75887); *Methanobacterium thermoautotrophicum*-1 (SW Q50785); *Methanococcus jannaschii* (SW Q58289); *Methanococcus thermolithotrophicus*-1 (SW P25767); *Methanococcus thermolithotrophicus*-2 (SW P08625); *Methanococcus thermolithotrophicus*-3 (SW P25767); *Methanococcus voltae* (SW P06119); *Methanosarcina barkeri*-1 (SW P54799); *Methanosarcina barkeri*-2 (SW P54800); *Nostoc* sp. 6720 (PIR S50135); *Nostoc commune* (SW P26250); *Parasponia rhizobium* (SW P00463); *Plectonema boryanum* (SW Q00240); *Rhizobium* sp. ANU240 (PIR JS0238); *Azorhizobium caulinodans* ORS571-A (SW P26252); *Azorhizobium caulinodans* ORS571-B (SW P26251); *Bradyrhizobium japonicum* (SW P06117); *Rhizobium meliloti* (SW P00460); *Rhizobium leguminosarum* bv. phaseoli (SW P00462); *Rhizobium leguminosarum* bv. trifolii (SW P00461); *Rhodobacter capsulatus*-1A (GB M15270); *Rhodobacter capsulatus*-1B (SW P08718); *Rhodobacter capsulatus*-1C (PIR JN0888); *Rhodobacter capsulatus*-2 (SW Q07942); *Rhodobacter capsulatus* BchL (SW P26237); *Rhodospirillum rubrum* (SW P22921); *Synechococcus* sp. (GB U22146); *Thiobacillus ferrooxidans* (SW P06661); and *Trichodesmium thiebautii* (GB U23507). The sequences were aligned using the program CLUSTAL-W (Thompson *et al.*, 1994), revealing a consensus core of 140 residues among 59 of the sequences. The four outliers were the *M. ivanovii*-1, *M. jannaschii*, *M. thermolithotrophicus*-2, and *M. voltae* nifH sequences.

Unless otherwise noted, the residue numbers for the Av2 sequence (Hausinger & Howard, 1982) are used; these are related to the Cp2 sequence (Tanaka *et al.*, 1977) through the following relationships; for Av2 residues 2 to 65, subtract 1 to get the equivalent Cp2 residue number; for Av2 residues 68 on, subtract 3 to get the equivalent Cp2 residue number. Molecular figures were prepared with the MOLSCRIPT (Kraulis, 1991), BOBSCRIPT (Esnouf, 1997) and Raster3D (Merritt & Murphy, 1994) programs.

Acknowledgments

This work was supported by USPHS grant GM45162. The advice and contributions of A. J. Chirino, M. M. Georgiadis, B. T. Hsu, C. Kisker, C. L. Kielkopf, H. Komiya, J. J. Kornuc, J. W. Peters, H. Schindelin and M. H. B. Stowell to this project are gratefully appreciated, as is the support of D. S. Eisenberg and the UCLA structural biology group during the early phases of this work. The advice of R. Bare (Exxon) concerning the *C. pasteurianum* preparations was invaluable. The rotation camera facility at SSRL is supported by the DOE Office of Basic Energy Sciences and the NIH Biomedical Research Technology Program, Division of Research Resources.

References

Abrahams, J. P. & Leslie, A. (1996). Methods used in the structure determination of bovine mitochondrial F_1 ATPase. *Acta Crystallog. sect. D,* **52**, 30–42.

Adman, E. T., Watenpaugh, K. D. & Jensen, L. H. (1975). NH–S hydrogen bonds in *Peptococcus aerogenes* ferredoxin, *Clostridium pasteurianum* rubredoxin and *Chromatium* high potential iron protein. *Proc. Natl Acad. Sci. USA,* **72**, 4854–4858.

Alberty, R. A. (1994). Thermodynamics of nitrogenase reactions. *J. Biol. Chem.* **269**, 7099–7102.

Allen, R. M., Chatterjee, R., Madden, M., Ludden, P. W. & Shah, V. K. (1994). Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Crit. Rev. Biotechnol.* **14**, 225–249.

Anderson, G. L. & Howard, J. B. (1984). Reactions with the oxidized iron protein of *Azotobacter vinelandii* nitrogenase: formation of a 2Fe center. *Biochemistry,* **23**, 2118–2122.

Angove, H. C., Yoo, S. J., Burgess, B. K. & Münck, E. (1997). Mössbauer and EPR evidence for an all-ferrous Fe_4S_4 cluster with $S = 4$ in the Fe protein of nitrogenase. *J. Am. Chem. Soc.* **119**, 8730–8731.

Bolin, J. T., Campobasso, N., Muchmore, S. W., Morgan, T. V. & Mortenson, L. E. (1993). The structure and environment of the metal clusters in the nitrogenase MoFe protein from *Clostridium pasteurianum*. In *Molybdenum Enzymes, Cofactors and Model Systems. ACS Symposium Series No. 535* (Stiefel, E. I., Coucouvanis, D. & Newton, W. E., eds), vol. 535, pp. 186–195, American Chemical Society, Washington, DC.

Brändén, C.-I. (1980). Relation between structure and function of α/β proteins. *Quart. Rev. Biophys.* **13**, 317–338.

Brünger, A. T. (1992a). Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature,* **355**, 472–475.

Brünger, A. T. (1992b). *X-PLOR version 3. 1–A system for X-ray crystallography and NMR*, Yale University Press, New Haven and London.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J.-F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. M., Weidman, J. F., Fuhrmann, J. L., Nguyen, D., Utterback, T. R., Kelley, J. M., Peterson, J. D., Sadow, P. W., Hanna, M. C., Cotton, M. D., Roberts, K. M., Hurst, M. A., Kaine, B. P., Borodovsky, M., Klenk, H.-P., Fraser, C. M., Smith, H. O., Woese, C. R. & Venter, J. C. (1996). Complete genome sequence of the methanogenic archaeon *Methanococcus jannaschii*. *Science,* **273**, 1058–1073.

Burgess, B. K. & Lowe, D. J. (1996). Mechanism of molybdenum nitrogenase. *Chem. Rev.* **96**, 2983–3011.

Burgess, B. K., Jacobs, D. B. & Steifel, E. I. (1980). Large scale purification of high activity *Azotobacter vinelandii* nitrogenase. *Biochim. Biophys. Acta,* **614**, 196–209.

Chang, C., Davis, L. C., Rider, M. & Takemoto, D. J. (1988). Characterization of *nifH* mutations of *Klebsiella pneumoniae*. *J. Bacteriol.* **170**, 4015–4022.

Chothia, C. & Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. *EMBO J.* **5**, 823–826.

Crowther, R. A. (1972). The fast rotation function. In *Molecular Replacement Method* (Rossmann, M. G., ed.), pp. 173–178, Gordon and Breach, New York.

Daesch, G. & Mortenson, L. E. (1967). Sucrose catabolism in *Clostridium pasteurianum* and its relation to N_2 fixation. *J. Bacteriol.* **96**, 346–351.

Eady, R. R. (1996). Structure-function relationships of alternative nitrogenases. *Chem. Rev.* **96**, 3013–3030.

Emerich, D. W., Ljones, T. & Burris, R. H. (1978). Nitrogenase: properties of the catalytically inactive complex between the *Azotobacter vinelandii* MoFe protein and the *Clostridium pasteurianum* Fe protein. *Biochim. Biophys. Acta,* **527**, 359–369.

Emerich, D. W., Hageman, R. V. & Burris, R. H. (1981). Interactions of dinitrogenase and dinitrogenase reductase. *Advan. Enzymol.* **52**, 1–22.

Esnouf, R. M. (1997). An extensively modified version of MOLSCRIPT that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 133–138.

Freymann, D. M., Keenan, R. J., Stroud, R. M. & Walter, P. (1997). Structure of the conserved GTPase domain of the signal recognition particle. *Nature,* **385**, 361–364.

Frishman, D. & Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins: Struct. Funct. Genet.* **23**, 566–579.

Gavini, N. & Burgess, B. K. (1992). FeMo cofactor synthesis by a nifH mutant with altered MgATP reactivity. *J. Biol. Chem.* **267**, 21179–21186.

Georgiadis, M. M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. J. & Rees, D. C. (1992). Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. *Science,* **257**, 1653–1659.

Hageman, R. V. & Burris, R. H. (1978). Nitrogenase and nitrogenase reductase associate and dissociate with each catalytic cycle. *Proc. Natl Acad. Sci. USA,* **75**, 2699–2702.

Hagen, W. R., Eady, R. R., Dunham, W. R. & Haaker, H. (1985). A novel $S = 3/2$ EPR signal associated with native Fe-protein of nitrogenase. *FEBS Letters,* **189**, 250–254.

Haselkorn, R. & Buikema, W. J. (1996). *Methanococcus* genome. *Science,* **274**, 901–902.

Hausinger, R. P. & Howard, J. B. (1982). The amino acid sequence of the nitrogenase iron protein from *Azotobacter vinelandii*. *J. Biol. Chem.* **257**, 2483–2490.

Holm, L. & Sander, C. (1998). Touring protein fold space with Dali/FSSP. *Nucl. Acids Res.* **26**, 316–319.

Howard, J. B. (1993). Protein component complex formation and adenosine triphosphate hydrolysis in nitrogenase. In *Molybdenum Enzymes, Cofactors and Model Systems. ACS Symposium Series No. 535* (Stiefel, E. I., Coucouvanis, D. & Newton, W. E., eds), pp. 271–289, American Chemical Society, Washington, D.C.

Howard, J. B. & Rees, D. C. (1994). Nitrogenase: a nucleotide-dependent molecular switch. *Annu. Rev. Biochem.* **63**, 235–264.

Howard, J. B. & Rees, D. C. (1996). Structural basis of biological nitrogen fixation. *Chem. Rev.* **96**, 2965–2982.

Howard, J. B., Davis, R., Moldenhauer, B., Cash, V. L. & Dean, D. (1989). Fe-S cluster ligands are the only cysteines required for nitrogenase Fe-protein activities. *J. Biol. Chem.* **264**, 11270–11274.

Huang, W., Lindqvist, Y., Schneider, G., Givson, K. J., Flint, D. & Lorimer, G. (1994). Crystal structure of an ATP-dependent carboxylase, dethiobiotin synthetase, at 1.65 Å resolution. *Structure,* **2**, 407–414.

Jacobson, M. R., Cantwell, J. S. & Dean, D. R. (1990). A hybrid *Azotobacter vinelandii-Clostridium pasteurianum* nitrogenase iron protein that has *in vivo* and *in vitro* catalytic activity. *J. Biol. Chem.* **265**, 19429–19433.

Jones, T. A. (1985). Interactive computer graphics: FRODO. *Methods Enzymol.* **115**, 157–171.

Jones, T. A., Zhou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110–119.

Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallog. sect. A,* **32**, 922–923.

Kim, J. & Rees, D. C. (1992a). Crystallographic structure and functional implications of the nitrogenase molybdenum-iron protein from *Azotobacter vinelandii*. *Nature,* **360**, 553–560.

Kim, J. & Rees, D. C. (1992b). Structural models for the metal centers in the nitrogenase molybdenum-iron protein. *Science,* **257**, 1677–1682.

Kim, J., Woo, D. & Rees, D. C. (1993). X-ray crystal structure of the nitrogenase molybdenum-iron protein from *Clostridium pasteurianum* at 3.0 Å resolution. *Biochemistry,* **32**, 7104–7115.

Kleywegt, G. J. & Jones, T. A. (1994). A super position. *ESF/CCP4 Newsletter (November, 1994),* **31**, 9–14.

Kleywegt, G. & Jones, T. A. (1995). Where freedom is given, liberties are taken. *Structure,* **3**, 535–540.

Kornuc, J. J. (1988). X-ray crystallographic studies of the nitrogenase iron protein. PhD thesis, University of California, Los Angeles.

Kraulis, P. J. (1991). MOLSCRIPT–a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.

Lanzilotta, W. N., Ryle, M. J. & Seefeldt, L. C. (1995). Nucleotide hydrolysis and protein conformational changes in *Azotobacter vinelandii* nitrogenase iron protein: defining the function of aspartate 129. *Biochemistry,* **34**, 10713–10723.

Laskowski, R. A., McArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK–a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

Lindahl, P. A., Day, E. P., Kent, T. A., Orme-Johnson, W. H. & Münck, E. (1985). Mösssbauer, EPR and magnetization studies of the *Azotobacter vinelandii* Fe protein. Evidence for a $[4Fe\text{-}4S]^{1+}$ cluster with spin $S = 3/2$. *J. Biol. Chem.* **260**, 11160–11173.

Lowe, D. J. & Thorneley, R. N. F. (1983). Nitrogenase of *Klebsiella pneumoniae*. Kinetics of the dissociation of oxidized iron protein from molybdenum-iron protein: identification of the rate-limiting step for substrate reduction. *Biochem. J.* **215**, 393–405.

Lowe, D. J. & Thorneley, R. N. F. (1984). The mechanism of *Klebsiella pneumoniae* nitrogenase action: the determination of rate constants required for the simulation of the kinetics of N_2 reduction and H_2 evolution. *Biochem. J.* **224**, 895–901.

Merritt, E. A. & Murphy, M. E. P. (1994). Raster3D Version 2.0–a program for photorealistic molecular graphics. *Acta Crystallog. sect. D,* **50**, 869–873.

Milner-White, E. J., Coggins, J. R. & Anton, I. A. (1991). Evidence for an ancestral core structure in nucleotide-binding proteins with the type-A motif. *J. Mol. Biol.* **221**, 751–754.

Montoya, G., Svensson, C., Luirink, J. & Sinning, I. (1997). Crystal structure of the NG domain from the

signal-recognition particle receptor FtsY. *Nature,* **385**, 365–368.
Morgan, T. V., McCracken, J., Orme-Johnson, W. H., Mims, W. B., Mortenson, L. E. & Peisach, J. (1990). Pulsed electron paramagnetic resonance studies of the interaction of Mg-ATP and D_2O with the iron protein of nitrogenase. *Biochemistry,* **29**, 3077–3082.
Muchmore, S. W. (1995). Structure determination of the molybdenum iron protein of nitrogenase at 2.2 Å resolution. PhD dissertation, Purdue University.
Muchmore, S. W., Jack, R. F. & Dean, D. R. (1996). Developments in the analysis of nitrogenase FeMo-cofactor biosynthesis. In *Mechanisms of Metallocenter Assembly* (Hausinger, R. P., Eichhorn, G. L. & Marzilli, L. G., eds), pp. 111–132, VCH Publishers, Inc., New York.
Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536–540.
Navaza, J. (1994). AMORE – an automated package for molecular replacement. *Acta Crystallog. sect. A,* **50**, 157–163.
Nelson, M. J., Levy, M. A. & Orme-Johnson, W. H. (1983). Metal and sulfur composition of iron molybdenum cofactor of nitrogenase. *Proc. Natl Acad. Sci. USA,* **80**, 147–150.
Normand, P. & Bousquet, J. (1989). Phylogeny of nitrogenase sequences in *Frankia* and other nitrogen-fixing microorganisms. *J. Mol. Evol.* **29**, 436–447.
Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
Peters, J. W., Fisher, K. & Dean, D. R. (1994). Identification of a nitrogenase protein-protein interaction site defined by residues 59 through 67 within the *Azotobacter vinelandii* Fe protein. *J. Biol. Chem.* **269**, 28076–28083.
Peters, J. W., Fisher, K. & Dean, D. R. (1995). Nitrogenase structure and function: a biochemical-genetic perspective. *Annu. Rev. Microbiol.* **49**, 335–366.
Peters, J. W., Stowell, M. H. B., Soltis, S. M., Finnegan, M. G., Johnson, M. K. & Rees, D. C. (1997). Redox-dependent structural changes in the nitrogenase P-cluster. *Biochemistry,* **36**, 1181–1187.
Poland, B. W., Silva, M. M., Serra, M. A., Cho, Y., Kim, K. H., Harris, E. M. S. & Honzatko, R. B. (1993). Crystal structure of adenylosuccinate synthetase from *Escherichia coli*. *J. Biol. Chem.* **268**, 25334–25342.
Richards, F. M. (1977). Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–176.
Ryle, M. & Seefeldt, L. C. (1996). Elucidation of a MgATP signal transduction pathway in the nitrogenase iron protein: formation of a conformation resembling the MgATP-bound state by protein engineering. *Biochemistry,* **35**, 4766–4775.
Schindelin, H., Kisker, C., Schlessman, J. L., Howard, J. B. & Rees, D. C. (1997). Structure of ADP-AlF_4^- stabilized nitrogenase complex and its implications for signal transduction. *Nature,* **387**, 370–376.
Schlesinger, W. H. (1991). *Biogeochemistry: An Analysis of Global Change*, Academic Press, San Diego.
Schulz, G. E. (1992). Binding of nucleotides by proteins. *Curr. Opin. Struct. Biol.* **2**, 61–67.
Seefeldt, L. C. & Dean, D. R. (1997). Role of nucleotides in nitrogenase catalysis. *Acc. Chem. Res.* **30**, 260–266.
Seefeldt, L. C. & Mortenson, L. E. (1993). Increasing nitrogenase catalytic efficiency for MgATP by changing serine 16 of its Fe protein to threonine: use of Mn^{2+} to show interaction of serine 16 with Mg^{2+}. *Protein Sci.* **2**, 93–102.
Seefeldt, L. C., Morgan, T. V., Dean, D. R. & Mortenson, L. E. (1992). Mapping the site(s) of MgATP and MgADP interaction with the nitrogenase of *Azotobacter vinelandii*. *J. Biol. Chem.* **267**, 6680–6688.
Simpson, F. B. & Burris, R. H. (1984). A nitrogen pressure of 50 atmospheres does not prevent evolution of hydrogen by nitrogenase. *Science,* **224**, 1095–1096.
Smith, B. E., Yousafzai, F., Eady, R. R., Gormal, C. A., Mayer, S., Roe, S. M., Lawson, D. M., Hasnain, S. S. & Grossmann, J. G. (1998). Structure of *Klebsiella pneumonia* nitrogenase. In *Biological Nitrogen Fixation for the 21st Century* (Elmerich, C., Kondorosi, A. & Newton, W. E., eds), pp. 43–46, Kluwer Academic Publishers, Dordrecht.
Sprang, S. R. (1997). G protein mechanisms: insights from structural analysis. *Annu. Rev. Biochem.* **66**, 639–678.
Tanaka, M., Haniu, M., Yasunobu, K. T. & Mortenson, L. E. (1977). The amino acid sequence of *Clostridium pasteurianum* iron protein component of nitrogenase. III. The NH_2-terminal and COOH-terminal sequences, tryptic peptides of large cyanogen bromide peptides, and the complete sequence. *J. Biol. Chem.* **252**, 7093–7100.
Thompson, J. D., Higgins, D. G. & Givson, T. J. (1994). Clustal-W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**, 4673–4680.
Thorneley, R. N. F. (1992). Nitrogenase of *Klebsiella pneumoniae* an MgATP hydrolysing energy transduction system with similarities to actomyosin and p21ras. *Phil. Trans. Roy. Soc. ser. B,* **336**, 73–82.
Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **8**, 945–981.
Watt, G. D. & Reddy, K. R. N. (1994). Formation of an all ferrous Fe_4S_4 cluster in the iron protein-component of *Azotobacter vinelandii* nitrogenase. *J. Inorg. Biochem.* **53**, 281–294.
Weston, M. F., Kotake, S. & Davis, L. C. (1983). Interaction of nitrogenase with nucleotide analogs of ATP and ADP and the effect of metal ions on ADP inhibition. *Arch. Biochem. Biophys.* **225**, 809–817.
Wolle, D., Dean, D. R. & Howard, J. B. (1992). Nucleotide-iron-sulfur cluster signal transduction in the nitrogenase iron-protein: The role of Asp^{125}. *Science,* **258**, 992–995.

Edited by I. A. Wilson

(Received 2 March 1998; received in revised form 20 April 1998; accepted 22 April 1998)