X-ray Crystal Structure of the Nitrogenase Molybdenum–Iron Protein from *Clostridium pasteurianum* at 3.0-Å Resolution[†,‡]

Jongsun Kim, D. Woo, and D. C. Rees*

Division of Chemistry and Chemical Engineering, 147-75CH, California Institute of Technology, Pasadena, California 91125

Received December 7, 1992; Revised Manuscript Received March 29, 1993

ABSTRACT: The crystal structure of the nitrogenase molybdenum–iron (MoFe) protein from *Clostridium pasteurianum* (Cp1) has been determined at 3.0-Å resolution by a combination of isomorphous replacement, molecular replacement, and noncrystallographic symmetry averaging. The structure of Cp1, including the two types of metal centers associated with the protein (the FeMo-cofactor and the P-cluster pair), is similar to that previously described for the MoFe-protein from *Azotobacter vinelandii* (Av1). Unique features of the Cp1 structure arise from the presence of an ~50-residue insertion in the α subunit and an ~50-residue deletion in the β subunit. As a consequence, the FeMo-cofactor is more buried in Cp1 than in Av1, since the insertion is located on the surface above the FeMo-cofactor. The location of this insertion near the putative nitrogenase iron protein binding site provides a structural basis for the observation that the nitrogenase proteins from *C. pasteurianum* have low activity with complementary nitrogenase proteins isolated from other organisms. Mechanistic implications of the Cp1 structure for substrate entry/product release, substrate binding to the FeMo-cofactor, and electron- and proton-transfer reactions of nitrogenase are discussed.

Biological nitrogen fixation (reduction of dinitrogen to ammonia) is carried out by a variety of free-living bacteria, cyanobacteria and symbiotic bacteria, in a reaction catalyzed by the nitrogenase enzyme system. The conventional nitrogenase enzyme system consists of two metalloproteins, the molybdenum–iron (MoFe) protein and the iron (Fe) protein [for a review, see Burgess (1984, 1990), Orme-Johnson (1985), Holm and Shimon (1985), Stiefel et al. (1988), Burris (1991), and Smith and Eady (1992)], although homologous alternate nitrogenase systems may be induced under the depletion of molybdenum [for a review, see Bishop et al. (1988) and Eady (1991)]. In addition to the nitrogenase proteins, a source of reducing equivalents (ferredoxin or flavodoxin *in vivo*), MgATP, and protons are required for nitrogen fixation. The overall stoichiometry of biological nitrogen fixation may be represented as (Simpson & Burris, 1984)

$$N_2 + 8H^+ + 8e^- + 16MgATP \rightarrow 2NH_3 + H_2 + 16MgADP + 16P_i \quad (1)$$

The nitrogenase MoFe-protein from *Clostridium pasteurianum* is an $\alpha_2\beta_2$ tetramer with a molecular mass of ~220 kDa, and the corresponding Fe-protein is a γ_2 dimer with a molecular mass of ~60 kDa (Tso, 1974). The MoFe-protein contains two copies of the FeMo-cofactor [M-center; reviewed in Burgess (1990) and Newton (1992)], which are believed to be the substrate binding and reduction sites, and two copies of the P-cluster pair [P-clusters; reviewed in Holm et al. (1990)], which may serve to mediate electron transfer between the Fe-protein and the FeMo-cofactor. The Fe-protein has a single 4Fe:4S cluster which transfers electrons to the MoFe-protein in an ATP-dependent manner. Electron transfer from the Fe-protein to the MoFe-protein involves a cycle of association and dissociation of the protein complex concomitant with ATP hydrolysis, and the dissociation has been identified as the rate-determining step (Lowe & Thorneley, 1983, 1984; Hageman & Burris, 1978).

The physicochemical properties and the primary structures of Fe- and MoFe-proteins are highly conserved among all nitrogen-fixing bacteria studied so far, but significant differences also exist, especially between the nitrogenase enzyme system of *C. pasteurianum* and that found in other bacteria. The *C. pasteurianum* MoFe-protein (designated Cp1) has the lowest sequence homology with other MoFe-proteins (Ioannidis & Buck, 1987; Wang et al., 1988). The amino acid sequence identity between Cp1 and the *Azotobacter vinelandii* MoFe-protein (designated Av1) is ~36%. Additionally, Cp1 has a long insertion (~50 amino acid residues) in the α subunit and a long deletion (~50 amino acid residues) in the β subunit (Wang et al., 1988). As a result of these sequence changes, the nitrogenase components from *C. pasteurianum* have a distinctly low capacity to form an active hybrid enzyme with the complementary components of other organisms (Emerich & Burris, 1978). Differences between Cp1 and Av1 in both the relative reduction sequence of the metal centers and their measured midpoint potentials have been observed when redox titrations were monitored by electron paramagnetic resonance (EPR) (Morgan et al., 1987). Since similar studies of the isolated cofactors from each species have shown them to be essentially identical (Morgan et al., 1987), the difference in the behavior of these two proteins probably results from the constraints imposed by each protein on an identical cofactor. The *C. pasteurianum* nitrogenase enzyme system is also less sensitive to H_2 as an inhibitor (Guth & Burris, 1983) and shows a higher specificity for nucleotides (Weston et al., 1983).

The crystal structure of nitrogenase MoFe-protein from *A. vinelandii*, including the structure of the FeMo-cofactor and the P-cluster pair, has been reported in Kim and Rees (1992a,b). In this paper, the tertiary and quaternary structures of the nitrogenase MoFe-protein from *C. pasteurianum* are

† This work was supported by NSF Grant DMB 91-18689.

‡ The crystallographic coordinates have been deposited in the Brookhaven Protein Data Bank under the file name 1MIO.

* To whom correspondence should be addressed [telephone, (818) 395-8393; FAX, (818) 568-9430; E-mail, REES@CITRAY.CALTECH.EDU].

0006-2960/93/0432-7104$04.00/0

presented, based on a 3.0-Å-resolution X-ray crystallographic analysis. The structural comparison of Av1 versus Cp1 and the functional implications of the nitrogenase MoFe-protein in dinitrogen reduction, including substrate entry/product release, proton-transfer pathways, electron-transfer pathways, and Fe-protein binding sites, are also discussed.

Structure Determination

The structure of Cp1 was solved by a combination of single isomorphous replacement (SIR), molecular replacement, and noncrystallographic symmetry (NCS) averaging both within and between two crystal forms (Rossmann, 1972; Bricogne, 1976). The purification, crystallization, and heavy atom derivative screening of nitrogenase MoFe-proteins have been previously described (Kim & Rees, 1992a). Two crystal forms of the MoFe-protein in space group $P2_1$ were prepared from *C. pasteurianum*, designated Mg1 and Cs1, with unit cell constants a = 70.0 Å, b = 151.3 Å, c = 121.9 Å, and β = 110.4° and a = 87.9 Å, b = 171.4 Å, c = 73.6 Å, and β = 91.5°, respectively. Both crystal forms contain one tetramer molecule in an asymmetric unit. The Mg1 crystal form is similar to a form previously described (Weininger & Mortenson, 1982). Each crystal form was derivatized using (ethylmercuri)thiosalicylate (EMTS), and four common EMTS binding sites which are related by the 2-fold NCS axis were found for both crystal forms (Table I). The NCS relationships both within and between crystal forms (Av1, Mg1, and Cs1) were determined from rotation functions (Rossmann & Blow, 1962; Crowther, 1972) and translation functions (Crowther & Blow, 1967) and were confirmed by the heavy atom locations and the FeMo-cofactor and P-cluster pair locations. Model phases from the Av1 structure oriented in the Cp1 unit cell were combined with the SIR phases of the Mg1 and Cs1 crystal forms, and the combined phases were subsequently refined by averaging both within and between the two different crystal forms (Bricogne, 1976) using the 2-fold NCS in both crystal forms. The final R factor of the 4-fold averaging was ~22% and that of the subsequent 2-fold averaging within the Mg1 crystal form was ~18%. The Mg1 crystal form was used for further crystallographic analysis; an atomic model of the Cs1 crystal form has not yet been built and refined. The averaged Mg1 electron density map was of sufficient quality to trace 1820 of the 1980 amino acid residues. The long inserted polypeptide in the α subunit (α375–α430) did not appear in the averaged map because the Av1 envelope was used for averaging. Hence this region was built following inspection of $2F_o - F_c$ and $F_o - F_c$ maps. The initial model was built into the averaged electron density map using the graphics program TOM/FRODO (Jones, 1985) implemented with the fragment fitting option (Jones & Thirup, 1986) and subsequently refined using the restrained least squares program TNT (Tronrud et al., 1987). The resultant $2F_o - 2F_c$ and $F_o - F_c$ maps and the electron density maps obtained by iterative cycles of model building, refinement, phase combination, and NCS averaging were used for further model building. The simulated annealing protocol in the X-PLOR (Brünger, 1988) program was used during the final stages of coordinate refinement.

The current model contains 1964 amino acid residues (of 1980 total residues) with 15 273 non-hydrogen atoms (99% complete). This model has presently been refined to a crystallographic R factor of 0.18 (10–3.0 Å) with root mean square deviations from ideal bond distances and angles of 0.018 Å and 3.9°, respectively (Table I). The correctness of the chain trace is further supported by the location of the heavy atom binding sites (Table I) in addition to the similar folding of both the α and β subunits. The EMTS sites are found coordinated to nonconserved Cys residues (α302 and β257). The model has also been examined using the 3D–1D profile method (Lüthy et al., 1992), and all the residues have reasonable average 3D–1D scores except for the region α380–α391 in which electron density is diffuse. Residue numbers are prefixed with either α or β to indicate the appropriate subunit, except where interactions in the tetramer are described, in which case the prefixes A and C designate the two distinct α subunits while the prefixes B and D designate the two distinct β subunits present in the tetramer.

Table I: Structure Determination and Refinement Statistics for the *C. pasteurianum* MoFe-Protein (Cp1)[a]

(A) Heavy Atom Binding Sites

derivative	x	y	z	rel occupancy	location
EMTS	0.356	0.073	0.858	0.23	Cys α302
(Mg1)	0.125	0.295	0.105	0.22	Cys α302 (NCS related)
	0.579	–0.264	0.451	0.15	Cys β257
	–0.348	–0.161	0.104	0.15	Cys β257 (NCS related)
EMTS	0.416	0.231	–0.105	0.17	[b]
(Cs1)	–0.033	–0.254	0.071	0.16	
	0.737	0.021	0.719	0.14	
	0.019	–0.008	0.924	0.11	

(B) Model Status (Mg1)

parameter	value
no. of amino acid residues	1964/1980
total non-hydrogen atoms	15273 (99% complete)
no. of water molecules	0
missing residues	α527–α534
no. of Ramachandran outliers	18/1960

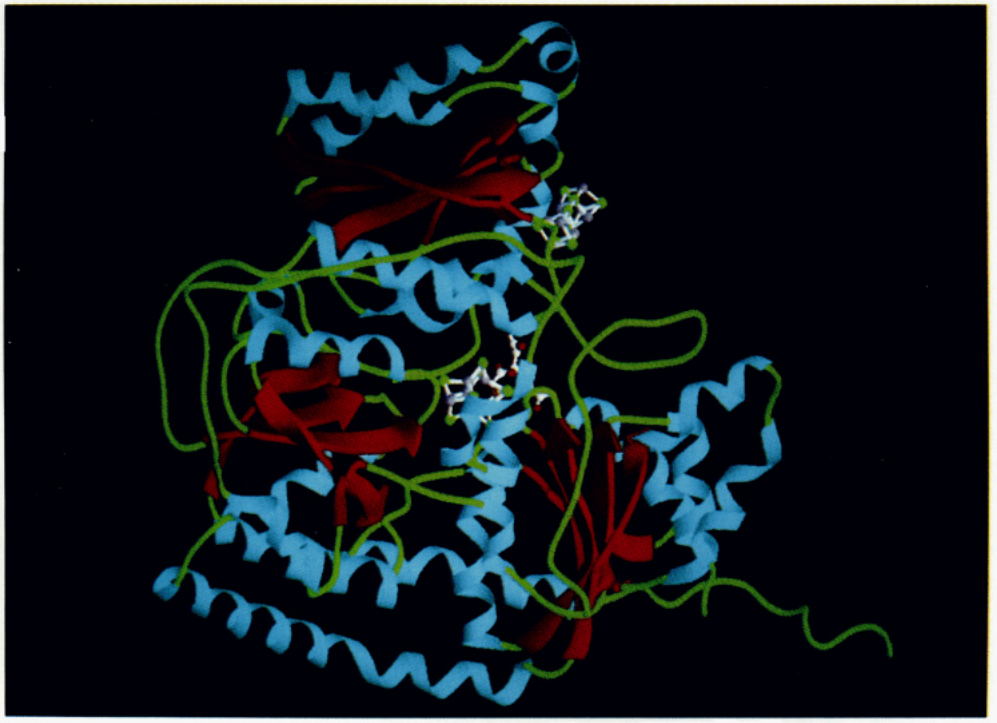
(C) Refinement Statistics

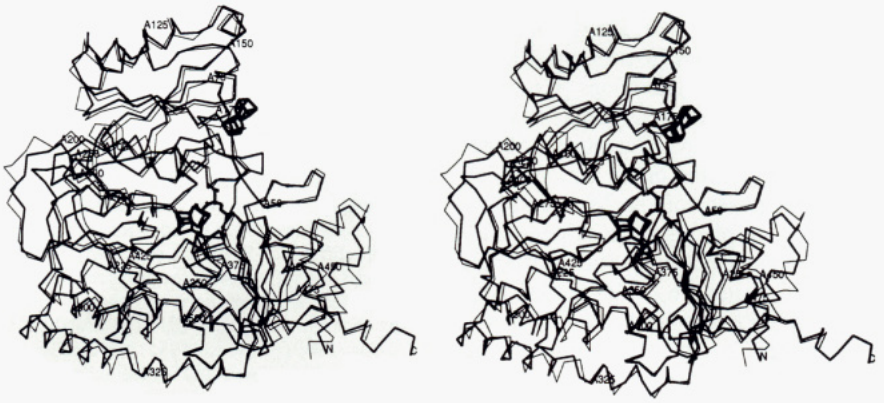
	before refinement	after refinement
R factor (10.0–3.0 Å)	0.39	0.18
rms deviation of		
bond length (Å)	0.024	0.018
bond angles (deg)	2.90	3.9
dihedral angles (deg)	28.3	25.3
improper torsion (deg)	1.66	1.50

[a] The Mg1 native data set collected by a Siemens multiwire area detector at room temperature contains 38 527 reflections and is 78% complete to 3.0-Å resolution with a merging R factor of 0.07. The Cs1 native data set collected by a Siemens area detector at room temperature contains 33 013 reflections and is 74% complete to 3.0-Å resolution with a merging R factor of 0.084. Heavy atom derivatives were prepared with (ethylmercuri)thiosalicylate (EMTS) for both the Mg1 and Cs1 crystal forms, and the derivative data sets were also collected by an area detector (Kim & Rees, 1992a). Locations of the heavy atoms were determined from difference Patterson maps with the aid of molecular replacement phases and refined with the program HEAVY (Terwilliger et al., 1987). The mean figure of merit and phasing power of both the Mg1 and Cs1 heavy atom derivatives (EMTS) are about 0.36 and 1.37, respectively. TOM/FRODO (Jones, 1985) was used for model building, and X-PLOR (Brünger, 1988) was used in the final refinements with PARAM19X.PRO as the parameter file. [b] The binding locations for the Cs1 heavy atom sites presumably correspond to the Mg1 sites, but this has not been directly established.

Description of the Protein Structure

The α subunit of Cp1 consists of 533 amino acid residues, which is the longest among all known MoFe-proteins (Wang et al., 1988). The three-dimensional structure of the α subunit is similar to that of Av1, with the exception of a long inserted loop region (α375–α430). The α subunit of Cp1 consists of three domains of the α/β type with some extra helices (Figure 1a,b). The overall shape of the α subunit may be described as a cloverleaf. Domain I is composed of seven helices, four





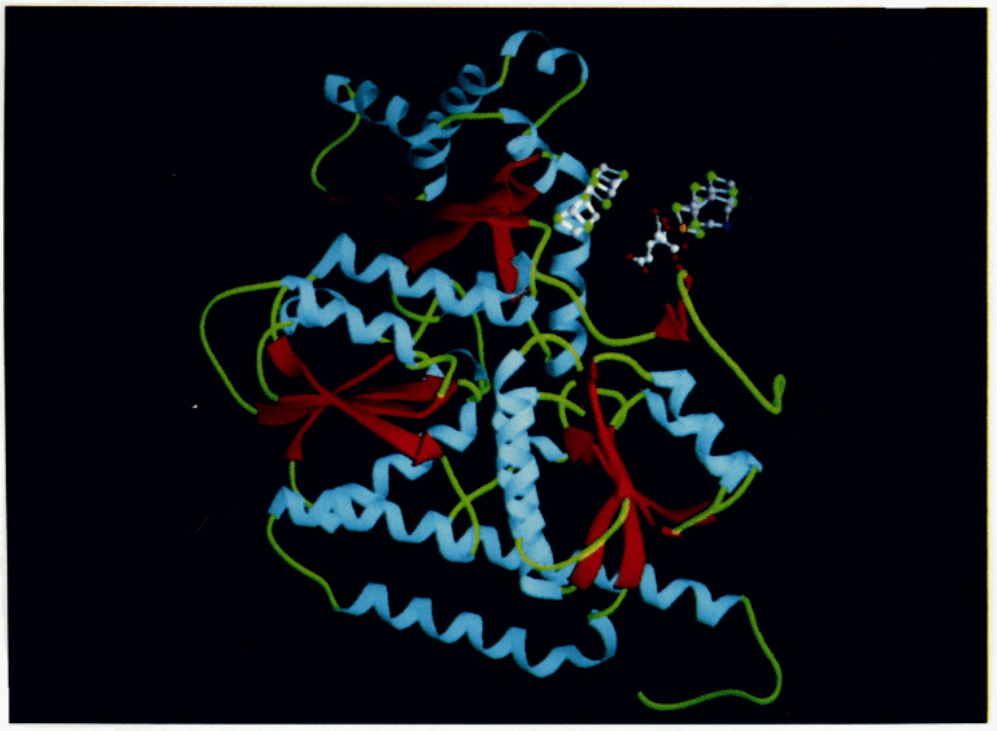


FIGURE 1: (a, left page, top) RIBBONS (Carson & Bugg, 1986) diagram of the polypeptide fold of the α subunit of Cp1. α-Helical regions are in light blue, β-sheet regions are in red, and the remaining regions of the polypeptide chain are in green. The three domains are designated I, II, and III, respectively, counterclockwise from the top. The FeMo-cofactor is located in the center of the figure, and the P-cluster pair is on the top right side. (b, left page, middle) Cα chain trace of the α subunit of Cp1 superimposed on that of Av1. The rms deviation between common Cα positions is 1.57 Å. The α subunits of Cp1 and Av1 are represented by the thick and thin lines, respectively. The view is approximately from the same orientation as that in (a). (c, left page, bottom) RIBBONS diagram of the polypeptide fold of the β subunit of Cp1. α-Helical regions are in light blue, β-sheet regions are in red, and the remaining regions of the polypeptide chain are in green. The three domains are designated I′, II′, and III′, respectively, counterclockwise from the top. The left-most metal center is the P-cluster pair, while the FeMo-cofactor is to the right of the P-cluster pair. (d, above) Cα chain trace of the β subunit of Cp1 superimposed on that of Av1. The rms deviation between common Cα positions is 1.61 Å. The β subunits of Cp1 and Av1 are represented by the thick and thin lines, respectively. The view is approximately from the same orientation as that in (c).

parallel β strands, and one antiparallel β strand which is provided by the β subunit; domain II is composed of six helices, four parallel β strands, and one antiparallel β strand; and domain III is composed of ten helices, five parallel β strands, and one antiparallel β strand which is from the N-terminus of the α subunit. There is a wide and shallow cleft between the three domains, and the FeMo-cofactor sits beneath the bottom of this cleft. This cleft has been proposed as a Fe-protein subunit binding site in Av1 (Kim & Rees, 1992b). Residues α375–α430, which are absent in Av1, are located between domains II and III; in particular, α383–α397 are located above the cleft providing an additional polypeptide environment in the vicinity of the FeMo-cofactor.

The β subunit of Cp1 consists of 457 amino acid residues, which is the shortest among all known MoFe-proteins (Wang et al., 1988). The structure of the β subunit is also similar to that of Av1 except for the N-terminus ($\sim$10 residues of Cp1) and loop regions at the surface where deletions or insertions occur (Figure 1c,d). The overall folding of the β subunit, which is composed of three α/β-type domains, is similar to that of the α subunit, as has been observed in the Av1 structure. However, significant differences also exist, in particular at the N-terminus, in the α375–α430 region, and in surface loop regions. In the β subunit, domain I′ is made up of seven helices and four parallel β strands; domain II′ is made up of six helices and four parallel β strands; and domain III′ is made up of six helices and five parallel β strands. Analogous to the α subunit, there is a cleft between the three domains, which has been proposed to provide part of the Fe-protein subunit binding site in Av1. The C-terminus of the β subunit, which is located at the interface of the B and D subunits, is not exposed to solvent and is very well ordered in both Cp1 and Av1, while the C-terminus of the α subunit is exposed to solvent and is not as well ordered in Cp1.

The $\alpha\beta$ subunit pair consists of six α/β-type domains which are arranged like a six-petaled flower, with the P-cluster pair located in the center of the dimer like a pistil (Figure 2). The α subunit and the β subunit of the MoFe-protein are related by an approximate 2-fold axis which passes through the center of the P-cluster pair, and there are two wide and shallow clefts around the P-cluster pair which may provide the binding site for the dimeric Fe-protein. The α and β subunits are in very close contact with each other. Domain I of the α subunit contacts domains I′ and III′ of the β subunit, and domain I′ of the β subunit contacts domains I and III of the α subunit. Domain I of the α subunit and domain I′ of the β subunit are bridged by the P-cluster pair. An antiparallel β-sheet arrangement between β15–β18 and α105–α109 also contributes to the $\alpha\beta$ subunit interface. In addition to these general contacts, salt bridge, hydrophobic, and hydrogen-bonding interactions between the two subunits are also important for dimerization, as is observed in Av1.

As mentioned above, the overall structure of Cp1 is similar to that of Av1. In particular, the tertiary structure around the metal centers is well conserved between the two MoFe-proteins. However, significant differences also exist. The overall dimensions of the $\alpha_2\beta_2$ Cp1 tetramer are $\sim$70 Å × 80 Å × 120 Å, and its general shape is slightly elongated along one axis compared to that of Av1 due to the $\sim$50-residue insertion in the α subunit (Figure 3). The two $\alpha\beta$ subunit pairs are related by the 2-fold NCS rotation axis used for initial phase refinement. Even though the α and β subunits in an $\alpha\beta$ subunit pair are also related by an approximate 2-fold rotation axis, the MoFe-protein tetramer does not exhibit 222 symmetry, which had been proposed by a low-resolution rotation function study (Yamane et al., 1982), because the $\alpha\beta$ 2-fold and the tetramer 2-fold axes are not perpendicular and do not intersect each other. Consequently, the organization of α and β subunits in the MoFe-protein tetramer differs from that of hemoglobin (Perutz et al., 1960), where the homologous subunits are arranged with approximate 222 symmetry.

The tetramer interface is generated by extensive interactions between domains II′ and III′ of the two β subunits, along with some additional interactions involving domain III of each α subunit. Packing between helices from the β subunit (B181–B194, B290–B310, B270–B284, B427–B447, and the corresponding regions of subunit D) dominates the interactions at the tetramer interface, along with additional contributions

FIGURE 2: (a, top) RIBBONS diagram of the polypeptide fold of an $\alpha\beta$ subunit pair. The α subunit is in red and the β subunit is in light blue. The long inserted loop in the α subunit, which is unique to Cp1, is in green. The view is down the 2-fold axis through the P-cluster pair that approximately relates the α and β subunits. The P-cluster pair is located in the center of the figure, and the FeMo-cofactor is located below the P-cluster pair. (b, bottom) Cα chain trace of the $\alpha\beta$ subunit pair of Cp1 superimposed on that of Av1. The rms deviation between 880 common Cα positions is 1.63 Å. The $\alpha\beta$ subunit pairs of Cp1 and Av1 are represented by the thick and thin lines, respectively. The view is approximately from the same orientation as that in (a).

from helices in the α subunit (A289–A302, A306–A334, A493–A507, and the corresponding regions of subunit C). The helical interactions appear to provide a major driving force for tetramerization. As a result of the six-helix barrel in the center of the tetramer, the MoFe-protein has a large channel ~8–10 Å wide and ~35 Å in length. In addition to these helical packing interactions, electrostatic, hydrophobic, and hydrogen-bonding interactions are important for tetramerization, as is observed in Av1.

The divalent cation sites found in the Av1 structure (Kim & Rees, 1992b) also exist in Cp1, and the tertiary structure around these sites is well conserved. On the basis of the electron density value, temperature factor, and the coordination environment, these sites could be assigned as Ca^{2+} (or, less likely, Mg^{2+}). This ion has an octahedral coordination environment provided by the carboxyl oxygens of Glu B62, Asp D301, Asp D305, the carbonyl oxygen of Lys B61, and probably two water molecules. Gln A472, Lys A473, Thr B212, Phe A469, Phe B60, Tyr B420, and Tyr B421 are located outside the immediate coordination sphere. The site is buried at the interface of two $\alpha\beta$ dimers and is ~25 and ~21 Å away from the P-cluster pair and FeMo-cofactor, respectively. This site does not appear to have a functional role, but rather may serve to stabilize the subunit associations in the tetramer. Neuraminidase (Varghese et al., 1983) and some spherical plant viruses (Hogle et al., 1983) also have divalent cations (Ca^{2+}) at subunit interfaces that function in the stabilization of subunit contacts.

No disulfide bonds are present in either the Cp1 or Av1 structures, and no ATP/ADP is found in Cp1 and Av1, although some experimental studies have indicated that oxidized MoFe-protein can bind MgADP (Miller & Eady, 1989).

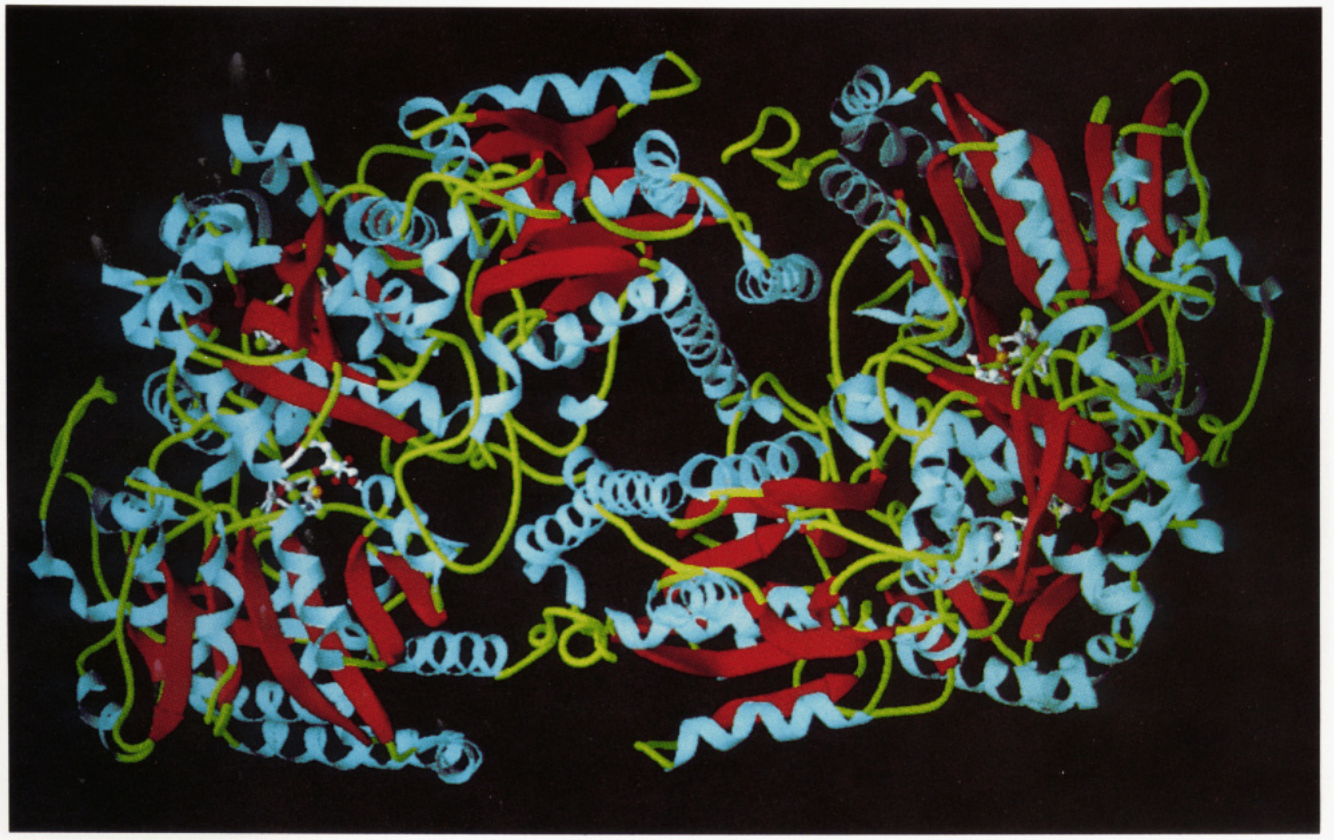


FIGURE 3: RIBBONS diagram of the polypeptide fold of the $\alpha_2\beta_2$ MoFe-protein tetramer. The view is down the tetramer 2-fold axis. α-Helical regions are in light blue, β-sheet regions are in red, and the remaining regions of the polypeptide chain are in green.

Structures of the Metal Centers

The structures of the metal centers in Av1 have been described in Kim and Rees (1992a) and verified by high-resolution X-ray diffraction analysis (Chan et al., 1993). The FeMo-cofactor and P-cluster pair structures of Cp1 appear identical to those of Av1. The FeMo-cofactor contains 4Fe:3S and 1Mo:3Fe:3S clusters that are bridged by three nonprotein ligands (Figure 4a). Two of the bridging ligands are assigned as sulfurs, while the chemical identity of the third ligand is still ambiguous and could be either a well-ordered O/N species or a less well ordered S species in Av1. The electron density of this site looks more like a sulfur in Cp1. The Mo site has the highest electron density value in the 3.0-Å-resolution $2F_o$–F_c map. Although the other atomic positions are not resolved, the FeMo-cofactor model generally matches the electron density well. Homocitrate, an essential component of FeMo-cofactor (Hoover et al., 1989), is coordinated through hydroxyl and carboxyl oxygens to the Mo site. The importance of the hydroxyl group in homocitrate had been implicated by studies in which the homocitrate was replaced with other carboxylic acids (Hoover et al., 1988). The FeMo-cofactor is attached to the α subunit through two protein ligands, Cys α262, which coordinates Fe1, and His α482, which coordinates Mo along with homocitrate.

A ball and stick model of the P-cluster pair is presented in Figure 4b. The P-cluster pair consists of two 4Fe:4S clusters, as had been suggested by Mössbauer and extrusion studies (Huynh et al., 1980; Kurtz et al., 1979), that are bridged by two cysteine thiol ligands. X-PLOR refinement results indicate that the two 4Fe:4S clusters are additionally linked by a disulfide bond (S–S bond distance = ~2.2–2.3 Å) formed between the sulfurs from each 4Fe:4S cluster. The existence of a disulfide bond in the P-cluster pair is consistent with the Av1 P-cluster pair structure (Chan et al., 1993). The P-cluster pair is attached between the α and β subunits through seven protein ligands: Cys α53, Cys α79, Cys α145, Cys β23, Cys β48, Cys β106, and Ser β141. Cys α79 bridges Fe4 and Fe5 and Cys β48 bridges Fe1 and Fe8. Cys α53, Cys α145, Cys β23, and Cys β106 coordinate the remaining Fe sites. Ser β141 is close to Fe6 and may coordinate this site along with Cys β106.

Environment of the FeMo-Cofactor

The FeMo-cofactor, which almost certainly provides the substrate binding and reduction site, can be isolated intact from the MoFe-protein (Shah & Brill, 1977). In its isolated form, however, it no longer catalyzes dinitrogen reduction (Shah & Brill, 1977). Because the FeMo-cofactor must be protein-bound in order to reduce substrate, the polypeptide environment must contribute to its substrate binding and reduction properties. The FeMo-cofactor and its surrounding residues are represented in Figure 5a. The FeMo-cofactor is buried at least 10 Å from the protein surface, and the polypeptide environment around the FeMo-cofactor is primarily provided by the α subunit. The FeMo-cofactor is buried more deeply in Cp1 than in Av1 because the α383–α397 loop, which is unique to Cp1, is located above the FeMo-cofactor and partially occupies the cleft as shown in Figures 5c and 7c. Cys α262 and His α482 are coordinated to the FeMo-cofactor and Ser α265 is hydrogen bonded to the Sγ of Cys α262. These residues are strictly conserved among all known MoFe-protein sequences and are structurally important, along with conserved Gly residues (Gly α344 and Gly α345) whose amide

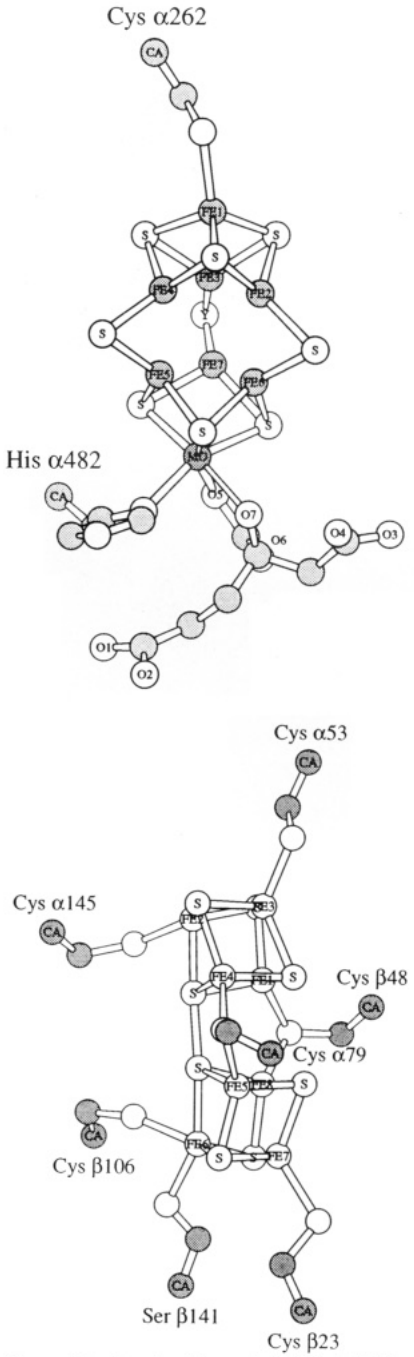


FIGURE 4: (a, top) Ball and stick model of the FeMo-cofactor with protein ligands, drawn with the MOLSCRIPT program (Kraulis, 1991). The Y ligand could be S or a well-ordered O/N species (less likely). (b, bottom) Ball and stick model of the P-cluster pair with protein ligands, drawn with the MOLSCRIPT program.

NH groups are hydrogen bonded to sulfurs in the FeMo-cofactor. Other highly conserved residues near the FeMo-cofactor include Arg α87 and Arg α347, which can potentially form hydrogen bonds to sulfurs in the FeMo-cofactor and may serve to electrostatically stabilize the FeMo-cofactor and/or partially reduced intermediate formed during substrate reduction; His α186, which is hydrogen bonded to a bridging sulfur and may participate in proton-transfer reactions; Gln α182, Glu α368, Glu α467, Gln α480, and His α482, which are near the homocitrate and interact with this group either directly or through water molecules; and aromatic and hydrophobic residues, such as Tyr α216, Phe α369, Val α61, and Ile α220.

Substrate Entry and Product Release

Even though the FeMo-cofactor is buried at least 10 Å below the protein surface and there are no permanent channels between the protein surface and the FeMo-cofactor (Kim & Rees, 1992b), there are two clefts which could be potentially utilized for substrate entry/product release and/or H_3O^+ transfer to the active site. The first cleft, which is made up of five stretches of polypeptide chain (α255–α275, α279–α301, α338–α359, α363–α378, and α422–α436), exists between domains II and III and has a funnel shape with an outer diameter of ~4–10 Å (Figure 5b). The second cleft, also made up of five stretches of polypeptide chain (α172–α199, α341–α352, α365–α400, α256–α276, and α34–α43), exists between the three domains of the α subunit and has a funnel shape with an outer diameter of ~3–10 Å (Figure 5c). These two clefts are near the one of the putative Fe-protein subunit binding sites (see below) and may also provide possible openings for cofactor insertion during biosynthesis of the MoFe-protein. However, neither of these two clefts is wide enough to allow free diffusion of either substrate/product or H_3O^+. Therefore, it seems reasonable to assume that there must be structural fluctuations or conformational changes to allow the diffusion of ligands into and away from the FeMo-cofactor, analogous to the binding and release of oxygen to the buried hemes in globins (Case & Karplus, 1979).

In addition to the physiological reactions of N_2 reduction and H_2 formation, nitrogenase catalyzes a wide variety of reactions involving small unsaturated molecules, such as azides, nitrous oxide, nitriles, isonitriles, and alkynes (Burgess, 1985). Functionally, the details of the interaction between various substrates and the FeMo-cofactor are central to understanding the catalytic properties of nitrogenase. While many modes of substrate binding to one or more of the Fe, Mo, and/or S sites are possible [see, for example, Deng and Hoffmann (1993)], the six trigonally coordinated iron sites of the FeMo-cofactor are particularly intriguing for substrate binding. On the basis of the FeMo-cofactor structure, three general types of binding modes of substrates to these coordinatively unsaturated irons may be envisioned: (i) Substrates could bind in an end-on fashion to one of the six central iron sites which are bridged by nonprotein ligands. It is also conceivable that some substrates could displace the bridging ligand Y, if this ligand is not a sulfur, and interact simultaneously with the two adjacent iron sites. (ii) Some small substrates, such as N_2 and/or H^+, could occupy the central cavity in the FeMo-cofactor, thereby replacing the weak iron–iron bonds with Fe–substrate bonds (Chan et al., 1993). (iii) Substrates could bind to the exterior surface of the FeMo-cofactor. Three sets of cyclic eight-membered rings occur on the exterior surface of the FeMo-cofactor, consisting of alternating S(Y) and Fe sites. These arrangements may create the equivalent of a small region of an iron surface, each containing four iron atoms, and substrates could bind to those surfaces so as to simultaneously interact with up to four irons. The presence of multiple, potential substrate binding sites in the FeMo-cofactor may be related to the noncompetitive kinetics observed between N_2 and other substrates (with the exception of N_2O) (Rivera-Ortiz & Burris, 1975).

Proton Transfer

In order to reduce dinitrogen to ammonia, a supply of protons is essential in addition to dinitrogen. In the vicinity of the

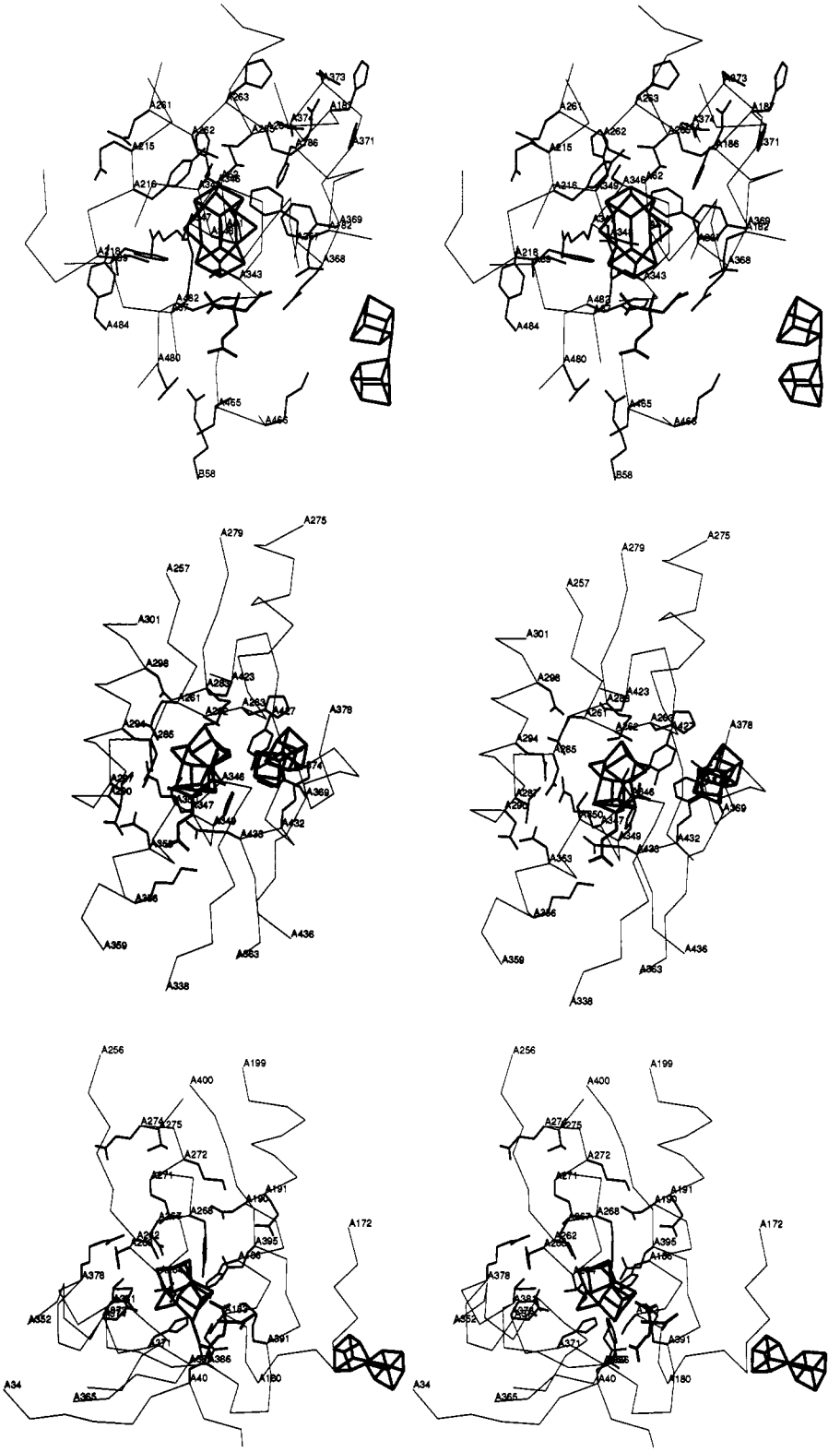


FIGURE 5: (a, top) Stereoview of the polypeptide environment surrounding the FeMo-cofactor. (b, middle) Stereoview of a cleft around the FeMo-cofactor that could be utilized for substrate entry/product release and/or H_3O^+ transport. (c, bottom) Stereoview of the other cleft around the FeMo-cofactor which could be utilized for substrate entry/product release and/or H_3O^+ transport.

FeMo-cofactor, several potential pathways occur which could transfer protons to the FeMo-cofactor: Glu α368 and Gluα467 to homocitrate to FeMo-cofactor, and His α186 to FeMo-cofactor. If an intermediate that is a sufficiently strong base is generated during dinitrogen reduction, it is possible that protons may also be transferred from Arg α347 and/or Argα87 to the FeMo-cofactor. The two potential channels around the FeMo-cofactor shown in Figure 5b,c could be utilized for the proton transfer to the vicinity of the cofactor. Charged or hydrophilic residues lining these potentials channels, such as Lys α432, Lys α283, His α263, His α186, His α349, Glu α433, Arg α378, His α187, Glu α294, Asn α285, Glu α271, Asp α298, His α371, Asn α267, Arg α264, Asp α373, and Asp α374, may facilitate proton transfer during the substrate reduction. Two patches of histidines at the entrance to these clefts might represent the initial site for proton binding; these

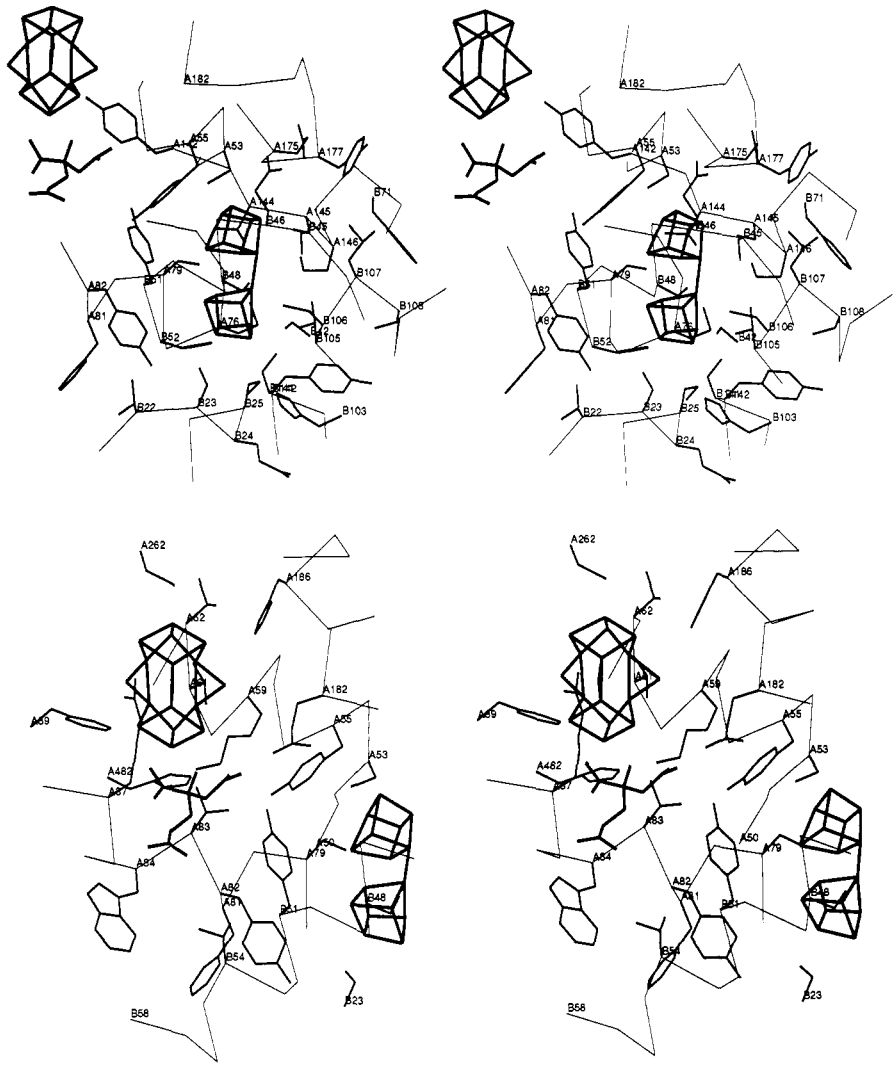


FIGURE 6: (a, top) Stereoview of the polypeptide environment surrounding the P-cluster pair. (b, bottom) Stereoview of the polypeptide environment between the FeMo-cofactor and the P-cluster pair.

patches contain the imidazole side chains of residues α187, α371, and α263; and of residues α349 and α435. The presence of multiple, potential proton-transfer routes suggests that there is not a unique pathway by which protons are shuttled from the surface to the active site. The proton-transfer pathway need not be unique, although there may be a major pathway,considering that at least eight protons are required for reduction of one dinitrogen (Simpson & Burris, 1984).

Environment of the P-Cluster Pair

The P-cluster pair, which may function in electron transfer between the 4Fe:4S cluster of Fe-protein and the FeMo-cofactor, is composed of two 4Fe:4S clusters bridged by two cysteine thiol ligands and one disulfide bond between two cluster sulfurs. In spite of the functional importance and novel structural features of the P-cluster pair, nitrogenase research has not generally focused on the P-cluster pair. Consequently, the understanding of how the P-cluster pair is synthesized and incorporated in MoFe-protein is quite limited, compared to that of the FeMo-cofactor. The P-cluster pair and its surrounding residues are shown in Figure 6a. The P-cluster pair bridges the α and β subunits and is completely buried at least 12 Å from the protein surface. The highly buried nature of the P-cluster pair and its polypeptide environment may provide an explanation for why the MoFe-protein requires a special electron donor, i.e., the Fe-protein. The polypeptide environment around the P-cluster pair is mainly provided by aromatic and hydrophobic residues, including Tyr α55, Pro α76, Tyr α82, Pro α146, Pro β25, Tyr α142, Tyr β51, Tyr β142, and Phe β71. These residues are highly conserved among all known MoFe-protein sequences. Cys α53, Cys α79, Cys α145, Cys β23, Cys β48, Cys β106, and Ser β141 are coordinated to the P-cluster pair, and these residues are strictly conserved, reflecting their structural importance. Strictly conserved Gly residues (Gly α78, Gly α176, and Gly β47) in the vicinity of the P-cluster pair are also structurally important to avoid steric hindrance with the P-cluster pair. Hydrophilic residues around the P-cluster pair, such as Ser α80, Thr α144, Glu α175, Ser β50, His β52, and Gln β24, are generally not conserved, with the exceptions of Gln β46 and Thr β105.

Mechanism of H_2 Evolution

Under optimal conditions, nitrogenase catalyzes the reaction $N_2 + 8H^+ + 8e^- \rightarrow 2NH_3 + H_2$. H_2 evolution is thought to arise from N_2 binding and displacing H_2 from the active site (Thorneley & Lowe, 1985). The proportion of electron flux which results in H_2 evolution can increase under conditions where the rate of electron flux decreases, i.e., limiting MgATP

or Fe-protein concentrations (Bishop et al., 1986), or where the pressure of N_2 is low (Hadfield & Bulen, 1969). These findings suggest that there may be an additional H_2 evolution mechanism. Recently, it has been proposed that the disulfide bond in the P-cluster pair may provide a site for H_2 evolution (Chan et al., 1993). Protonation of the doubly reduced P-cluster pair may generate a species that can produce H_2 upon disulfide formation. We propose that there are two H_2 evolution sites in MoFe-protein: the N_2 binding site in the FeMo-cofactor and the disulfide bond site in the P-cluster pair. The aromatic and hydrophobic residues around the P-cluster pair may function in suppressing H_2 evolution activity of nitrogenase by inhibiting proton transfer to the P-cluster pair.

Electron Transfer from the P-Cluster Pair to the FeMo-Cofactor

The polypeptide environment between the FeMo-cofactor and the P-cluster pair is illustrated in Figure 6b. The edge–edge distance of the FeMo-cofactor to the P-cluster pair is ~14 Å, which is consistent with low-resolution studies of the metal center positions in Cp1 (Bolin et al., 1990). Four helices (α54–α64, α77–α83, α182–α196, and β46–β60) oriented in parallel starting from the P-cluster pair toward the FeMo-cofactor may play an important role in electron transfer. In particular, helices α54–α64 and α77–α83, adjacent to the P-cluster pair ligands Cys α53 and Cys α79, provide the most direct polypeptide connection between the P-cluster pair and the FeMo-cofactor. Potential proton-transfer pathways, such as salt bridges and hydrogen-bonding networks, have not been found that could permit the coupling of electron and proton transfer between the P-cluster pair and the FeMo-cofactor. Since the homocitrate is located on the side of the FeMo-cofactor that faces the P-cluster pair, it is possible that the electrons are transferred from the P-cluster pair to the FeMo-cofactor through the homocitrate. Indeed, the importance of homocitrate to the substrate reduction mechanism may arise from its function in the protonation of intermediates and/or participation in the electron-transfer pathway between the P-cluster pair and the FeMo-cofactor, as well as from modulation of the redox properties of the coordinated FeMo-cofactor (Kim & Rees, 1992a).

Fe-Protein Binding Sites

The α and β subunits of MoFe-protein are related by an approximate 2-fold axis that passes through the center of the P-cluster pair, and there are two wide and shallow clefts related by this pseudo-2-fold rotation. As the Fe-protein dimer also has a 2-fold NCS axis (Georgiadis et al., 1992), a plausible model for docking the two proteins involves superposition of the Fe-protein's 2-fold axis with the approximate 2-fold axis of MoFe-protein passing through the P-cluster pair (Kim & Rees, 1992b). This docking model, with a stoichiometry of 1:2 MoFe-protein:Fe-protein, is consistent with both mutagenesis and cross-linking results (Govenzensky & Zamir, 1989; Willing & Howard, 1990; Wolle et al., 1992; Thorneley et al., 1993). Three potential contact regions, based on the structures of the MoFe-protein and Fe-protein, include two Fe-protein subunit binding sites (designated A and B in Figure 7a) and a 4Fe:4S cluster binding site (designated C in Figure 7a). Interestingly, the long inserted sequence of Cp1 (α375–α430) is located around the A-subunit binding site. In particular, residues α383–α397 span the cleft above the FeMo-cofactor, and this region may directly interact with the Fe-protein. Lys α385, Asp α387, Asp α389, and Asn α392 are exposed to water, and these residues may recognize the *C. pasteurianum* Fe-protein (Cp2) and discriminate against other Fe-proteins such as the *A. vinelandii* Fe-protein (Emerich & Burris, 1978). The location of the long inserted sequences in Cp1 may explain why Cp1 cannot form an active complex with any Fe-proteins other than Cp2 (Emerich & Burris, 1978), and this finding indirectly supports the assignment of Fe-protein binding sites in MoFe-protein and the proposed docking model (Kim & Rees, 1992b).

The potential 4Fe:4S cluster binding site (site C in Figure 7a) and surrounding residues are shown in Figure 7b. Four short helices (α112–α117, α146–α150, β106–β111, and β73–β78), which are related by the approximate 2-fold axis, are oriented in parallel from the P-cluster pair toward the protein surface, forming an approximate four-helical bundle upon which the 4Fe:4S cluster of Fe-protein is likely to sit. The edge–edge distance from the P-cluster pair to the end of these helices is about 12 Å; thus the edge–edge distance from the P-cluster pair to the 4Fe:4S cluster of Fe-protein may be about 15 Å. It seems likely that the electron is transferred through these helices. The polypeptide environment around this electron-transfer pathway is primarily provided by aromatic or hydrophobic residues such as Tyr α177, Tyr β142, Leu α149, Leuβ107, Ile α150, Leu β111, Phe α117, Phe β78, and Pro α146, and they are highly conserved among the known MoFe-protein sequences. There are no unbroken hydrogen-bonding/salt bridge networks along the electron-transfer pathway; therefore, it seems likely that the electron transfer from Fe-protein to the P-cluster pair is not coupled with proton transfer. Polar and charged residues outside the four-helical bundle may be important for interaction with Fe-protein through formation of hydrogen bonds and salt bridges. Asp α153 and Asp β114, which are related by the pseudo-2-fold axis, are completely exposed to solvent and are about 18 Å away from the P-cluster pair. His β136 is exposed to solvent and is about 16 Å from the P-cluster pair. Asn α121, Glu α112, and Glu β73 are also exposed to solvent.

The putative Fe-protein subunit binding sites (sites A and B in Figure 7a), which are related by the pseudo-2-fold rotation, and the exposed residues around these sites are shown in Figures 7c,d. It is intriguing that the FeMo-cofactor is located under one of these possible binding sites. This observation may be relevant to the observation that the Fe-protein is involved in FeMo-cofactor insertion during the MoFe-protein maturation (Robinson et al., 1987). The polypeptide environment around the Fe-protein subunit binding sites is primarily provided by charged or hydrophilic residues located on the MoFe-protein surface. Thus, electrostatic interactions and hydrogen bond formation appear to be important factors in the association of MoFe-protein and Fe-protein, as had been implicated by salt effects on nitrogenase activity (Deits & Howard, 1990).

ACKNOWLEDGMENT

We thank A. J. Chirino, B. T. Hsu, H. Komiya, M. K. Chan, and J. B. Howard for contributions to this project. X-PLOR calculations were performed on the CRAY-YMP at the San Diego Supercomputer Center, supported by the NSF.

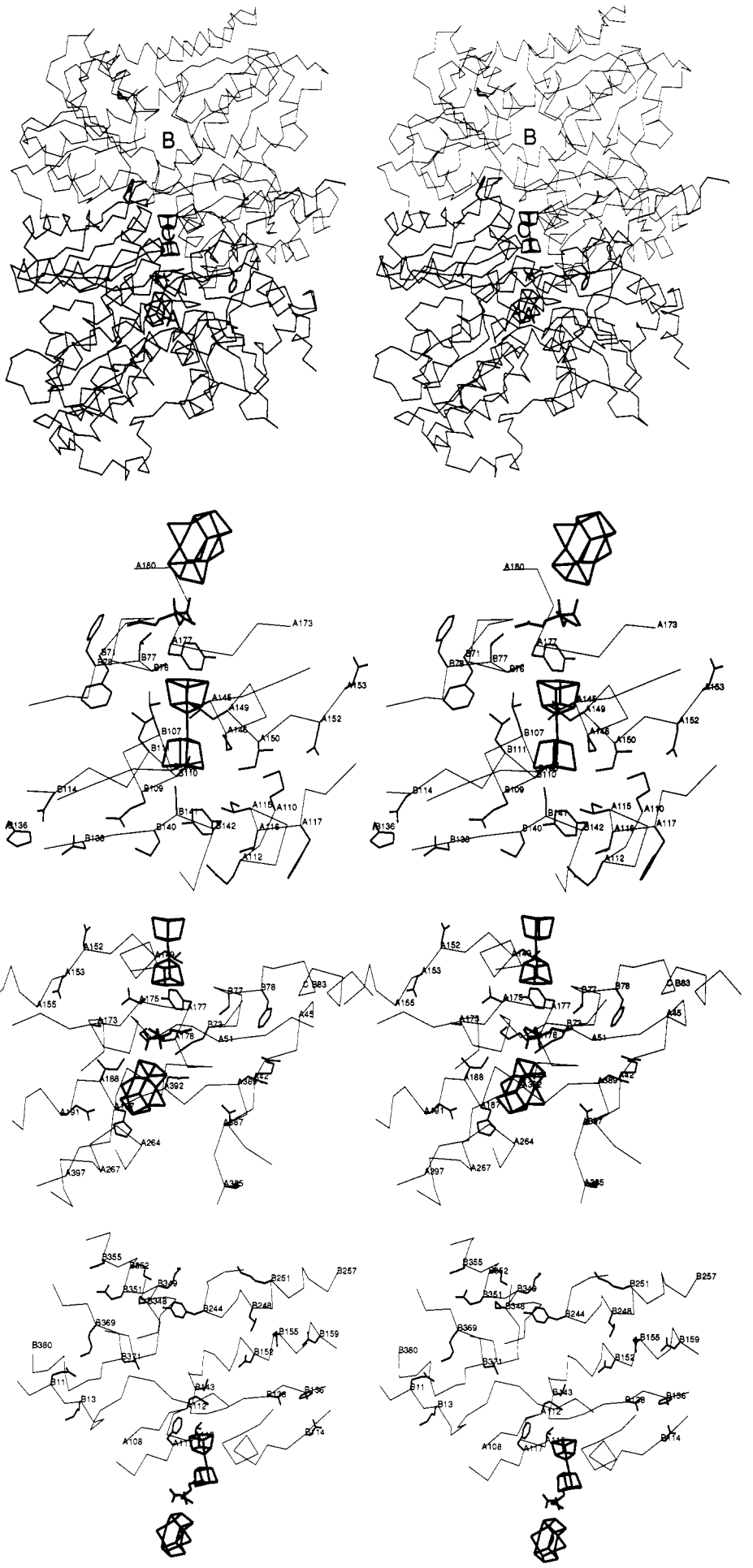

B
B
A180
A173
A177
B71
B77
B78
A153
A145
A149
B107
B111
A150
A152
B110
B114
B109
A115
A110
B136
B140
B142
A117
A112
A152
A153
A155
A175
A177
B77
B78
B83
A173
A178
B73
A45
A51
A188
A392
A264
A397
A267
B355
B352
B349
B351
B348
B251
B257
B244
B248
B369
B155
B159
B380
B371
B152
B11
B13
B143
B136
A112
A108
B174

FIGURE 7: From top to bottom: (a) Overview of the Fe-protein binding sites. Three potential contact regions are designated A, B, and C, and side chains of Phe α117, Phe β78, Asp α153, Asp β114, and Met β348 (Lys β400 in Av1), which correspond to residues implicated in Fe-protein binding in Av1, are also shown. The α and β subunits are represented by the thick and thin lines, respectively. (b) Stereoview of the polypeptide environment in the putative Fe-protein cluster binding site (C), near the P-cluster pair. (c) Stereoview of the polypeptide environment in the putative Fe-protein subunit binding site (A) on the α subunit, near the FeMo-cofactor. (d) Stereoview of the polypeptide environment in the putative Fe-protein subunit binding site (B) on the β subunit.

REFERENCES

Bishop, P. E., Hawkins, M. E., & Eady, R. R. (1986) *Biochem. J. 238*, 437–442.

Bishop, P. E., et al. (1988) in *Nitrogen Fixation: Hundred Years After* (Bothe, H., deBruijn, R. J., & Newton, W. E., Eds.) pp 71–79, Gustav Fischer, Stuttgart.

Bolin, J. T., Ronco, A. E., Mortenson, L. E., Morgan, T. V., Williamson, M., & Xuong, N.-H. (1990) in *Nitrogen Fixation: Achievements and Objectives* (Gresshoff, P. M., Roth, L. E., Stacey, G., & Newton, W. E., Eds.) pp 117–124, Chapman and Hall, New York.

Bricogne, G. (1976) *Acta Crystallogr. A32*, 832–847.

Brünger, A. T. (1988) *J. Mol. Biol. 203*, 803–816.

Burgess, B. K. (1984) in *Advances in Nitrogen Fixation* (Veeger, C., & Newton, W. E., Eds.) pp 103–114, Martinus Nijhoff, Boston.

Burgess, B. K. (1985) in *Molybdenum Enzymes* (Spiro, T. G., Ed.) p 161, Wiley & Sons, New York.

Burgess, B. K. (1990) *Chem. Rev. 90*, 1377–1406.

Burris, R. H. (1991) *J. Biol. Chem. 266*, 9339–9342.

Carson, M., & Bugg, C. E. (1986) *J. Mol. Graphics 4*, 121–122.

Case, D. A., & Karplus, M. (1979) *J. Mol. Biol. 132*, 343–368.

Chan, M. K., Kim, J., & Rees, D. C. (1993) *Science 260*, 792–794.

Crowther, R. A. (1972) in *The Molecular Replacement Method* (Rossmann, M. G., Ed.) pp 173–178, Gordon & Breach, New York.

Crowther, R. A., & Blow, D. M. (1967) *Acta Crystallogr. 23*, 544–548.

Deits, T. L., & Howard, J. B. (1990) *J. Biol. Chem. 265*, 3859–3867.

Deng, H., & Hoffmann, R. (1993) *Angew. Chem., Intl. Ed. Engl.* (in press).

Eady, R. R. (1991) *Adv. Inorg. Chem. 36*, 77–102.

Emerich, W. D., & Burris, R. H. (1978) *J. Bacteriol. 134*, 936–943.

Georgiadis, M. M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. J., & Rees, D. C. (1992) *Science 257*, 1653–1659.

Govenzensky, D., & Zamir, A. (1989) *J. Bacteriol. 171*, 5729–5735.

Guth, J. H., & Burris, R. H. (1983) *Biochemistry 22*, 5111–5122.

Hadfield, K. L., & Bulen, W. A. (1969) *Biochemistry 8*, 5103–5108.

Hageman, R. V., & Burris, R. H. (1978) *Biochemistry 17*, 4117–4124.

Hogle, J., Kirchhausen, T., & Harrison, S. C. (1983) *J. Mol. Biol. 171*, 95–100.

Holm, R. H., & Simhon, E. D. (1985) in *Molybdenum Enzymes* (Spiro, T. G., Ed.) Chapter 1, Wiley-Interscience, New York.

Holm, R. H., Ciurli, S., & Weigel, J. A. (1990) *Prog. Inorg. Chem. 38*, 1.

Hoover, T. R., Imperial, J., Liang, J., Ludden, P. W., & Shah, V. K. (1988) *Biochemistry 27*, 3647–3652.

Hoover, T. R., Imperial, J., Ludden, P. W., & Shah, V. K. (1989) *Biochemistry 28*, 2768–2771.

Huynh, B. H., Henzel, M. T., Christner, J. A., Zimmermann, R., Orme-Johnson, W. H., & Münck, E. (1980) *Biochim. Biophys. Acta 623*, 124–138.

Ioannidis, I., & Buck, M. (1987) *Biochem. J. 247*, 287–291.

Jones, T. A. (1985) *Methods Enzymol. 115*, 151–171.

Jones, T. A., & Thirup, S. (1986) *EMBO J. 5*, 819–822.

Kim, J., & Rees, D. C. (1992a) *Science 257*, 1677–1682.

Kim, J., & Rees, D. C. (1992b) *Nature 360*, 553–560.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Kurtz, D. M., McMillan, R. S., Burgess, B. K., Mortenson, L. E., & Holm, R. H. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4986–4989.

Lowe, D. J., & Thorneley, R. N. F. (1983) *Biochem. J. 215*, 393–405.

Lowe, D. J., & Thorneley, R. N. F. (1984) *Biochem. J. 224*, 895–901.

Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992) *Nature 356*, 83–85.

Miller, R. W., & Eady, R. R. (1989) *Biochem. J. 263*, 725–729.

Morgan, T. V., Mortenson, L. E., McDonald, J. W., & Watt, G. D. (1987) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 46*, 2241.

Newton, W. E. (1992) in *Biological Nitrogen Fixation* (Stacey, G., Burris, R. H., & Evans, H. J., Eds.) p 877, Chapman and Hall, New York.

Orme-Johnson, W. H. (1985) *Annu. Rev. Biophys. Biophys. Chem. 14*, 419–459.

Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., & North, A. C. T. (1960) *Nature 185*, 416–422.

Rivera-Ortiz, J. M., & Burris, R. H. (1975) *J. Bacteriol. 123*, 537.

Robinson, A. C., Dean, D. R., & Burgess, B. K. (1987) *J. Biol. Chem. 262*, 14327–14332.

Rossmann, M. G. (1972) *The Molecular Replacement Method*, Gordon and Breach, New York.

Rossmann, M. G., & Blowe, D. M. (1962) *Acta Crystallogr. 15*, 24–31.

Shah, V. K., & Brill, W. J. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3249–3253.

Simpson, F. B., & Burris, R. H. (1984) *Science 224*, 1095–1097.

Smith, B. E., & Eady, R. R. (1992) *Eur. J. Biochem. 205*, 1–15.

Stiefel, E. I., et al. (1988) in *Metal Clusters in Proteins* (Que, L., Jr., Ed.) ACS Symposium Series 372, pp 372–389, American Chemical Society, Washington, DC.

Terwilliger, T. C., Kim, S.-H., & Eisenberg, D. S. (1987) *Acta Crystallogr. A43*, 1–5.

Thorneley, R. N. F., & Lowe, D. J. (1985) in *Molybdenum Enzymes* (Spiro, T. G., Ed.) pp 221–284, Wiley & Sons, New York.

Thorneley, R. N. F., Asby, G. A., Fisher, K., & Lowe, D. J. (1993) in *Molybdenum Enzymes, Cofactors and Models* (Stiefel, E., Coucouvanis, D., & Newton, W. E., Eds.) American Chemical Society, Washington, DC (in press).

Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489–501.

Tso, M.-Y. W. (1974) *Arch. Microbiol. 99*, 71–80.

Varghese, J. N., Laver, W. G., & Colman, P. M. (1983) *Nature 303*, 35–40.

Wang, S.-Z., Chen, J.-S., & Johnson, J. L. (1988) *Biochemistry 27*, 2800–2810.

Weininger, M. S., & Mortenson, L. E. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 378–380.

Weston, M. F., Kotake, S., & Davis, L. C. (1983) *Arch. Biochem. Biophys. 225*, 809–817.

Willing, A., & Howard, J. B. (1990) *J. Biol. Chem. 265*, 6596–6599.

Wolle, D., Kim, C.-H., Dean, D., & Howard, J. B. (1992) *J. Biol. Chem. 267*, 3667–3673.

Yamane, T., Weininger, M. S., Mortenson, L. E., & Rossmann, M. G. (1982) *J. Biol. Chem. 257*, 1221–1223.