Crystallographic Structures and Functional Implications of Nitrogenase Molybdenum-Iron Proteins from *Azotobacter vinelandii* and *Clostridium pasteurianum*

Thesis by
Jongsun Kim

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1993
(submitted May 4, 1993)

To my family

Acknowledgments

I would like to first and most foremost thank my research advisor, Doug Rees, for not only his support, encouragement, advice, but for also allowing me to do this research as independently as possible. If it were not for his understanding and patience, this research would not have been possible. I would also like to thank all the members of the Rees laboratory for their support and friendship. I am especially grateful to Dr. Millie Georgiadis for teaching me how to purify nitrogenase proteins and for providing me with cell cultures; Debbie Woo for her support and for providing me with the MoFe-protein from *C. pasteurianum*; Dr. Mike Chan for collecting a higher resolution native data set; Barbara Hsu for her support and for helping me with data collection and data processing; Dr. Hiromi Komiya for teaching me the fine art of model building and the evaluation of electron density maps; Art Chirino for introducing me to the X-PLOR and RIBBONS programs; and Mike Day and Jamie Schlessman for their friendship and contributions to the nitrogenase work.

Discussions with Professor H. B. Gray, Professor J. E. Bercaw, Professor J. B. Howard, Professor R. H. Holm, Dr. Mike Chan, Mike Day and Michael Stowell, as well as the assistance of Kyoung-Hee Kim, Tim McPhillips, Dr. Leemor Joshua-Torr, Dr. G. Santillan, Xiaotian Zhu, Dave Malerba and Phoebe Ray are gratefully appreciated. I would also like to thank Mike Day, Barbara Hsu and Professor D. C. Rees for reading my thesis carefully.

And most importantly, I would like to thank my wife Eunsook and my parents, brother and sisters, whose continued and unwavering support made the completion of my graduate work possible. I would also like to thank the members of the Caltech Korean community for their friendship and encouragement.

Abstract

Three-dimensional structures of the nitrogenase molybdenum-iron (MoFe-) proteins from *Azotobacter vinelandii* and *Clostridium pasteurianum* have been determined by X-ray crystallography. The structure of MoFe-protein from *A. vinelandii* (Av1) was determined at 2.7Å by the method of multiple isomorphous replacement (MIR) and noncrystallographic symmetry (NCS) averaging both within and between crystal forms. The Av1 model has been refined to a crystallographic R factor of 19% with good geometry. The root mean square (rms) deviation of bond lengths and bond angles are 0.016Å and 3.3°, respectively. The structure of MoFe-protein from *C. pasteurianum* (Cp1) was determined at 3.0Å by a combination of molecular replacement, single isomorphous replacement (SIR) and NCS averaging both within and between crystal forms. The Cp1 model has been refined to a crystallographic R factor of 18% with good geometry. The rms deviation of bond lengths and bond angles are 0.018Å and 3.9°, respectively.

The MoFe-protein, which is an $\alpha_2\beta_2$ tetramer with a total molecular weight of ~240 kD, contains two types of metal centers: the FeMo-cofactor and the P-cluster pair. The FeMo-cofactor is believed to represent the site of substrate reduction and the P-cluster pair may function in electron transfer between iron (Fe-) protein and the FeMo-cofactor. The FeMo-cofactor contains two clusters of composition 4Fe:3S and 1Mo:3Fe:3S that are bridged by three non-protein ligands. Two of the bridging ligands are assigned as sulfurs, while the chemical identity of the "Y" ligand is still ambiguous, but it could be sulfur. The Fe-Fe distance between bridged iron sites average ~2.5Å, suggesting that there may be some iron-iron bonding interactions, which could contribute the fourth coordination for the bridging irons. Ignoring the partial iron-iron bonding interactions between bridged irons, six of the seven Fe atoms in the FeMo-cofactor have trigonal coordination geometry, are coordinatively unsaturated, and are potential sites for N_2 activation. The N_2 binding site in FeMo-cofactor may be relevant to the H_2 binding site in the H-cluster of Fe-hydrogenases

and the O_2 binding site in the Mn-center of PSII. Homocitrate, an essential component of FeMo-cofactor, is coordinated through a hydroxyl and carboxyl oxygen to the molybdenum site. The FeMo-cofactor is attached to the α subunit through two protein ligands, Cys α275 and His α442. The P-cluster pair consists of two 4Fe:4S clusters that are bridged by two cysteine thiol ligands and a disulfide bond between two of the cluster sulfurs. The P-cluster pair is attached at the interface between the α and β subunits through seven protein ligands: Cys α62, Cys α88, Cys α154, Cys β70, Cys β95, Cys β153 and Ser β188. The structure of the P-cluster pair indicates that the P-cluster pair can act as a two-electron redox group, involving cleavage and reformation of the μ_3-disulfide bridge coupled to the transfer of electrons into the FeMo-cofactor. This disulfide bond may also provide a site for H_2 evolution.

The α and β subunits in the $\alpha_2\beta_2$ MoFe-protein tetramer exhibit similar polypeptide folds consisting of three domains of α/β type with some extra helices. The α and β subunits of MoFe-protein are related by an approximate two-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. Docking studies between the Fe-protein and MoFe-protein suggest a possible interaction mode between the two proteins that involves the surface of the MoFe-protein near the approximate two-fold axis passing through the P-cluster pair, and the surface of the Fe-protein near the 4Fe:4S cluster. The overall dimensions of the $\alpha_2\beta_2$ MoFe-protein tetramer are ~70Å x 80Å x 110Å. The two $\alpha\beta$ subunit pairs are related by a two-fold NCS axis. Even though the α and β subunits in an $\alpha\beta$ subunit pair are also approximately related by a two-fold rotation, the MoFe-protein does not exhibit 222 symmetry. The MoFe-protein tetramer interface is stabilized by packing of helices primarily provided by two β subunits, with some contribution from the α subunit, and further stabilized by divalent cation binding.

The FeMo-cofactor is buried at least 10Å below the protein surface. No permanent channels between the protein surface and the FeMo-cofactor are present, however, there are two potential clefts which could be utilized for substrate entry/product release and/or H_3O^+ transport. The protein environment of the FeMo-cofactor indicates that there are multiple potential transfer pathways. The P-cluster pair is also buried about 12Å below the protein surface and the environment of the P-cluster pair is primarily provided by hydrophobic residues. The edge-edge distance of the FeMo-cofactor to the P-cluster pair is about 14Å. Four helices are oriented in parallel between the two metal centers and could play a role in electron transfer. In particular, the helices α63-74 and α88-92 provide the most direct structural connection between the P-cluster pair and FeMo-cofactor.

The structure of Cp1, including the two types of metal centers associated with the protein, are similar to that of Av1. Unique features of the Cp1 structure arise from the presence of a ~50 residue insertion in the α subunit and a ~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 Fe-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.

Table of Contents

Chapter 6. Similarities with Other Electron Transfer Systems

List of Illustrations

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Chapter 1

Chapter 2

Chapter 4

Chapter 5

Chapter 1. General Introduction to the Nitrogenase Enzyme System

Nitrogen is an essential constituent of vital macromolecules such as proteins and nucleic acids. Although dinitrogen is abundant in the earth's atmosphere, it is essentially inert at room temperature in the absence of a suitable catalyst. In the nitrogen cycle, several groups of bacteria reduce nitrate to dinitrogen - a process which amounts to a net loss of biologically usable nitrogen. Loss of biologically usable nitrogen into the atmosphere would bring life on this planet to an end as the result of nitrogen starvation, but nitrogen fixation replenishes the nitrogen cycle with biologically usable nitrogen.

Dinitrogen requires either a strong reagent (molten potassium or red hot magnesium) or a special catalyst, as well as anoxic anhydrous conditions, to react chemically. In some cases, the inert nature of dinitrogen is due to thermodynamic considerations; for example, all the nitrogen oxides have positive standard free energies of formation and thus are intrinsically unstable with respect to N_2 and O_2. In other instances, however, the inert character of N_2 is a result of kinetic factors, since it reacts quite slowly with many reagents. The formation of ammonia from N_2 and H_2 is such an example. The activation energy required to reduce N_2 is very large, as has been clearly shown in the industrial fixation of nitrogen by the Haber-Bosch process in which N_2 and H_2 are reacted at temperatures between 300-500°C and pressures over 300 atm using Fe-based catalysts. This creates a paradox which nature solved ingeniously at the outset of life. In fact, only a small group of microorganisms is capable of utilizing dinitrogen for biological processes. These organisms represent the oldest 'cells' and are called diazotrophs[1]. All other organisms depend directly or indirectly upon diazotrophs for their supply of nitrogenous compounds that are subsequently utilized for the synthesis of nucleic acids, proteins, etc. via complex arrays of biochemical pathways. It is impressive to note that diazotrophs annually add

about 60% of the earth's newly fixed nitrogen, while industrially fixed nitrogen contributes only about 25% of the bulk (the other 15% come from lightening, UV radiation, etc.)[1]. In other words, diazotrophs are the chief agent for global soil fertility and help maintain the nitrogen cycle in the earth's biosphere.

Although diazotrophs display a wide spectrum of habitats that range from free forms to association with various plants[2,3] (Table 1-1 and 1-2), they all utilize the same machinery for N_2 fixation which is accomplished by the nitrogenase enzyme system. The nitrogenase enzyme system consists of two metallo-protein components, the iron (Fe-) protein and the molybdenum-iron (MoFe-) protein (reviewed in references 4-11). In the depletion of molybdenum, alternative nitrogenase system homologous to the molybdenum-containing "conventional" nitrogenase system may be induced (reviewed in references, 12-13). Nitrogenase catalyzes not only the reduction of dinitrogen to ammonia, but also the reduction of protons to hydrogen, which appears to be an obligatory part of nitrogen fixation[14], and the reduction of small unsaturated molecules such as acetylene, azide or cyanide[15]. Substrate reduction by nitrogenase involves three basic types of electron transfer reactions: (i) the reduction of Fe-protein by electron carriers such as flavodoxin or ferredoxin *in vivo* and dithionite *in vitro*; (ii) transfer of single electrons from Fe-protein to MoFe-protein in a Mg-ATP dependent process; and (iii) electron transfer to the substrate, which is almost certainly bound to the active site within the MoFe-protein. The overall stoichiometry of the nitrogenase-catalyzed reaction under optimal conditions has been established as[16]:

$$N_2 + 8H^+ + 8e^- + 16\text{Mg-ATP} \rightarrow 2NH_3 + H_2 + 16\text{Mg-ADP} + 16P_i$$

The kinetics of dinitrogen reduction have been extensively studied[17-19] and show that nitrogenase is a relatively slow enzyme, with a turnover time per electron of ~5sec^{-1}. Each electron transfer reaction between Fe-protein and MoFe-protein involves an obligatory cycle of association and dissociation of the protein complex, with the dissociation step having been identified as rate determining for the overall reaction[17-19].

Table 1-1. List of diazotrophs: free-living nitrogen-fixers

Genus or type	Species
Obligate aerobes	
Azotobacter	*vinelandii, chroococcum, paspali, beijerinckii*
Beijerinckia	*indica*
Derxia	*gummosa*
Azotococcus	*agilis*
Obligate aerobes that fix nitrogen only at low oxygen tensions	
Azospirillum	*brasilense, lipoferum*
Xanthobacter	*autotrophicus, flavus*
Thiobacillus	*ferro-oxidans*
Rhizobium	*cowpea group*
Methylosinus	*sporium*
Methylococcus	*capsulatus*
Facultative anaerobic bacteria that fix nitrogen only under anaerobic conditions	
Klebsiella	*pneumoniae*
Bacillus	*polymyxa, macerans*
Propionibacterinum	*shermanii, petersonii*
Escherichia	*intermedia*
Citrobacter	*freundii*
Enterobacter	*cloacae, agglomerans*
Erwinia	*herbicola*
Obligate anaerobes	
Clostridium	*pasteurianum, butyricum*
Desulfovibrio	*gigas, desulfuricans, vulgaris*
Desulfotomaculum	*ruminis*
Phototrophic bacteria	
Rhodospirillum	*rubrum, tenue, fulvum, molischianum, photometricum*
Rhodopseudomonas	*palustries, viridis, capsulata, spherioides*
Chromatium	*vinosum, gracile, minus, violacea*
Thiocystis	*violacea*
Thiocapsa	*roseopersicina, pfennigii*
Amoebobacter	*roseus*
Ectothiorhodospira	*spaposhnikovii*
Blue-green algae or Cyanobacteria	
Gloeothece	*alpicola*
Aphanothece	
Nostoc	*muscorum, commune*
Anabaena	*cylindrica, variabilis*
Aphanizomenon	*floe-aquae*
Cylindrosperum	various strains
Calothrix	various strains
Plectonema	*boryanum*
Oscillatoria	various strains
Pseudanbaena	various strains
Lyngbya	various strains
Phormidium	various strains

Table 1-2. List of diazotrophs: Symbiotic nitrogen-fixing systems and associations

Host family	Host genus	N_2-fixing microorganism
Rhizobium symbioses		
Leguminosae	Most species	*Rhizobium spp.*
Ulmaceae	*Parasponia*	*Rhizobium*
Non-rhizobium symbiosis		
Betulaceae	*Alnus*	
Myricaceae	*Myrica, Comptonia*	
Eleagenaceae	*Eleagnus, Hippophae, Shepherdia*	
Rhamnaceae	*Ceanothus,*	*Actinomycetes*
	Trevoa, Discaria	*Frankia*
Rosaceae	*Dryas, Cercocarpus, purshia*	
Coriariaceae	*Coriaria, Colleta*	
Casuarinaceae	*Casuarina*	
Lichens	*Collema, Peltiger,*	
	Dendriscocaulon	*Nostoc scytonema*
Liverworts	*Anthoceros, Blasia,*	
	Cavicularia	*Nostoc sphaericum*
Waterfern	*Azola*	*Anabaena azollae*
Cycads	*Cycads, Ceratozamia,*	*Nostoc, Anabaena*
	Encephalartos, Macrozamia,	
	Dioon etc.	
Higher plants		
Haloragaceae	*Gunnera*	*Nostoc punctiforme*
Associative symbioses and casual associations		
Phyllosphere	Leaves	*Azotobacter spp.*
Rhizosphere	Roots of grasses	
	Paspalum notatum, Zea mays	*Azotobacter paspali,*
		Azospirillum

Studies on the structure and function of nitrogenase proteins have benefited significantly from information provided by studies on the molecular genetics of nitrogen fixation. The genetics of nitrogen fixation is complex (Figure 1-1)[10]. In *Klebsiella pneumoniae*, the most intensively studied diazotroph and one for which there is no evidence for the presence of alternative nitrogenases, 20 *nif* genes are required solely for nitrogen fixation and its regulation and they are clustered on a 23-kb region of the chromosome. The entire *nif* gene cluster has been sequenced[20]. In addition to the structural genes of nitrogenase (*nifHDK*), genes involved in the maturation of nitrogenase Fe-protein (*nifM*), FeMo-cofactor biosynthesis (*nifBQVSENH*), electron donation (*nifJF*), several genes of unknown function (*nifTYXUWZ*) and the regulatory genes (*nifAL*) are located in this cluster[21,22] (Figure 1-1). *Azotobacter vinelandii* and *Azotobacter chroococcum* have a comparable gene cluster, except that they contain additional open reading frames of unknown function and *nifABQ* genes outside the major *nif* gene cluster. The *Azotobacter* nitrogen fixation genes associated with alternative nitrogenases have also been cloned and sequenced[23]. They include the structural genes of the Vanadium-nitrogenase and, in *A. vinelandii* only, the structural genes of the Iron-nitrogenase. In addition to these structural genes, specific regulatory *nifA*-like genes and a reiteration of *nifEN*-like genes have been identified.

The Fe-protein is a dimer of two identical subunits, which is encoded by the *nifH* gene, with a total molecular weight of ~60 kD. The amino acid sequence of over 20 different Fe-proteins have been determined, and indicate that this protein is highly conserved in both the conventional and alternative nitrogenases[24]. The Fe-protein dimer contains one 4Fe:4S cluster, which is coordinated to each subunit through two cysteine residues[25-27]. Unlike the cluster in ferredoxins, the 4Fe:4S cluster in Fe-protein is exposed to the solvent[27]. The Fe-protein dimer has two nucleotide binding sites for Mg-ADP and/or Mg-ATP located in the cleft formed between the two subunits[27-28]. In the absence of MoFe-protein and reductant, Fe-protein exhibits no catalytic ATPase activity[28]. Changes in the EPR spectra[29] and the ability of α,α-dipyridyl to chelate iron from the 4Fe:4S cluster[30] after

nucleotide binding indicate that nucleotide binding is accompanied by changes in the Fe-protein structure. Since the phosphate groups of this nucleotide are ~20Å from the 4Fe:4S cluster, it is unlikely that ATP hydrolysis and electron transfer are directly coupled[27]. Instead, it appears that interactions between the nucleotide and cluster sites must be indirectly coupled by allosteric changes occurring at the subunit interface. In addition to its mechanistic role in nitrogenase enzymatic function, Fe-protein also participates at several stages in the biosynthesis of the nitrogenase proteins. Fe-protein is essential for the production of active MoFe-protein, and is involved in both the synthesis of FeMo-cofactor and its insertion into cofactor deficient MoFe-protein[31-33]. Fe-protein may also function as an activator for the expression of alternative nitrogenases[34]. The significant sequence conservation observed in the Fe-protein family may reflect the structural constraints associated with these diverse aspects of Fe-protein function.

The MoFe-protein is an $\alpha_2\beta_2$ tetramer with a total molecular weight of ~240 kD. The α and β subunits, which are encoded by the *nifD* and *nifK* genes, are of similar size; for example, the isolated α and β subunits of *Azotobacter vinelandii* MoFe-protein have 491 and 522 amino acids, respectively[35]. In general, the amino acid sequences of MoFe-proteins are less well conserved than are Fe-protein sequences, so that the MoFe-protein sequences from *A. vinelandii* and *Clostridium pasteurianum* are only ~36% identical[36] (Figure 1-2). The MoFe-protein contains two types of metal centers, the FeMo-cofactor and the P-cluster pair. The structures and properties of these centers have been extensively probed by a wide variety of techniques, but the structures of these metal centers were completely unknown.

The FeMo-cofactor (reviewed in references 37-38), also referred to as the M-center, was first identified by Shah and Brill[39] as a stable metallocluster isolated from acid denatured MoFe-protein. Although isolated cofactor exhibits no catalytic activities, it can fully activate a defective MoFe-protein obtained from the mutant *A. vinelandii* strain UW45 that contains P-cluster pairs, but not the FeMo-cofactor. Intense interest has been focused

on the FeMo-cofactor since it contains molybdenum in a biologically unprecedented form, and is believed to represent the site of substrate reduction[40,41]. The FeMo-cofactor has been reported to have a composition of 1Mo atom, 6-8 Fe atoms, 8-9 inorganic sulfurs and 1 homocitrate group[37,38,42-44]. The redox properties and oxidation states of the cofactor are intimately associated with the ability of this center to function in substrate reduction. Three oxidation states differing by one electron, M(r), M(s-r), and M(ox), appear to be accessible to the FeMo-cofactor within the MoFe-protein (reviewed in 38). M(s-r) is the semi-reduced form of the FeMo-cofactor that is normally isolated in the presence of excess dithionite, and is characterized by a unique S=3/2 EPR signal. The Mo may formally exhibit the +4 oxidation state, while Mössbauer studies suggest that the average oxidation state of the Fe is +2.67. The isolated cofactor has a net negative charge, however, which if this reflects the charge on the core of the protein bound cluster, would indicate that formally, the Fe must be primarily in the ferrous (+2) state. A satisfactory description of the electronic properties and oxidation states of the FeMo-cofactor remains an important problem.

The P-cluster pair (reviewed in 44), also referred to as the P-clusters, may function in electron transfer between the Fe-protein and the FeMo-cofactor. Mössbauer, EPR and extrusion studies indicated that the P-cluster pair contained 4Fe:4S type-clusters[45-50] that were in close-proximity[51-53] which was also suggested from an analysis of anomalous scattering effects of the metal centers from Cp1[54]. These studies also demonstrated, however, that the detailed properties of this cluster were distinct from better characterized proteins that contain one or more mononuclear 4Fe:4S clusters. Mössbauer studies indicated that the iron atoms in each center could be assigned to three distinct types, designated D, Fe^{2+} and S, in the approximate ratio 5:2:1. The iron isomer shifts are consistent with all irons in the dithionite reduced form of the P-cluster pair formally having the ferrous oxidation state, which is unprecedented in biological 4Fe:4S clusters. Redox titrations of the P-cluster pair have identified the oxidation states P_N, P_{Ox1}, P_{Ox2} and

higher oxidation states[53]. P_N denotes the P-cluster pair state observed in dithionite reduced MoFe-protein. The oxidation of P_N to P_{Ox} has been described as a $2e^-$ process, although evidence for intermediates generated by $1e^-$ steps has been presented[55]. More reduced forms of the P-cluster pair have not been described, although it is possible that they could be produced, perhaps transiently, under turnover conditions.

A comparison of the DNA sequences of the genes encoding the α and β subunits (*nifDK*) of MoFe-proteins of some 10 diazotrophic species reveals a strong interspecies conservation of derived amino acid sequence and predicted secondary structure (reviewed in 10 and 21). Attention has focused on the five conserved cysteine residues (Cys α62, Cys α88, Cys α154, Cys α183 and Cys α275)of the α subunit, and the three conserved cysteine residues (Cys β70, Cys β95 and Cys β153) of the β subunit. Structurally similar domains are found in both α and β subunits around Cys α62, Cys α88, Cys α154, Cys β70, Cys β95 and Cys β153. Secondary structure predictions are consistent with Cys α62, Cys α88, Cys β70, and Cys β95 having a near to surface location in the α and β subunits providing possible subunit/subunit contact regions. Cys α275 and His α195 have been proposed to have a role in the binding of FeMo-cofactor, and the other conserved cysteine residues, Cys α62, Cys α88, Cys α154, Cys β70, Cys β95 and Cys β153 have been indicated as ligands to the P-cluster pair by site-directed mutagenesis studies[56-62].

Despite the accomplishments over the past century in characterizing the biochemical, spectroscopic and genetic properties of nitrogenase, a detailed molecular description of the enzyme mechanism has been elusive, in part due to the absence of structural information. We undertook the crystallographic structure determination of the nitrogenase MoFe-protein to understand its unique properties and functions in nitrogen fixation, and to provide the necessary framework for addressing outstanding issues related to the assembly and mechanism of the nitrogenase enzyme system. The determination of a crystal structure of a gigantic oxygen-sensitive protein has provided additional challenges to those inherent in

macromolecular crystallography. This work is described in five chapters which follow. Chapter 2 describes the cell growth, purification of MoFe-proteins, crystallization of MoFe-proteins, and the details of the structure determination including data collection, heavy atom derivative preparation, multiple isomorphous replacement phase calculation, non-crystallographic symmetry averaging within and between crystal forms, model building, and refinement. The structures and functions of the metal centers in MoFe-protein are described in Chapter 3, and the protein structure and functional implications of MoFe-protein from *A. vinelandii* (Av1) are discussed in Chapter 4. Chapter 5 includes the three-dimensional structure of the MoFe-protein from *C. pasteurianum* (Cp1), structural comparisons of Av1 and Cp1, and functional and mechanistic implications of MoFe-protein. Structural similarities between nitrogenase and other electron transfer systems, including hydrogenases, plant photosynthetic system II (PSII), and bacterial photosynthetic reaction center are discussed in Chapter 6.

References

1. Burns, R. C. & Hardy, R. W. F. (1975) *Nitrogen Fixation in Bacteria and Higher Plants*, Springer-Verlag, Berlin.

2. Brill, W. J. (1977) *Scientific America* **236**, 68-81.

3. Bothe, H., Yates, M. G. & Cannon, F. C. (1983) *Encyclopedia of Plant and Physiology New series* (Lauchli, A & Bieleski, R. L., eds.) Vol. **15A**, pp. 241-285. Springer-Verlag, Berlin.

4. Ansari, S. A., Khan, F. A. & Khan, N. A. (1989) *Sci. Prog. Oxf.* **73**, 287-300.

5. Burgess, B. K. (1984) in *Advances in Nitrogen Fixation* (Veeger, C. and Newton, W. E., eds.) p. 103, Martinus Nijhoff, Boston.

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

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

8. Stiefel, E. I., Thomann, H., Jin, H., Bare, R. E., Morgan, T. V., Burgmayer, S. J. N. & Coyle, C. L. (1988) in *Metal clusters in Proteins* (Que, L. eds.) p. 372. American Chemical Society, Washington, D. C.

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

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

11. Stacey, G. Burris, R. H. & Evans, H. J., ed. (1992) *Biological Nitrogen Fixation*, Chapman & Hall, New York.

12. Bishop, P. E. & Joerger, R. D. (1990) *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **41**, 109-125.

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

14. Hadfield, K. L. & Bulen, W. A. (1969) *Biochem.* **8**, 5103-5108.

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

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

17. Lowe, D. J. & Thorneley, R. N. F. (1983) *Biochem. J.* **215**, 393-403.

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

19. Hageman, R. V. & Burris, R. H. (1978) *Biochem.* **17**, 4117-4124.

20. Arnold, W., Rump, A., Klipp, W., Priefer, U. B. & Puhler, A. (1988) *J. Mol. Biol.* **203**, 715-738.

21. Dean, D. R. & Jacobson, M. R. (1992) in *Biological Nitrogen Fixation* (Stacey, G., Evans, B. J. & Burris, R. H., eds.) pp. 763-834, Chapman & Hall, New York.

22. Elmerich, C. (1991) in *Biology and Biochemistry of Nitrogen Fixation* (Dilworth, M. J. & Glenn, A. R., eds.) pp. 103-141, Elsevier, Amsterdam, New York.

23. Robson, R. L., Woodley, P. R., Pau, R. & Eady, R. R. (1988) in *Nitrogen Fixation: Hundred Years After* (Bothe, H., de Bruijn, W. J., Newton, W. E. eds.) p. 323. Gustav

Fischer Verlag, Stuttgart; Joerger, R. D., Jacobson, M. R., Premakumar, R., Wolfinger, E. D. & Bishop, P. E. (1989) *J. Bacteriol.* **171**, 1075-1086.

24. Normand, P. & Bousquet, J. (1989) *J. Mol. Evol.* **29**, 436-447.

25. Hausinger, R. P. & Howard, J. B. (1983) *J. Biol. Chem.* **258**, 13486-13492.

26. Howard, J. B., Davis, R., Moldenhauer, B., Cash, V. L. & Dean, D. R. (1989) *J. Biol. Chem.* **264**, 11270-11274.

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

28. Yates, M. G. (1992) in *Biological Nitrogen Fixation* (Stacey, G., Burris, R. H. & Evans, H. J., eds.) p. 685, Chapman & Hall, New York.

29. Watt, G. D., Wang, Z. -C. & Knotts, R. R. (1986) *Biochem.* **25**, 8156-8162.

30. Deits, T. L. & Howard, J. B. (1989) *J. Biol. Chem.* **264**, 6619-6628.

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

32. Robinson, A. C., Chu, W., Li, J. -G. & Burgess, B. K. (1989) *J. Biol. Chem.* **264**, 10088-10095.

33. Tal, S., Chun, T. W., Gavini, N. & Burgess, B. K. (1991) *J. Biol. Chem.* **266**, 10654-10657.

34. Joerger, R. D., Wolfinger, E. D. & Bishop, P. E. (1991) *J. Bacteriol.* **173**, 4440-4446.

35. Brigle, K. E., Newton, W. E. & Dean, D. R. (1985) *Gene* **37**, 37-44.

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

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

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

39. Shah, V. K. & Brill, W. J. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3249-3253.

40. Hawkes, T. R., McLean, P. A. & Smith, B. E. (1984) *Biochem. J.* **217**, 317-321.

41. Imperial, J., Hoover, T. R., Madden, M. S., Ludden, P. W. & Shah, V. K. (1989) *Biochem.* **28**, 7796-7799.

42. Nelson, M. J., Levy, M. A. & Orme-Johnson, W. H. (1983) *Proc.Natl.Acad. Sci. USA* **80**, 147-150.

43. Hoover, T. R., Robertson, A. D., Cerny, R. L., Hayes, R. N., Imperial, J., Shah, V. K. & Ludden, P. W. (1987) *Nature* **329**, 855-857.

44. Holm, R. H., Ciurli, S. & Weigel, J. A. (1990) *Prog. Inorg. Chem.* **38**, 1-74.

45. Smith, B. E. & Lang, G. (1974) *Biochem. J.* **137**, 169-180.

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

47. Zimmermann, R., Münck, E. Brill, W. J., Shah, V. K., Henzl, M. T., Rawlings, J., Orme-Johnson, W. H. (1978) *Biochem. Biophys. Acta* **537**, 185-207.

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

49. McLean, P. A., Papaefthymiou, V., Orme-Johnson, W. H. & Münck, E. (1987) *J. Biol. Chem.* **262**, 12900-12903.

50. Lindahl, P. A., Papaefthymiou, V., Orme-Johnson, W. H. & Münck, E. (1988) *J. Biol. Chem.* **263**, 19412-19418.

51. Hagen, W. R., Wassink, H., Eady, R. R., Smith, B. E. & Haaker, H. (1987) *Eur. J. Biochem.* **169**, 457-465.

52. Surerus, K. K., Hendrich, M. P., Christie, P. D., Rottgardt, D., Orme-Johnson, W. H. & Münck, E. (1992) *J. Am. Chem. Soc.* **114**, 8579-8590.

53. Pierik, A. J., Wassink, H., Haaker, H., Hagen, W. R. (1993) *Eur. J. Biochem.* **212**, 51-61.

54. 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.) p. 117. Chapman & Hall, New York.

55. Tittsworth, R., & Hales, B. J. (1992) *9th International Congress on Nitrogen Fixation*, Cancún, Mexico, Abs. #307.

56. Kent, H. K., Ioannidis, I., Gormal, C., Smith, B. E. & Buck, M. (1989) *Biochem. J.* **264**, 257-264.

57. Kent, H. M., Baines, M., Gormal, C., Smith, B. E. & Buck, M. (1990) *Mol. Microbiol.* **4**, 1497-1504.

58. Scott, D. J., May, H. D., Newton, W. E., Brigle, K. E. & Dean, D. R. (1990) *Nature* **343**, 188-190.

59. Dean, D. R., Setterquist, R. A., Brigle, K. E., Scott, D. J., Laird, N. F. & Newton, W. E. (1990) *Mol. Microbiol.* **4**, 1505-1512.

60. May, H. D., Dean, D. R. & Newton, W. E. (1991) *Biochem. J.* **277**, 457-464.

61. Thomann, H., Morgan, T. V., Jin, H., Burgmayer, S. J. N., Bare, R. E. & Stiefel, E. I. (1987) *J. Am. Chem. Soc.* **109**, 7913-7914.

62. Thomann, H., Bernardo, M., Newton, W. E. & Dean, E. R. (1991) *Proc. Natl. Acad. Sci. USA.* **88**, 6620-6623.

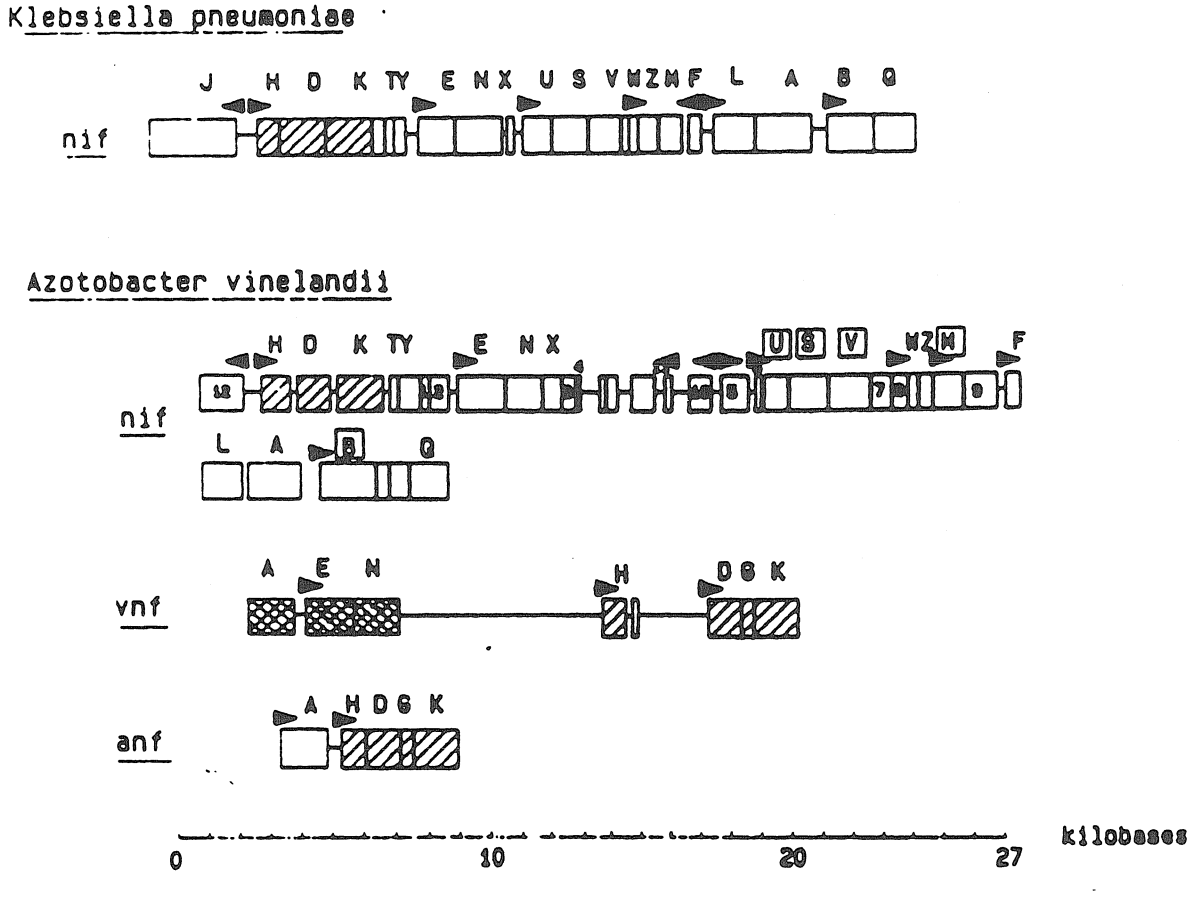


Figure 1-1. The organization of nitrogen fixation genes. *nif* encodes conventional nitrogenase, *vnf* encodes Vanadium-nitrogenase and *anf* encodes the third (Iron only-) nitrogenase.

Figure 1-2a. Amino acid sequence alignment of the alpha subunit of MoFe-proteins. Avnifd:MoFe-protein from Azotobacter vinelandii; Avvnfd:VFe-protein from A. vinelandii; Avanfd:FeFe-protein from A. vinelandii; Acvnfd: VFe-protein from A. chroococcum; Annifd: MoFe-protein from Anabaena 7120; Cpnifd:MoFe-protein from Clostridium pasteurianum; Kpnifd:MoFe-protein from Klebsiella pneumoniae; Prnifd:MoFe-protein from Parasponia Rhizobium; Tfnifd:Thiobacillus ferrooxidans; Mtnifd:MoFe-protein from Methanococcus thermolithotrophicus; Rcnifd:MoFe-protein from Rhizobium cowpea; Rjnifd: MoFe-protein from Rhizobium japonicum.

```
         1                                                  50
Avnifd   .......... .MTGMSREEV ESLIQEVLEV YPEKARKDRN KHLAVNDPAV
Avvnfd   .......... .......... ..MPMVLLEC ..DKDIPERQ KHIYLKAPNE
Avanfd   .......... .......... ..MPHHEFEC ..SKVIPERK KHAVIKGKGE
Acvnfd   .......... .......... ..MPMVLLEC ..DKDIPERQ KHIYLKAPNE
Annifd   .......MTP PENKNLVDEN KELIQEVLKA YPEKSRKKRE KHLNVHEENK
Cpnifd   .......... ........VS ENLKDEILEK YIPKTKKTRS GHIVIKTEET
Kpnifd   .......... MMTNATGERN LALIQEVLEV FPETARKERR KHMMVSDPKM
Prnifd   MSLATTQSIA EIRARN.... KELIEEVLKV YPEKTAKRRA KHLNVHQAGK
Tfnifd   MSISAEDLST QPQRRKLPEI AELIDETLKA YPEKFAKRRA KHLNVYEEGK
Mtnifd   .......... .......... ..MPYILLDC ..DKFIPERM KHTYVYDPEE
Rcnifd   ....MAKDHA GGPEDLERLV RDLIAEVLEA YPAKAQKKRA KHLSVAGATS
Rjnifd   MSLATTNSVA EIRARN.... KELIEEVLKV YPEKTAKRRA KHLNVHQAGK

         51                                                 100
Avnifd   TQSK...... ..KCIISNKK SQPGLMTIRG CAYAGSKGVV WGPIKDMIHI
Avvnfd   DTREFLP... .....IANAA TIPGTLSERG CAFCGAKLVI GGVLKDTIQM
Avanfd   TLADALP... .....QGYLN TIPGSISERG CAYCGAKHVI GTPMKDVIHI
Acvnfd   DTREFLP... .....IANAA TIPGTLSERG CLL.RRKLVI GGVLKDTIQM
Annifd   SDCG...... ....VKSNIK SVPGVMTARG CAYAGSKGVV WGPIKDMIHI
Cpnifd   PNPE...... ....IVANTR TVPGIITARG CAYAGCKGVV MGPIKDMVHI
Kpnifd   KSVG...... ..KCIISNRK SQPGVMTVRG CAYAGSKGVV FGPIKDMAHI
Prnifd   SDCG...... ....VKSNIK SIPGVMTIRG CAYAGSKGVV WGPIKDMVHI
Tfnifd   SECD...... ....VKSNIK SVPGVMTIRG CAYAGSYGVV WSPVKDMIHI
Mtnifd   ...NILP... .....ACNTN TVPGDMTERG CAFAGSRGVV GGPIKDAIHM
Rcnifd   EDADASRHRV EMRHVKSNIK SVPGVLTIRG CAYAGSKGVV WGPIKDMVHI
Rjnifd   SDCG...... ....VKSNIK SIPGVMTIRG CAYAGSKGVV WGPIKDMVHI

         101                                                150
Avnifd   SHGPVGCGQY SRAGRRNYYI GTTGVN.AFV TMNFTSDFQE KDIVFGGDKK
Avvnfd   IHGPLGCAYD TWHTKR.YPT DNGHFNMKYV ...WSTDMKE SHVVFGGEKR
Avanfd   SHGPVGCTYD TWQTKR.YIS DNDNFQLKYT ...YATDVKE KHIVFGAEKL
Acvnfd   IHGPLGCAYD TWHTKR.YPT DNGHFNMKYV ...WSTDMKE SHVVFGGEKR
Annifd   SHGPVGCGYW SWSGRRNYYV GVTGIN.SFG TMHFTSDFQE RDIVFGGDKK
Cpnifd   THGPIGCSFY TWGGRRFKSK PENGTGLNFN EYVFSTDMQE SDIVFGGVNK
Kpnifd   SHGPAGCGQY SRAERRNYYT GVSGVD.SFG TLNFTSDFQE RDIVFGGDKK
Prnifd   SHGPVGCGQY SWGSRRNYYV GTTGVD.SFV TLQFTSDFQE KDIVFGGDKK
Tfnifd   SHGPVGCGHY ARAGRRAYYI GTTGVD.TYT TMHFTSDFQE KDIVFGGDKK
Mtnifd   VHGPIGCAYY TWGTRR.ALS DNE.FHRRYC ...FCTDMQE SDIVYGGEKT
Rcnifd   SHGPVGCGTY SWSQRRNYYT GKTGVD.SFV TMQFTTDFQE KDIVFGGDKK
Rjnifd   SHGPVGCGQY SWGSRRNYYV GTTGID.SFV TLQFTSDFQE KDIVFGGDKK

         151                                                200
Avnifd   LAKLIDEVET LFPLNKGISV QSECPIGLIG DDIESVSKVK GAEL.SKTIV
Avvnfd   LEKSMHEAFD EMPDIKRMIV YTTCPTALIG DDIKAVAKKV MKDRPDVDVF
Avanfd   LKQNIIEAFK AFPQIKRMTI YQTCATALIG DDINAIAEEV MEEMPEVDIF
Acvnfd   LEQRMHEAFD EMPDIKRMIV YTTCPTALIG DDIKAVAKKV MKERPDVDVF
Annifd   LTKLIEELDV LFPLNRGVSI QSECPIGSIG DDIEAVAKKT SKQI.GKPVV
Cpnifd   LKDAIHEAYE MFH.PAAIGV YATCPVGLIG DDILAVAATA SKEI.GIPVH
Kpnifd   LSKLIEEMEL LFPLTKGITI QSECPVGLIG DDISAVANAS SKAL.DKPVI
Prnifd   LIKVLDEIQE LFPLNNGITI QSECPIGLIG DDIEAVSRSK SKEYGGKTIV
Tfnifd   LAKLMDELEE LFPMSKGITV QSECPIGLIG DDIEAVSKKK AAEF.GKPVV
Mtnifd   LEKASLEVME EFPEASGTFI YTTCPTALIG DNVDAIARNI EKATKKPAI.
Rcnifd   LEKTIDEINE LFPLSKGITI QSECPIGLIG DDIEAVSKKK NKEI.NKTIV
Rjnifd   LDKILDEIQE LFPLNNGITI QSECPVGLIG DDIEAVSRAK SKEYGGKTIV
```

```
          201                                                       250
Avnifd    PVRCEGFRGV SQSLGHHIAN DAVRDWVL.. .GK..R..DE DTTFASTPYD
Avvnfd    TVECPGFSGV SQSKGHHVLN IGW..... ...INEKVET MEKEITSEYT
Avanfd    VCNSPGFAGP SQSGGHHKIN IAW..... ...INQKVGT VEPEITGDHV
Acvnfd    TVECPGFSGV SQSKGHHVLN IGW..... ...INEKVET MEKEITSEYT
Annifd    PLRCEGFRGV SQSLGHHIAN DAIRDWIFPE YDKLKK..ET RLDFEPSPYD
Cpnifd    AFSCEGYKGV SQSAGHHIAN NTVMTDIIGK GNK..... ....EQKKYS
Kpnifd    PVRCEGFRGV SQSLGHHIAN DVVRDWIL.. .NN..R..EG Q.PFETTPYD
Prnifd    PVRCEGFRGV SQSLGHHIAN DAVRDWIF.. .DKLEP..EG EPKFQPTPYD
Tfnifd    PNRCEGFRGV SQSLGHHIAN DSIRDWVL.. .DP.AA..DK HPDFESTPYD
Mtnifd    AINSPGFCGV SQSKGHHVFN MTFYKWLKLK RKEFPEK..C MPEEEKTPYD
Rcnifd    PVRCEGFRGV SQSLGHHIAN DAVRDWIF.. .EQPES..EA TKAFEPGPYD
Rjnifd    PVRCEGFRGV SQSLGHHIAN DAVRDWIF.. .GHIEA..EG KPKFEPTPYD

          251                                                       300
Avnifd    VAIIGDYNIG GDAWSSRILL EEMGLRCVAQ WSGDGSISEI ELTPKVKLNL
Avvnfd    MNFIGDFNIQ GDTQLLQTYW DRLGIQVVAH FTGNGTYDDL RCMHQAQLNV
Avanfd    INYVGEYNIQ GDQEVMVDYF KRMGIQVLST FTGNGSYDGL RAMHRAHLNV
Acvnfd    MNFIGDFNIQ GDTQLLQTYW DRLGIQVVAH FTGNGTYDDL RCMHQAQLNV
Annifd    VALIGDYNIG GDAWASRMLL EEMGLRVVAQ WSGDGTLNEL IQGPAAKLVL
Cpnifd    INVLGEYNIG GDAWEMDRVL EKIGYHVNAT LTGDATYEKV QNADKADLNL
Kpnifd    VAIIGDYNIG GDAWASRILL EEMGLRVVAQ WSGDGTLVEM ENTPFVKLNL
Prnifd    VAIIGDYNIG GDAWSSRILL EEMGLRVIAQ WSGDGSLAEL EATPKAKLNI
Tfnifd    VTLLGDYNIG GD.WGSRIIL EEMGLRVIAQ WSGDAPSRSS TASSKSKLNL
Mtnifd    VALIGDYNMD WDVAVIKPLL EKIGCRYVTT FTGNASLDEL FQLMDVKLNI
Rcnifd    VNIIGDYNIG GDAWASPILL EEIGLNVI.. WSGDATLAEM ERAPKAKLNL
Rjnifd    VAIIGDYNIG GDAWSSRILL EEMGLRVIAQ WSGDGSLAEL EATPKAKLNI

          301                                                       350
Avnifd    VHCYRSMNYI SRHMEEKYGI PWMEYNFFGP TKTIESLRAI AAKFDE.SIQ
Avvnfd    VNCARSSGYI ANELKKRYGI PRLDIDSWGF NYMAEGIRKI CAFFG..I.E
Avanfd    LECARSAEYI CNELRVRYGI PRLDIDGFGF KPLADSLRKI GMFFG..I.E
Acvnfd    VNCARSSGYI ANELKKRYGI PRLDIDSWGF SYMAEGIRKI CAFFG..I.E
Annifd    IHCYRSMNYI CRSLEEQYGM PWMEFNFFGP TKIAASLREI AAKFDS.KIQ
Cpnifd    VQCHRSINYI AEMMETKYGI PWIKCNFIGV DGIVETLRDM AKCFDDPELT
Kpnifd    VHCYRSMNYI ARHMEEKHQI PWMEYNFFGP TKIAESLRKI ADQFDD.TIR
Prnifd    LHCYRSMNYI SRHMEEKFGI PWCEYNFFGP SKIAESLRKI AGYFDD.KIK
Tfnifd    LHCYRSVNYI TRHMEEKYGI PYIEFNFFGP TKIKESLRQI AAFFDE.SIQ
Mtnifd    VHCQRSAEYI AQMINDGFDI PYTRATFFGL SDIAESLYDV AKALD..LPK
Rcnifd    IHCYRSMNYI CRYMEEKYSI GWMEYNFFGP TQIEASLR.. IGKFDE.TIQ
Rjnifd    LHCYRSMNYI SRHMEEKFGI PWCEYNFFGP SKIADSLRRI AGYFDD.KIK

          351                                                       400
Avnifd    KKCEEVIAKY KPEWEAVVAK YRPRLEGKRV MLYIGGLRPR HVIGAYE.DL
Avvnfd    EKGEELIAEE YAKWKPKLDW YKERLQGKKM AIWTGGPRLW HWTKSVEDDL
Avanfd    DRAKAIIDEE VARWKPELDW YKERLMGKKV CLWPGGSKLW HWAHVIEEEM
Acvnfd    EKGERLIAEE YAKWKPKLDW YKERLQGKKM AIWTGGPRLW HWTKSVEDDL
Annifd    ENAEKVIAKY TPVMNAVLDK YRPRLEGNTV MLYVGGLRPR HVVPAFE.DL
Cpnifd    KRTEEVIAEE IAAIQDDLDY FKEKLQGKTA CLYVGGSRSH TYMNMLK.SF
Kpnifd    ANAEAVIARY EGQMAAIIAK YRPRLEGRKV LLYMGGLRPR HVIGAYE.DL
Prnifd    EGAERVIEKY QPLVDAVIAK YRPRLEGKTV MLYVGGLRPR HVIGAYE.DL
Tfnifd    EKAEKAIAKY QPQWDAVVEK FRPRLEGKKV MLFVGGLRPG HTIGAFE.DL
Mtnifd    ERVDQVIKEE MEAIQPKLDY YKSKLEGKTC MVYVGGPRTW HWIKAM.KDL
Rcnifd    ANVEKVIAKY RPLVDGILAK YKPRLEGKSV MLYVRP.RSA PRRHAYD.DL
Rjnifd    EGAERVIEKY QPLVDAVIAK YRPRLEGKTV MLYVGGLRPR HVIGAYE.DL

          401                                                       450
Avnifd    GMEVVGTGYE FAHNDDYD.. .......... .......... ..........
Avvnfd    GVQVVAMSSK FGHEEDFE.. .......... .......... ..........
Avanfd    GLKVVSVYIK FGHQGDME.. .......... .......... ..........
Acvnfd    GIQVVAMSSK FGHEEDFE.. .......... .......... ..........
Annifd    GIKVVGTGYE FAHNDDYK.. .......... .......... ..........
Cpnifd    GVDSLVAGFE FAHRDDYEGR EVIPTIKIDA DSKNIPEITV TPDEQKYRVV
Kpnifd    GMEIIAAGYE FAHNDDYD.. .......... .......... ..........
Prnifd    GMEVVGTGYE FGHNDDYQ.. .......... .......... ..........
Tfnifd    GMEVIGTGYE FGHNDDYQ.. .......... .......... ..........
```

```
Mtnifd  GVEYVAACCT FSHTDDYE.. .......... .......... ..........
Rcnifd  GMVIAGTGYE FAHNDDYK.. .......... .......... ..........
Rjnifd  GMDVIGTGYE FGHNDDYQ.. .......... .......... ..........

        451                                                  500
Avnifd  .......... .........R TMKE.MGDST LLYDDVTGYE FEEFVKRIKP
Avvnfd  .......... ......KVIA RGKEGT.... YYIDDGNELE FFEIIDLVKP
Avanfd  .......... ......KGIA RCGEGT.... LAIDDPNELE GLEALEMLKP
Acvnfd  .......... ......KVIA RGKEGT.... YYIDDGNELE FFEIIDLVKP
Annifd  .......... .........R TT.HYIDNAT IIYDDVTAYE FEEFVKAKKP
Cpnifd  IPEDKVEELK KAGVPLSSYG GMMKEMHDGT ILIDDMNHHD MEVVLEKLKP
Kpnifd  .......... .........R TLPD.LKEGT LLFDDASSYE LEAFVKALKP
Prnifd  .......... .........R TAQHYVKDST LIYDDVNGYE FERFVEKVQP
Tfnifd  .......... .........R TTHE.IKGNT LIYDDVTGYE FEKFAEKLRP
Mtnifd  .......... ......KMNK NFKEAGIKDI LVIDAPNEPE LEEAVKTLDP
Rcnifd  .......... .........R TG.HYVKEGT .IYDDVTGYE LEKFIEKIRP
Rjnifd  .......... .........R TAQHYVKDST LIYDDVNGYE FERFVERLQP

        501                                                  550
Avnifd  DLIGSGIK.E KFIFQKMGIP FREMHSWDYS GPYHGF...D GFAIFARDMD
Avvnfd  DVIFTGPRVG E.LVKKLHIP YVNGHGYH.N GPYMGF...E GFVNLARDMY
Avanfd  DIILTGKRPG E.VAKKVRVP YLNAHAYH.N GPYKGF...E GWVRFARDIY
Acvnfd  DVIFTGPRVG E.LVKKLHIP YVNGHGYH.N GPYMGF...E GFVNLARDTY
Annifd  DLIASGIK.E KYVFQKMGLP FRQMHSWDYS ELGDGVQMSD EVRFFCEGRK
Cpnifd  DMFFAGIK.E KFVIQKGGVL SKQLHSYDYN GPYAGFR... GVVNFGHELV
Kpnifd  DLIGSGIK.E KYIFQKMGVP FRQMHSWDYS GPYHGY...D GFAIFARDMD
Prnifd  DLVGSGIK.E KYVFQKMGVP FPEMHSWDYS GPYHGY...D GFAIFARDMD
Tfnifd  DLVASGVK.E KYIFQKMGFP FRQMHSWDYS GPYHGP...D GFAIFARDMD
Mtnifd  DFMLVGLKER Y.LFRKYGVP TINSHSYE.E GPYAGY...R GFVNFARDVY
Rcnifd  DLVASGIKEE KYPVQKMGIP FRQMHSWDYS GPYHGY..WA SPILLDAPWD
Rjnifd  DLVGSGIK.E KYVFQKMSVP FRQMHSWDYS GPYHGY...D GFAIFARDMD

        551                                                  600
Avnifd  MTLNNPCWKK LQ.APWEASE GAEKVAASA* .......... ..........
Avvnfd  NAVHNPLRHL AAVDI..... ..RDKSQTTP VIVRGAA*.. ..........
Avanfd  NAIYSPIHQL SGIDITKDNA PEWGNGFRTR QMLSDGNLSD AVRNSETLRQ
Acvnfd  NAVHNPLRHL AAVDI..... ..RDSSQTTP VIVRGAA... ..........
Annifd  KSLFL..... .......... .......... .......... ..........
Cpnifd  NGIYTPAWKM ITPPWKKASS ESKVVVGGEA *......... ..........
Kpnifd  MTLNNPAWNE LT.APWLKSA *LPTHCPVCS PICGAGGEHH ..........
Prnifd  MAVNSPIWKK TK.APWKEAA KPKLLAAE*. .......... ..........
Tfnifd  MAVNNPVWGL TQ.APWK*.. .......... .......... ..........
Mtnifd  KAVCHPVWNV LKEGEDKFKN FKGDLNE... .......... ..........
Rcnifd  KPHRASSWKI TPAAPWKAAS EA*....... .......... ..........
Rjnifd  MAVNSPIWKR TK.APWKDAE RQDSRLQNNA TRLALRESPG IPI*......

        601                    625
Avnifd  .......... .......... .....
Avvnfd  .......... .......... .....
Avanfd  YTGGYDSVSK LREREYPAFE RKVG*
Acvnfd  .......... .......... .....
Annifd  .......... .......... .....
Cpnifd  .......... .......... .....
Kpnifd  .......... .......... .....
Prnifd  .......... .......... .....
Tfnifd  .......... .......... .....
Mtnifd  .......... .......... .....
Rcnifd  .......... .......... .....
Rjnifd  .......... .......... .....
```

Figure 1-2b. Amino acid sequence alignment of the betta subunit of MoFe-proteins. Avnifk:MoFe-protein from Azotobacter vinelandii; Avvnfk:VFe-protein from A. vinelandii; Avanfk:FeFe-protein from A. vinelandii; Acvnfk: VFe-protein from A. chroococcum; Annifk:MoFe-protein from Anabaena 7120; Cpnifk:MoFe-protein from Clostridium pasteurianum; Kpnifk:MoFe-protein from Klebsiella pneumoniae; Prnifk:MoFe-protein from Parasponia Rhizobium; Tfnifk:MoFe-protein from Methanococcus thermolithotrophicus; Frnifk:MoFe-protein from Frankia

```
         1                                                  50
Avnifk   MSQQVDKIKA SYPLFLDQDY KDMLAKKRDG FEEKYPQDKI DEVFQWTTTK
Avvnfk   .......... .......... .......... ........MS NCELTVLKPA
Avanfk   .......... .......... .......... .......... .......MTC
Acvnfk   .......... .......... .......... ........MS NCELTVLKPA
Annifk   MPQNPERTVD HVDLFKQPEY TELFENKRKN FEGAHPPEEV ERVSEWTKSW
Cpnifk   .......... .......... .......... MLDATPKEIV E.........
Kpnifk   MSQTIDKINS CYPLFEQDEY QELFRNKRQ. LEEAHDAQRV QEVFAWTTTA
Prnifk   MAQSADHVLD HLELFRGPEY QQMLADK.KM FENPRDPAEV ERIRAVTKTP
Tfnifk   MSQNADKIVD HFNLFKQPDY QEMFKNKQKT FENRLPADQV ARGQEWTKTW
Frnifk   MSQTIDKINS CYPLFEQDEY QELVRNKRQ. LEERHDAQRV QEVFAWTTTA
         51                                                 100
Avnifk   EYQELNFQRE ALTVNPAKAC QPLGAVLCAL GFEKTMPYVH GSQGCVAYFR
Avvnfk   EVKLSPRDRE GI.INPMYDC QPAGAQYAGI GIKDCIPLVH GGQGCTMFVR
Avanfk   EVK..EKGRV GT.INPIFTC QPAGAQFVSI GIKDCIGIVH GGQGCVMFVR
Acvnfk   EVKLVKRERE GI.INPMYDC QPAGAQYAGI GVKDCIPLVH GGQGCTMFVR
Annifk   DYREKNFARE ALTVNPAKGC QPVGAMFAAL GFEGTLPFVQ GSQGCVAYFR
Cpnifk   ........RK ALRINPAKTC QPVGAMYAAL GIHNCLPHSH GSQGCCSYHR
Kpnifk   EYEALNFQRE ALTVDPAKAC QPLGAVLCSL GFANTLPYVH GSQGCVAYFR
Prnifk   EYREKNFA.E ALAVNPAKAC QPLGAVFVSV GFEGTLPFVH GSQGCVAYYR
Tfnifk   EYREKNFARE ALSVNPDKAC QPLGRIFAR. GFDGTRPFVH GSQGCVAYFR
Frnifk   EYEALNFH.E ALTVDPAKAC HALGAVLCSL GFANTLPYVH GSQGCVAYFG
         101                                                150
Avnifk   SYFNRHFREP VSCVSDSMTE DAAVFGGQQN MKDGLQNCKA TY.KPDMIAV
Avvnfk   LLFAQHFKEN FDVASTSLHE ESAVFGGAKR VEEGVLVLAR RYPNLRVIPI
Avanfk   LIFSQHYKES FELASSSLHE DGAVFGACGR VEEAVDVLLS RYPDVKVVPI
Acvnfk   LLFAQHFKEN FDVASTSLHE ESAVFGGAKR VEEGVLVLAR RYPELRLIPI
Annifk   THLSRHYKEP CSAVSSSMTE DAAVFGGLNN MIEGMQVSYQ LY.KPKMIAV
Cpnifk   TVLSRHFKEP AMASTSSFTE GASVFGGGSN IKTAVKNIFS LY.NPDIIAV
Kpnifk   TYFNRHFKEP IACVSDSMTE DAAVFGGNNN MNLGLQNASA LY.KPEIIAV
Prnifk   SHLSRHFKEP SSCVSSSMTE DAAVFGGLNN MIDGLANSYN MY.KPKMIC.
Tfnifk   SHFNRHFKEP SSCVSSSMTE DPAVFGGLNN MIDGLAISYS LY.KPKMIAV
Frnifk   TYFNRHFKEP IACVSDSMTE DAAVFGGNNN MNLGLQNASA LY.KPEIIAV
         151                                                200
Avnifk   STTCMAEVIG DDLNAFINNS KK....EGFI PDEFPVPFAH TPSFVGSHVT
Avvnfk   ITTCSTEVIG DDIEGSIRVC NRALEAE.FP DRKIYLAPVH TPSFKGSHVT
Avanfk   ITTCSTEIIG DDVDGVIKKL NEGLLKEKFP DREVHLIAMH TPSFVGSMIS
Acvnfk   ITTCSTEVIG DDIEGTINVC NRALAAE.FP ERKIYLAPVH TPSFKGSHVT
Annifk   CTTCMAEVIG DDLGAFITNS KN....AGSI PQDFPVPFAH TPSFVGSHIT
Cpnifk   HTTCLSETLG DDLPTYISQM ED....AGSI PEGKLVIHTN TPSYVGSHVT
Kpnifk   STTCMAEVIG DDLQAFIANA KK....DGFV DSSIAVPHAH TPSFIGSHVT
Prnifk   STTCMAEVIG DDLNAFIKTS KE....KGSV RRS.STPFAH TPAFVGSHVT
Tfnifk   STTCMAEVIG DDLNAFIKTA KE....KGNV PESFDVPFAH TPSFVGSHIT
Frnifk   STTCMAEVIG DDLQRFIANA KK....DGFV DSSIAWPHAH TPSFIGSHVT
         201                                                250
Avnifk   GWDNMFEGIA RYFTLKSMDD KVVG...SNK KINIVPGFET YL..GNFRVI
Avvnfk   GYAECVKSVF KTITDAHGKG Q......PSG KLNVFPGW.. .VNPGDVVLL
Avanfk   GYDVAVRDVV RHFAK...RE A......PND KINLLTGW.. .VNPGDVKEL
Acvnfk   GYAECVKSMF KTITEVHGKG Q......PSG KLNVFPGW.. .VNPGDVVLL
Annifk   GYDNMMKGIL SNLTEGKKKA T......SNG KINFIPGFDT Y..VGNNREL
Cpnifk   GFANMVQGIV NYLSENTGAK ........NG KINVIPGF.. .VGPADMREI
Kpnifk   GWDNMFEGFA KTFTADYQGQ PGKL...P.. KLNLVTGFET YL..GNFRVL
Prnifk   GYDNALKGIL EHFWNGKAGT APKLERKPNE AINIIGGFDG N.TVGNLREI
Tfnifk   GYDNMMKGIL THFWDGKAGT VPALERKPDE KINFIGGFDG Y.TVGNMREI
Frnifk   GWDNMFEGFA KTFTADYQGQ PGKL...PKL KLNLVPGFET YLGTGNFRVL
         251                                                300
Avnifk   KRMLSEMGVG YSLLSDPEEV LDTPADGQFR MYA.GGTTQE EMKDAPNALN
```

```
Avvnfk  KRYFKEMDVE ANIYMDTED. FDSPMLPNKS IETHGRTTVE DIADSANALA
Avanfk  KHLLGEMDIE ANVLFEIES. FDSPILPDGS AVSHGNTTIE DLIDTGNARA
Acvnfk  KRYFKEMGVD ATVFMDTED. FDSPMLPNKS IETHGRTTVE DIADSANALA
Annifk  KRMMGVMGVD YTILSDSSDY FDSPNMGEYE MYP.SGTKLE DAADSINAKA
Cpnifk  KRLFEAMDIP YIMFPDTSGV LDGPTTGEYK LYPEGGTKIE DLKDTGNSDL
Kpnifk  KRMMEQMAVP CSLLSDPSEV LDTPADGHYR MYS.GGTTQQ EMKEAPDAID
Prnifk  KRILALMGIK HTILADNSEV FDTPTDGEFR MYD.GGTHVE DTANAIHAKA
Tfnifk  KRLFSLMNVD YTILGDGSDV WDTPADGEFR MYD.GGTTFA EAEAALNAKA
Frnifk  KRMMEQMAVP CSLLSDPSEV LDTPADGQYR MYS.GGTTQQ EMKEAPDAID
        301                                                  350
Avnifk  TVLLQPWHLE KTKKFVEGTW KHEVPKLNIP MGLDWTDEFL MKVSEISGQP
Avvnfk  TLSLARYEGN TTGELLQKTF AVPNALVNTP YGIKNTDDML RKIAEVTGKE
Avanfk  TFALNRYEGT KAAEYLQKKF EIPAIIGPTP IGIRNTDIFL QNLKKATGKP
Acvnfk  TLALARYEGA TTGEYLEKTF AVPNSLVNTP YGIKNTDDML RKIAEITGKE
Annifk  TVALQAYTTP KTREYIKTQ. WKQETQVLRP FGVKGTDEFL TAVSELTGKA
Cpnifk  TLSLGSYASD LGAKTLEKKC KVPFKTLRTP IGVSATDEFI MALSEATGKE
Kpnifk  TLLLQPWQLL KSKKVVQEMW NQPATEVAIP LGLAATDELL MTVSQLSGKP
Prnifk  TISMQQWCTE KTLPFVSEH. GQDVVSFNYP VGVSATDDLL VALSRISGKE
Tfnifk  TVCMQGISTE KTMAYIQEK. GQEVVALHCP IGVTGTDHFL QEVSGISGKP
Frnifk  TLLLQPWQLL KSKKVVQEMW NQPATEVAIP LGLAATDELL MTVSQLSGKP
        351                                                  400
Avnifk  IPASLTKERG RLVDMMTD.S HTWLHGKRFA LWGDPDFVMG LVKFLLELGC
Avvnfk  IPESLVRERG IALDALADLA HMFFANKKVA IFGHPDLVLG LAQFCMEVEL
Avanfk  IPQSLAHERG VAIDALADLT HMFLAEKRVA IYGAPDLVIG LAEFCLDLEM
Acvnfk  IPESLVREPR IAWIALADLA HMFFANKKVA IFGHPDLVLG LAQFCLEVEL
Annifk  IPEELEIERG RLVDAITD.S YAWIHGKKFA IYGDPDLIIS ITSFLLEMGA
Cpnifk  VPASIEEERG QLIDLMID.A QQYLQGKKVA LLGDPDEIIA LSKFIIELGA
Kpnifk  IADALTLERG RLVDMMLD.S HTWLHGKKFG LYGDPDFVMG LTRFLLELGC
Prnifk  IPEQLARERG RLVDAIAD.S SAHIHGKKFA IYGDPDLCYG LAAFLLELGA
Tfnifk  ISEELKKERG RLVDAIGT.S ISYLHGKKFA IKGDPDFCLG VAGFLLELGA
Frnifk  IADALTLERG RLVDMMLD.S HTWLHGKKFG LYGDPDFVMG LTRFLLELGC
        401                                                  450
Avnifk  EPVHIL.CHN GNKRWKKAVD AILAASPYGK NATVYIGKDL WHLRSLVFTD
Avvnfk  EPVLLLIGDD QGNKYKKDPR IEELKNTAHF DIEIVHNADL WELEKRI.NA
Avanfk  KPVLLLLGDD .NSKYVDDPR IKALQENVDY GMEIVTNADF WELENRIKNE
Acvnfk  EPVLLLIGDD QGSKYKKDPR LQELKDAAHF DMEIVHNADL WELEKRI.ND
Annifk  EPVHIL.CNN GDDTFKKEME AILAASPFGK EAKVWIQKDL WHFRSLLFTE
Cpnifk  IPKYVV.TGT PGMKFQKEID AMLAEAGIE. GSKVKVEGDF FDVHQWIKNE
Kpnifk  EPTVIL.SHN ANKRWQKAMN KMLDASPYGR DSKVFINCDL WHFRSLMFTR
Prnifk  EPTHVL.STN GNNV.AGENA TLFAGSPFG. ELPAYPGRDL WHMRSLLFTE
Tfnifk  EPVHVV.CTR GNKDWAEKMN ALFASSPFGT GCHAYPGKDL WHMRSLVFTE
Frnifk  EPTVIL.SHN ANKRWQKAMN KMLDASPNRA RYSVFINCDL WHFRSLMFTR
        451                                                  500
Avnifk  ..KPDFMIGN SYGKFIQRDT LHKGKEFEVP LIRIGFPIFD RHHLHRSTTL
Avvnfk  GLQLDLIMGH SKGRYVAIEA .......NIP MVRVGFPTFD RAGLYRKPSI
Avanfk  GLELDLILGH SKGRFISIDY .......NIP MLRVGFPTYD RAGLFRYPTV
Acvnfk  GLQLDLIMGH SKGRYVAIEA .......NIP MVRVGFPTFD RAGLYRKPSI
Annifk  ..PVDFFIGN SYGKYLWRDT .......SIP MVRIGYPLFD RHHLHRYSTL
Cpnifk  ..GVDLLISN TYGKFIAREE .......NIP FVRFGFPIMD RYGHYYNPKV
Kpnifk  ..QPDFMIGN SYGKFIQRDT LAKGKAFEVP LIRLGFPLFD RHHLHRQTTW
Prnifk  ..PVDFLIGN THGKYLERDT .......GTP LIRIGFPIFD RHHHHRFPVW
Tfnifk  ..PVDFLIGN SYTKYLERDT .......GTP MIHIGFPIHD RHHHHRYPIW
Frnifk  ..QPDFMIGN SYGKFIQRVT LAKGKAFEVP LIRLGFPLFD RHHLHRQTTW
        501                                          539
Avnifk  GYEGAMQILT TLVNSILERL DEETRGMQAT DYNHDLVR*
Avvnfk  GYQGAMELGE MIANAMFAHM EYTRNKEWIL NTW*.....
Avanfk  GYGGAIWLAE QMANTLFADM EHKKNKEWVL NVW*.....
Acvnfk  GYQGAMELGE MIANAMFAHM EYTRNKEWIL NTW......
Annifk  GYQGGLNILN WVVNTLLDEM DRSTNITGKT DISFDLIR.
Cpnifk  GYKGAIRLVE EITNVILDKI EREC.....T EEDFEVVR*
Kpnifk  GYEGAMNIVT TLVNAVLEKL DSDTSQLGKT DYSFDLVR.
Prnifk  GYQGGLNVLV KILDKIFDEI DKKTSVLGKT DYSFDIIR*
Tfnifk  GYQGAVNVLV WILDRIHLEL DRNTLGIGTT DFSYDLVR*
Frnifk  GYEGAMNIVT TLVNAVLEKL DSDTSQLAKT DYSFDLVR*
```

Chapter 2. Structure Determination of Nitrogenase Molybdenum-Iron protein from *Azotobacter vinelandii*.

2-1. Cell growth

Azotobacter vinelandii strain ATCC 13705 was obtained from the American Type Culture Collection (ATCC). The ATCC cells were started in liquid cultures of modified Burk's nitrogen free minimal media[1] and grown at 30°C. The media used for all cell growths contained the following: 1.67 mM $MgSO_4$, 36.0 μM Fe_2SO_4, 2.0 μM Na_2MoO_4, 0.9 mM $CaCl_2$, 7.75 mM K_2HPO_4, 2.45 mM KH_2PO_4, 36.0 μM sodium citrate, 2% (w/v) sucrose. Large-scale cell cultures were grown in a 20 L carboy or a 180 L fermenter (New Brunswick model I250) at the UCLA fermentation facility[2].

Clostridium pasteurianum W5, ATCC catalog number 6013, was rehydrated in 1 ml of nitrogen free medium containing 1 mM $MgSO_4$, 0.2 mM Fe_2SO_4, 20μM Na_2MoO_4, 0.2 mM sodium citrate, 0.5 mM $CaCl_2$, 135 mM K_2HPO_4, 15 mM KH_2PO_4, 10 ng/ml biotin and 2% (w/v) sucrose. The cell slurry was then inoculated into potato medium[3] in test tubes and incubated at 37°C until growth was observed by the evolution of hydrogen gas. Cells were inoculated from the potato media into 125 ml nitrogen-free medium to initiate liquid culture. Large-scale cell cultures were grown in a 180 L fermenter at the UCLA fermentation facility[4].

2-2. Preparation of MoFe-proteins

The nitrogenase MoFe-protein from *A. vinelandii* was purified anaerobically as described by Burgess *et al.*[5] with a few modifications. Since the nitrogenases are extremely oxygen sensitive, all the experiments and the manipulation of crystals must be done under strictly anaerobic conditions. Approximately 200 g of frozen *A. vinelandii* cells

were lysed by osmotic shock, and the lysate was centrifuged for 40 min at 20,000g. The supernatant was heated to 55°C and incubated for 5 min at this temperature to precipitate heat unstable proteins (the nitrogenase proteins are somewhat heat stable), and the mixture was centrifuged again for 30 min at 12,000g. The supernatant was loaded on a DEAE-Sepharose CL-6B ion exchange column equilibrated with a loading buffer containing 50 mM Tris·HCl, pH 7.75, 0.1M NaCl, and the MoFe-protein and Fe-protein were eluted with a gradient of 0.1 - 0.6 M NaCl in 50 mM Tris.HCl, pH 7.75. The MoFe-protein fraction was diluted with an equal volume of no salt buffer (50 mM Tris·HCl, pH 8.0) and loaded on a second DEAE ion exchange column equilibrated with the loading buffer. The MoFe-protein was eluted with a more gradual salt gradient (0.1 - 0.4 M NaCl in 50 mM Tris·HCl, pH 8.0). The MoFe-protein fraction was dialyzed against a low-salt buffer (0.025M Tris.Hcl, pH 7.4, containing 0.028 M NaCl) and simultaneously concentrated using the Micro-Prodicon vacuum dialysis concentrator (Bio-Molecular Dynamics, Oregon). As the solubility of MoFe-protein is low in the low-salt buffer[6], MoFe-protein is selectively precipitated by the above method. The concentrated sample was heated to 38°C in a water bath for 1 hour and centrifuged at 20,000g for 20 min at 38°C. The supernatant was discarded and the pellet was resuspended in 50 mM Tris·HCl, pH 8.0 containing 0.35 M NaCl. The final concentration of MoFe-protein was ~20 mg/ml and it was used for crystallization experiments without further manipulation. The MoFe-protein was sometimes stored in liquid nitrogen in a plastic serum vial.

The MoFe-protein from *C. pasteurianum* was purified anaerobically as described by Nelson *et al.*[7] with a few modifications. About 300 g of frozen cells were lysed anaerobically under an atmosphere of 20% CO in Ar by addition of buffer containing 200 mg of lysozyme, 200 mg of DNase I in 1 L 50 mM Tris·HCl, pH 8.5. After 1.5 hr at 37°C the cell debris was pelleted by centrifugation, and the supernatant was loaded onto a DEAE-Sepharose CL-6B column equilibrated with a loading buffer containing 50 mM Tris·HCl, pH 7.5. After loading, the column was washed with an additional 1 L of the loading buffer

and the nitrogenase proteins were eluted subsequently with a high-salt buffer containing 50 mM Tris·HCl, pH 7.5, 0.45 M NaCl. Subsequently, the nitrogenase fraction was diluted with two volumes of no salt buffer (50 mM Tris·HCl, pH 7.5) and was loaded onto the second DEAE-Sepharose CL-6B column equilibrated with the loading buffer. The MoFe-protein and Fe-protein were eluted with a 2 liter linear gradient of 0.15 - 0.6 M NaCl in 50 mM Tris·HCL, pH 7.5. The MoFe-protein fraction[8] was concentrated with a Micro-Prodicon vacuum dialysis concentrator in the nitrogen tent. The concentrated sample was loaded on a Sephacryl S-300 sizing column equilibrated with 50 mM Tris·HCl, pH 8.0, and the MoFe-protein was eluted with the same buffer. The MoFe-protein fraction was again concentrated with a Micro-Prodicon and the concentrated sample was used for crystallization experiments. The final concentration of MoFe-protein was ~20 mg/ml.

2-3. Crystallizations and preliminary diffraction studies

The MoFe-protein from *A. vinelandii* (Av1) was crystallized in 1.5 mm melting point capillaries by the microcapillary batch method[9,10] from five different conditions (Table 2-1): (1) 18% polyethylene glycol (PEG) 4000, 0.14 M NaCl, 0.10 M Na_2MoO_4, 80 mM Tris·HCl, pH 8.0, and ~8 mg/ml protein (crystal form designated Av1); (2) 16% PEG 4000, 0.14 M NaCl, 0.2 M $MgCl_2$, 80 mM Tris·HCl, pH 8.0, and ~8 mg/ml protein (designated MgAv1); (3) 18% PEG 4000, 0.14 M NaCl, 0.3 M CsCl, 80 mM Tris·HCl, pH 8.0, and ~8 mg/ml protein (designated CsAv1); (4) 16% PEG 4000, 0.14 M NaCl, 0.12 M Ammonium phosphate, 80 mM Tris·HCl, pH 8.0, and ~8 mg/ml protein (designated ApAv1); and (5) 16% PEG 4000, 0.14 M NaCl, 0.3 M Sodium citrate, 80 mM Tris·HCl, pH 8.0, and ~8 mg/ml protein (designated CitAv1). A typical Av1 crystal is shown in Figure 2-1. The Av1 crystals are plate-shaped with maximum dimensions of 1.5 mm x 1.0 mm x 0.2 mm. MoFe-protein crystals are brown due to the presence of the Fe-S clusters. MoFe-protein is hard to manipulate because it is very oxygen sensitive; however, the crystals are relatively easy to handle because they are large and colored.

Table 2-1. Crystallization conditions of Av1. The Av1 was crystallized by the micro-capillary batch method[9,10].

Crystal form	Crystallization condition	Space group	Unit cell constants
Av1	PEG/MoO_4	$P2_1$	a=108.4Å, b=130.5Å, c=81.5Å, β=110.8°
MgAv1	PEG/$MgCl_2$	$P2_1$	same as above
CsAv1	PEG/CsCl	$P2_1$	same as above
APAv1	PEG/$(NH_4)_2PO_4$	$P2_1$ or P2	a=153.6Å, b=73.92Å, c=210.6Å, β=104.2°
CitAv1	PEG/Citrate	?	?

Precession photographs were taken on an Enraf Nonius precession camera to determine the space group and unit cell constants of the crystal forms. The crystal to film distance was 100 mm and the exposure time was 4-8 hour with 50 kV x 50 mA power. Figure 2-2 shows the precession pictures of the hk0 zone and 0kl zone of the Av1 crystal form. The spindle angle between the two zones was 110.8°. Similar precession photographs were obtained for the MgAv1 and CsAv1 crystal forms (data not shown). The Av1, MgAv1 and CsAv1 crystal forms belong to space group $P2_1$ with same unit cell constants; a = 108.4Å, b = 130.5Å, c = 81.5Å, β = 110.8°. Based on the Matthews coefficient[11] of 2.4$Å^3$/dalton for a typical protein crystal, they contain one $\alpha_2\beta_2$ tetramer per asymmetric unit. Only the Av1 crystal form was used for further x-ray diffraction analysis, although MgAv1 and CsAv1 crystal forms are also of sufficient quality. The Av1 crystal form diffracts at least to 2Å resolution and it lasts 2-3 days in the x-ray beam at room temperature. The ApAv1 crystal form belongs to space group P2 or $P2_1$ with unit cell constants a = 210.6Å, b = 73.9Å, c = 153.3Å, β = 104.2°, and contains two $\alpha_2\beta_2$ tetramer molecules per asymmetric unit based on the Matthews coefficient. The ApAv1 crystal form was not used for further

x-ray diffraction analysis because the crystals only diffract to ~4Å resolution. The CitAv1 crystal form was not characterized, since it was difficult to reproduce the crystallization.

The MoFe-protein from *C. pasteurianum* was also crystallized in 1.5 mm melting point capillaries by the microbatch method from three different conditions (Table 2-2): (1) 15% PEG 4000, 0.21 M $MgCl_2$, 80 mM Tris·HCl, pH 8.0, and ~8 mg/ml protein (crystal form designated Mg1); (2) 18% PEG 4000, 0.3 M CsCl, 80 mM Tris·HCl, pH 8.0, and ~8 mg/ml protein (designated Cs1); and (3) 18% PEG 4000, 0.10 M Na_2MoO_4, 80 mM Tris·HCl, pH 8.0, and ~8 mg/ml protein (designated MoCp1). Typical crystals of each crystal form are dark-brown and plate shaped with maximum dimensions of 1.0 mm x 0.5 mm x 0.2 mm.

Table 2-2. Crystallization conditions of Cp1. The Cp1 was also crystallized by the micro-capillary batch method.

Crystal form	Crystallization condition	Space group	Unit cell constants
Mg1	PEG/$MgCl_2$	$P2_1$	a=69.96Å, b=151.3Å, c=121.9Å, β=110.4°
Cs1	PEG/CsCl	P21	a=87.9Å, b=171.4Å, c=73.6Å, b=91.5Å
MoCp1	PEG/MoO_4	?	?

Precession photographs were taken on an Enraf Nonius precession camera, to determine the space group and unit cell constants of the crystal forms. The crystal to film distance was 100 mm and the exposure time was ~4 hour with 50 kV x 50 mA power. Figure 2-3 shows the precession pictures of the Mg1 and Cs1 crystal forms. The space group and the unit cell constants of the Mg1 and Cs1 crystal forms were determined by the XENGEN[15] autoindexing program with the aid of the precession photographs. The Mg1 crystal form belongs to space group $P2_1$, with unit cell constants a= 70.0Å, b = 151.3Å, c = 121.9Å, β = 110.4°, and the Cs1 crystal form belongs to space group $P2_1$, with unit cell constants a =

87.9Å, b = 171.4Å, c = 73.6Å, β = 91.5°. Each crystal form contains one $\alpha_2\beta_2$ tetramer molecule per asymmetric unit based on the Matthews coefficient. Both crystal forms diffract to ~2.5Å resolution and they can last 1-2 days in the x-ray beam at room temperature. The Mg1 crystal form is similar to a crystal form previously described[12]. The MoCp1 crystal form was not characterized, since it was difficult to reproduce the crystallization.

2-4. Crystal mounting procedure

The crystals were mounted in 1.5 mm x-ray capillaries in the anaerobic chamber. The melting point (mp) capillaries containing the crystals were scored approximately 10 mm from the top of the liquid layer in the capillary using a triangular file. The top end of the mp capillary was then detached by breaking the glass. The mp capillary was next scored approximately 1 mm from the crystal which was going to be mounted and detached as described above. The broken mp capillary which contains the crystal(s) of interest with native mother liquor was inserted into the funnel of the x-ray capillary, and the crystal was gently pushed out using a 20 µl pipetman. This procedure is possible because the standard yellow pipet tip fits into the 1.5 mm mp capillary and the mp capillary fits into the funnel of the 1.5 mm x-ray capillary. The bottom of the x-ray capillary was then cut approximately 6 cm from the funnel, filled with ~5 µl of synthetic mother liquor and sealed with wax using a soldering iron. Synthetic mother liquor was prepared by using ~110% concentrations of precipitants for each respective crystallization condition Most of the liquid around the crystal was then removed from the x-ray capillary by drawing it up in a drawn-out Pasteur pipet. Care was taken not to touch the crystal with the pipet. Filter paper wicks were used to remove the remainder of the liquid around the crystal, and a plug of synthetic mother liquor was then placed in the x-ray capillary starting approximately 1 mm below the funnel and filling about half of the funnel. The capillary was then sealed with wax.

2-5. Heavy atom derivative screening

In order to obtain phase information by the multiple isomorphous replacement (MIR) method[13], heavy atom derivatives were screened by soaking MoFe-protein crystals in heavy atom solutions. Table 2-3 shows a list of heavy atom reagents that were used in soaking experiments and the results of heavy atom derivative screening. The Av1 crystal form was derivatized by mercury compounds (EMTS and EMP) and platinum compounds (PIP, $Pt(Cl)_4$), while the Mg1 and Cs1 crystal forms were derivatized with a mercury compound (EMTS). The heavy atom solutions were prepared by dissolving heavy atom compounds in synthetic mother liquor solutions. The conventional soaking method of transferring a crystal into the heavy atom solution was not suitable for these plate-shaped MoFe-protein crystals, because the crystal usually became more brittle and fragile after derivatization by heavy atom compounds. To decrease handling of the crystal, it was mounted in an x-ray capillary first, then the heavy atom solutions were injected into the capillary. This 'inverse' method works particularly well for small crystals or plate-shaped crystals. In addition to soaking experiments, cocrystallization experiments were also done to obtain heavy atom derivatives. However, cocrystallization experiments produced crystals that were too small for data collection or that did not diffract well (Table 2-3).

Table 2-3. The results of heavy atom derivative screening. Soaking experiments and cocrystallization experiments were performed to obtain heavy atom derivative.

Heavy atom compound	Concentration	Soaking time	R-factor	Result
Gold-thiomaleate	1.5 mM	14 h	14%	not isomorphous
$IrCl_6$	5 mM	14 h	13%	not isomorphous
UO_2F_5	saturated	14 h	6.5%	not derivative
$SmAc_4$	saturated	14 h	6%	not derivative
$Pt(CN)_4$	5 mM	6 h	6 %	?

$PtCl_4$	5 mM	6 h	12.1%	good derivative
$PtCl_4$	5 mM	12 h	13.5%	good derivative
PIP	4 mM	12 h	14.0%	good derivative
EMP	1.5 mM	6 h	13.9%	good derivative
EMTS	5 mM	6 h	13.5 %	best derivative
TMLC	saturated	24 h	7%	not derivative
WO_4	100 mM	24 h	7.5%	not derivative
WSe_4	2 mM	24 h	7%	not derivative
WO_4	100 mM	cocrystallization	-	tiny crystals
WSe_4	10 mM	cocrystallization	-	tiny crystals
$Hg(CN)_2$	1 mM	cocrystallization	8%	?
$IrCl_6$	2 mM	cocrystallization	-	did not diffract well
$Pt(CN)_4$	1 mM	cocrystallization	-	did not diffract well
$AuCl_3$	1 mM	cocrystallization	-	not crystallized
$PtCl_4$	1 mM	cocrystallization	-	not crystallized

EMTS: Ethyl mercuri thiosalicylate

EMP: Ethyl mercuri phosphate

PIP: di-μ-iodobis(ethylenediamine)-di-platinum (II) nitrate

TMLC: tri-methyl lead chloride

Ac: Acetate

Preliminary screening for heavy atom derivatives was carried out using precession photography by visually comparing intensity changes after heavy atom derivatization. However, no significant intensity change was detected by this method. It might be rationalized that the MoFe-protein is so big that intensity changes brought about by the introduction of a few heavy atoms cannot be detected readily by visual inspection of

precession photographs. Based on this assumption, complete heavy atom derivative data sets were collected using the area detector, and then isomorphous difference Pattersons were calculated to identify useful derivatives.

2-6. Intensity data collection

Native data sets and heavy atom derivative data sets were collected on a Siemens X-1000 area detector with a crystal to detector distance of 14-16 cm depending on the unit cell constants. CuKα (λ = 1.5418Å) radiation was generated by a Siemens rotating anode generator with a graphite monochromator operating at 4.5 kW (50 kV x 90 mA). All data sets were collected at room temperature. The cryocrystallography techniques of Hope[14] were attempted for the Av1 and CsAv1 crystal forms, but the crystals no longer diffracted after cooling to ~161°C with a stream of cold nitrogen generated by a Simens LT-2A low temperature device.

The crystals were mounted into x-ray capillaries in a random orientation (section 2-4), and Omega scans on a Siemens 3-axis camera were used for all data collection. Data collection started from a random Omega angle and was stopped after scanning ~200° using frame widths of 0.15-0.2° depending on the mosaic spread of the crystals. The typical mosaic spread was ~0.6°. One or more additional Omega scans with different Φ-angles were carried out as necessary and as permitted by crystal life time to complete the data sets.

A high resolution Av1 native data set (MOFE27) was collected to 2.7Å resolution using two crystals, and a low resolution native data set (MOFE48) was also collected to 4.8Å resolution using two crystals (Table 2-4). The latter was collected for the purpose of using anomalous scattering for the phase determination. Systematic error in the measurement of anomalous differences may be minimized by collecting a low resolution data set with 2θ = 0° so that data for the anomalous pairs are collected on the same frame and are the same distance from the center of the detector. The MOFE27 data set is 93% complete to 2.7Å with a merging R factor of 0.085 and the MOFE48 data set is 95% complete to 4.8Å with a

merging R factor of 0.062. Four heavy atom derivative data sets were collected for the Av1 crystal form. The EMTS1 data set was collected to 3.2Å resolution using one crystal and the EMTS3 data set was collected to 2.8Å resolution using three crystals. The EMTS1 data set is 83% complete to 3.2Å with a merging R factor of 0.057 and the EMTS3 data set is 91% complete to 3.0Å with a merging R factor of 0.086. The PTCL3 and PIP3 data sets were collected to 3.0Å resolution using three crystals each. The PTCL3 data set is 89% complete to 3.0Å with a merging R factor of 0.09 and the PIP3 data set is 92% complete to 3.0Å with a merging R factor of 0.111. The data sets were autoindexed, integrated and merged with the XENGEN[15] software package. The ROCKS[16] crystallographic package was used for scaling and reduction of data sets. All of the Av1 data sets shown in Table 2-4 were locally scaled[17] to the MOFE27 native data set to decrease systematic error.

A Mg1 native data set (MgCp1) was collected to 3.0Å resolution using three crystals, and a heavy atom derivative data set (MgEMTS) was collected to 3.2Å resolution using one crystal (Table 2-4). A Cs1 native data set (CsCp1) was collected to 3.0Å resolution using three crystals, and a heavy atom derivative data set (CsEMTS) was collected to 3.2Å using one crystal (Table 2-4). The MgCp1 data set is 78% complete to 3.0Å with a merging R factor of 0.07 and the CsCp1 data set is 74% complete to 3.0Å with a merging R factor of 0.084. The MgEMTS data set is 67% complete to 3.2Å with a merging R factor of 0.09 and the CsEMTS data set is 64% complete to 3.2Å with a merging R factor of 0.087. The data sets were autoindexed, integrated and merged with the XENGEN software package, while the derivative data sets (MgEMTS and CsEMTS) were locally scaled to each native data set using ROCKS.

When the structure determination of Av1 was completed, a higher resolution Av1 native data set (MOFE22) was collected to 2.2Å resolution[18] with a MAR Research imaging plate detector at the Stanford Synchrotron Radiation Laboratory. The images were processed with the MOSFLM and CCP4 packages[19]. The MOFE22 data set is 90% complete to 2.2Å with a merging R factor of 0.136.

Table 2-4. Data collection statistics of the Av1 and Cp1 crystal forms. Native and heavy atom derivative data sets were collected on a Siemens area detector with CuKα (λ = 1.5418Å) radiation at room temperature. The XENGEN[15] software package was used for data processing.

Crystal form/ Data set	Number of crystals used	Resolution (Å)	Completeness (%)	R_{merge}
Av1				
MOFE27	3	2.7	94	0.085
EMTS1	1	3.2	83	0.057
EMTS3	3	3.0	91	0.086
PTCL3	3	3.0	89	0.090
PIP3	3	3.0	92	0.111
Mg1				
MGCP1	3	3.0	78	0.070
MgEMTS	1	3.2	67	0.090
Cs1				
CSCP1	3	3.0	74	0.084
CsEMTS	1	3.2	64	0.087

2-7. Heavy atom binding sites in the Av1 crystal form

Heavy atom binding sites in Av1 crystal form were determined by inspection of the Harker section[20] of the isomorphous difference Patterson[21] map. The Harker sections (v = 1/2 for space group $P2_1$) of the difference Pattersons calculated for EMTS1, PTCL3 and PIP3 data are shown in Figure 2-4. The EMTS1 isomorphous difference Patterson was solved, and the sites found were (x,y,z) = (0.423, 0.000, 0.124), (0.144, -0.263, 0.124),

(0.715, 0.011, 0.028) and (-0.197, -0.274, 0.028). Two additional mercury binding sites were found by isomorphous difference Fourier using SIR (single isomorphous replacement) phases calculated by the program PHARE. The $PtCl_4$ and PIP sites were determined by isomorphous difference Fourier using the EMTS SIR phases to determine the location of those sites relative to the EMTS sites. Heavy atom binding sites of three derivatives are summarized in Table 2-5. Six mercury sites were found in the EMTS derivative and 10 platinum binding sites were found in the PTCL derivative. PIP has seven binding sites in common with the PTCL derivative.

In order to find metal cluster sites in MoFe-protein, an anomalous Patterson and native Patterson were calculated using the ROCKS program. Figure 2-5 shows the Harker sections of the native anomalous Patterson and native Patterson. There are several high peaks which were expected to be self vectors of the metal clusters, but the Patterson could not be solved because of the lack of cross vectors. Later the cluster sites were determined by anomalous Fourier using MIR phases and by inspection of electron density map.

Table 2-5. Heavy atom binding sites and metal center sites in the Av1 crystal form. The heavy atom binding sites were determined from isomorphous difference Pattersons and isomorphous difference Fourier, and refined by program HEAVY[26]. The metal center sites were determined by native anomalous Fourier and inspection of electron density map.

derivative	x	y	z	relative occupancy
EMTS	0.423	0.000	0.124	0.39
	0.144	-0.265	0.124	0.41
	-0.198	-0.272	0.029	0.14
	0.714	0.008	0.028	0.14
	-0.172	-0.335	-0.352	0.11
	0.487	0.070	-0.347	0.06
$PtCl_4$	0.292	-0.335	0.323	0.36

	0.377	0.061	0.314	0.39
	0.250	-0.167	0.184	0.33
	0.350	-0.092	0.190	0.30
	-0.198	-0.277	0.022	0.28
	0.711	0.018	0.016	0.25
	-0.240	-0.271	-0.293	0.11
	0.583	0.008	-0.300	0.13
	-0.024	-0.335	0.209	0.11
	0.618	0.067	0.171	0.11
Metal centers	-0.105	-0.183	-0.106	
	-0.063	-0.171	-0.071	
	0.579	-0.204	0.010	
	0.561	-0.223	0.045	
	0.551	-0.081	-0.105	
	0.529	-0.093	-0.072	
	-0.073	-0.060	0.008	
	-0.036	-0.040	0.043	

Table 2-6. Heavy atom binding sites and metal center sites in the Mg1 crystal forms. The heavy atom binding sites were determined from the isomorphous difference Patterson with the aid of molecular replacement phases, and were refined by program HEAVY. The metal center sites were determined by inspection of averaged electron density map.

derivative	x	y	z	relative occupancy
EMTS	0.356	0.073	0.858	0.23
	0.579	-0.264	0.451	0.15
	0.125	0.295	0.105	0.22

	-0.348	-0.161	0.104	0.15
Metal centers	0.377	-0.038	0.070	
	0.357	-0.055	0.668	
	-0.211	0.076	0.112	
	-0.227	0.054	0.133	
	0.035	0.147	0.131	
	0.037	0.116	0.140	
	0.569	-0.085	0.653	
	0.581	-0.092	0.620	

Table 2-7. Heavy atom binding sites and metal center sites in the Cs1 crystal form. The heavy atom binding sites were determined from the isomorphous difference Patterson with the aid of molecular replacement phases, and refined by program HEAVY. The metal center sites were determined by inspection of averaged electron density map.

derivative	x	y	z	relative occupancy
EMTS	0.416	0.231	-0.105	0.17
	0.737	0.021	0.719	0.14
	-0.033	-0.254	0.071	0.16
	0.019	-0.008	0.924	0.11
Metal center	0.478	0.161	0.183	
	0.489	0.148	0.224	
	-0.076	-0.106	0.498	
	-0.064	-0.080	0.507	
	0.034	-0.170	0.333	
	0.039	-0.157	0.374	
	0.643	0.100	0.286	

0.635 0.074 0.313

2-8. Determination of non-crystallographic symmetry (NCS) relationships within and between crystal forms

2-8-1. NCS relationship within Av1 crystal form

Since the Av1 crystal form contains an $\alpha_2\beta_2$ tetramer in the asymmetric unit, there is a 2-fold NCS axis which relates the two $\alpha\beta$ dimers. The 2-fold NCS relationship within the Av1 crystal form was determined from the heavy atom binding sites, which are related by 2-fold rotation, using the program KABROT[39], and refined by density correlation studies[22]. The 2-fold NCS relationship was consistent with the self-rotation function studies (Figure 2-7) and was further supported by the location of the metal centers that were determined from anomalous Fourier maps calculated with MIR (multiple isomorphous replacement) phases. The relation between the two $\alpha\beta$ dimers in $\alpha_2\beta_2$ tetramer is described by the spherical angles $\phi = 270.0°$, $\psi = 90.0°$ and $\kappa = 180.0°$, and the molecular center was chosen to be (x,y,z) = (-0.25, -0.13, 0.00) (Table 2-8).

2-8-2. NCS relationships between crystal forms

The NCS relationships between the Av1 and Cp1 crystal forms (Mg1 and Cs1) were established by rotation function studies[23,24] and translation function studies[25] with the use of structure factors calculated from averaged electron density maps of Av1 masked to contain primarily one molecule. The NCS relationships determined by this method were confirmed by the location of heavy atom binding sites and the metal centers (the FeMo-cofactor and P-cluster pair) in MoFe-protein. The rotation and translation parameters were refined by density correlation studies[22]. The NCS relationships between crystal forms are summarized in Table 2-8.

2-8-3. NCS relationships within Cp1 crystal forms

The NCS relationships within both the Mg1 and Cs1 crystal forms were determined by three methods. First, the two-fold NCS rotation axes within both the Mg1 and Cs1 crystal forms were calculated from the NCS relationships within the Av1 crystal form and the relationship between the Av1 and Cp1 crystal forms. Second, the two-fold NCS axes within both the Mg1 and Cs1 crystal forms were established by self-rotation function studies and translation function studies. Third, the two-fold NCS axes within both the Mg1 and Cs1 crystal forms were also determined by the heavy atom binding sites using KABROT. Heavy atom binding sites in both the Mg1 and Cs1 crystal forms (Table 2-6 and 2-7) were determined by inspection of the Harker section from the isomorphous difference Patterson maps (Figure 2-6) with the aid of molecular replacement phases which were calculated from the averaged Av1 electron density maps oriented in the Mg1 or Cs1 unit cell. The NCS relationships determined by the three methods were all consistent and were further supported by the location of the metal centers in MoFe-protein. The NCS relationships within both the Mg1 and Cs1 crystal forms were refined by density correlation studies. The refined NCS parameters are shown in Table 2-8.

Table 2-8. The NCS relationships within and between crystal forms.

Crystal form	Space group	Unit cell constants a	b	c	β	Spherical angle φ	ψ	Molecular center x	y	z
Av1	P21	108.4	130.5	81.5	110.8	270.005	90.079	0.2511	-0.1309	0.0000
Mg1	P21	69.96	151.3	121.9	110.4	269.794	21.795	0.1806	-0.0035	0.3844
Cs1	P21	87.9	171.4	73.6	91.5	103.476	92.516	0.2870	-0.0028	0.3711

From	To	φ	ψ	κ
Mg1	Av1	219.466	133.247	-221.839

		150.871	54.310	-231.358
Cs1	Av1	93.475	107.537	-314.394
		277.254	91.458	-223.273
Cs1	Mg1	354.245	110.565	-237.869
		71.818	38.955	-209.681

2-9. Isomorphous replacement phase calculation

Refinement of the heavy atom parameters (coordinates, occupancy and temperature factor) and isomorphous replacement phase calculations were performed with two programs, PHARE and HEAVY[26]. PHARE refines the heavy atom parameters by a "phased" refinement that minimizes the lack of closure between observed and calculated derivative amplitudes, while HEAVY refines the heavy atom parameters against an origin-removed Patterson function and should provide a less biased estimate of heavy atom occupancies. Isomorphous replacement phases calculated by HEAVY were used for the subsequent NCS averaging of the electron density map. Statistics of the phase calculation by PHARE are listed in Tables 2-9, 2-10 and 2-11 for the sake of comparison. The MIR phases calculated by PHARE and HEAVY were of similar quality and the correlation coefficients between the two phase set was 92%.

Four heavy atom derivatives (EMTS1, EMTS3, PTCL3 and PIP3) were used to calculate the initial MIR phases for the Av1 crystal form. The phasing power of the EMTS1 derivative, EMTS3 derivative, PTCL3 derivative and PIP3 derivative are 1.75, 1.70, 1.35, and 1.21, respectively, and the mean figure of merit is 0.62 to 3.5Å resolution. One heavy atom derivative (MgEMTS) was used to calculate the initial isomorphous replacement phases for the Mg1 crystal form. The phasing power of the MgEMTS derivative is 1.40 and the mean figure of merit is 0.37 to 3.5Å resolution. The phasing

statistics of the Cs1 crystal form are similar to those of Mg1 crystal form (Tables 2-10 and 2-11).

Table 2-9. Phasing statistics of the Av1 crystal form. MIR phases were calculated with PHARE.

Shell (Å)	Number of reflections	mean figure of merit	Phasing power EMTS1	EMTS3	PTCL3	PIP3
18.79	136	0.62	1.94	2.01	1.22	0.87
11.57	591	0.67	2.31	2.14	1.27	1.30
8.36	1251	0.69	2.34	2.15	1.60	1.44
6.54	2154	0.69	2.64	2.35	1.78	1.64
5.37	3315	0.70	2.55	2.49	1.76	1.65
4.56	4695	0.63	1.85	1.85	1.30	1.27
3.96	6323	0.60	1.52	1.50	1.12	1.02
3.50	8182	0.55	1.33	1.29	0.95	0.89
Totals	26647	0.62	1.79	1.71	1.26	1.17

Table 2-10. Phasing statistics of the Mg1 crystal form. SIR phases were calculated with PHARE.

Shell (Å)	Number of reflections	mean figure of merit	Phasing power EMTS
18.79	160	0.34	1.60
11.57	624	0.36	1.61
8.36	1367	0.33	1.53
6.54	2331	0.34	1.56
5.37	3478	0.37	1.56
4.56	4628	0.38	1.43

3.96	5122	0.37	1.26
3.50	5374	0.34	1.11
Totals	23084	0.36	1.37

Table 2-11. Phasing Statistics of the Cs1 crystal form. SIR phases were calculated with PHARE.

Shell (Å)	Number of reflections	Mean figure of merit	Phasing power EMTS
18.79	124	0.35	1.56
11.57	525	0.33	1.60
8.36	1103	0.35	1.59
6.54	1836	0.35	1.59
5.37	2700	0.35	1.55
4.56	3756	0.36	1.40
3.96	4431	0.36	1.28
3.50	4771	0.34	1.16
Totals	19246	0.35	1.37

2-10. Molecular envelope generation

MIR phases obtained from the Av1 derivative data were used to define molecular envelopes. The initial envelope was generated using averaged 5Å resolution electron density maps by a modification of the Wang algorithm[27]. Using this envelope, 3.5Å MIR phases of the Av1 crystal form were iteratively refined by molecular averaging and solvent flattening[28]. The updated envelope was then generated using this averaged 3.5Å electron density map. The molecular envelope generated by this algorithm is a contour map. The size of the envelope can be controlled by choosing a suitable contour level depending on the

solvent content. A low solvent content (~20%) was used for the initial stages of NCS averaging, while a high solvent content (~45%) was used for the final stages of NCS averaging.

2-11. Structure determination of Av1

2-11-1. Six-fold averaging

The structure of Av1 was determined by the method of MIR and NCS averaging both within and between crystal forms[28,29]. Although the Av1 and Cp1 sequences are only ~36% identical[30,31], averaging between crystal forms was invaluable in the initial stages of the analysis. Since three crystal forms (Av1, Mg1 and Cs1) were used for the NCS averaging and each crystal form contains an $\alpha_2\beta_2$ tetramer in the asymmetric unit, the NCS averaging both within and between crystal forms virtually represents a six-fold averaging. The process of the six-fold averaging is shown in Figure 2-8a and the subsequent two-fold averaging within the Av1 crystal form is shown in Figure 2-8b.

Electron density maps were calculated by fast Fourier transform (FFT) using MIR phases for the Av1 crystal form and SIR phases for the Mg1 and Cs1 crystal forms. Each electron density map was skewed into an orthogonal cell (a = 150Å, b = 150Å, c = 150Å, $\alpha = 90°$. $\beta = 90°$, $\gamma = 90°$) and two-fold averaged[28] in the skewed cell. The three skewed maps were then averaged by a program, ADDMAP and the averaged map was skewed back into the original cells. New structure factor amplitudes and phases were calculated in the original cells by Fourier inversion, and the calculated phases were combined with the MIR or SIR phases. Electron density maps were then calculated for each crystal form using the combined phases. The process was iterated until convergence occurred (Figure 2-8a). After six cycles of phase combination, phase transfer and $2F_O$-F_C Fourier synthesis, the final R-factor was ~30% and the correlation coefficient was ~75%. The relatively poor averaging statistics may reflect the structural differences between Av1 and Cp1 which have been indicated by low amino acid sequence homology and many insertions and

deletions[31], and may be caused in part by the fact that the Av1 envelope was used for the NCS averaging.

The averaged map was significantly improved and it showed many secondary structural features such as α-helices and β-sheets, although the map was still not interpretable in part due to broken densities. Considering the high R-factor and low correlation coefficients, the averaged map was very promising. The averaged phases were combined with the MIR phases again and subsequently two-fold averaging was performed within the Av1 crystal form (Figure 2-8b). The two-fold averaging converged very fast. After 4 cycles of phase combination, phase transfer and $2F_O$-F_C Fourier synthesis, the final R-factor dropped to ~10% and the correlation coefficient increased to ~96%. The resultant averaged map (designated 6x Averaged map I) was of sufficient quality to build ~80% of polyalanine model. Almost all secondary structural elements were revealed in this map (Figure 2-9a, b and c), although there was some broken density and missing side chain density.

2-11-2. Model building and refinement

Initially, a polyalanine backbone model was built into the averaged electron density map using the graphics program TOM/FRODO[32] implemented with a fragment fitting option, DIGNL[33]. Standard α-helix and β-strand models were used for the α-helical and β-sheet regions, and the DIGNL option was used to build the model around loop regions. The initial polyalanine model contained about 80% of the residues. The model was not built around regions in which the electron density was ambiguous or poorly determined. Only a single αβ dimer was built at each cycle of model building, with the second αβ dimer generated by the two-fold noncrystallographic symmetry. Two kinds of electron density maps were generated using this partial model. First, the model was refined with the restrained least square program TNT[34], and calculated phases from the model were combined with MIR phases and then re-averaged within the Av1 crystal form. The combined and averaged map (designated Combined map I) was used for the next cycle of

model building. Second, the model was refined in the Av1 crystal form with TNT and the refined model was transformed to and further refined in the Mg1 and Cs1 crystal forms. Calculated phases from each refined model (Av1, Mg1 and Cs1 models) were combined with the MIR and SIR phases of each crystal form, and the combined phases were subsequently refined by six-fold NCS averaging as described in section 2-10-1. The resultant map (designated 6x Averaged map II) and the Combined map I were improved substantially and allowed the building of ~88% complete polyalanine model. Subsequent maps, 6x Averaged map III and Combined map II, were generated by a similar process of refinement, phase combination and NCS averaging as described above. The side chains were built into these averaged maps with the aid of the ROTAMER[35] option in TOM/FRODO.

The electron density map was significantly improved by iterative cycles of model building, refinement, phase combination and NCS averaging. During successive cycles of phase refinement, the resolution was gradually extended from 3.5Å to 2.7Å. The program TNT was used for coordinate refinement during the initial stages of modeling, while the simulated annealing program X-PLOR[36] was used during the final stages of refinement. Residues were rebuilt following inspection of $2F_O$-F_C and F_O-F_C maps, and examination of Eisenberg's 3D-1D profile[37] calculated for each model. The current model contains 1980 amino acid residues (of 2026 total residues) with 15,858 non-hydrogen atoms (~98% complete). This model has been refined to a crystallographic R-factor of 0.188 (8-2.7Å resolution) with root mean square (rms) deviations from ideal bond distances and angles of 0.016Å and 3.3°, respectively (Table 4-1). Refinement of this model at 2.2Å resolution with the program X-PLOR, followed by TNT, has progressed to R = 0.202, with rms deviations from ideal bond distances and angles of 0.020Å and 2.98°, respectively.

2-11-3. Correctness of the chain trace

The correctness of the chain trace is supported by the positions of the heavy atom binding sites (the mercury sites are found coordinated to nonconserved Cys residues in the Av1 and Cp1 crystal forms, while the platinum sites are coordinated to Met ligands in Av1; Table 4-2 and 5-2. Cys and Met residues are good ligands for heavy atom compounds[38]) and the similarity in chain folds of the α and β subunits, which were modeled independently (see Chapter 4). The model has also been checked using the 3D-1D profile method[37] and all the residues have reasonable 3D-1D scores (data not shown).

References

1. Newton, J. W., Wilson, P. W. & Burris, R. H. (1953) *J. Biol. Chem.* **204**, 445-451.

2. The fermentation was done by T. Sutherland & M. M. Georgiadis, and cell cultures were kindly provided by M. M. Georgiadis.

3. Kornuc, J. K. (1988) Ph.D. Thesis, University of California, Los Angeles. *X-ray Crystallographic studies of the Nitrogenase Iron Protein.*

4. The fermentation was done by T. Sutherland & D. Woo.

5. Burgess, B. K., Jacobs, D. B. & Stiefel, E. I. (1980) *Biochem. Biophys. Acta* **614**, 196-209.

6. Shah, V. K. & Brill, W. J. (1973) *Biochem. Biophys. Acta.* **305**, 445-454.

7. Nelson, M. J., Levy, M. A. & Orme-Johnson, W. H. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 147-150.

8. The MoFe-protein fraction was kindly provided by D. Woo.

9. Drenth, J., Hol, W. G. J. & Wierenga, R. K. (1975) *J. Biol. Chem.* **250**, 5268-5269.

10. Georgiadis, M. M. (1990) Ph.D. Thesis, University of California, Los Angeles. *Crystal structure of the Nitrogenase Iron Protein from Azotobacter vinelandii.*

11. Matthews, B. W. (1985) *Meth. Enzymol.* **114**, 176-187.

12. Weininger, M. S. & Mortenson, L. E. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 378-380.

13. Harker, D. (1956) *Acta Cryst.* **9**, 1-9.

14. Hope, H. (1988) *Acta Cryst.* **B44**, 22-26.

15. Howard, A. J., *et al.* (1987) *J. Appl. Cryst.* **20**, 383-387.

16. Reeke, G. N. (1984) *J. Appl. Cryst.* **17**, 125-130; Bethe, D. H. (1984) *J. Appl. Cryst.* **17**, 215-215.

17. Matthews, B. W. & Czerwinski, E. W. (1975) *Acta Cryst.* **A31**, 480-487.

18. The 2.2Å resolution data set was collected by M. K. Chan.

19. CCP4, The SERC (U.K.) Collaborative Computing Project N0. 4, A Suite of Programs for Protein Crystallography, distributed by Daresbury Laboratory, Warrington WA4 4AD U.K.

20. Harker, D. (1936) *J. Chem. Phys.* **4**, 381-390.

21. Patterson, A. L. (1935) *Z. Kristallogr.* **90**, 517-542.

22. Cox, J. M. (1967) *J. Mol. Biol.* **28**, 151-156.

23. Rossmann, M. G., Blow, D. M. (1962) *Acta Cryst.* **15**, 24-31.

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

25. Crowther, R. A. & Blow & D. M. (1967) *Acta. Cryst.* **23**, 544-548.

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

27. Wang, B. C (1985) *Meth. Enzymol.* **115**, 90-112.

28. Bricogne, G. (1976) *Acta Cryst.* **A32**, 832-847.

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

30. Brigle, K. E., *et al.* (1985) *Gene* **37**, 37-44.

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

32. Jones, T. A. (1985) *Meth. Enzymol.* **115**, 151-171.

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

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

35. Summers, N. L., Carson, W. D. & Karplus, M. (1987) *J. Mol. Biol.* **196**, 175-198.

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

37. Bowie, J. U., Lüthy, R. & Eisenberg, D. (1991) *Science* **253**, 164-170; Lüthy, R., Bowie, J. U. & Eisenberg, D. (1992) *Nature* **356**, 83-85.

38. Blundell, T. L. & Johnson, L. N. (1976) in *Protein Crystallography* (Horecker, B., Kaplan, N. O., Marmur, J. & Scheraga, H. A., Eds.) pp. 183-239. Academic Press, New York.

39. Kabsch, W. (1976) *Acta Cryst.* **A32**, 922-932; Kabsch, W. (1978) *Acta Cryst* **A34**, 827-828.

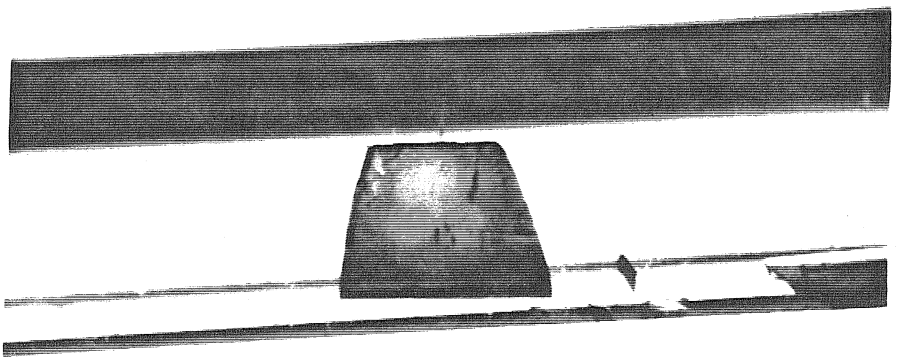
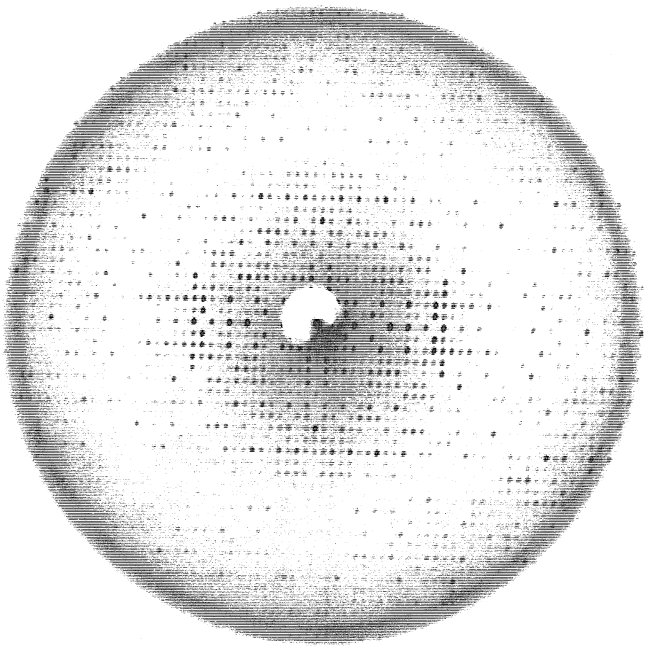
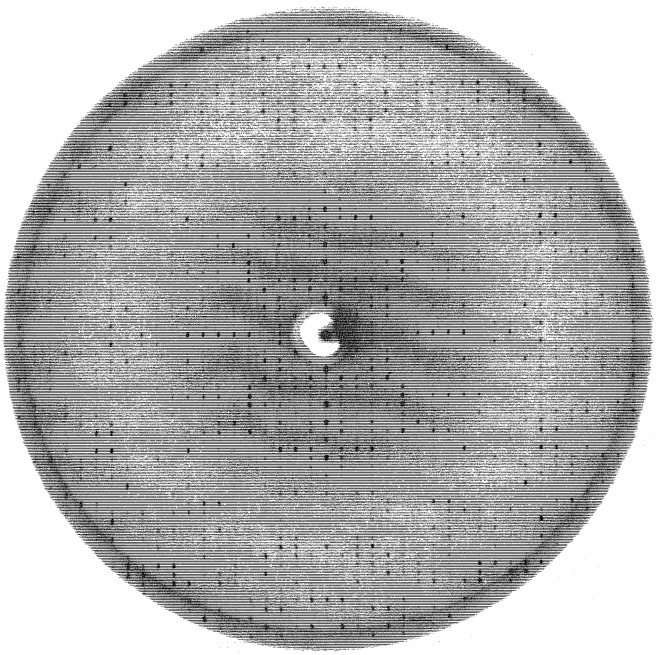


Figure 2-1. A photograph of a typical Av1 crystal. The Av1 crystals are plate-shaped with maximum dimensions of 1.5 mm x 1.0 mmx 0.2 mm and brown due to the presence of the Fe-S clusters.

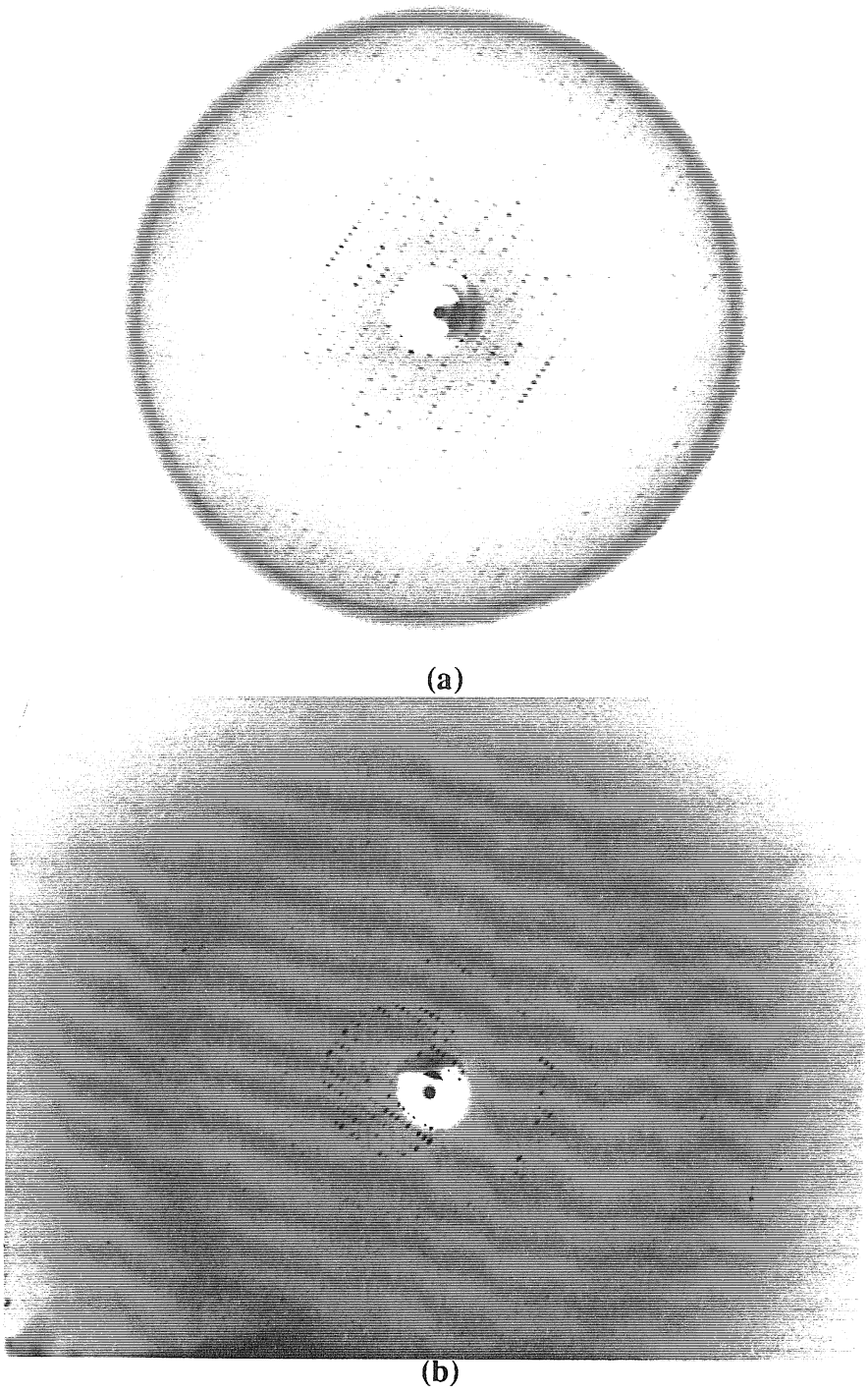


(a)



(b)

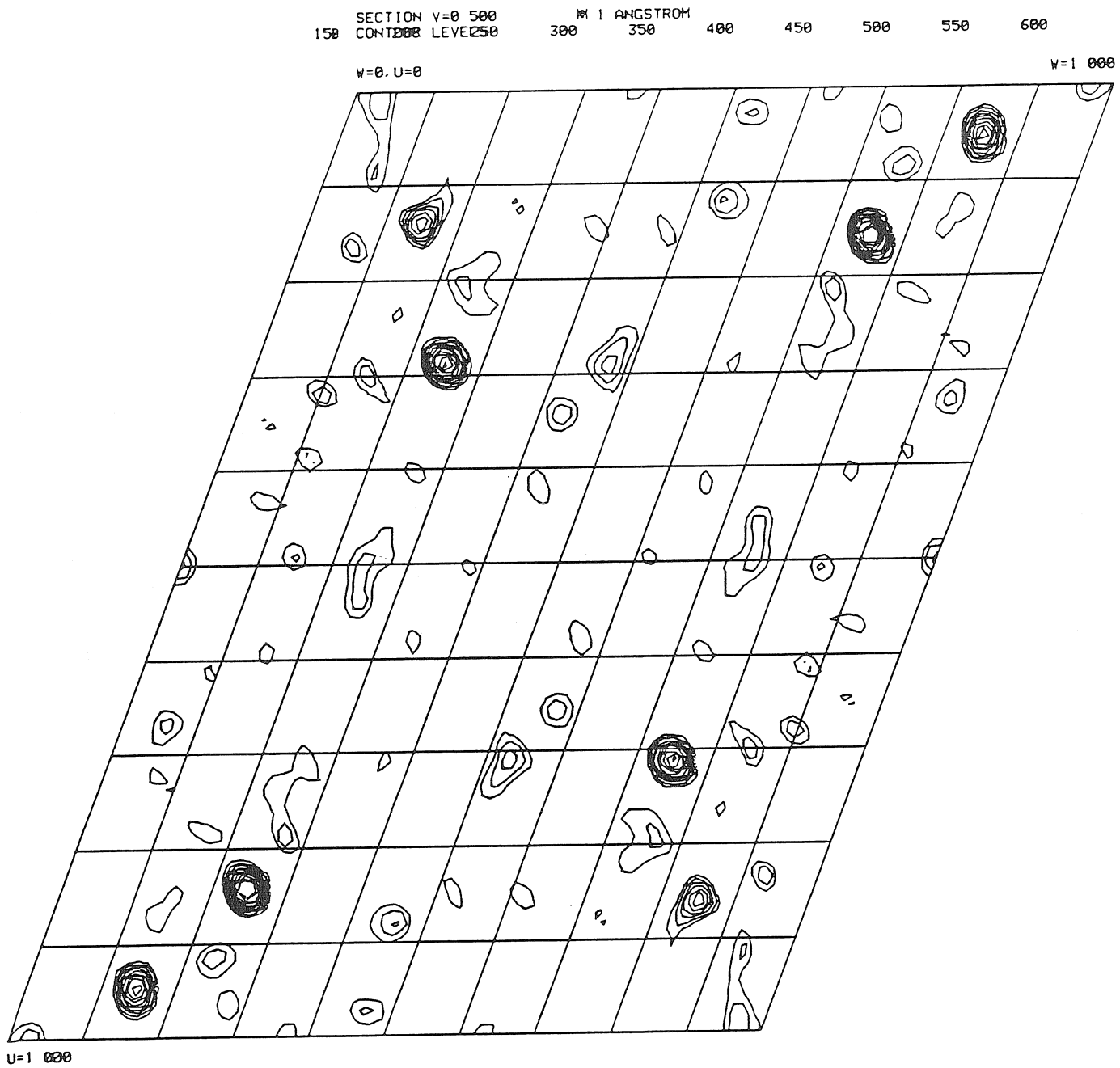
Figure 2-2. Precession pictures of the Av1 crystal form. (a) hk0 zone and (b) 0kl zone. The Av1 crystal form belongs to space group $P2_1$ with unit cell constants; a= 108.4Å, b = 130.5Å, c = 81.5Å, β = 110.8°.



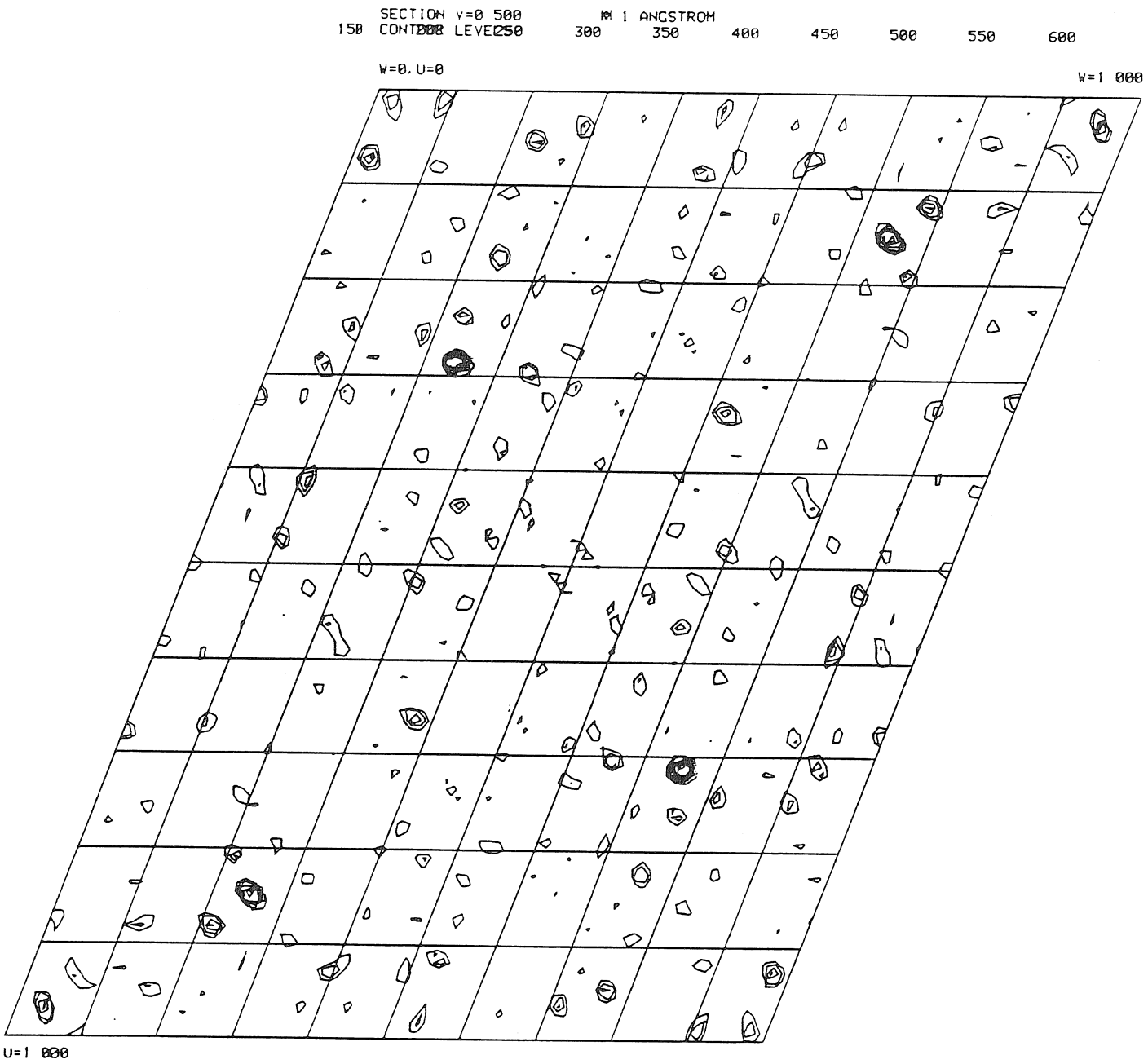
(a)

(b)

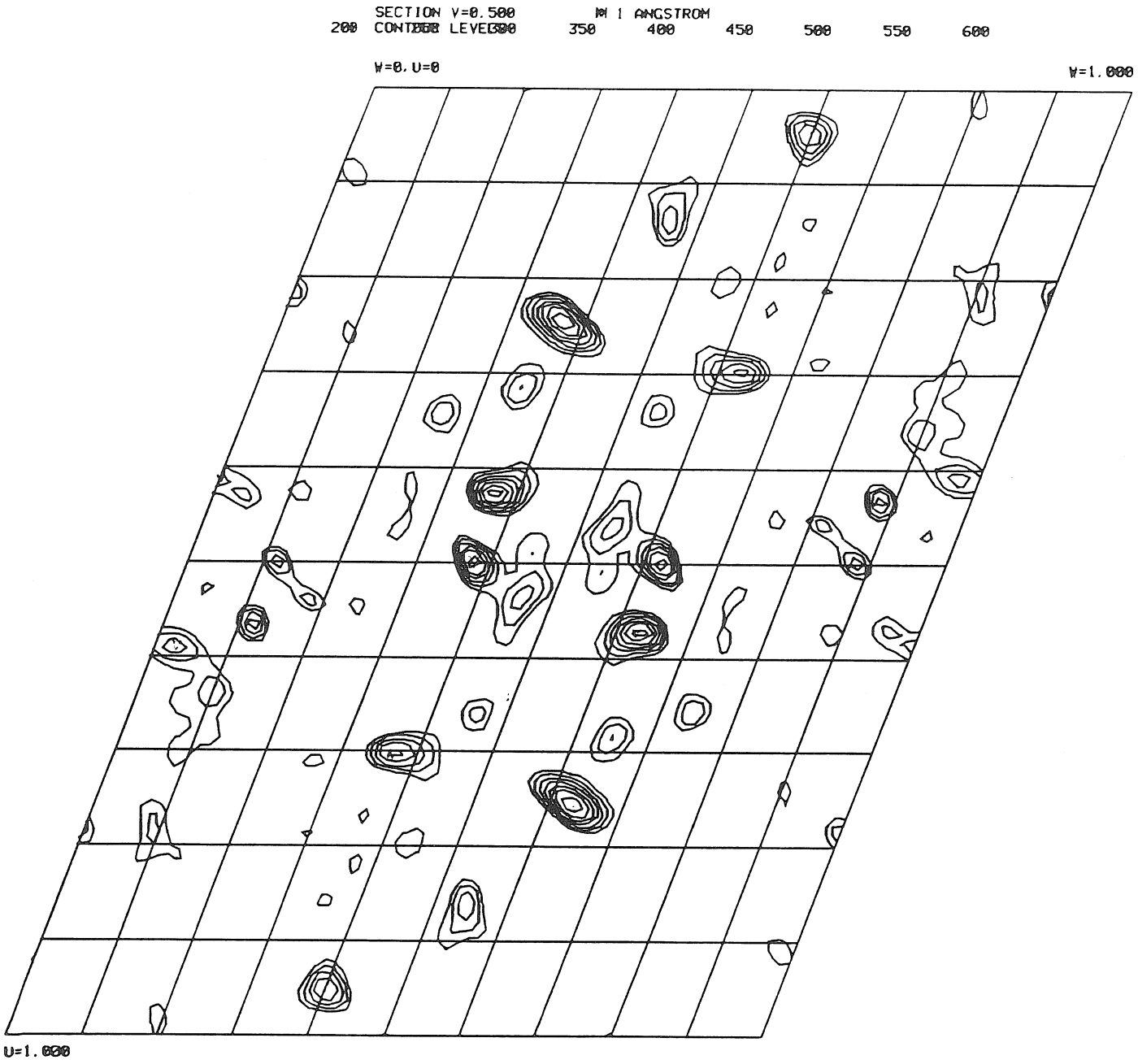
Figure 2-3. Precession pictures of the Cp1 crystal forms. (a) hk0 zone of the Mg1 crystal form, and (b) hk0 zone of the Cs1 crystal forms. The Mg1 crystal form belongs to space group $P2_1$ with unit cell constants a = 70.0Å, b = 151.3Å, c = 121.9Å, β = 110.4°; and the Cs1 crystal form belongs to space group $P2_1$ with unit cell constants a = 87.9Å, b = 171.4Å, c = 73.6Å, β = 91.5°.


SECTION V=0 500
158 CONTOUR LEVELS
1 ANGSTROM
300
350
400
450
500
550
600
W=0, U=0
W=1 000
U=1 000

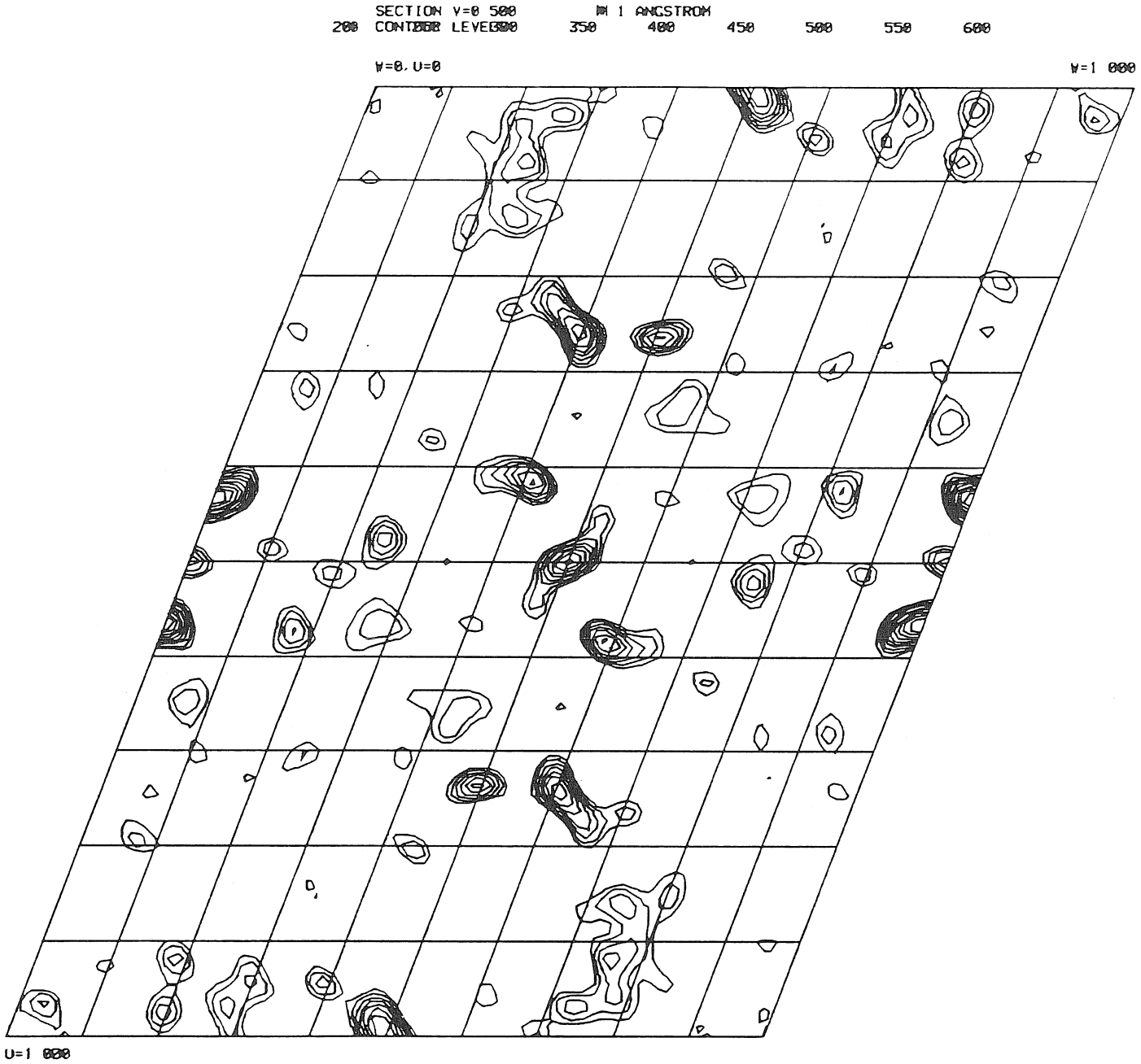
(a)


SECTION V=0 500
1 ANGSTROM
150 CONTOUR LEVELS 250 300 350 400 450 500 550 600
W=0, U=0
W=1 000
U=1 000

(b)

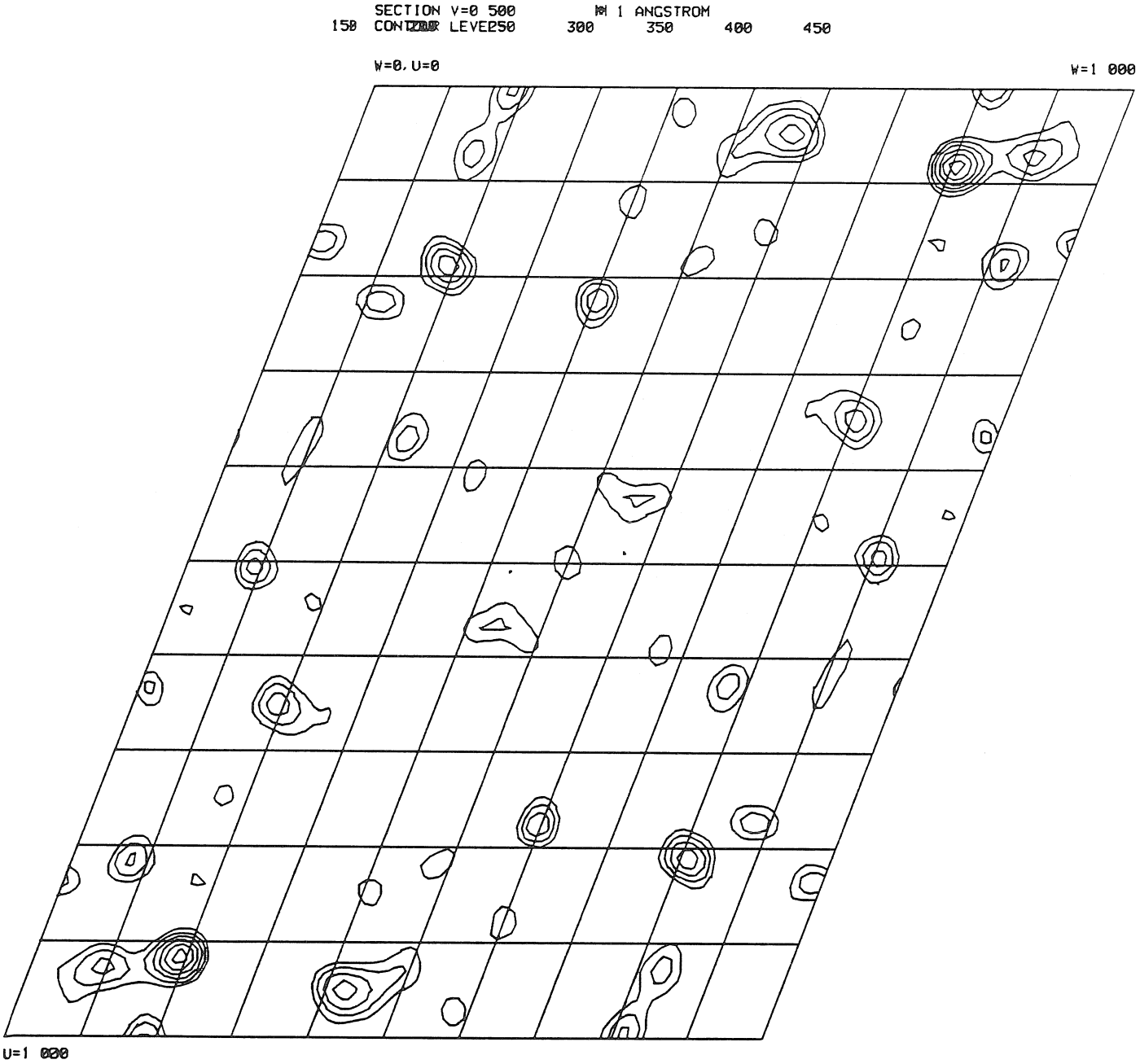

SECTION V=0.500
200 CONTOUR LEVELS
1 ANGSTROM
350
400
450
500
550
600
W=0. U=0
W=1.000
U=1.000

(c)

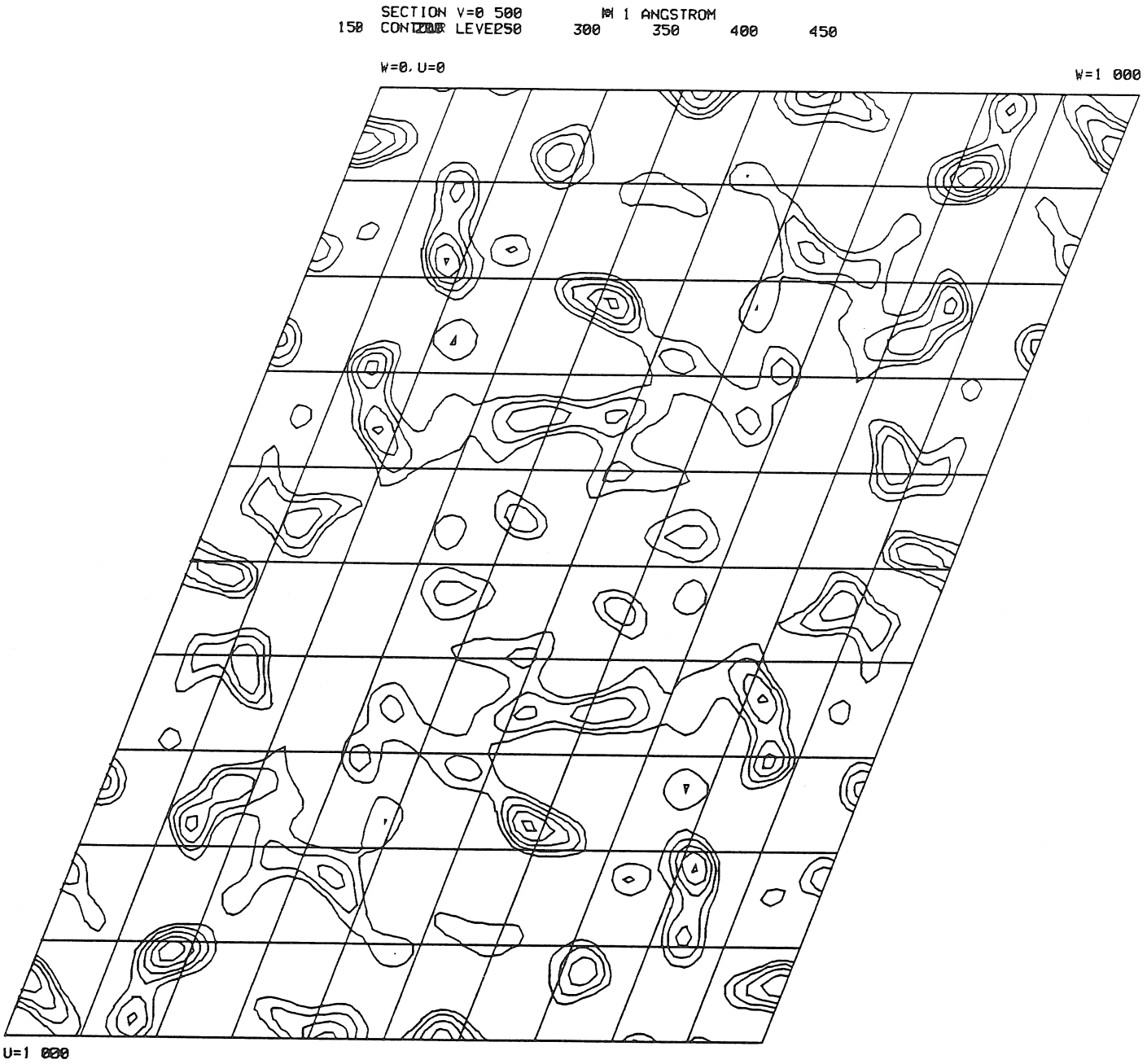


(d)

Figure 2-4. The Harker sections (v = 1/2) of isomorphous difference Pattersons for the Av1 crystal form. (a) EMTS, 50-5Å shell; (b) EMTS, 5-3.2Å shell; (c) PTCL3, 50-5Å shell; and (d) PIP 50-5Å shell. The EMTS isomorphous difference Patterson was solved, and the sites found were (x,y,z) = (0.423,0.000,0.124), (0.144,-0.263,0.124), (0.715,0.011,0.028), and (-0.197,-0.274,0.028).

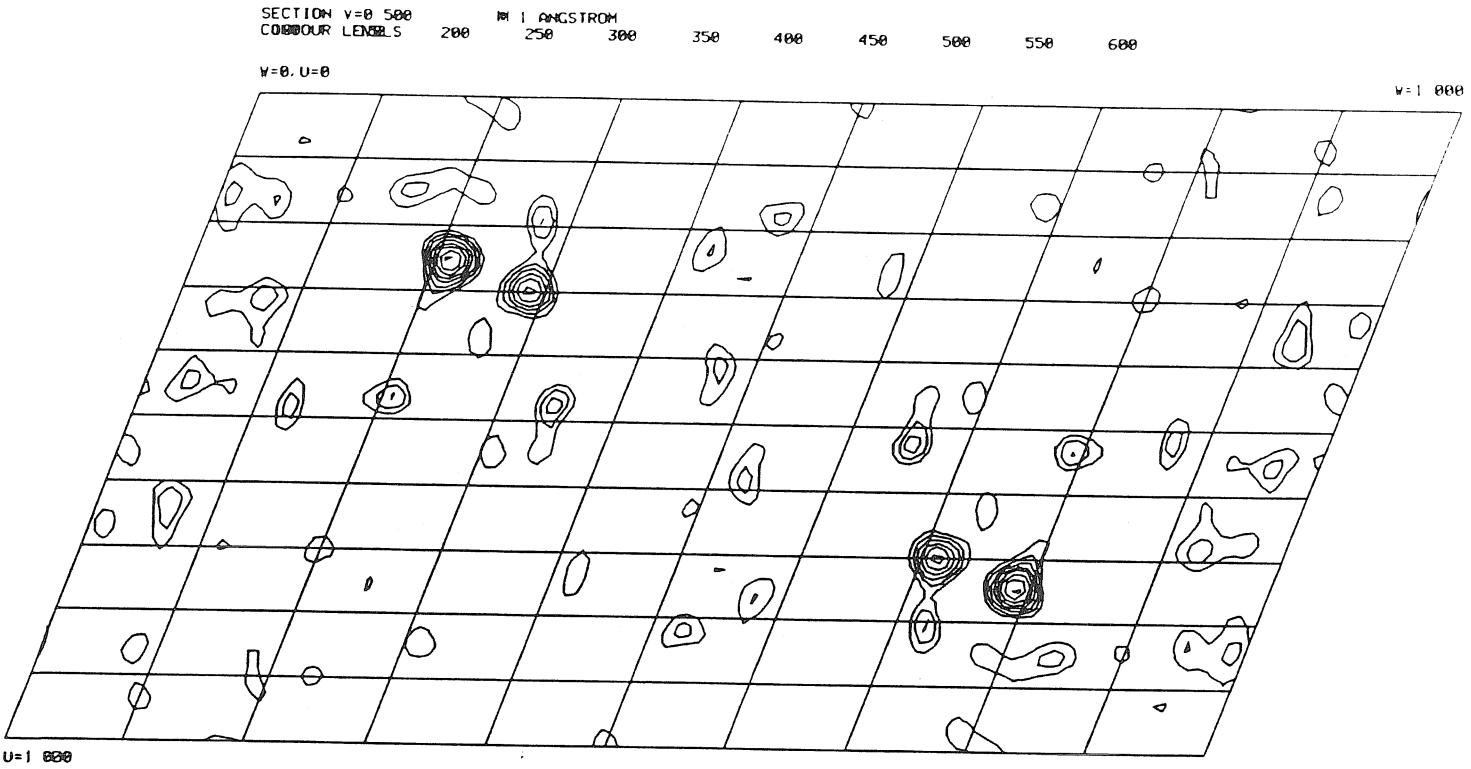

SECTION V=0 500
1 ANGSTROM
150 CONTOUR LEVELS 250 300 350 400 450
W=0, U=0
W=1 000
U=1 000

(a)

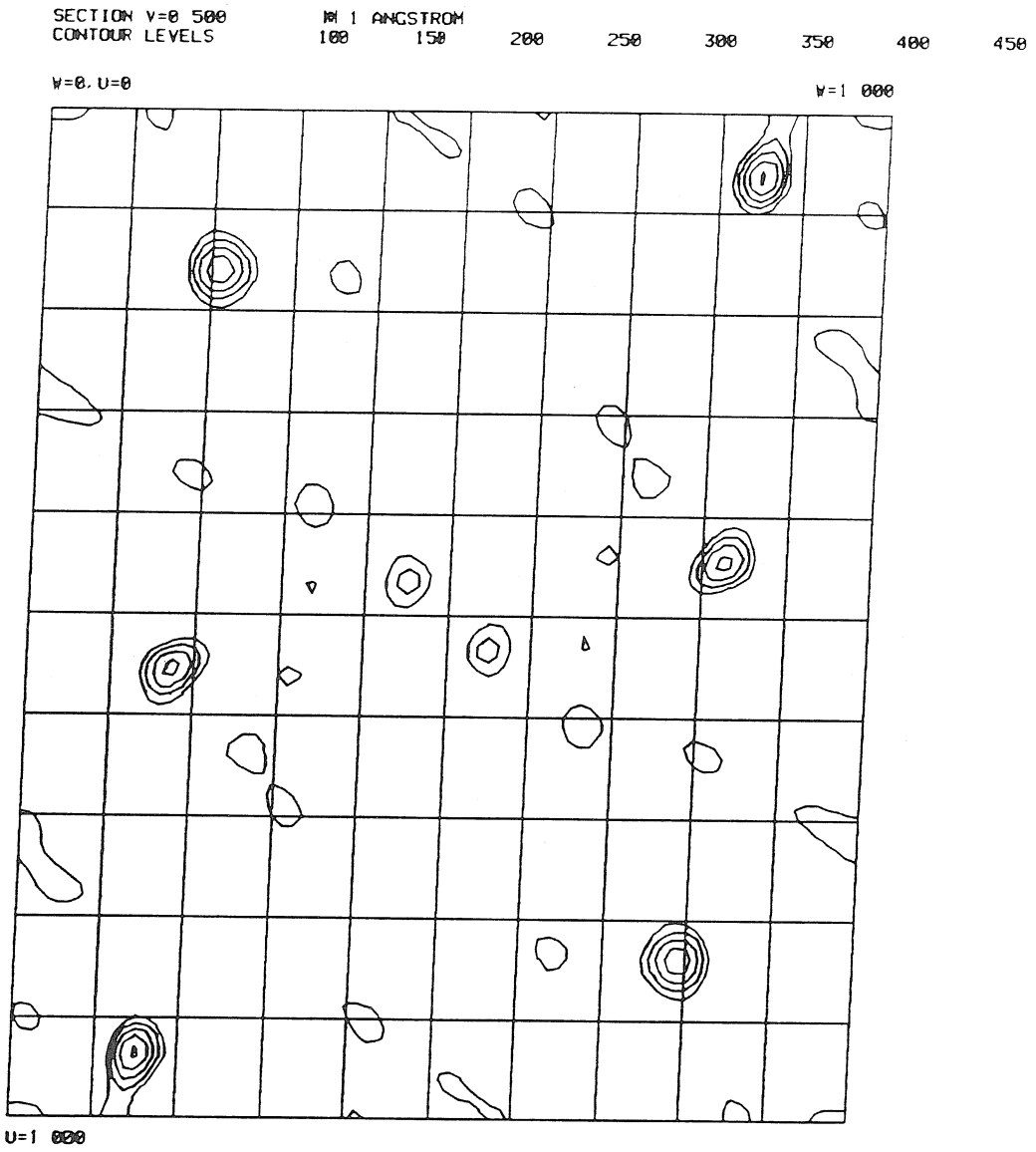


(b)

Figure 2-5. The Harker sections (v = 1/2) of the native anomalous Patterson (a) and native Patterson (b). There are several high peaks which were expected to be self-vectors of the metal centers, but the Patterson could not be solved because of the lack of cross vectors.

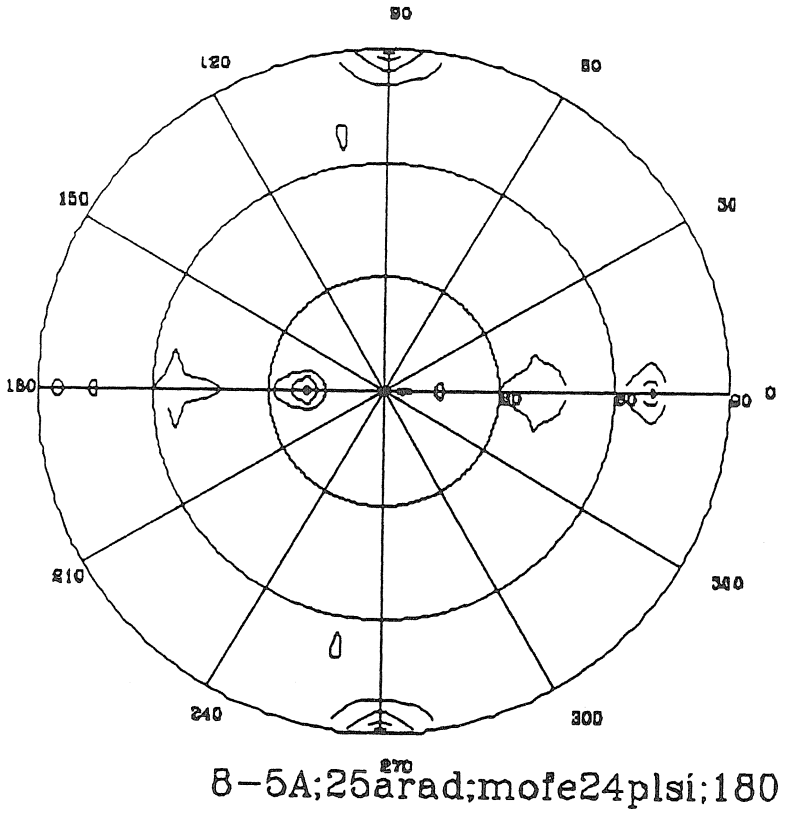

SECTION V=0 500
1 ANGSTROM
CONTOUR LEVELS 200 250 300 350 400 450 500 550 600
W=0, U=0
W=1 000
U=1 000

(a)



(b)

Figure 2-6. The Harker sections (v = 1/2) of isomorphous difference Pattersons for the Cp1 crystal forms. (a) MgEMTS, 50-5Å shell and (b) CsEMTS, 50-5Å shell. The Pattersons were solved with the aid of molecular replacement phases which were calculated from the Av1 electron density map oriented in the Mg1 or Cs1 unit cell.



P21 CELL (Z axis unique)			P21 CELL(Y axis unique)		
PHI	PSI	KAPPA	PHI	PSI	KAPPA
90	90	180	0	0	180
0	70	180	-90	90	180
180	20	180	0	90	180

Figure 2-7. The κ = 180° section of the Crowther self-rotation function of the Av1 crystal form. Three peaks, which are orthogonal each other, are found in this section. One (Y axis) is a 2_1 screw axis which is a crystallographic symmetry element. One of the other two is true 2-fold NCS axis and the other one is generated by the 2_1 screw axis and 2-fold axis because they are orthogonal each other.

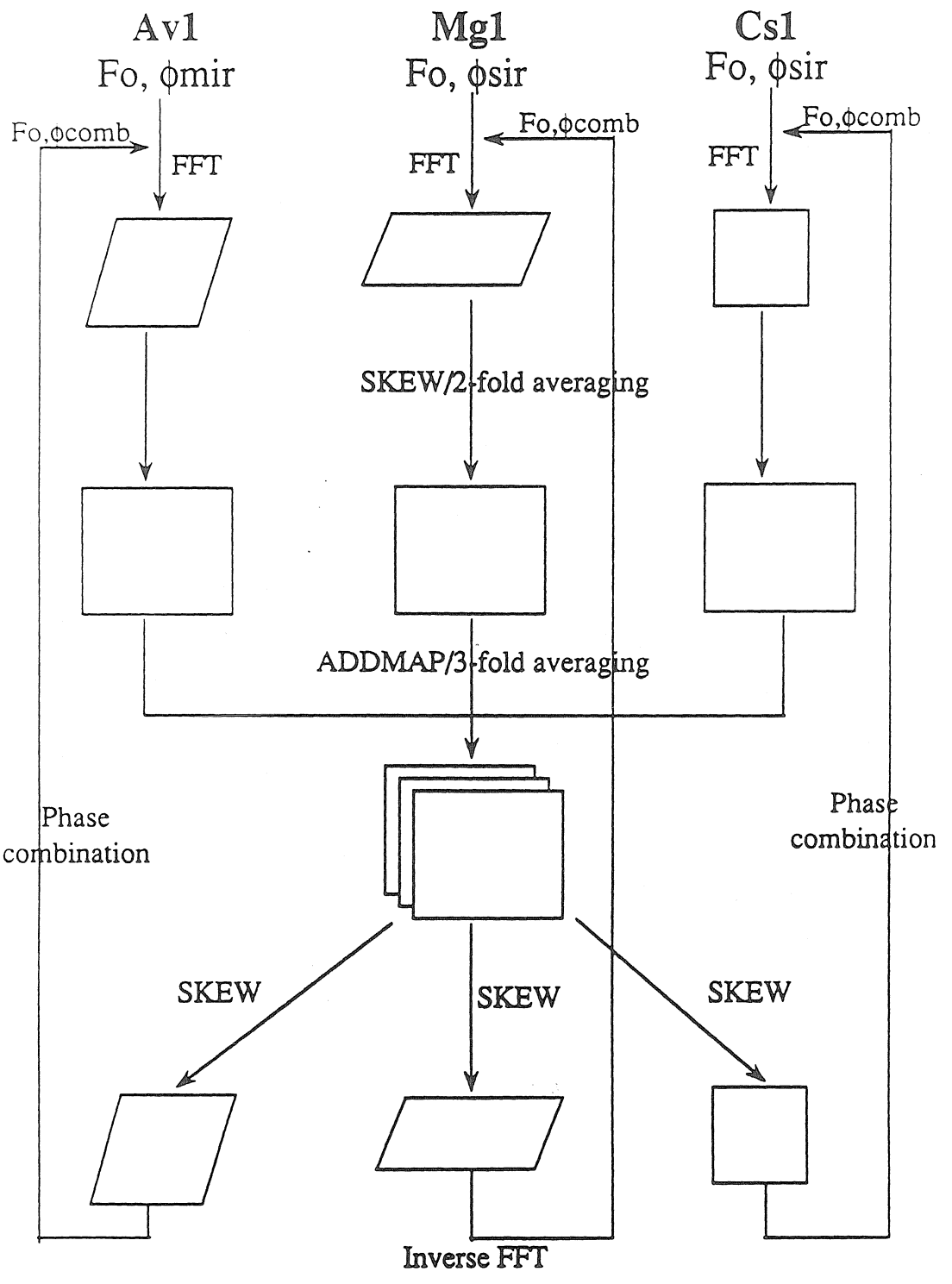


Figure 2-8a. Schematic representation of the process of the 6-fold averaging within and between crystal forms.

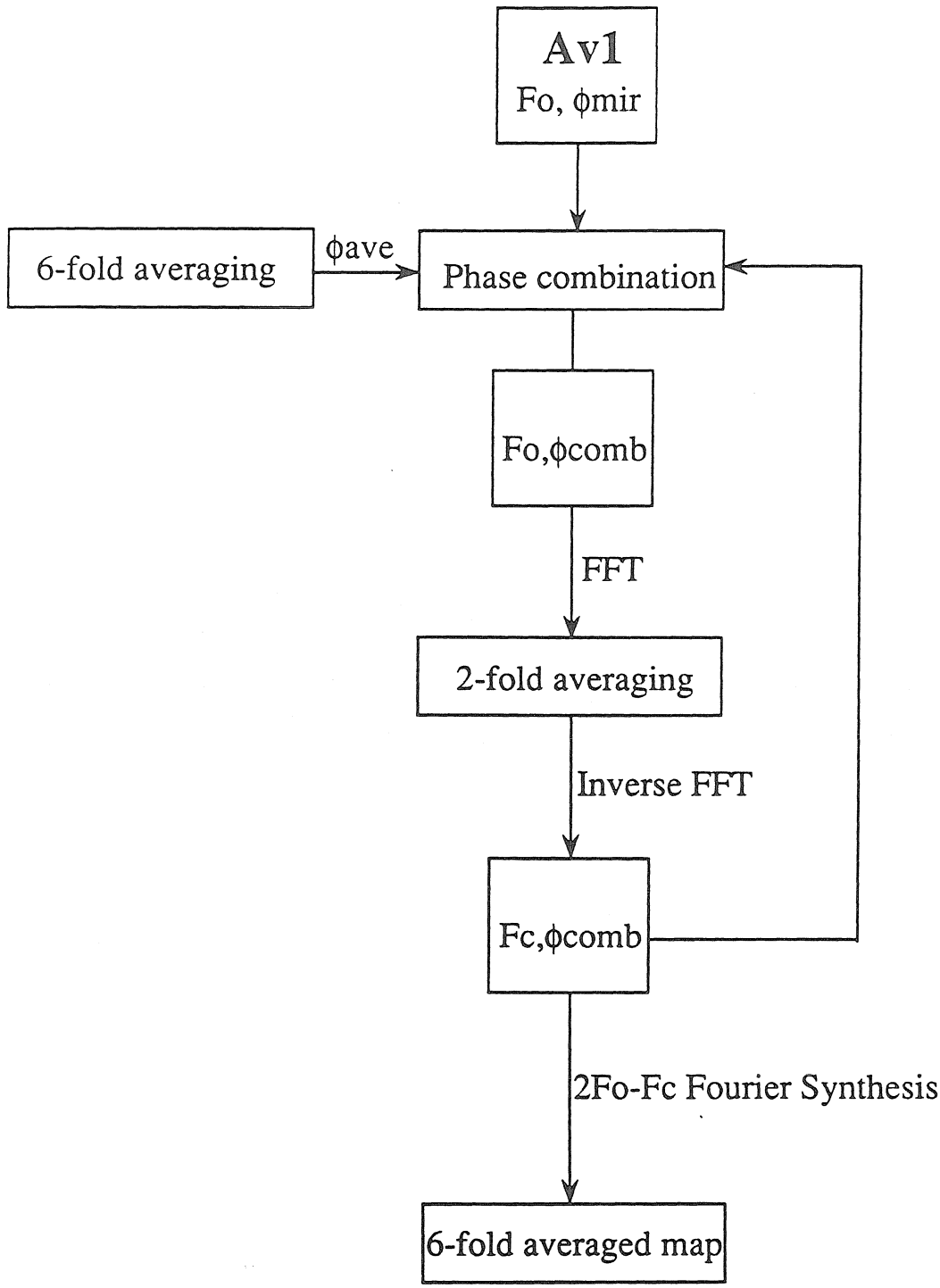


Figure 2-8b. Schematic representation of the process of the subsequent 2-fold averaging within the Av1 crystal form.

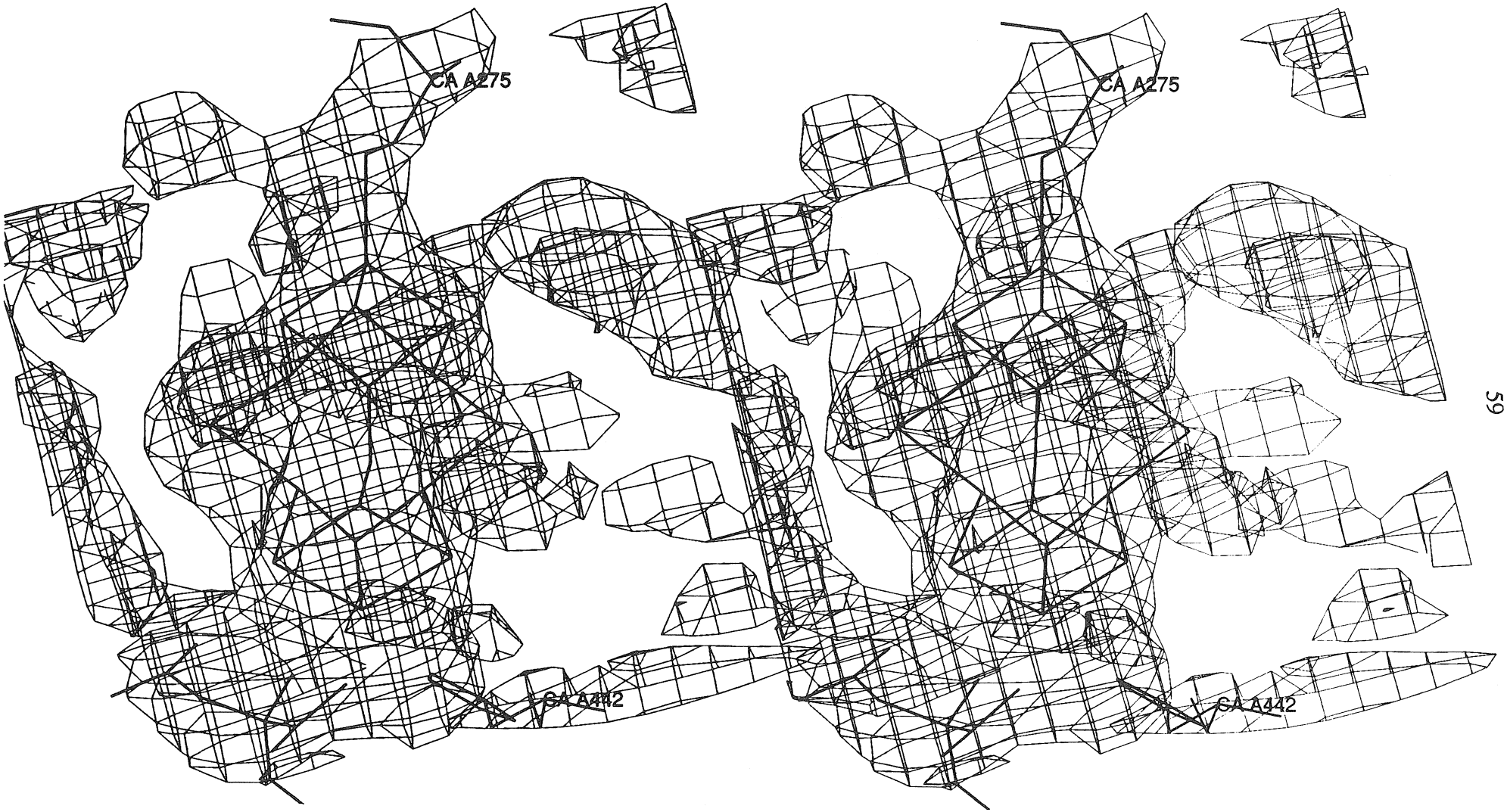


Figure 2-9a. A section of the 6x Averaged map I around the FeMo-cofactor. It is contoured at 2.0 sigma of the electron density map.

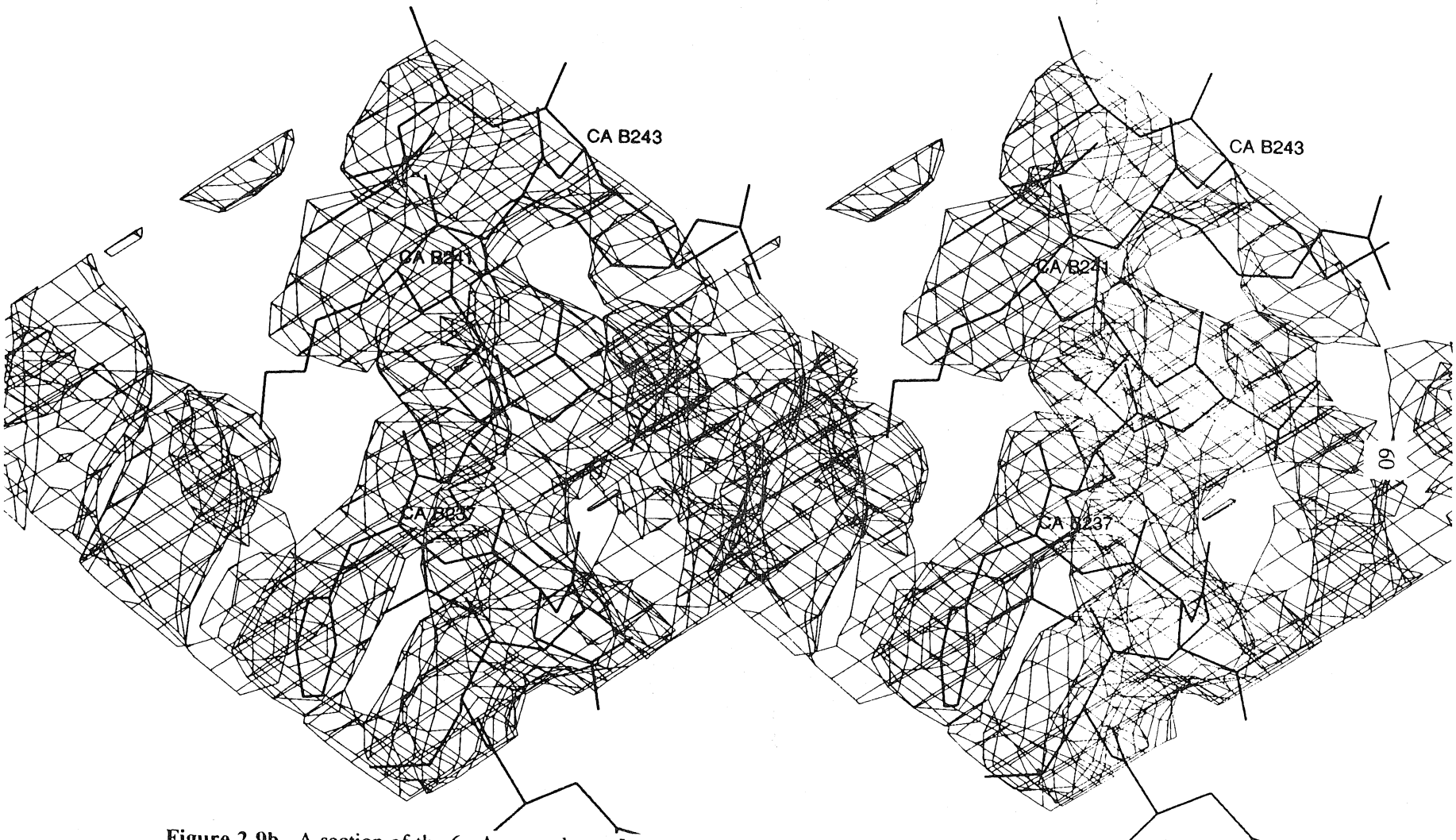


Figure 2-9b. A section of the 6x Averaged map I around a α helix. It is contoured at 1.0 sigma of the electron density map. Refined model of residues, B234-B244 is shown.

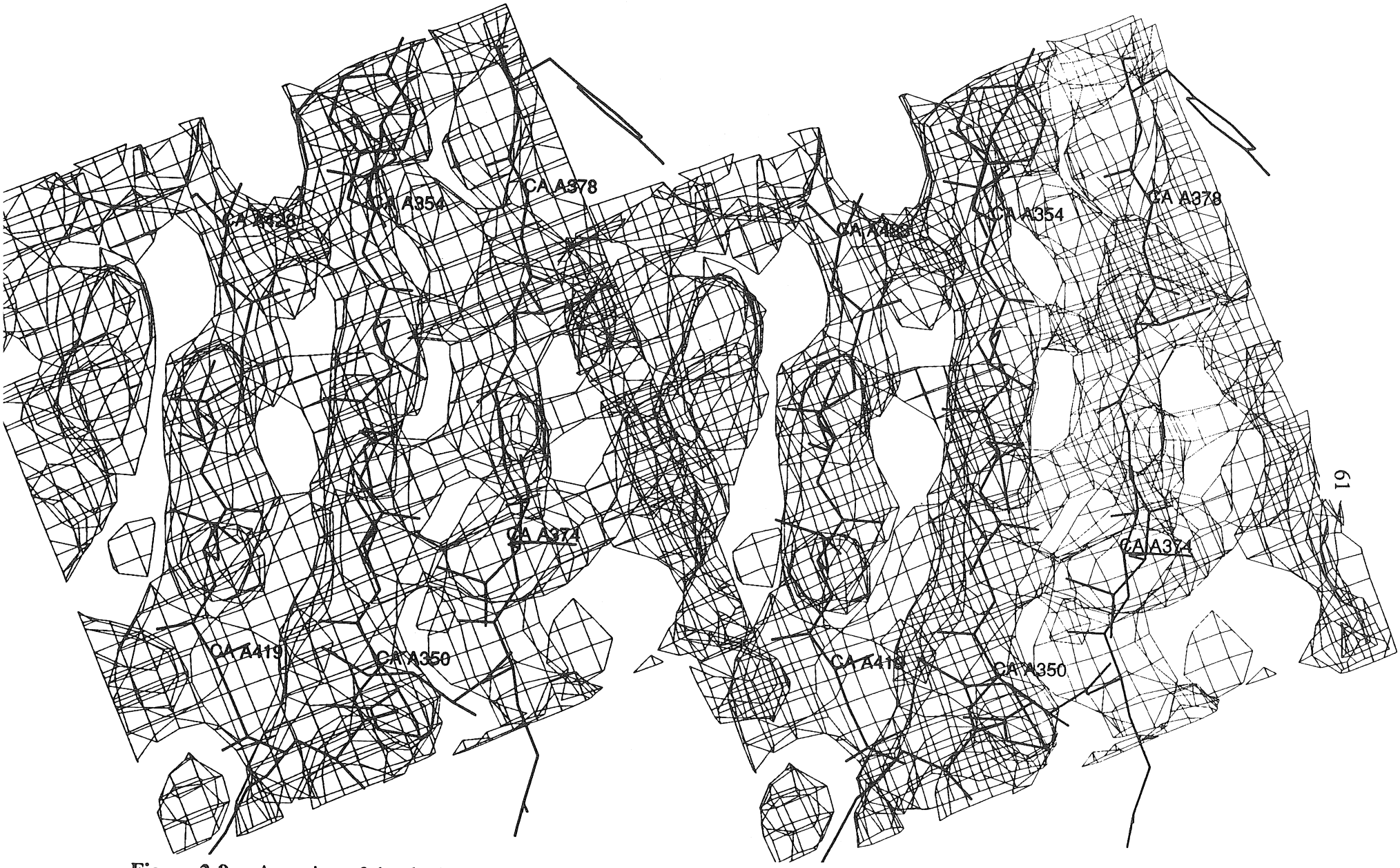


Figure 2-9c. A section of the 6x Averaged map I around a β-sheets. It is contoured at 1.0 sigma of the electron density map. Refined model of residues A418-A423, A349-A355, and A373-A379 is shown.

Chapter 3. Structures of the Metal Centers in the Nitrogenase Molybdenum-Iron Protein

3-1. Introduction

Reduction of dinitrogen to ammonia by the nitrogenase enzyme system requires two metalloproteins, the MoFe-protein and the Fe-protein (reviewed in 1-7). The MoFe-protein is an $\alpha_2\beta_2$ tetramer with a total molecular weight of ~240 kD, and the Fe-protein is a γ_2 dimer with a molecular weight of ~60 kD. Three distinct types of redox centers are associated with these proteins: the MoFe-protein contains two types of novel metal centers, the FeMo-cofactor (also called the M center, reviewed in 7-9) and the P-cluster pair (also called the P-clusters, reviewed in 7,10), and the Fe-protein dimer, whose structure has been solved recently[11], contains one 4Fe:4S cluster . Because the active site of nitrogenase is provided by the MoFe-protein, the redox centers of this protein have attracted considerable attention. The FeMo-cofactor, first identified by Shah and Brill[12], most likely represents the site of substrate reduction and was believed to have the approximate composition of one Mo atom, six to eight Fe atoms, eight to nine S atoms and one homocitrate molecule (1Mo:6-8 Fe:8-9 S:1 homocitrate)[8,9,13-15]. The P-cluster pair was believed to contain two 4Fe:4S clusters[16-21] in close proximity[22-25], although the detailed properties of this center are distinct from better characterized proteins that contain one or more mononuclear 4Fe:4S clusters. The overall metal composition of the MoFe-protein, approximately 2Mo:30Fe:30S, is consistent with the presence of two copies each of the FeMo-cofactor and P-cluster pair per MoFe-protein tetramer.

Despite extensive studies, the detailed structures of the FeMo-cofactor and P-cluster pair were not established. In this chapter, crystallographic structures and some mechanistic features of the FeMo-cofactor and P-cluster pair are described. The structures of the metal centers were originally proposed based on x-ray diffraction data collected to 2.7Å

resolution[26], however, this chapter will also include the results obtained from 2.2Å resolution x-ray diffraction analysis[27].

3-2. Structure determination of the metal centers

The metal centers in the nitrogenase MoFe-protein were identified as the strongest features in the electron density map and are consistent with the highest peaks in both native anomalous difference Patterson and Fourier maps. As native anomalous scattering effects were not used in the phasing of the MoFe-protein structure, the initial electron density maps are not influenced by any specific model for the metal centers. Models for the centers were built into the electron density maps with fragments of 4Fe:4S clusters as the basic building blocks. Metal-ligand distances were restrained to values observed in model compound structures[28,29].

The FeMo-cofactor and P-cluster pair were identified from the positions of amino acid residues that have been deduced from amino acid sequence comparisons and mutagenesis studies to be in the cofactor environment, including residues Cys α275 and His α195 for the FeMo-cofactor and residues Cys α62, Cys α88, Cys α154, Cys β70, Cys β95 and Cys β153 for the P-cluster pair[30-36]. At 2.7Å resolution, atoms are not resolved, so the identities of the various sites were inferred from the available analytical and spectroscopic information. The Mo site in the FeMo-cofactor was assigned to the position in the more electron dense cluster with approximate octahedral coordination geometry, as observed in extended X-ray absorption fine structure (EXAFS) studies[37,38], although octahedral iron coordination in 4Fe:4S clusters has also been observed[39]. An organic molecule, homocitrate, was unambiguously located due to the knowledge of the chemical composition and structure[40]. At 2.2Å resolution, all the Fe atoms and the Mo atom in the metal centers are well resolved in the $2F_O$-F_C electron density map and the S atoms are resolved in the F_O-F_C difference map calculated by omitting S atoms from the metal centers. The 2.2Å

resolution X-ray diffraction analysis has confirmed that the structures of the metal centers at 2.7Å resolution are correct.

3-3. Structure of the FeMo-cofactor

The 2.2Å resolution $2F_O$-F_C electron density map surrounding the FeMo-cofactor is illustrated in Figure 3-1. The Fe atoms and Mo atom are well resolved, although the S atoms are not resolved in this map. The S atoms are resolved in the F_O-F_C map in which F_C values were obtained from a refinement containing the metal positions, but with the cluster sulfur atoms omitted (data not shown). The ball and stick model and the schematic drawing of the FeMo-cofactor are illustrated in Figure 3-2. The FeMo-cofactor contains two clusters of composition 4Fe:3S and 1Mo:3Fe:3S that are bridged by three non-protein ligands. Compounds containing the 4Fe:3S cluster have been described[41], although the terminal ligation of the Fe sites in these molecules differs significantly from that observed in the FeMo-cofactor. Based on the electron density values at the position of the bridging ligands, two of these ligands are assigned as sulfur (presumably S^{2-}), while the third ligand has ~20% lower electron density in the electron density maps calculated at 2.2Å resolution for Av1. This third bridging ligand, designated "Y", is most likely a less-well ordered sulfur, although it might also be a well-ordered O/N species or even a disordered chloride. Homocitrate, which is an essential component of FeMo-cofactor[14-15], coordinates the Mo through a hydroxyl and carboxyl oxygen. There is no evidence in the electron density maps for a hexacoordinate sulfur atom at the center of the cluster. Consequently, a cavity appears inside the cofactor between the two cluster fragments. The longest metal-metal distance in the FeMo-cofactor is ~7.4Å between sites Fe1 and Mo. The Fe-Fe separation distance between bridged iron sites (such as Fe2 and Fe6) is ~2.5Å, while the distance between non-bridged iron sites (such as Fe2 and Fe5, etc.) is ~3.8Å. The distance between the Mo and the three closest irons of the 4Fe:3S fragment averages

~5.2Å, which is approximately the same as the distance between Fe1 and the three irons of the 1Mo:3Fe:3S fragment.

If the three-fold axes of the isolated 4Fe:3S and 1Mo:3Fe:3S clusters are superimposed, then the cluster is liganded to the protein through the metals (Fe1 and Mo) located on the three-fold axis. The protein environment around the FeMo-cofactor is primarily provided by the α subunit. Cysteine α275 coordinates Fe1, while the Mo is liganded by the side chain of His α442. Site directed mutagenesis studies had implicated Cys α275 as a cofactor ligand[30,31,33]. Based on sequencing of Nif$^-$ mutants altered in the α subunit, Zamir had proposed that His α442 could coordinate the FeMo-cofactor[42]. The side chains of two other residues proposed to be near the cofactor, His α195[32,35,36] and Gln α191[43], are within 4Å of Fe2 and Fe6, respectively, but these residues are not directly liganded to the metals. Site directed substitution of His α195 results in the loss of a nitrogen ESEEM signal from the MoFe-protein[36], which may originate from the coupling of that residue to the S = 3/2 spin system of the cofactor through a side-chain hydrogen bond to one of the bridging sulfurs. The side chain of Gln α191 interacts with one of the carboxyl groups of homocitrate, which is consistent with a proposal based on site-directed mutagenesis studies[43]. Cys α275, His α442, His α195 and Gln α191 are conserved in all known MoFe-protein sequences.

The tetrahedral coordination geometry of Fe1 and the octahedral coordination geometry for Mo are typical of the coordination environments for these metals observed in model compounds and Fe:S proteins[10,28,29]. An unusual feature of the FeMo-cofactor model, however, is the trigonal ligand geometry of the six Fe sites that bind the bridging sulfurs. Trigonal coordination geometry for Fe is not unprecedented, however, and has been described in a small molecule structure of an iron-thiolate species[44]. In this case, the coordinating thiols contained bulky substitutents, and it was proposed that the low coordination number of the Fe atoms reflected the effects of steric crowding between the ligands. No such corresponding features in the FeMo-cofactor environment are evident,

however. Solvent molecules that might possibly serve as fourth ligands have not been observed. It is also unlikely, although not impossible, that a less electron dense group (such as O or N) could be present in the cluster interior. The separation distance of ~2.5Å between bridged irons suggest that there might be some Fe--Fe bonding that could serve as a fourth coordination interaction.

The structural features of the FeMo-cofactor model are generally consistent with the results of analytical and spectroscopic (EXAFS, ENDOR and Mössbauer) studies of the cofactor. The composition of the non-protein part of the FeMo-cofactor model, 1Mo:7Fe:9-8S:1 homocitrate, is within the range of values that have been reported[8,13-15]. The absence of protein bound bridging ligands between the two clusters in FeMo-cofactor is consistent with the ability to extract the intact cofactor from MoFe-protein. EXAFS studies of the Mo environment in both the MoFe-protein and isolated cofactor indicate that 2-3 O(N) and 3-5 S are directly coordinated to Mo[37,38], with 3-4 Fe present in the second coordination shell of Mo; the crystallographic model contains 3 O(N), 3 S and 3 Fe. Studies of Fe EXAFS on the isolated cofactor[45] indicate that, on average, each Fe is surrounded by ~3 S, 0.3-0.8 Mo (~2.7Å), with ~2.2 and ~1.3 Fe atoms located at ~2.7Å and ~3.8Å, respectively. Assuming Y is a sulfur, the average Fe environment in the present FeMo-cofactor model bound to the MoFe-protein contains ~3.1 S and ~0.4 Mo with ~3.4 and 0.9 Fe at 2.7Å and 3.8Å, respectively. ^{57}Fe ENDOR studies[46] have indicated that five magnetically inequivalent iron species are present; neglecting the asymmetric protein environment and the chiral homocitrate ligand and assuming that the Y ligand is distinct from the other two bridging ligands, five different types of Fe sites are observed in the structural model (Fe1; Fe3; Fe7; Fe2 and Fe6; and Fe4 and Fe5). Mössbauer studies[18,19] identified 5-7 iron sites in the FeMo-cofactor that could be grouped into two sets; however, assignment of these sites to particular atoms in the FeMo-cofactor model cannot be made at present.

3-4. Binding of substrates to the FeMo-cofactor

In addition to the physiological reaction 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[47]. Functionally, the details of the interaction between various substrates and the FeMo-cofactor are central to understanding the catalytic properties of nitrogenase. Based on the FeMo-cofactor structure, three possible binding modes of substrates to the FeMo-cofactor 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 (Figure 3-3a). 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) Substrates could bind in a side-on fashion to the exterior surface of the FeMo-cofactor (Figure 3-3b). 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 iron surface, each containing four iron atoms, and substrates could bind to those surfaces so as to simultaneously interact with up to four irons. (iii) Some small substrates, such as N_2 and/or $H^+(H_2)$ could occupy the central cavity in the FeMo-cofactor, thereby replacing the weak iron-iron bonds with Fe--substrate bonds (Figure 3-3c). 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)[48]. The octahedral coordination of Mo in the FeMo-cofactor suggests that the Mo does not directly participate in substrate binding, without the occurrence of either a change in coordination number or a change in liganding groups. The function of Mo in FeMo-cofactor could be the modulation of reduction potential adequate for N_2 binding and reduction, or perhapse a structural role to anchor the cofactor and/or homocitrate in proper position.

3-5. Structure of the P-cluster pair

The 2.2Å resolution $2F_O$-F_C electron density map surrounding the P-cluster pair is illustrated in Figure 3-4. The Fe atoms are well resolved, although the S atoms are not resolved in this map. The S atoms are resolved in the F_O-F_C map in which F_C values were obtained from a refinement containing the metal positions, but with the cluster sulfur atoms omitted. The ball and stick model and the schematic drawing of the P-cluster pair are shown in Figure 3-5. The P-cluster pair consists of two 4Fe:4S clusters that are bridged by two cysteine thiol ligands (residues Cys α88 and Cys β95). The 4Fe:4S clusters are further connected by a disulfide bond between two of the cluster sulfurs. This disulfide bridge is particularly intriguing and it is located on the side of the P-cluster pair closest to proposed binding site for Fe-protein (see Chapter 4 and 5). To our knowledge this is the first example of a 4Fe:4S cluster containing a $(\mu_3\text{-S})_2$ disulfide, although an intermolecular disulfide bridged Fe_2S_2 cluster has been synthesized previously[49-51]. Singly coordinating cysteine thiols (from residues Cys α62, Cys α154, Cys β70 and Cys β153) ligate the remaining four irons, such that non-bridging cysteines coordinated to a specific 4Fe:4S cluster are from the same subunit. In addition to the cysteine ligands, Ser β188 appears to coordinate Fe6 along with Cys β153. The coordination environments of both Fe6 and Fe2 are distorted from ideal tetrahedral geometry, due to interactions with Ser β188 and the main chain of Gly α185, respectively.

As had been proposed from sequence comparisons and mutagenesis experiments[30,31,33], the P-cluster pair is located at the interface between the α and β subunits. The approximate two-fold symmetry of the P-cluster pair is reflected in the sequence similarities between the polypeptide chains for the α and β subunits in this region[52]. The cysteine ligands to the P-cluster pair had been correctly identified by mutagenesis experiments[30,31,33,34]. Individual replacement of any of the six cysteines by alanine eliminates diazotropic growth of the mutant strains. However, some non-alanine substitutions for Cys α88 and Cys β153 have been described that can still produce nitrogenase with low activity, possibly as a consequence of the substituted residue

functioning as a cluster ligand, through a structural rearrangement that allows a previously non-liganding group to coordinate the cluster, or simply because no ligand is actually needed. Additionally, substitution of Gly α185 with Asp has been found in a Nif$^-$ mutant of nitrogenase[42].

Mössbauer and extrusion studies indicated that the P-cluster pair contained 4Fe:4S clusters[16-24]. The proximity of two 4Fe:4S clusters in the P-cluster pair was suggested from an analysis of anomalous scattering effects from Cp1[25]. The properties of the 4Fe:4S groups in the P-cluster pair differ in significant ways from "typical" 4Fe:4S clusters, however. The close spatial proximity of two 4Fe:4S clusters in MoFe-protein suggests that interactions between clusters might confer distinctive properties. Mössbauer studies[20,21,23] have identified three distinct classes of iron sites in the P-cluster pair, designated Fe^{2+}, D and S, that are present in the approximate ratio of 4:10:2 in the MoFe-protein. Because the environment of each metal site is necessarily unique when coordinated to asymmetric protein, it is not clear what effects distinguish the different Mössbauer classes, so that the assignment between P-cluster pair iron sites, and Mössbauer classes has not yet been determined.

3-6. Role of the P-cluster pair

The presence of the disulfide bond in the P-cluster pair suggests that this center can act as a two electron redox group, involving cleavage and reformation of the μ_3-disulfide bridge coupled to the transfer of pairs of electrons into the FeMo-cofactor. The role of the P-cluster pair could then be to mediate the electron transfer processes within the nitrogenase proteins by taking one electron reducing equivalents provided by the Fe-protein and converting them into a two electron transfer to the FeMo-cofactor. In addition to a purely electron transfer function, this disulfide bond may also provide a site for H_2 evolution by nitrogenase. Hydrogen production has been suggested to accompany disulfide bond formation between two Fe_2S_2 clusters in model systems[49]. In this case, reaction between

two $Fe_2S(SH)$ clusters generates H_2 concomitantly with formation of the disulfide bridged cluster dimer. It is possible that protonation of the doubly reduced P-cluster pair in nitrogenase may generate a similar species that can produce H_2 upon disulfide bond formation.

The disulfide bridge also provides a mechanism for coupling conformational changes to redox reactions at the P-cluster pair. Breakage of the disulfide bond would result in an increased separation distance between the 4Fe:4S clusters in the P-cluster pair, because the bonded S-S distance is >~1Å shorter than the non-bonded distance. To accommodate the increased cluster separation produced by disulfide bond cleavage, some change in the coordination geometry of the two clusters by the cysteine ligands must occur. These changes would serve to link the oxidation state of the P-cluster pair and the conformation of the MoFe-protein. As the P-cluster pair is at the αβ subunit interface in the MoFe-protein, and because this region of the MoFe-protein appears to interact with the Fe-protein, it is likely that conformational alterations associated with changes in P-cluster pair oxidation state could be propagated to, and influence the properties of, both the FeMo-cofactor and Fe-protein. Conversely, interactions between MoFe-protein and nucleotide bound Fe-protein would influence the redox behavior of the P-cluster pair, and this coupling between protein conformation and oxidation state may be related to the role of MgATP in the nitrogenase reaction.

3-7. Organization of the metal centers in MoFe-protein

As had been indicated by the overall metal composition data, MoFe-protein contains two copies each of the FeMo-cofactor and P-cluster pair. The organization of the metal centers in MoFe-protein is illustrated in Figure 3-6. As demonstrated by Bolin *et al.*[25], these centers are separated by ~19Å, whereas the distances between centers on the other αβ subunit pair of the MoFe-protein are ~70Å. The closest distance between the two FeMo-cofactors, ~64Å, suggests that the two αβ dimers function independently. The distance of

closest approach between metal sites in FeMo-cofactor and P-cluster pair is ~14Å. This suggests that the electron transfer rate[53,54] between these centers could be faster than the rate of nitrogenase turnover (~5 sec^{-1})[55-58].

References

1. Burgess, B. K. (1984) in *Advances in Nitrogen Fixation* (Veeger, C. & Newton, W. E. eds.) pp. 103-114, Mortinus Nijoff, Boston.

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

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

4. Stiefel, E. I., *et al.* (1988) in *Metal Clusters in Proteins* (Que, L. eds.) pp. 372-389, American Chemical Society, Washington, D.C.

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

6. Stacey, G., Burris, R. H. & Evans, H. J., eds. (1992) *Biological Nitrogen Fixation*, Chapman & Hall, New York.

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

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

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

10. Holm, R. H., Ciurli, S. & Weigel, J. A. (1990) *Prog. Inorg. Chem.* **38**, 1-75.

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

12. Shah, V. K. & Brill, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3249-3253.

13. Nelson, M. J., Levy, M. A. & Orme-Johnson, W. H. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 147-150.

14. Hoover, T. R., Robertson, A. D., Cerny, R. L., Hayes, R. N., Imperial, J., Shah, V. K. & Ludden, P. W. (1987) *Nature* **329**, 855-857.

15. Madden, M. S., Kindon, N. D., Ludden, P. W. & Shah, V. K. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6517-6521.

16. Smith, B. E. & Lang, G. (1974) *Biochem. J.* **137**, 169-180.

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

18. Zimmermann, R., Münck, E., Brill, W. J., Shah, V. K., Henzl, M. T., Rawlings, J. & Orme-Johnson, W. H. (1978) *Biochem. Biophys. Acta* **537**, 185-207.

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

20. McLean, P. A., Papaefthymiou, V., Orme-Johnson, W. H. & Münck, E. (1987) *J. Biol. Chem.* **262**, 12900-12903.

21. Lindahl, P. A., Papaefthymiou, V., Orme-Johnson, W. H. & Münck, E. (1988) *J. Biol. Chem.* **263**, 19412-19418.

22. Hagen, W. R., Wassink, H., Eady, R. R., Smith, B. E. & Haaker, H. (1987) *Eur. J. Biochem.* **169**, 457-465.

23. Surerus, K. K., Hendrich, M. P., Christie, P. D., Rottgardt, D., Orme-Johnson, W. H. & Münck, E. (1992) *J. Am. Chem. Soc.* **114**, 8579-8590.

24. Pierik, A. J., Wassink, H., Haaker, H. & Hagen, W. R. (1993) *Eur. J. Biochem.* **212**, 51-61.

25. Bolin, J. T., Ronco, A. E., Mortenson, L. E., Morgan, T. V., Williamson, M. & Xuoung, 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 & Hall, New York.

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

27. Chan, M. K., Kim, J. & Rees, D. C. (1993) *Science* in press.

28. Herskovitz, T., Averil, B. A., Holm, R. H., Ibers, J. A., Phillips, W. D. & Weiher, J. F. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 2437-2441.

29. Coucouvanis, D. (1991) *Acct. Chem. Res.* **24**, 1-8.

30. Kent, H. K., Ioannidis, I., Gormall, C., Smith, B. E. & Buck, M. (1989) *Biochem. J.* **264**, 257-264.

31. Kent, H. K., Baines, M., Gormal, C., Smith, B. E. & Buck, M. (1990) *Mol. Microbiol.* **4**, 1497-1504.

32. Scott, D. J., May, H. D., Newton, W. E., Brigle, K. E. & Dean, D. R. (1990) *Nature* **343**, 188-190.

33. Dean, D. R., Setterquist, R. A., Brigle, K. E., Scott, D. J., Laird, N. F. & Newton, W. E. (1990) *Mol. Microbiol.* **4**, 1505-1512.

34. May, H. D., Dean, D. R., Newton, W. E. (1991) *Biochem. J.* **277**, 457-464.

35. Thomann, H., Morgan, T. V., Jin, H., Burgmayer, S. J. N., Bare, R. E. & Stiefel, E. I. (1987) *J. Am. Chem. Soc.* **109**, 7913-7914.

36. Thomann, H., Bernardo, M., Newton, W. E. & Dean, D. R. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 6620-6623.

37. Eidness, M. K., Flank, A. M., Smith, B. E., Flood, A. C., Garner, C. D. & Cramer, S. P. (1986) *J. Am. Chem. Soc.* **108**, 2746-2747.

38. Conradson, S. D., Burgess, B. K., Newton, W. E., Mortenson, L. E. & Hodgson, K. O. (1987) *J. Am. Chem. Soc.* **109**, 7507-7515.

39. Lauble, H., Kennedy, M. C., Beinert, H. & Stout, C. D. (1992) *Biochem.* **31**, 2735-2748.

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

41. Johansson, G. & Libscomb, W. N. (1958) *Acta Cryst.* **11**, 594-598; Chu, C. T. -W. & Dahl, F. (1977) *Inorg. Chem.* **16**, 3245-3251.

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

43. Scott, D. J., Dean, D. R. & Newton, W. E. (1990) in *Nitrogen Fixation: Achievements and Objectives* (Gresshoff, P. M., Roth, L. E., Stacey, G., Newton, W. E., eds.) p. 169. Chapman & Hall, New York.

44. Power, P. P. & Shoner, S. C. (1991) *Angew. Chem. Int. Ed. Engl.* **30**, 330-332.

45. Arber, J. M., Flood, A. C., Garner, C. D., Gorman, C. A., Hasnain, S. & Smith, B. E. (1988) *Biochem. J.* **252**, 421-425.

46. True, A. E., Nelson, M. J., Venters, R. A., Orme-Johnson, W. H. & Hoffmann, B. M. (1988) *J. Am. Chem. Soc.* **110**, 1935-1943.

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

48. Rivera-Ortiz, J. M. & Burris, R. H. (1975) *J. Bacteriol.* **123**, 537-545.

49. Bose, K. S., Sinn, E. & Averill, B. A. (1984) *Organometallics* **3**, 1126-1128.

50. Seyferth, D., Kiwan, A. M. & Sinn, E. (1985) *J. Organomet. Chem.* **281**, 111-118.

51. Blonk, H. L., *et al.* (1992) *Inorg. Chem.* **31**, 962-968.

52. Lammers, P. J. & Haselkorn, R. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4723-4727.

53. Marcus, R. N. & Sutin, N. (1985) *Biochim. Biophys. Acta* **811**, 265-322.

54. Wuttke, D. S., Bjerrum, M. J., Winkler, J. R. & Gray, H. B. (1992) *Science* **256**, 1007-1009.

55. Thorneley, R. N. F. & Lowe, D. J. (1984) *Biochem. J.* **224**, 887-894.

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

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

58. Hageman, R. V. & Burris, R. H. (1978) *Biochem.* **17**, 4117-4124.

59. Kraulis, P. J. (1991) *J. Appl. Cryst.* **24**, 946-950.

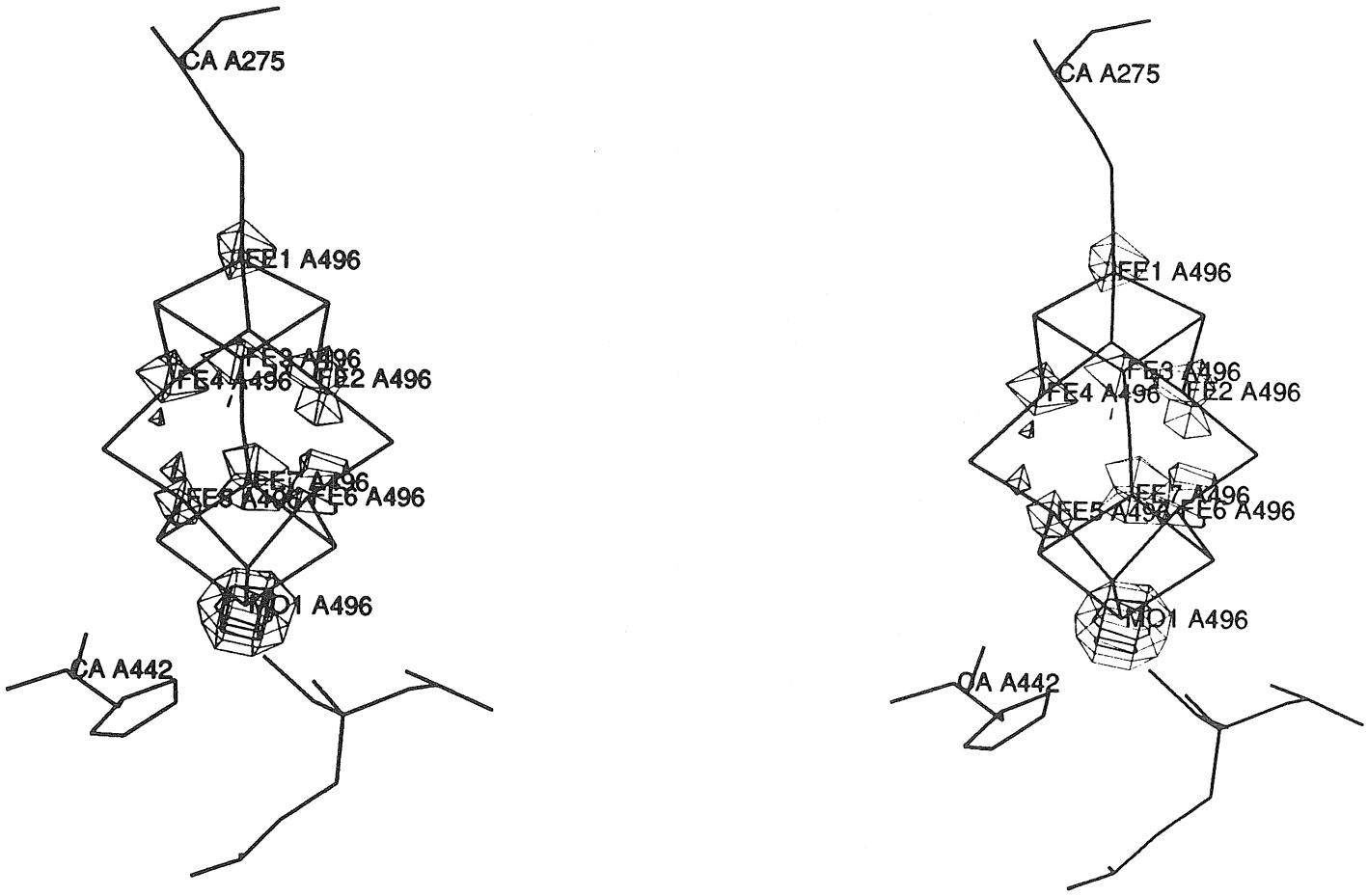


Figure 3-1. A 2.2Å $2F_o$-F_c electron density map around the FeMo-cofactor. It is contoured at 9 sigma and 18 sigma of the electron density map. All the iron and molybdenum atoms are well resolved.

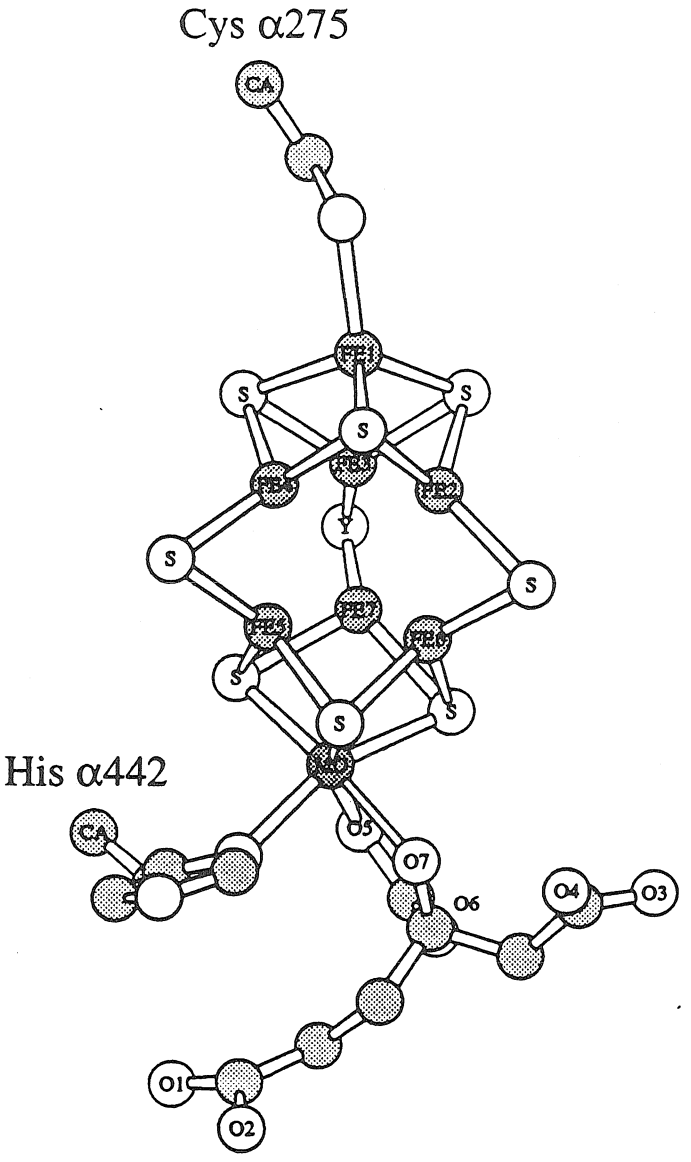


Figure 3-2a. Ball and stick model of the FeMo-cofactor with protein ligands. The MOLSCRIPT[59] program was used to draw the model. The FeMo-cofactor contains 4Fe:3S and 1Mo:3Fe:3S clusters that are bridged by three non-protein ligands, and an organic molecule, homocitrate. The "Y" ligand could be disordered S or a well-ordered O or N species.

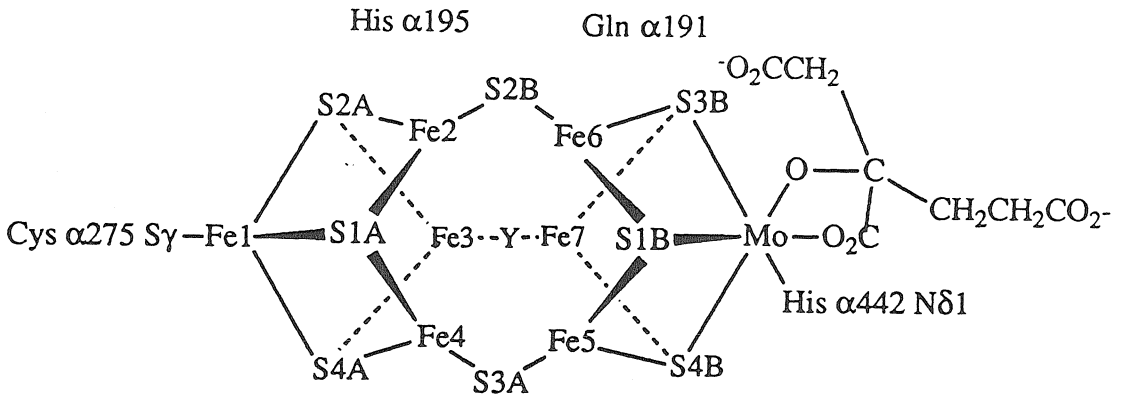


Figure 3-2b. Schematic representation of the FeMo-cofactor with protein ligands. The FeMo-cofactor is attached to the α subunit through two protein ligands, Cys α275 and His α442. Homocitrate, an essential component of the FeMo-cofactor, is coordinated through hydroxyl and carboxyl oxygens to the molybdenum site.

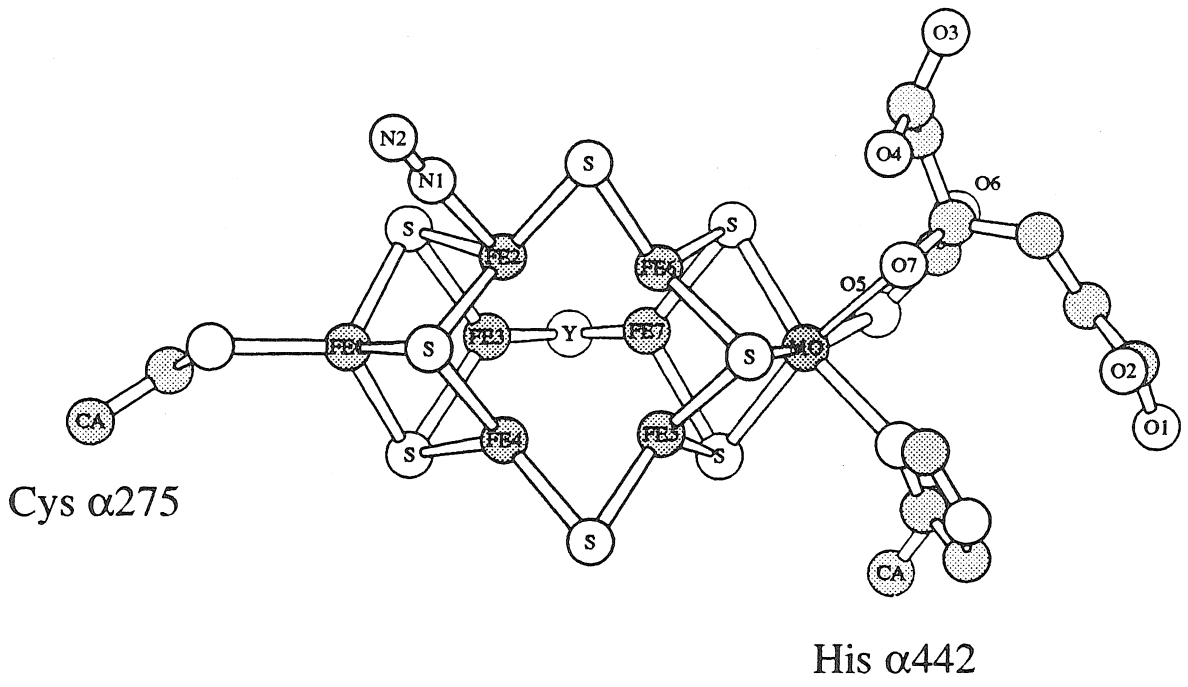


Figure 3-3a. Substrate binding mode I. Some substrates could bind in an end-on fashion to one of the six central iron sites which are bridged by non-protein ligands.

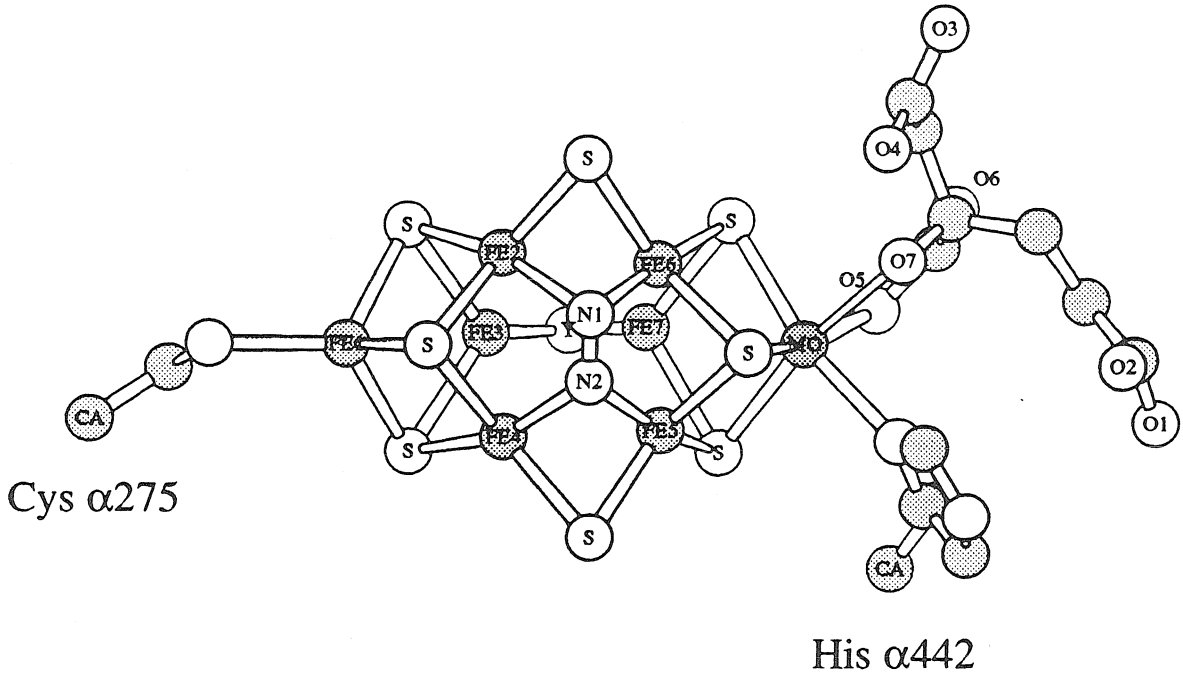


Figure 3-3b. Substrate binding mode II. Some substrates could bind to the exterior surface of the FeMo-cofactor, which is composed of 4 irons and 4 sulfurs.

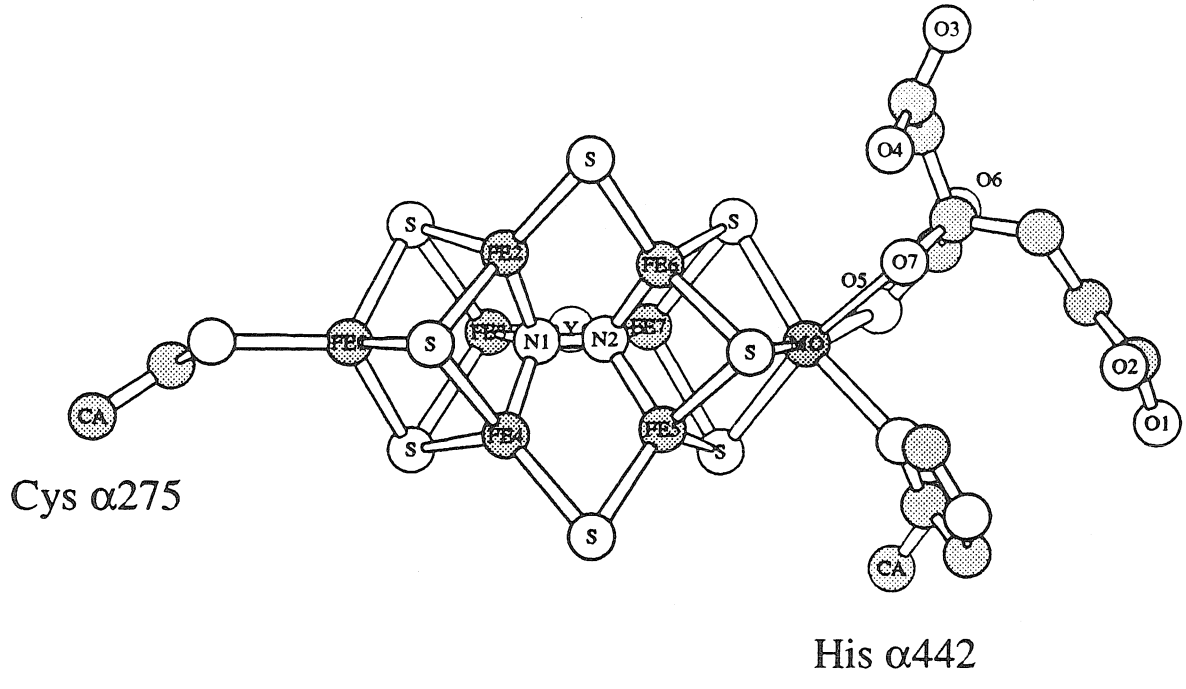


Figure 3-3c. Substrate binding mode III. Some small substrates, such as N_2 and/or H_2 could occupy the central cavity in the FeMo-cofactor. However, model studies indicate that the FeMo-cofactor should open a little to accomodate N_2 in the central cavity.

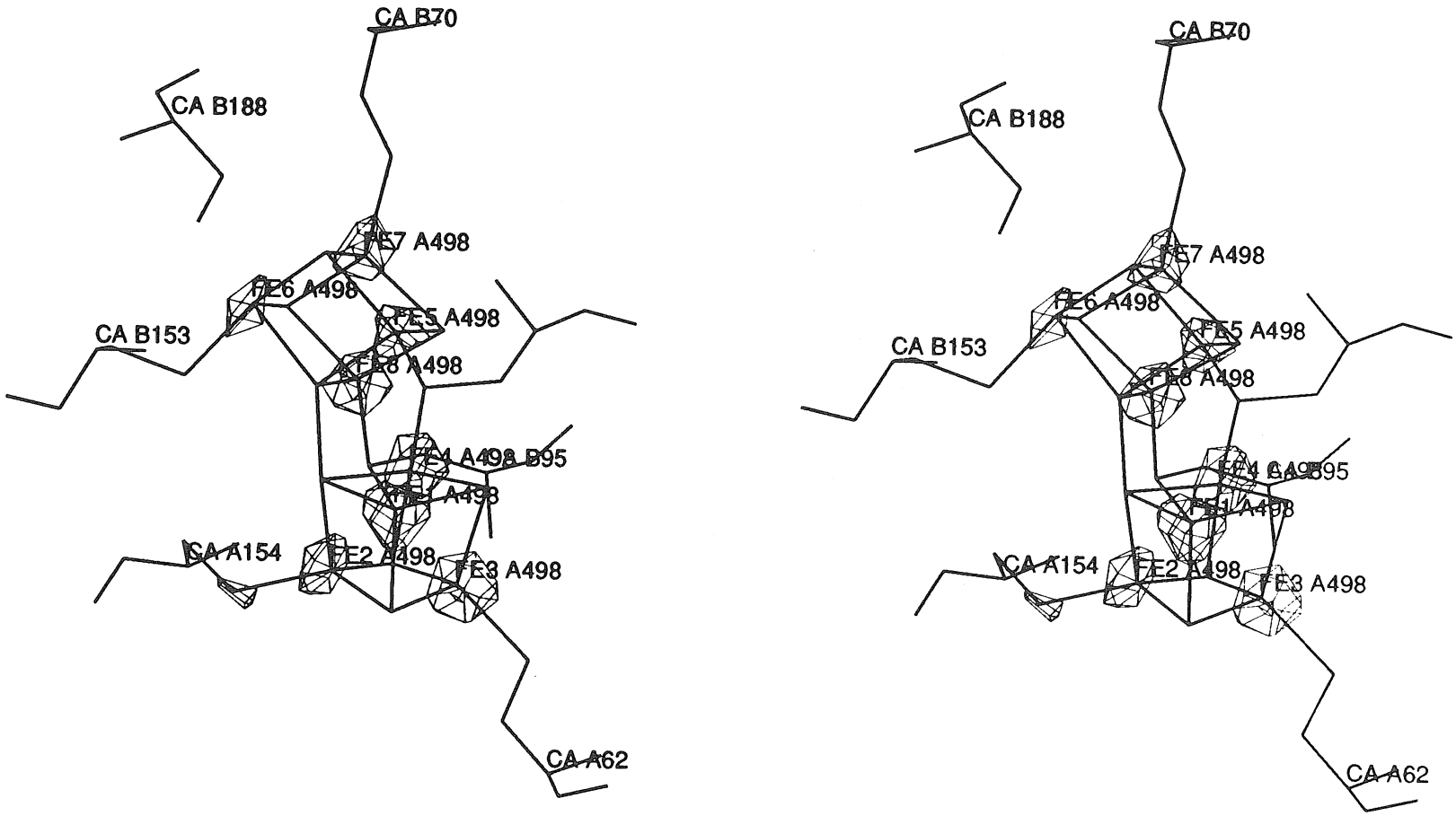


Figure 3-4. A 2.2Å 2Fo-Fc electron density map around the P-cluster pair. It is contoured at 9 sigma of the electron density map. All the iron atoms are well resolved.

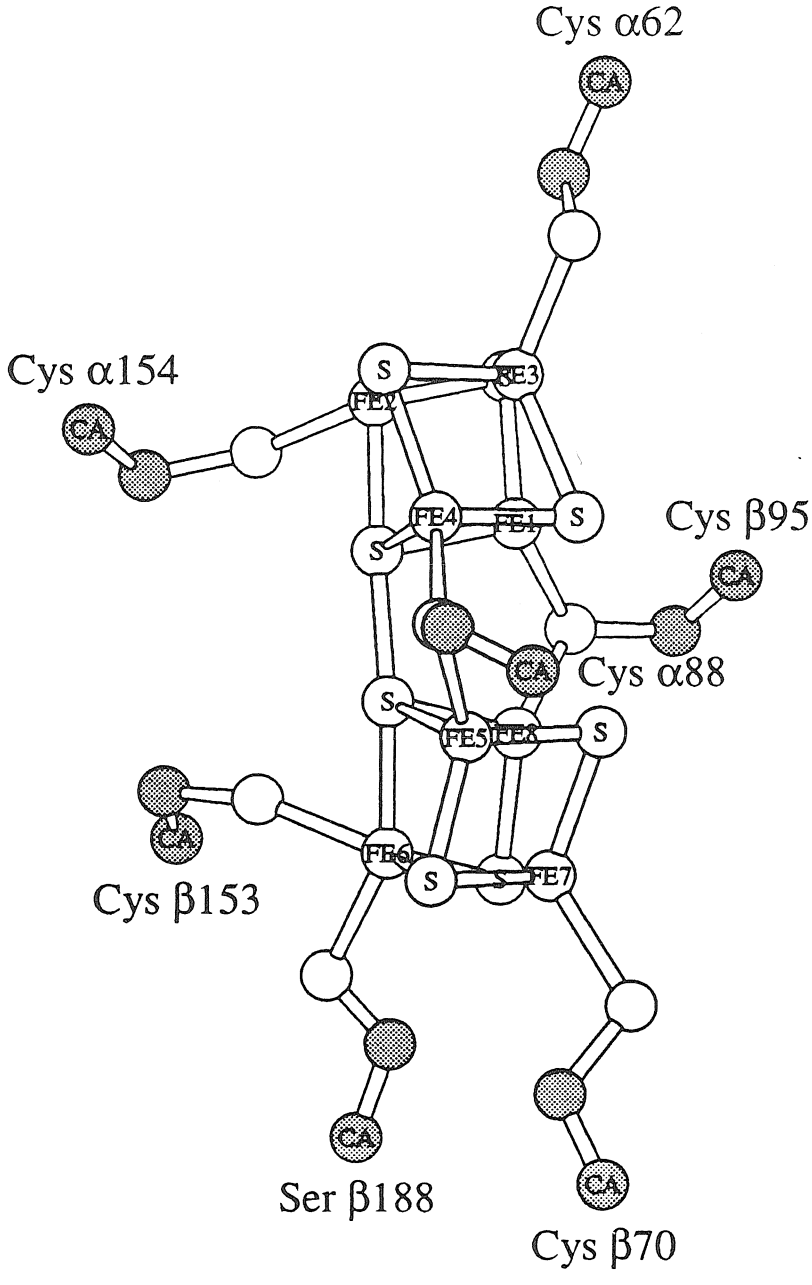


Figure 3-5a. Ball and stick model of the P-cluster pair with protein ligands. The MOLSCRIPT program was used to draw the model. The P-cluster pair consists of two 4Fe:4S clusters that are bridged by two cysteine ligands and a disulfide bond between the two clusters.

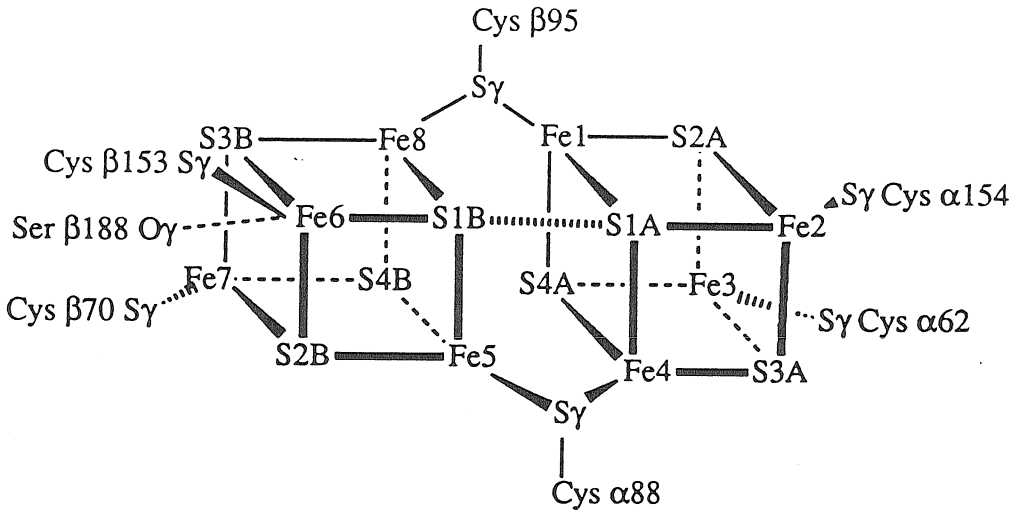


Figure 3-5b. Schematic representation of the P-cluster pair with protein ligands. The P-cluster pair is attached at the interface between the α and β subunits through seven protein ligands: Cys α62, Cys α88, Cys α154, Cys β70, Cys β95, Cys β153 and Ser β188.

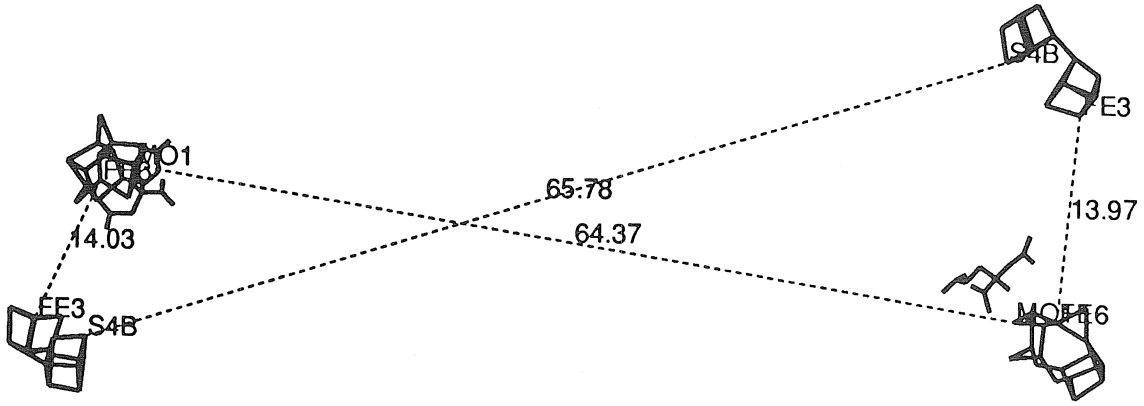


Figure 3-6. Organization of the metal centers in MoFe-protein. The distance of closest approach between metal sites in FeMo-cofactor and P-cluster pair is ~14Å and the closest distance between the two FeMo-cofactors is ~64Å.

Chapter 4. Crystallographic Structure and Functional Implications of the Nitrogenase Molybdenum-Iron Protein from *Azotobacter vinelandii*

4-1. Introduction

The enzymatic reduction of dinitrogen to ammonia during biological nitrogen fixation is essential for maintaining the nitrogen cycle on earth. The biochemical machinery for nitrogen fixation, provided by the nitrogenase system[1-7], occurs throughout a diverse range of prokaryotes. The conventional nitrogenase enzyme system consists of two component metalloproteins, the molybdenum-iron (MoFe-) protein and the iron (Fe-) protein, although molybdenum may be replaced by other metals in alternative nitrogenase systems[8,9]. The overall stoichiometry of the biological nitrogen fixation reaction[10]:

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

reflects the requirements for reducing equivalents, MgATP, and protons, in addition to the two nitrogenase proteins. The Fe-protein is a dimer of two identical subunits (total molecular weight of ~60 kD), whose crystallographic structure has been recently described[11]. The Fe-protein is initially reduced by ferredoxin or flavodoxin, *in vivo*, and subsequently transfers electrons to the MoFe-protein in a process that is coupled to the hydrolysis of MgATP. The MoFe-protein is an $\alpha_2\beta_2$ tetramer with a total molecular weight of ~240 kD. The two subunits are of similar size; for example the isolated α and β subunits of *Azotobacter vinelandii* MoFe-protein have 491 and 522 amino acids, respectively[12]. The MoFe-protein contains two types of metal centers, the FeMo-cofactor and P-cluster pair, for which structures are described in Chapter 3. As detailed in Chapter 3, the FeMo-cofactor has 4Fe:3S and 1Mo:3Fe:3S clusters bridged by three non-protein ligands, with homocitrate coordinated to the Mo site, while the P-cluster pair contains two

4Fe:4S clusters bridged by two cysteine ligands and one disulfide bond. The FeMo-cofactor almost certainly represents the site of dinitrogen binding and reduction, while the P-cluster pair appears to function in electron transfer between Fe-protein and the FeMo-cofactor. Both metal centers are completely surrounded by the polypeptide chains of the MoFe-protein, which must serve to organize the redox centers in the proper spatial relationships, control entry and exit of substrates and products, and are essential for the biosynthesis and assembly of the cofactors. In this chapter, the tertiary and quaternary structures of the MoFe-protein from *A. vinelandii* are presented, based on a 2.7Å resolution x-ray crystallographic analysis. The crystallographic model provides a structural basis for interpreting the functional role of the nitrogenase MoFe-protein in dinitrogen reduction.

4-2. Structure determination

The purification, crystallization and structure determination of the nitrogenase MoFe-protein are described in Chapter 2. Briefly, three crystal forms were prepared of the MoFe-protein from *A. vinelandii* (abbreviated Av1) and *Clostridium pasteurianum* (abbreviated Cp1). Each crystal form contained one tetramer per asymmetric unit. Phases were initially determined by the method of multiple isomorphous replacement (MIR)[13], and were subsequently refined by averaging both within and between different crystal forms[14,15]. The most detailed structural analysis has been currently achieved for an Av1 crystal form, space group $P2_1$ with unit cell constants of a = 108.4Å, b = 130.5Å, c = 81.5Å, and β = 110.8°. Initially, a polyalanine backbone model was built into the averaged electron density map using the graphics program TOM/FRODO[16] implemented with a fragment fitting option, and refined with the restrained least square program TNT[17]. Phases calculated from the model were combined with the MIR phases and re-averaged. The electron density map was significantly improved by iterative cycles of model building, refinement, phase combination and noncrystallographic symmetry (NCS) averaging. During successive

cycles of phase refinement, the resolution was gradually extended from 3.5Å to 2.7Å. The program TNT was used for coordinate refinement during the initial stages of modeling, while the simulated annealing program X-PLOR[18] was used during the final stages of refinement. Residues were rebuilt following inspection of $2F_O$-F_C and F_O-F_C maps, and examination of Eisenberg's 3D-1D profile[19] calculated for each model.

The current model contains 1980 amino acids residues (of 2026 total residues) with 15,858 non-hydrogen atoms (~98% complete). This model has been refined to a crystallographic R-factor of 0.188 (8-2.7Å resolution) with root mean square deviations from ideal bond distances and angles of 0.016Å and 3.3°, respectively (Table 4-2). A Ramachandran plot of this model is presented in Figure 4-1. Residue numbers of the MoFe-protein are prefixed with either α or β to indicate the appropriate subunit number, unless 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. No subunit prefix is included for residues in the nitrogenase Fe-protein.

Table4-1. Heavy atom binding sites in crystals of A. vinelandii MoFe-protein. Locations for the heavy atoms were determined from difference Patterson maps, and were used to calculate multiple isomorphous replacement (MIR) phases with the program HEAVY[56].

Derivative	x	y	z	relative occupancy	location
EMTS	0.144	-0.265	0.124	0.41	Cys α472
	0.423	0.000	0.124	0.39	Cys α472 (NCS related)
	-0.198	-0.272	0.029	0.14	Cys α45
	0.714	0.008	0.028	0.14	Cys α45 (NCS related)
	-0.172	-0.335	-0.352	0.11	Cys α324, Glu α325
	0.487	0.070	-0.347	0.06	Cys α324, Glu α325 (NCS related)

PtCl4	0.292	-0.321	0.323	0.36	Met β330
	0.377	0.061	0.314	0.39	Met β330 (NCS related)
	0.250	-0.167	0.184	0.33	Met β491
	0.350	-0.092	0.190	0.30	Met β491 (NCS related)
	-0.198	-0.277	0.022	0.28	Met α391
	0.711	0.018	0.016	0.25	Met α391 (NCS related)
	-0.240	-0.271	-0.293	0.11	Met α295, Lys α315
	0.583	0.008	-0.300	0.13	Met α295, Lys α315 (NCS related)
	-0.024	-0.335	0.209	0.11	Met α4? (not identified yet)
	0.618	0.067	0.171	0.11	Met α4? (NCS related, not identified yet)

PIP has seven common binding sites with the PtCl4 derivative

Table 4-2. Refinement statistics of the Av1 structure. The protein model was built with the graphics program TOM/FRODO[16]. The electron density map was improved by iterative cycles of model building, refinement, phase combination and NCS averaging. The program TNT[17] was used for coordinate refinement during the initial stages of modeling, whereas the simulated annealing program X-PLOR[18] was used during the final stages of refinement with the PARAM19X.PRO parameter file.

Model status

Number of amino acid residues	1980/2026
Total non-hydrogen atoms	15,858 (98% complete)
Number of water molecules	4 (ligands to divalent cation sites)
Missing residues	α2-α4, α36-α44, α482-α492
Number of Ramachandran outliers	10/1980

Refinement statistics

	Before refinement	After refinement (X-PLOR)

R factor (8.0 - 2.7Å)	0.368	0.188
R.M.S. deviation of		
bond lengths (Å)	0.020	0.016
bond angles (°)	2.77	3.29
dihedral angles (°)	27.95	24.15
improper torsion (°)	1.50	1.41

4-3. Description of the protein structure

4-3-1. Structures of the α and β subunits

The α and β subunits exhibit similar polypeptide folds consisting of three domains (domain I, II and III) of the α/β type, with some extra helices (Figures 4-2a, b and c). The overall shape of these subunits may be described as a clover leaf. Similarities between the sequences of the amino terminal regions of the α and β subunits had been previously noted[20]. Although the α/β type fold is a common motif in protein structure, no significant homology to other proteins has been identified for the α and/or β subunits by either sequence comparisons or 3D-1D profile methods[19] (D. S. Eisenberg, personal communication), with the exception of component 1 proteins of alternative nitrogenases, and the protein products of the *nifE* and *nifN* genes. These latter two proteins, which function in the incorporation of the FeMo-cofactor into the MoFe-protein, had been previously recognized[21] to have similar sequences to the α and β subunits, respectively. In the α subunit, domain I is composed of 4 parallel β strands, 1 antiparallel β strand which is provided by the β subunit and 7 helices; domain II is composed of 4 parallel β strands and 5 helices; and domain III is composed of 5 parallel β strands, 1 antiparallel β strand which is from N terminus of the α subunit and 8 helices (Figure 4-2c). The three domains of the β subunit (Figures 4-3a, b and c) are designated I', II' and III', and have similar folds to the corresponding domains in the α subunit (Figure 4-3c), with the

exception that the antiparallel β strands present in domains I and III are absent in domains I' and III'. In each subunit, there is a wide, shallow cleft between the three domains; in the α subunit, the FeMo-cofactor occupies the bottom of this cleft. This location for the FeMo-cofactor at the boundary between three domains is reminiscent of the site for the iron-sulfur cluster in aconitase[22]. The site in the β subunit corresponding to location of the FeMo-cofactor in the α subunit contains the side chains of residues His β193, Gln β294, His β297 and Asp β372.

4-3-2. Structure of αβ dimer

The αβ dimer of MoFe-protein consists of 6 α/β domains (three from each subunit) which are arranged like a flower with 6 petals (Figures 4-4a and b). Extensive contacts occur between the α and β subunits; this is especially evident in the bridging of domain I of the α subunit and domain I' of the β subunit by the P-cluster pair. Other examples of metal centers symmetrically bridging either identical or homologous subunits include the nitrogenase Fe-protein[11] and bacterial photosynthetic reaction centers[23,24]. The α and β subunits are approximately related by a two-fold rotation axis (refined rotational angle = 185°) that passes through the P-cluster pair. Additional inter-subunit contacts occur between domain I of α subunit and domains I' and III' of the β subunit, along with the two-fold related set of contacts between domain I' of the β subunit and domains I and III of the α subunit. An antiparallel β sheet arrangement between β61-β65 and α112-α117 also contributes to the αβ subunit interface. The N-terminus of the β subunit (about 50 residues), which is absent in the *C. pasteurinum* MoFe-protein[25] and alternative nitrogenase MoFe-proteins, extends from the β subunit and wraps around the α subunit. In addition to these general contacts, salt bridges at the interface of the α and β subunits are formed between the side chains of Asp α116-Lys β68, Lys α23-Asp β133, Lys α51-Asp β121, Lys α76-Glu β32, Arg α210-Glu β33, Glu α261-Lys β27 and Lys α433-Glu β109. Six buried, hydrophobic interface regions occur between the α and β subunits,

containing regions: (1) Tyr α20, Tyr α407, Leu α56, Met α57, Val α114, Tyr β142, Met β271 and Met β86; (2) Tyr α64, Tyr β98, Tyr α91, Phe β99, Tyr β102 and Leu β73; (3) Leu α141, Phe α142, Tyr β52, Leu β55, Trp β428, Pro α143, Leu α144, Tyr β35, Val β43, Ile β40, Trp β46, Met α112, Phe α114, Leu β62 and Val β64; (4) Phe β189, Phe α118, Val β157, Ile α123, Met β154, Phe α186, Ile α159 and Leu α158; (5) Trp α236, Phe β15, Leu β24, Met β23, Ile α101, Ile α240 and Tyr β20; and (6) Phe α429, Ile α430, Pro β110, Phe β269, Met β271 and Met α434. Hydrogen bonding pairs between sidechain groups (Arg α93-Asn β65, Glu α334-Ser β2, Ser α52-Gln β93 and Asp α454-Tyr β20) have also been identified at the αβ subunit interface.

4-3-3. Structure of $\alpha_2\beta_2$ tetramer

The overall dimensions of the $\alpha_2\beta_2$ MoFe-protein tetramer are ~70Å x 80Å x 110Å (Figures 4-5a, b and c). The two αβ dimers are related by the NCS two-fold rotation axis that was used for phase refinement during the crystallographic analysis. Even though the α and β subunits in an αβ dimer are also approximately related by a two-fold rotation, the MoFe-protein does not exhibit 222 symmetry as had been earlier proposed[26] in a low resolution rotation function study. At the positions of closest approach, the αβ two-fold and the tetramer two-fold axes are separated by ~12Å; furthermore the two axes are not exactly perpendicular as their orientations differ by ~97°. The general shape of MoFe-protein may be described as either a parallepiped along the direction of the tetramer two-fold axis (Figure 4-5a) or as a butterfly in a perpendicular direction (Figure 4-5b). 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 (B234-B247, B342-B362, B323-B336, B488-B510 and the corresponding regions of subunit D) dominate the interactions at the tetramer interface (Figure 4-6), with some contributions from helices in the α subunit (A302-A315, A318-A345, A453-A467, and the corresponding regions of subunit C).

These helical interactions seem to provide a major driving force for tetramerization. Intriguingly, the center of the six α-helical barrel surrounding the tetramer two-fold axis is not filled with sidechains; rather, an open channel of ~8-10Å diameter and length ~35Å exists that may serve as a model for helical pores in membrane-spanning proteins.

In addition to these helical packing interactions, salt bridge, hydrophobic and hydrogen bonding interactions are also important for tetramerization. Salt bridges at the interface of the two αβ dimers are formed between the terminal oxygen of Arg D523 and the side chains of Arg B108 and His B477; the side chains of the Arg B476-Asp D506-Lys B449 trio; Lys A433-Asp D353; Lys A474-Asp D323; and Glu B259-Lys D346. Hydrophobic regions at this interface are formed by residues (1) Met B330, Leu B329, Ile B340, Trp C479, Tyr b487 and Leu B344, and (2) Tyr A99, Tye D517, Trp A236, Phe B15, Tyr B12, Leu B16 and Ile A101. The side chains of Gln A432-Asp D357, Trp C479-Thr B345, Thr C466-His B359, His B477-Ser D358, Gln D37-Gln B513-Lys D34, Asp B520-Tyr C99-Asn B518 form hydrogen bonds at the interface. In addition to these interactions, the tetramer interface may be stabilized by cation binding to a site created by ligands from both β subunits and one α subunit (Figure 4-7). The atomic identity of this site, which was initially identified as the strongest feature in difference maps, has not yet been unambiguously established, but based on the electron density value, temperature factor and the coordination environment appears to be either Ca^{2+} or Mg^{2+} (less likely). This ion has an octahedral coordination environment provided by the carboxyl oxygens of Glu B109, Asp D353 and Asp D357, the carbonyl oxygen of Arg B108 and probably two water molecules. Gln A432, Lys A433, His B478 and His B480 are located outside the immediate coordination sphere. The site is buried at the interface of the two αβ dimers, and is about 25Å and 21Å 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 association in the tetramer. Some spherical plant viruses also have divalent cations (Ca^{2+}) at subunit interfaces that function in the stabilization of subunit contacts[27].

4-4. Unusual structural features

In addition to regular secondary structural elements, i.e., α-helix, β-strand and turns, MoFe-protein exhibits some interesting but unusual structural features, such as a π-helix, a left-handed α-helix (α_L), cis-peptides, and an interrupted α-helix. Although these structural features are unusual, their appearances are not statistically improbable, considering that the MoFe-protein is composed of more than one thousand unique residues. Interestingly, these unusual structural features are well conserved between the two MoFe-proteins (MoFe-protein from *A.vinelandii* (Av1) and *C. pasteurianum* (CP1)) suggesting that these features are structurally important.

A π-helix observed in Av1 structure is shown in Figure 4-8. The right-handed π-helix described by Lowe & Grenville-Wells[28] had not yet been found in protein structures[29,30]. The helix (α63-α74) starts from a α-helical conformation, but the last part of the helix shows π-helical conformation based on the hydrogen bonding pattern and the periodicity of the helix. The carbonyl oxygen of α64 shows a bifurcated hydrogen bond to both the amide nitrogen of α68 and α69. This suggests that the π-helix starts at residue α64. In residues α65-α73, the carbonyl oxygen of every (N)th residue is hydrogen bonded to the amide nitrogen of (N+5)th residue and the period of a turn is 4.4 residues. This is consistent with theoretical works on the π-helix[28]. The π-helix is disrupted after Gly α73 and ends in a α-helical conformation. Interestingly, the corresponding residues in the β subunit have a regular α-helical conformation, although the corresponding residues in the α subunit of Cp1 also have the π-helical conformation.

A short left-handed α-helix (α_L) observed in the Av1 structure is shown in Figure 4-9. The left-handed α-helices are energetically unfavourable[29,31], and only isolated short segments of this structure are found in proteins[32]. Residues α444-α446, near the Mo ligand His α442, form one turn of a left-handed α-helix. The amino acid sequence around the α_L-helix is His α442-Ser α443-Trp α444-Asp α445-Tyr α446-Ser α447-Gly α448.

This α_L-helix is also found in the Cp1 structure but only one of the corresponding residues (α444) in the β subunit has α_L-conformation.

Figures 4-10a and b show two cis-peptides observed in the two MoFe-proteins. One cis-peptide between Phe β471 and Pro β472 is found in the Av1 structure. This cis-peptide is also conserved in the Cp1 structure. In addition to this cis-peptide, Cp1 contains an additional cis-peptide between Phe β201 and Pro β202. Pro β202 is not a conserved residue between the two MoFe-proteins. Cis-peptides are frequently observed in Xxx-Pro or Pro-Xxx sequences[33], where Xxx represents any amino acid. About 25% of the Tyr-Pro sequences and 10% of the Phe-Pro sequences exhibit cis-peptide configuration in the survey of 154 protein structures[33].

An unusual interrupted helix in the Cp1 structure (α370-α430), in which the helical hydrogen bonds continue to a final turn that is formed after insertion of ~50 residues, is illustrated in Figure 4-11. These helical stretches are connected in Av1 forming a normal helix, but are interrupted in Cp1 due to the insertion of the residues, α375-α380. Such interrupted helices (broken on one side of the double helix) are apparently a fundamental feature of nucleic acid structure as illustrated by tRNA[34], but are rare in protein structure although they have been observed in subtilisin, thymidylate synthetase and bacterial reaction center structures[23,24,35].

4-5. Environment of metal centers

The FeMo-cofactor, first identified by Shah and Brill[36], almost certainly provides the substrate binding site[37]. A model has been recently proposed for the FeMo-cofactor which consists of two clusters, of composition 4Fe:3S and 1Mo:3Fe:3S, bridged by three non-protein ligands (Chapter 3). The Mo site is coordinated by homocitrate in a bidentate fashion, and the entire cofactor is coordinated to the α subunit through two residues, Cys α275 and His α442. Two of the bridging ligands are assigned as sulfurs, while the third ligand (designated "Y") has lower electron density and could be a well-ordered O/N species

or a less well-ordered S. The FeMo-cofactor and the surrounding residues are represented in Figure 4-12. The FeMo-cofactor is buried at least 10Å below the protein surface. Cys α275 and His α442, as well as Ser α278 which is hydrogen bonded to the Sγ of Cys α275, are strictly conserved in all known MoFe-protein sequences. Other highly conserved residues in the vicinity of the FeMo-cofactor include: Gly α356 and Gly α357, which are required to avoid steric interference with the cofactor; Arg α96 and Arg α359, that can potentially form hydrogen bonds to cluster sulfurs in the FeMo-cofactor, and may serve to stabilize the FeMo-cofactor and/or partially reduced intermediates formed during substrate reduction; His α195, that can potentially form hydrogen bonds to bridging sulfur in the FeMo-cofactor, and may function in proton transfer reactions; and three residues, Gln α191, Glu α440 and Glu α427, which are near the homocitrate and interact with this group either directly or through water molecules. The protein environment around the FeMo-cofactor is primarily provided by hydrophilic residues, although there are some hydrophobic residues such as Tyr α229, Ile α231, Val α70, Phe α381, Leu α358 and Ile α355. The function of these hydrophobic residues, if any, is not clear, but it is interesting that conserved residues (Tyr α229, Ile α231, Val α70, Phe α381) are located on the side of the FeMo-cofactor that faces toward the P-cluster pair site or Fe-protein binding sites (see below), while non-conserved residues, such as Leu α358 and Ile β355 , are positioned towards the protein surface. The positioning of the FeMo-cofactor near the N-terminal ends of helices α280-α290 and α359-α369 may serve to stabilize the cluster electrostatically[38].

The P-cluster pair, which has the structure of two bridged 4Fe:4S clusters, may function in electron transfer between the 4Fe:4S cluster of Fe-protein and the FeMo-cofactor. The P-cluster pair is located on the two-fold axis that approximately relates the α and β subunits. The P-cluster pair and surrounding residues are shown in Figure 4-13. The protein environment around the P-cluster pair is mainly provided by hydrophobic residues, including Tyr α64, Pro α85, Tyr α91, Pro α155, Phe α186, Pro β72, Phe β99,

Tyr β98, Met β154 and Phe β189. The location of the P-cluster pair near the N-terminal ends of six helices (α63-α74; α88-α92; α155-α159; β71-β81; β93-β106; and β153-β158) may serve to provide an electrostatic contribution to cluster stability. With the exception of the cluster ligands and Gln β93 and Thr β152, hydrophilic residues around the P-cluster pair, such as Glu α153, Glu α184, Ser α92, Ser α152 and Ser β92 are generally not conserved in different MoFe-protein sequences. The side chains of residues Cys α62, Cys α88, Cys α154, Cys β70, Cys β95, Cys β153 and Ser β188 are coordinated to the P-cluster pair, and are strictly conserved. Strictly conserved Gly residues (α87, β94 and α185) around the P-cluster pair are also structurally important to avoid steric interference with the P-cluster pair. Substitution of Gly α185 with Asp has been found in a Nif$^-$ mutant of nitrogenase[39].

4-6. Mechanistic Features

4-6-1. Substrate entry and Product release

No permanent channels between the protein surface and the FeMo-cofactor are present in the MoFe-protein structure that could represent the entry/exit pathway for dinitrogen and ammonia. However, there are two clefts (Figures 4-14a and b) which could be potentially utilized for substrate entry/product release and/or H_3O^+ transfer to the active site. The first cleft (Figure 4-14a) is made up of four stretches of polypeptide chain (α269-α288, α293-α315, α350-α368, and α378-α394) between domains II and III and is funnel-shaped with an outer diameter of ~4-10Å. The second cleft (Figure 4-14b) is made up of five stretches of polypeptide chain (α185-α203, α274-α285, α377-α389, α354-α362 and α45-α52) between the three domains of the α subunit and is also funnel-shaped with an outer diameter of ~3-10Å. These two clefts are near 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 clefts are wide enough to allow free diffusion of either substrate/product or H_3O^+. By analogy to the

binding and release of oxygen to the buried hemes in globins[40], it seems reasonable that structural fluctuations can create the transient cavities required for the diffusion of ligands into and away from the FeMo-cofactor, especially considering the rather slow turnover time (~5 s^{-1}) of nitrogenase[41-44]. Although covering a longer distance, it is also possible that a pathway for substrates and products could exist between the FeMo-cofactor and the protein surface lining the channel between the two β subunits.

4-6-2. Proton transfer

In addition to dinitrogen, ammonia formation also requires the delivery of protons to the active site. As no permanent channels exist by which H_3O^+ may diffuse from the protein surface to the FeMo-cofactor, it is possible that proton transfer could employ a "bucket-brigade" mechanism of the type envisioned for protonation of reduced quinones in the photosynthetic reaction center[45]. In this mechanism, protons are transferred by shuttling from one amino acid to another; the side chains of Asp, Glu and His residues, as well as water molecules, would seem to be particularly suited for this function. Two patches of histidines occur on the α subunit surface that might represent the initial site for proton binding (Figure 4-12); these patches contain the imidazole side chains of α196 and α383; and α274, α362 and α451. As several of these residues have been substituted either in naturally occurring MoFe-proteins or by site-directed mutagenesis without serious loss of activity[46], however, either these residues are not involved in proton transfer, or else a multiplicity of proton transfer pathways exist that can tolerate some alterations. In the interior of the protein, several potential pathways (Figure 4-15) occur which could funnel protons to the FeMo-cofactor: Glu α440, Glu α427 and/or Gln α191 to homocitrate, and His α195 to the 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 along Asp α234 to Arg α359 to the FeMo-cofactor, or by Arg α96 to the FeMo-cofactor.

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.

4-6-3. Electron transfer from the P-cluster pair to the FeMo-cofactor

Although no rigorous experimental demonstration has been provided, it appears plausible that the P-cluster pair functions as an intermediate electron carrier between the Fe-protein and the FeMo-cofactor. Consequently, the region between the two metal centers of the MoFe-protein could critically influence the mechanism and kinetics of intramolecular electron transfer. The protein environment between the FeMo-cofactor and the P-cluster pair is illustrated in Figure 4-16. The edge to edge distance from the FeMo-cofactor to the P-cluster pair is about ~14Å. Four helices (α63-α74; α88-α92; α191-α209; and β93-β106) are oriented in parallel between the two metal centers and may play an important role in electron transfer. In particular, the helices α63-α74 and α88-α92 adjacent to the P-cluster pair ligands Cys α62 and Cys α88 provide the most direct structural connection between the P-cluster pair and FeMo-cofactor. Potential electron transfer pathways between the P-cluster pair and FeMo-cofactor are (i) from Cys α88 along the helix to Arg α96; (ii) from Cys α88 via Gly α95 to homocitrate; (iii) from Cys α62 along the helix to Val α72; (iv) from Cys α62 via Ala α65 to homocitrate; and (v) from Cys β95 via Tyr β98 to homocitrate (Figure 4-17). No complete covalent or hydrogen bonded network exists along the most direct pathway between the two centers (although Arg α96 is hydrogen bonded to the FeMo-cofactor), suggesting that either some through-space jumps may occur during electron transfer[47], or that structural fluctuations/alterations occur during substrate reduction that permit the formation of more favorable electron transfer pathways. Additionally, potential proton transfer pathways, such as salt-bridge and hydrogen bonding networks, have not been found that could permit the coupling of electron and proton transfer between the P-cluster pair and FeMo-cofactor.

4-6-4. Interaction between MoFe-protein and Fe-protein

The Fe-protein contains a single 4Fe:4S cluster symmetrically coordinated to two identical subunits that transfers electrons one at a time to the MoFe-protein in a process coupled to the hydrolysis of two MgATP. Not only does Fe-protein represent the initial source of electrons for dinitrogen reduction within the nitrogenase system, but the details of the Fe-protein - MoFe-protein interaction contribute significantly to the overall kinetics of substrate reduction by nitrogenase. Each electron transfer between Fe-protein and MoFe-protein appears to involve a cycle of obligatory association and dissociation of the protein complex, with the dissociation step having been identified as rate determining for the overall reaction[42-44]. With the availability of the crystal structures for both proteins isolated from *A. vinelandii*, some general features of the complex between the two proteins may be addressed.

Two residues of Fe-protein that have been identified as interacting with the MoFe-protein, Arg 100[48] and Glu 112[49], are located on the same side of the protein as the 4Fe:4S cluster[11] (Figure 4-18). Hence, this surface almost certainly includes at least part of the interaction region between the two proteins. Relevant features of the interaction between Fe-protein and MoFe-protein that have been established biochemically include the ability of Glu 112 to crosslink with Lys β400[49], and the likely occurrence of salt bridges in the interface region[50], as established by salt effects on nitrogenase activity and the effects of replacing Arg 100 with other residues[51]. As the Fe-protein dimer has a two-fold symmetry axis[11], a plausible model for docking the two proteins involves superposition of the Fe-protein two-fold axis with the two-fold axis passing through the P-cluster pair that relates the α and β subunits of MoFe-protein (Figure 4-19). The surfaces of the two proteins are complementary in this region; the MoFe-protein surface near the P-cluster pair has a convex shape, while the Fe-protein surface is concave about the 4Fe:4S cluster. To either side of the P-cluster pair are two wide and shallow clefts related by the pseudo-twofold axis, that could accommodate the two Fe-protein subunits (Figure 4-4). With this

orientation, the side chains of Glu 112 and Lys β400 can be positioned sufficiently close to permit crosslinking.

The potential 4Fe:4S cluster binding site and the surrounding residues are shown in Figure 4-20. Four short helices (α155-α159; α120-α125; β153-β158; and β120-β125) which are related by a pseudo two-fold axis, are oriented in parallel from the P-cluster pair toward surface forming a four-helical bundle, and the 4Fe:4S cluster of Fe-protein could bind to the top surface of these helices. Kinetic studies of site-directed mutants at Phe β125 indicate that these alterations interfere with Fe-protein - MoFe-protein interaction[52]. Additionally, substitutions of Glu α120 and Gly α160 in or near these helices have been identified in Nif$^-$ mutants[39], possibly due to disruption of the Fe-protein - MoFe-protein interaction. The edge-edge distance from the P-cluster pair to the end of these helices is ~12Å, thus the edge-edge distance from the P-cluster pair to the 4Fe:4S cluster of Fe-protein may be about 15Å. These helices provide potential electron transfer pathways between the 4Fe:4S cluster and P-cluster pair. The protein environment around this electron transfer pathway is primarily occupied by hydrophobic residues. Phe α125, Phe β125, Phe α186, Phe β189, Leu α158, Val β157, Ile β158 and Ile β123 form the interface between the P-cluster pair and the proposed Fe-protein cluster binding site, and they are generally conserved in different MoFe-protein sequences. There are no channels or unbroken hydrogen bonded/salt bridge networks leading from this surface to the P-cluster pair site. Therefore, it seems likely that electron transfer from Fe-protein to the P-cluster pair is not coupled with proton transfer. Charged and polar residues on the MoFe-protein surface in this region may be important for the formation of salt bridges/hydrogen bonds with Fe-protein. Asp α162 and Asp β161, which are related by a pseudo twofold axis, are completely exposed to the solvent and are located about 15Å away from the P-cluster pair. Other solvent exposed side chains near the Fe-protein binding site include His β185, Asp α128, Asp α161 and Glu β156.

Kinetic studies indicate that nucleotide hydrolysis precedes electron transfer in the Fe-protein - MoFe-protein complex[42,43,53]. The nucleotide binding sites in the Fe-protein dimer are located at the interface between the two subunits[11]. Although the detailed structural consequences of nucleotide binding on the Fe-protein structure have not been established, it is likely that MgATP hydrolysis is accompanied by a change in Fe-protein structure, such as an alteration in the relative orientations of the two subunits. Presumably, this transition leads to the formation of an activated species that is competent for electron transfer from the Fe-protein to MoFe-protein (in particular, to the P-cluster pair in this model). As the nucleotide binding site is on the opposite surface of the Fe-protein to the interaction regions with the MoFe-protein, it should be possible for MgATP to exchange with MgADP without dissociation of the two proteins, as has been proposed from kinetic analyses of the nitrogenase reaction[53]. Kinetic studies also indicate that reduction of oxidized Fe-protein requires dissociation of complex. Since the 4Fe:4S cluster of Fe-protein appears to be buried in the interface region, reduction of oxidized Fe-protein would seem to require dissociation of the two proteins. Experimental studies have also indicated that oxidized MoFe-protein can bind MgADP[54], although it is not clear from the present studies where this site would be located, and what the relationship of this site might be to the MoFe-protein metal centers and the proposed Fe-protein binding site.

References

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

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

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

4. Steifel, E. I., *et al.* (1988) in *Metal Clusters in Proteins* (Que, L., eds.) pp. 372-389. American Chemical Society, Washington, D.C..

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

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

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

8. 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, 1988.

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

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

11. Georgiadis, M. M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. J. & Rees, D. C. (1992) *Science* **259**, 1653-1659.

12. Brigle, K. E., Newton, W. E. & Dean, D. R.(1985) *Gene* **37**, 37-44.

13. Harker, D. (1956) *Acta Cryst.* **9**, 1-9.

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

15. Bricogne, G. (1976) *Acta Cryst.* **A32**, 832-847.

16. Jones, T. A. (1985) *Meth. Enzym.* **115**, 151-171.

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

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

19. Bowie, J. U., Lüthy, R. & Eisenberg, D. (1991) *Science* **253**, 164-170; Lüthy, R., Bowie, J. U. & Eisenberg, D. (1992) *Nature* **356**, 83-85.

20. Lammers, P. J. & Haselkorn, R. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4723-4727.

21. Brigle, K. E., Weiss, M. C., Newton, W. E. & Dean, D. R. (1987) *J. Bact.* **169**, 1547-1553.

22. Robbins, A. H. & Stout, C. D. (1989) *Proteins: Struc. Func. Gen.* **5**, 289-312.

23. Deisenhofer, J., Epp, O., Mike, K., Huber, R. & Michel, H. (1985) *Nature* **318**, 618-624.

24. Allen, J. P., Fehr, G., Yeates, T. O., Komiya, H. & Rees, D. C. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5730-5734.

25. Wang, S. -Z., Chen, J. -S. & Johnson, J. L. (9188) *Biochem.* **27**, 2800-2810.

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

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

28. Low, B. W. & Grenville-Wells, H. J. (1953) *Proc. Natl. Acad. Sci. USA* **39**, 785-801.

29. Schulz, G. E. & Schirmer, R. H. (1979) *Principles of Protein Structure*. Springer Verlag, New York.

30. Barlow, D. J. & Thorneley, J. M. (1988) *J. Mol. Biol.* **201**, 601-619.

31. Chou, K. C., Nemethy, G. & Scheraga, H. A. (1990) *Accts. Chem. Res.* **23**, 134-141.

32. Kabsch, W. & Sander, C. (1983) *Biopolymers* **22**, 2577-2637.

33. Stewart, D. E., Sarker, A. & Wampler, J. E. (1990) *J. Mol. Biol.* **214**, 253-260.

34. Holbrook, S. R., Sussmann, J. L., Warrant, R. W. & Kim, S. -H. (1978) *J. Mol. Biol.* **123**, 631-660.

35. Wright, C. S., Alden, R. A. & Kraut, J. (1969) *Nature* **221**, 235-242.

36. Shah, V. K. & Brill, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3249-3253.

37. Hawkes, T. R., McLean, P. A. & Smith, B. E. (1984) *Biochem. J.* **217**, 317-321.

38. Adman, E. T., Watenpaugh, K. D. & Jensen, L. H. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4854-4858.

39. Govenzensky, D. & Zamir, A. (1989) *J. Bact.* **171**, 5729-5735.

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

41. Thorneley, R. N. F. & Lowe, D. J. (1987) *Biochem. J.* **224**, 887-894.

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

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

44. Hageman, R. V. & Burris, R. H. (1978) *Biochem.* **17**, 4117-4124.

45. Fehr, G., Allen, J. P., Okamura, M. Y. & Rees, D. C. (1989) *Nature* **339**, 111-116.

46. Thomann, H., Bernardo, M., Newton, W. E. & Dean, D. R. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 6620-6623.

47. Wuttke, D. S., Bjerrum, M. J., Winkler, J. R. & Gray, H. B. (1992) *Science* **256**, 1007-1009.

48. Murrel, S. A., Lowery, R. G. & Ludden, P. W. (1988) *Biochem. J.* **251**, 609-612.

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

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

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

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

53. Thorneley, R. N. F. (1990) in *Nitrogen Fixation: Achievements and Objectives* (Gresshoff, P. M., Roth, L. E., Stacey, G. & Newton, W. E., eds.) pp. 103-109. Chapman & Hall, New York.

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

55. Kraulis, P. J. (1991) *J. Appl. Cryst.* **24**, 946-950.

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

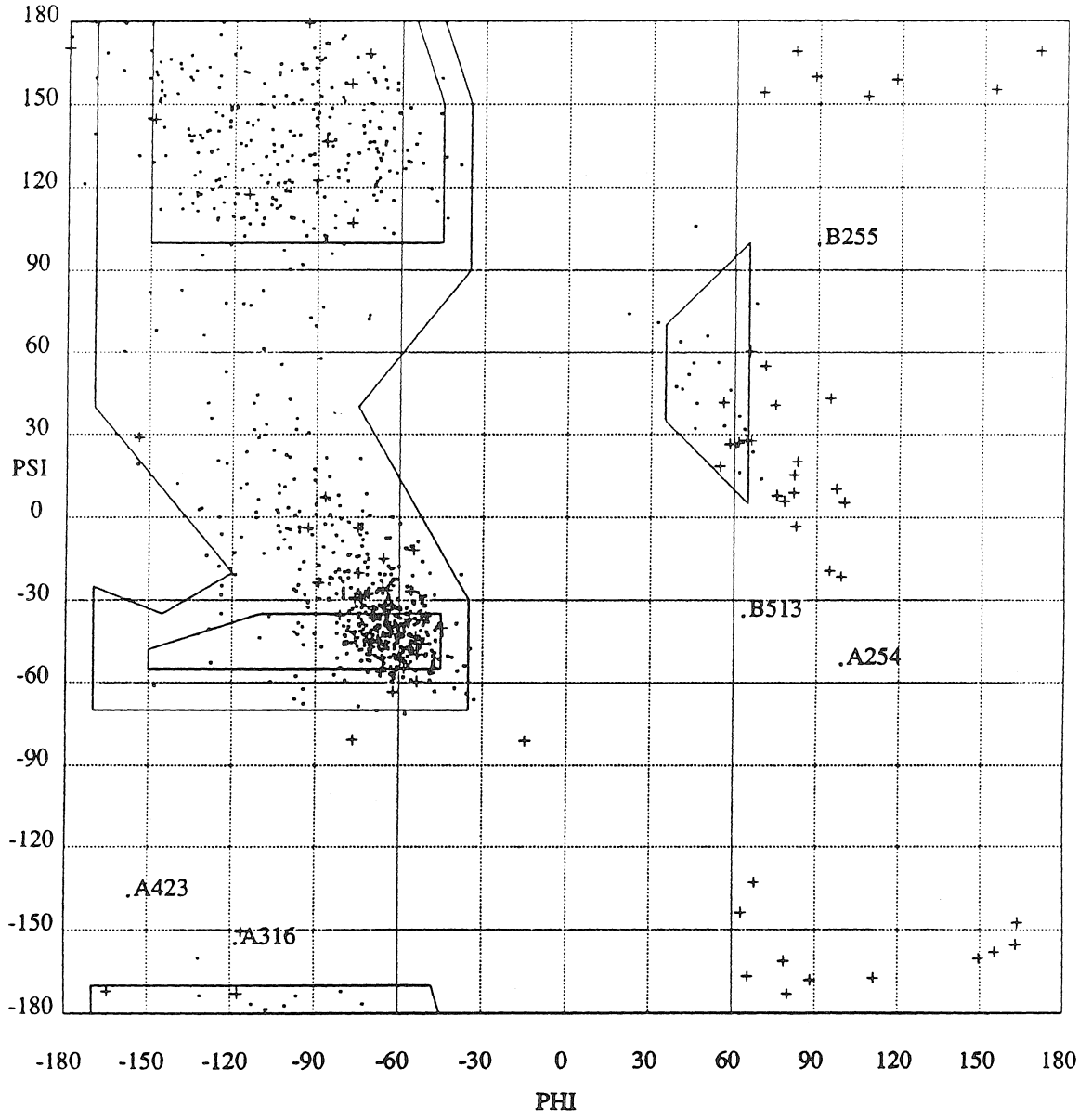


Figure 4-1. Ramachandran plot of the αβ subunit pair in MoFe-protein. Glycine residues are represented as crosses.

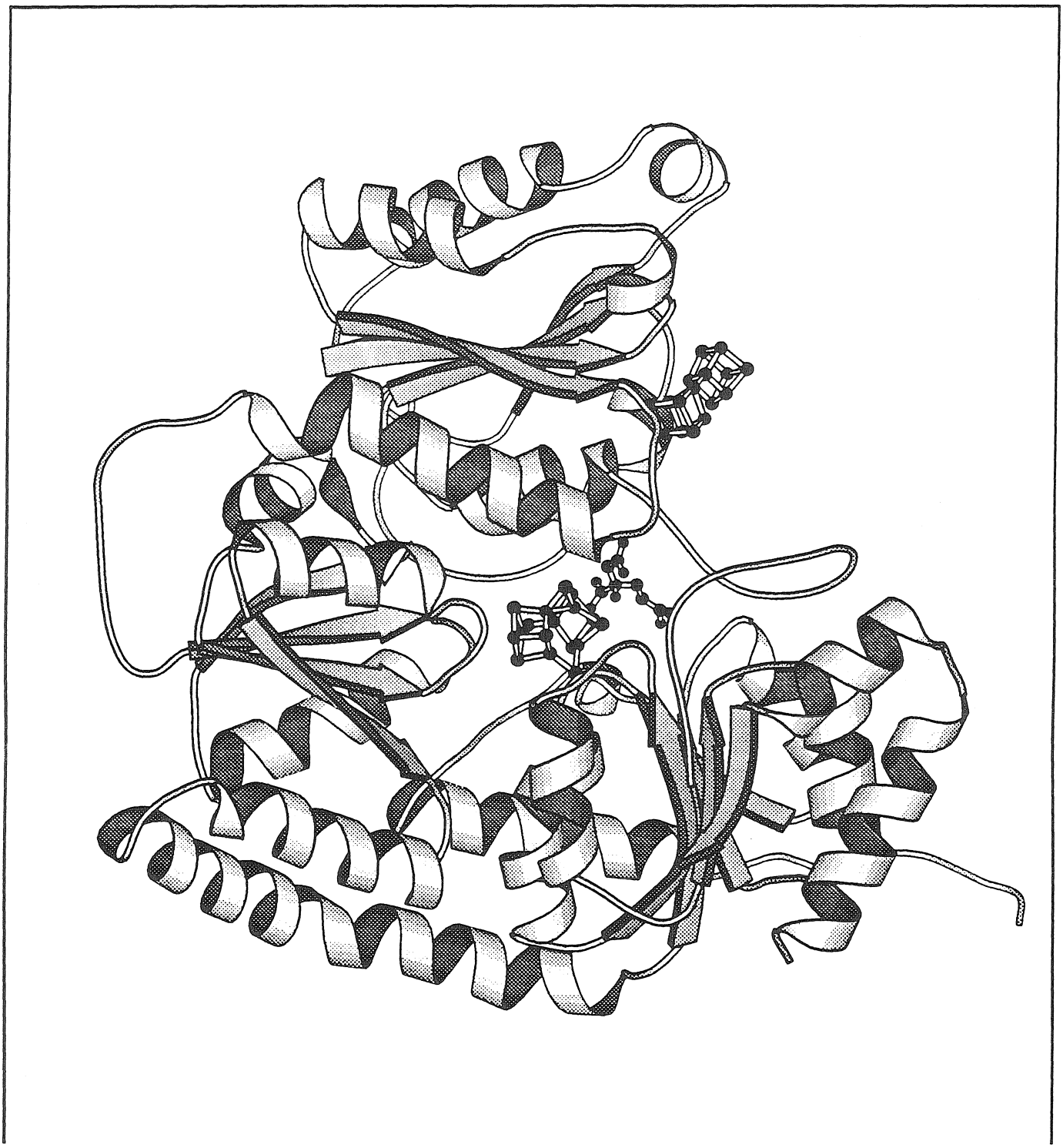


Figure 4-2a. Ribbons[55] diagram of the polypeptide fold of the α subunit. The FeMo-cofactor is located in the center of the figure and the P-cluster pair is on the top right.

Figure 4-2b. Stereoview of the Cα chain trace of the α subunit.

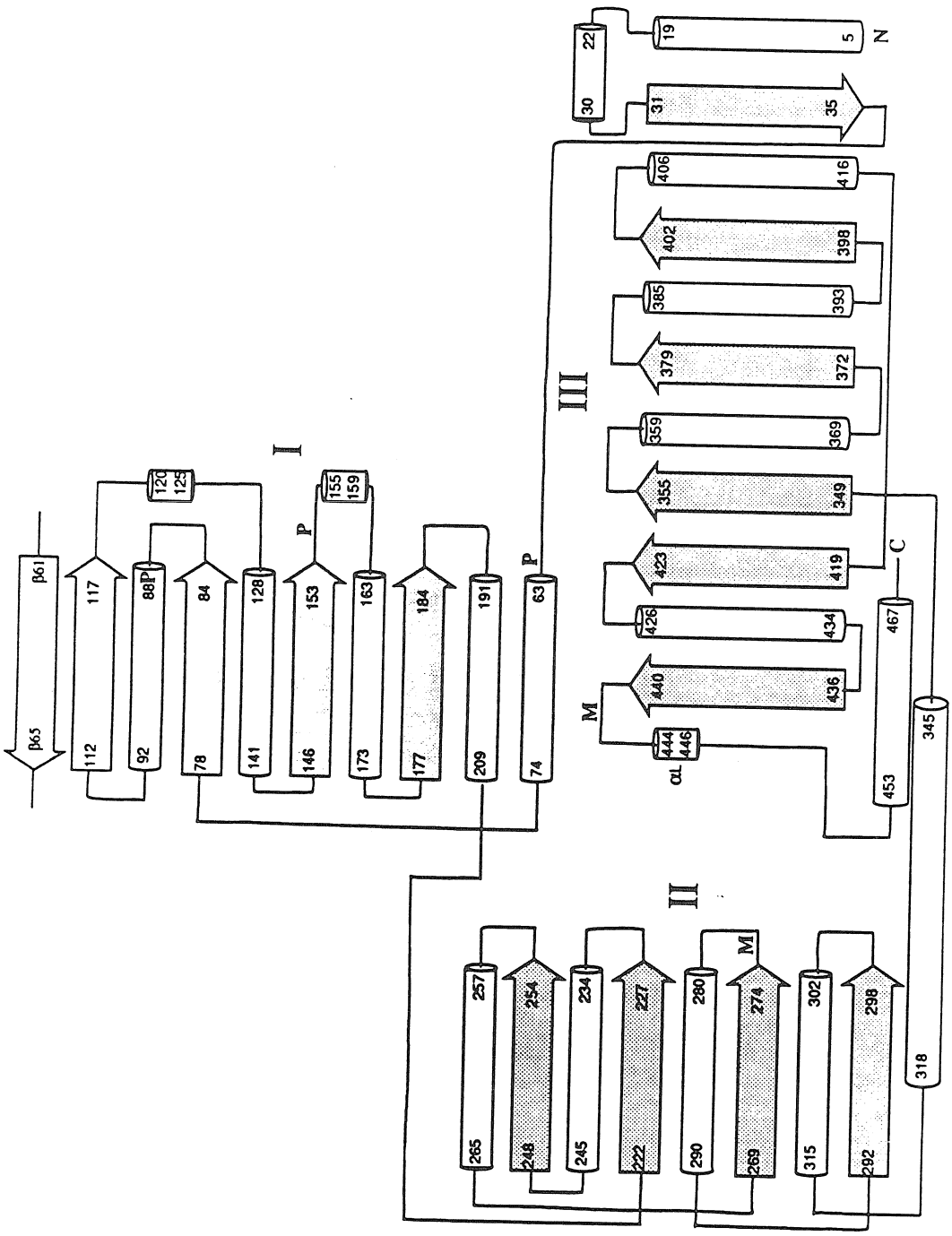


Figure 4-2c. Schematic representation of the α subunit secondary structure. The α subunit consists of three α/β type domains (designated I, II and III). Residues α444-α446 form one turn of a left-handed α-helix (α_L). The positions of FeMo-cofactor and P-cluster pair ligands are designated M and P, respectively.

Figure 4-3a. Ribbons diagram of the polypeptide fold of the β subunit. The left-most metal center is the P-cluster pair, whreas the FeMo-cofactor is to the right of the P-cluster pair.

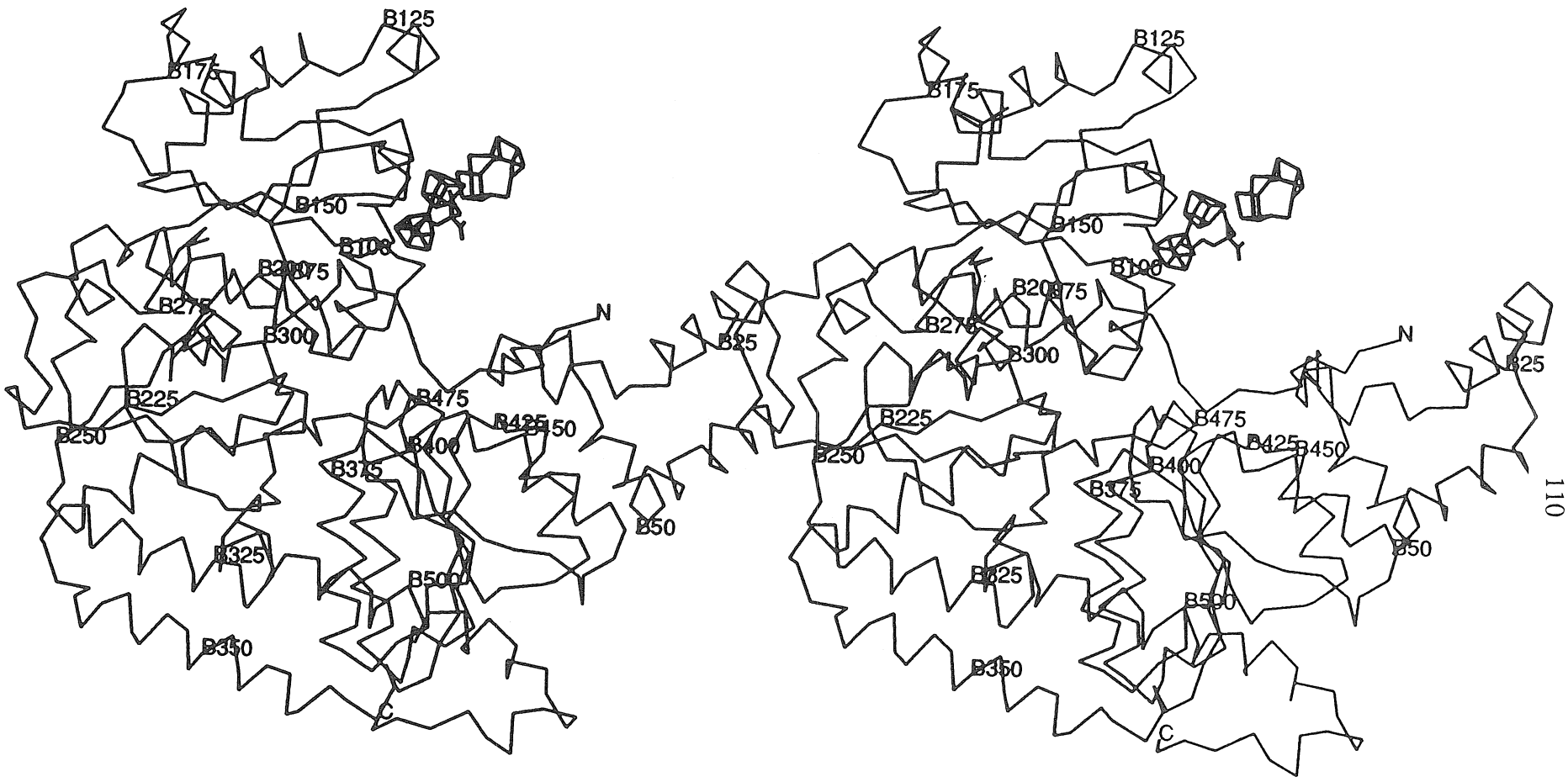


Figure 4-3b. Stereo view of the Cα chain trace of the β subunit.

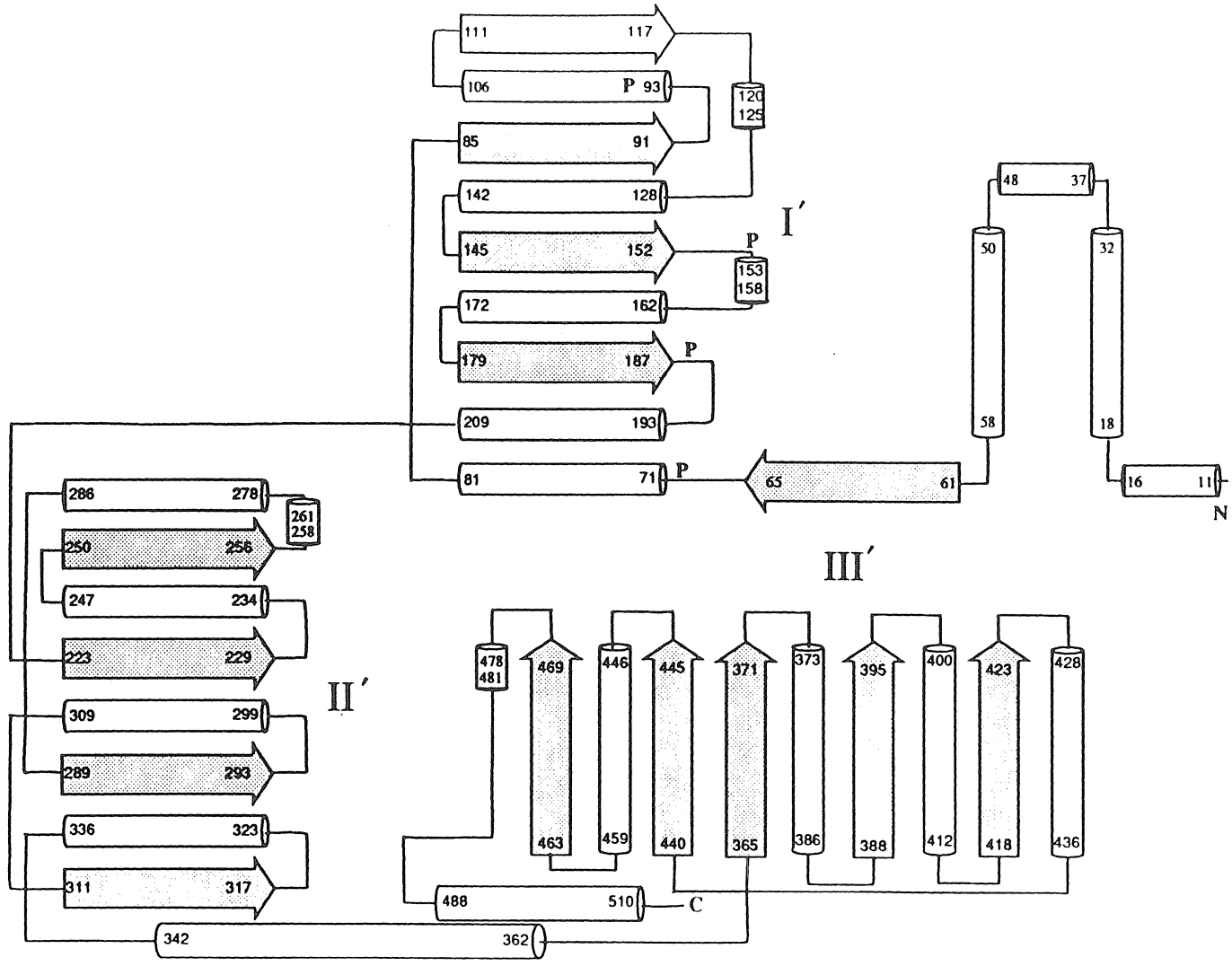


Figure 4-3c. Schematic representation of the β subunit secondary structure. The β subunit consists of three α/β type domains (designated I', II' and III'). The positions of the P-cluster pair ligands are designated P.

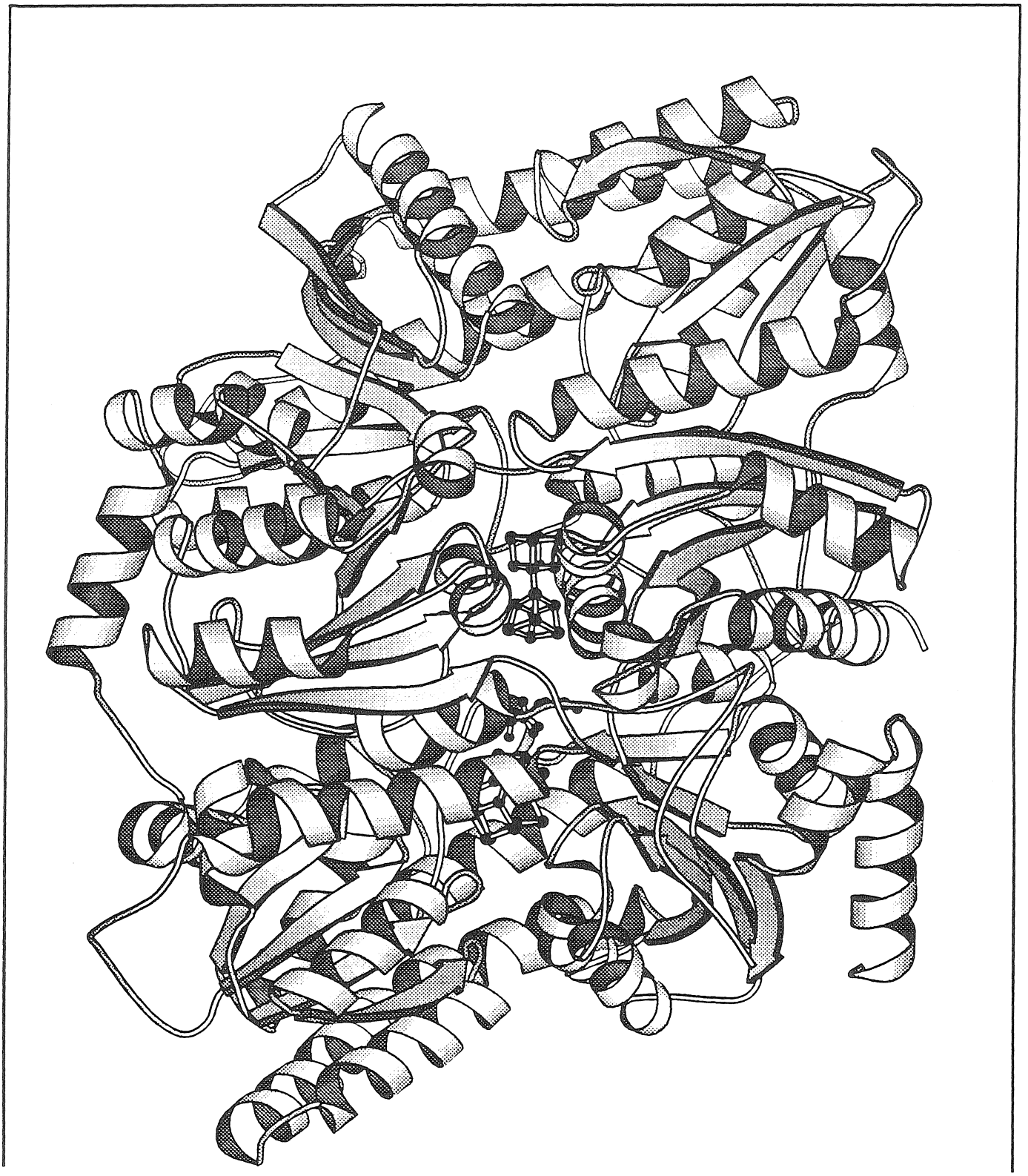


Figure 4-4a. Ribbons diagram of the polypeptide fold of an αβ subunit pair. The view is down the 2-fold axis through the P-cluster pair that roughly relates the α and β subunits.

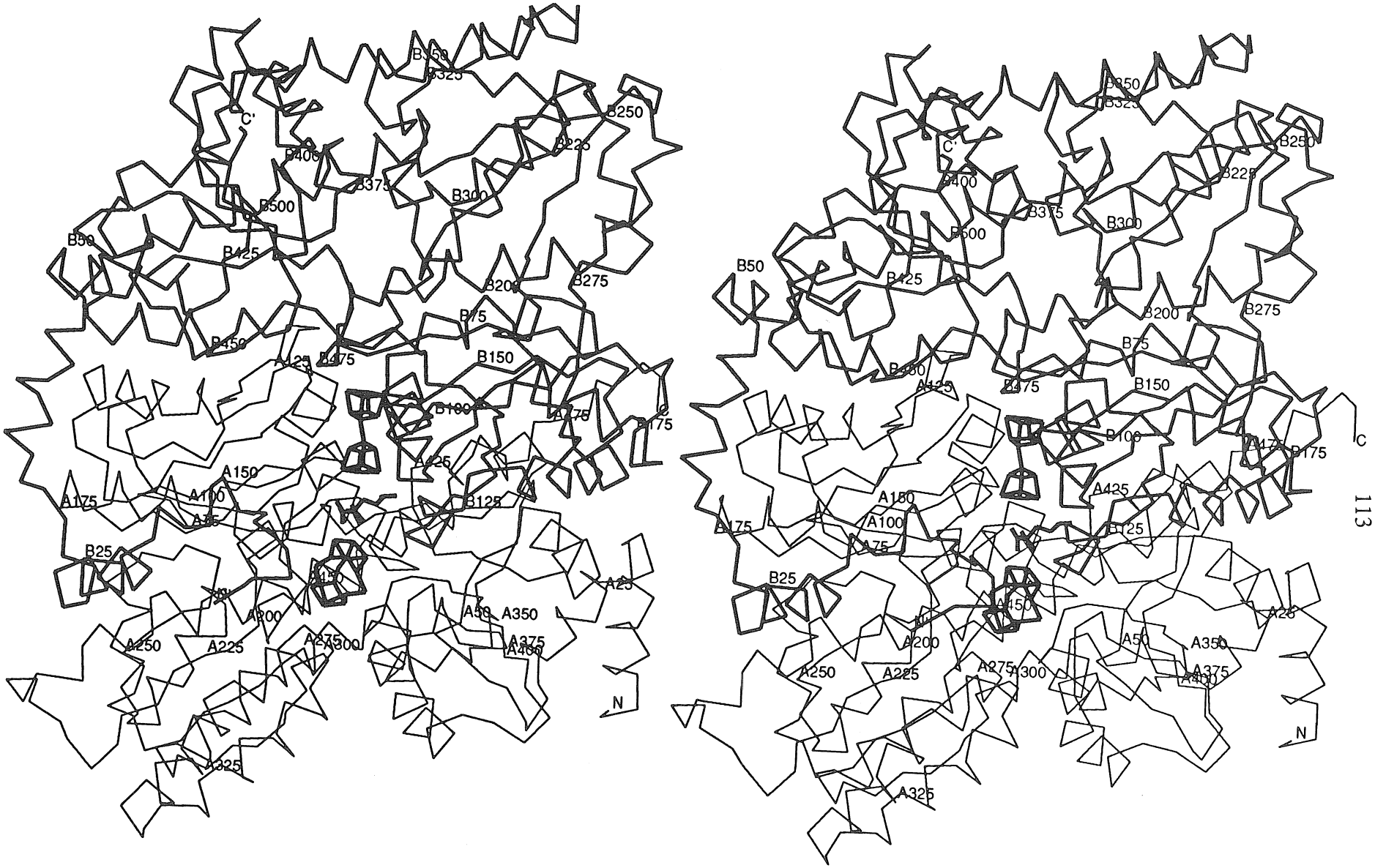


Figure 4-4b. Stereoview of the Cα chain trace of an αβ subunit pair. The view is down the 2-fold axis through the P-cluster pair that roughly relates the α and β subunits. The α and β subunits are represented by the thin and thick lines, respectively.

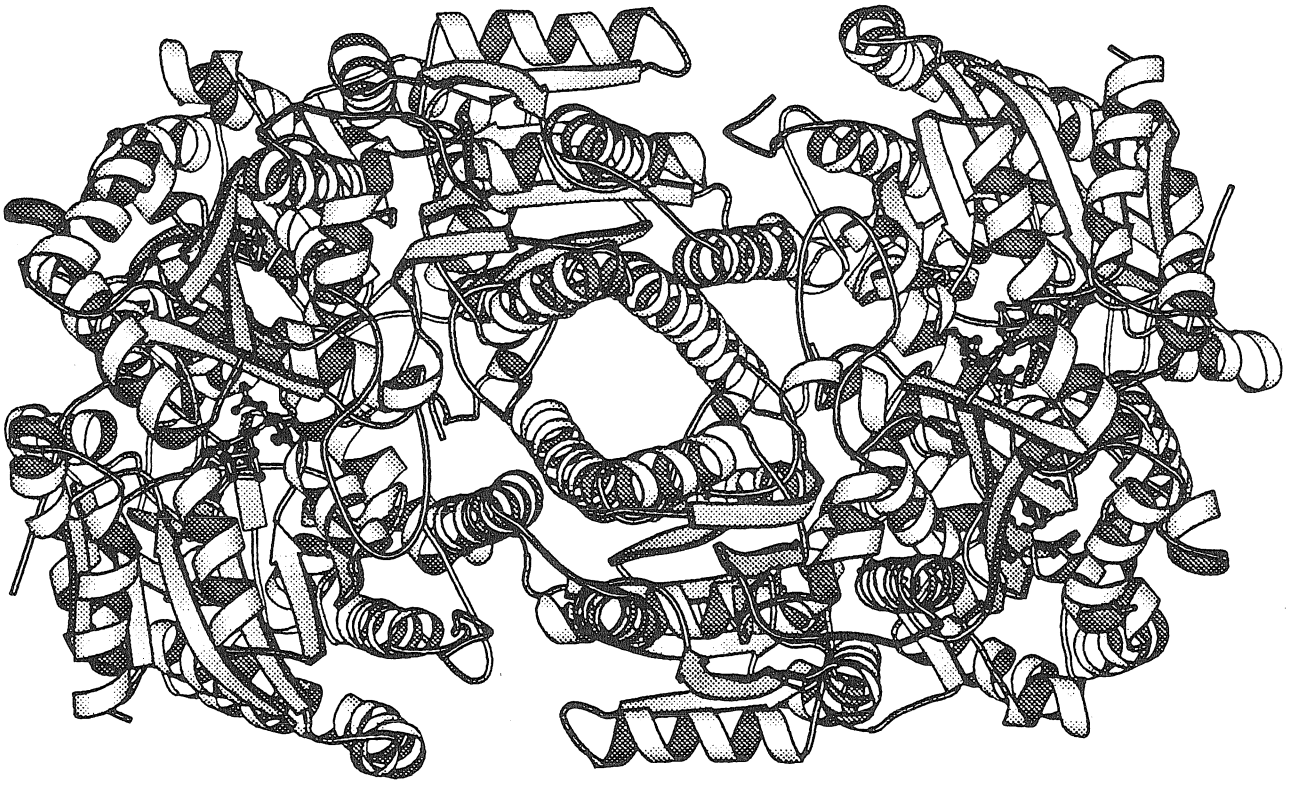


Figure 4-5a. Ribbons diagram of the polypeptide fold of the $\alpha_2\beta_2$ MoFe-protein tetramer I. The view is down the tetramer 2-fold axis.

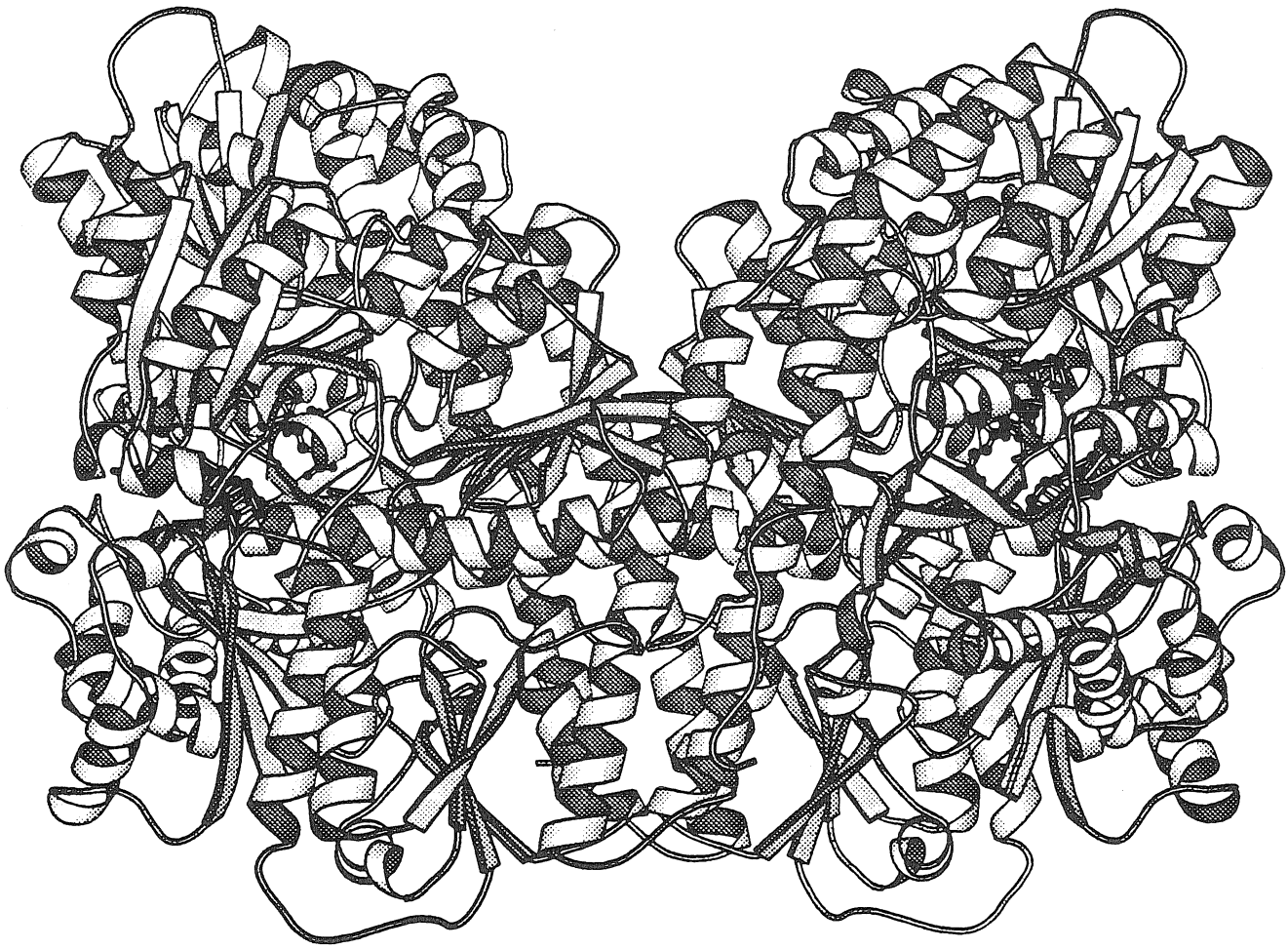


Figure 4-5b. Ribbons diagram of the polypeptide fold of the $\alpha_2\beta_2$ MoFe-protein tetramer II. The view is perpendicular to the tetramer 2-fold axis.

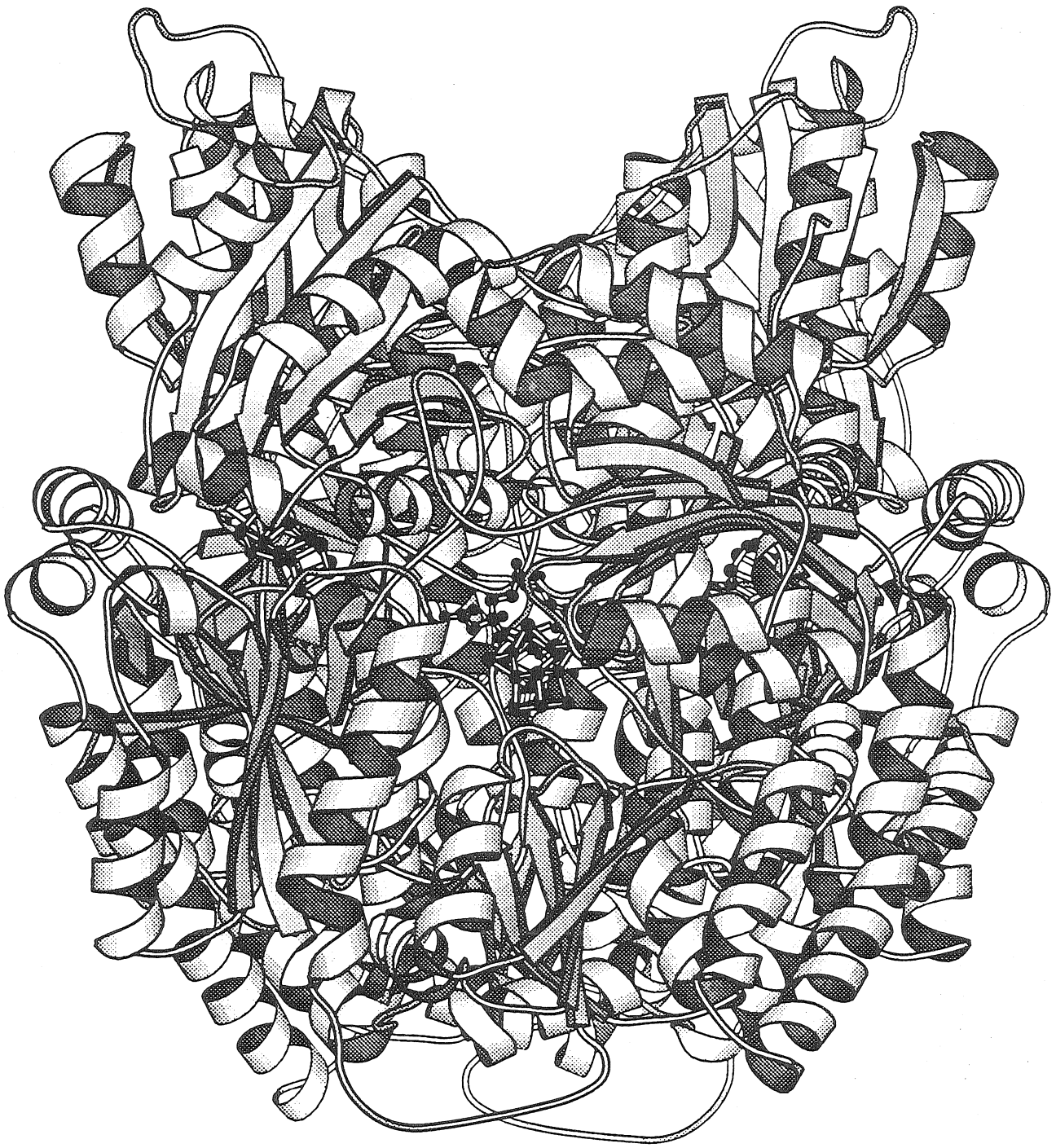


Figure 4-5c. Ribbons diagram of the polypeptide fold of the $\alpha_2\beta_2$ MoFe-protein tetramer III. The view is perpendicular to the tetramer 2-fold axis. The views of the three figures (Figure 4-5a, b and c) are perpendicular each other.

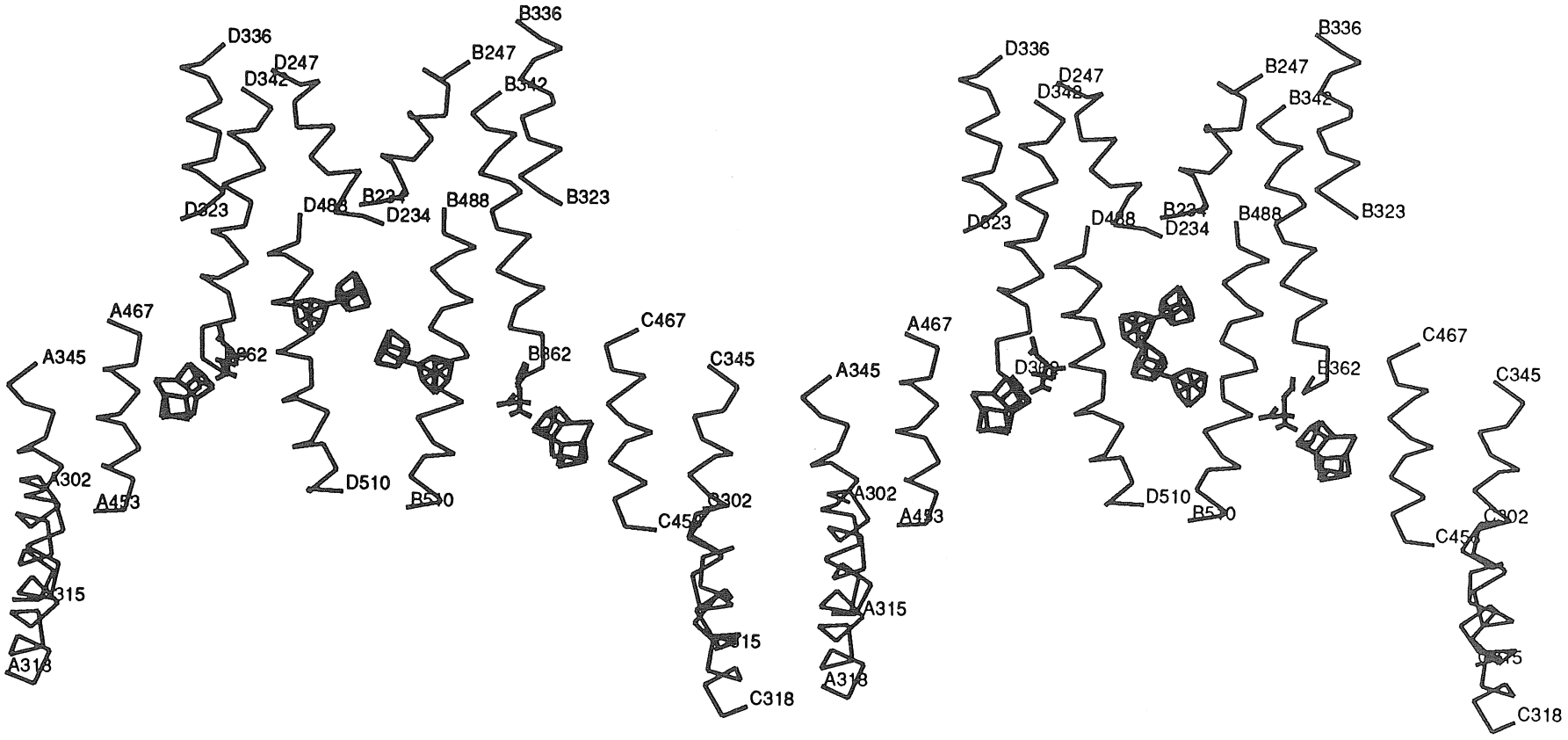


Figure 4-6. Stereoview of the Cα chain trace of α-helices stabilizing the tetramer interface. The view is perpendicular to the tetramer 2-fold axis. The positions of the metal centers are provided for reference.

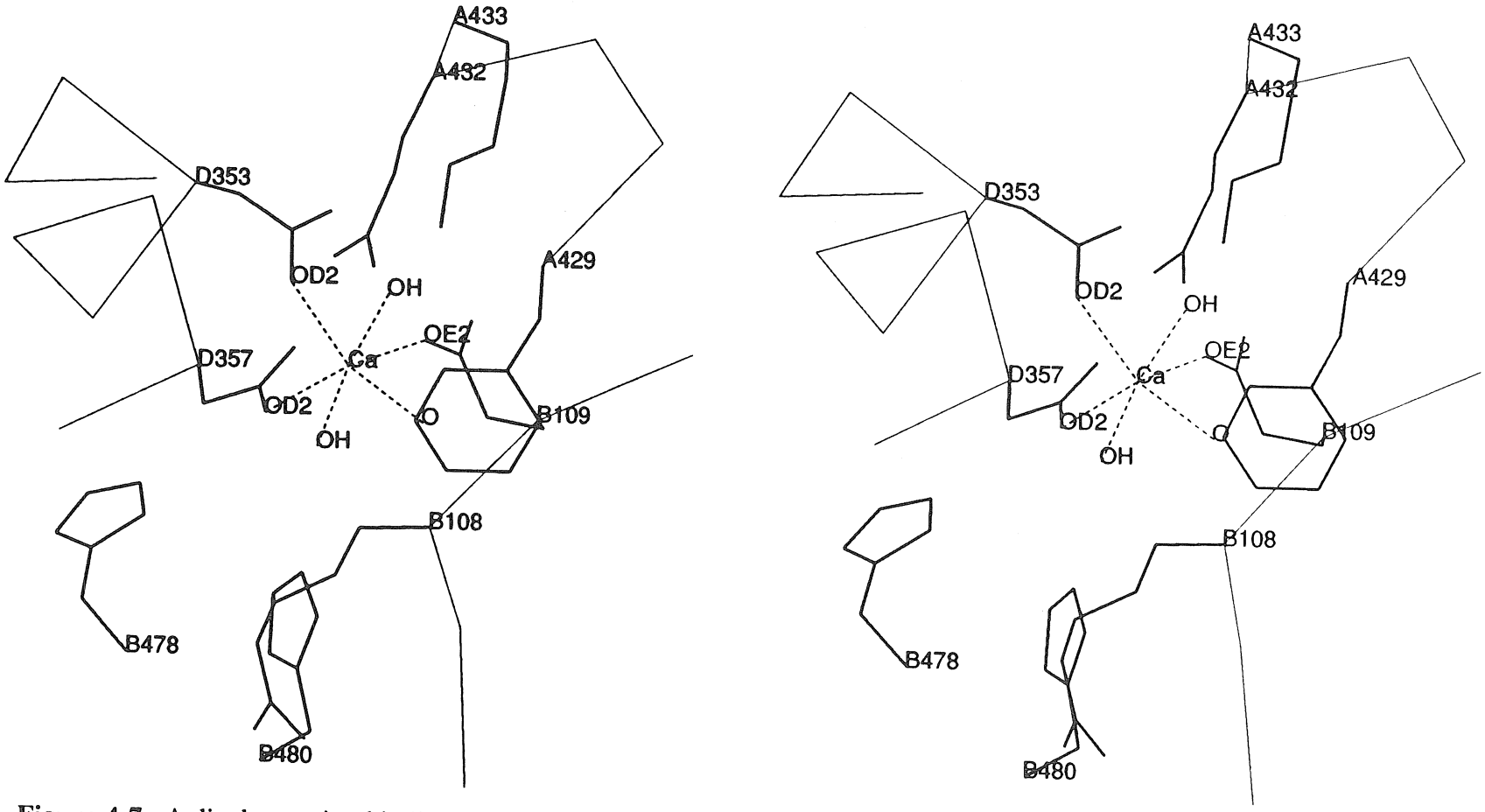


Figure 4-7. A divalent cation binding site and its surrounding residues. This ion has an octahedral coordination environment provided by the carboxyl oxygen of Glu B109, Asp D353 and Asp D357, the carbonyl oxygens of Arg B108 and probably two water molecules.

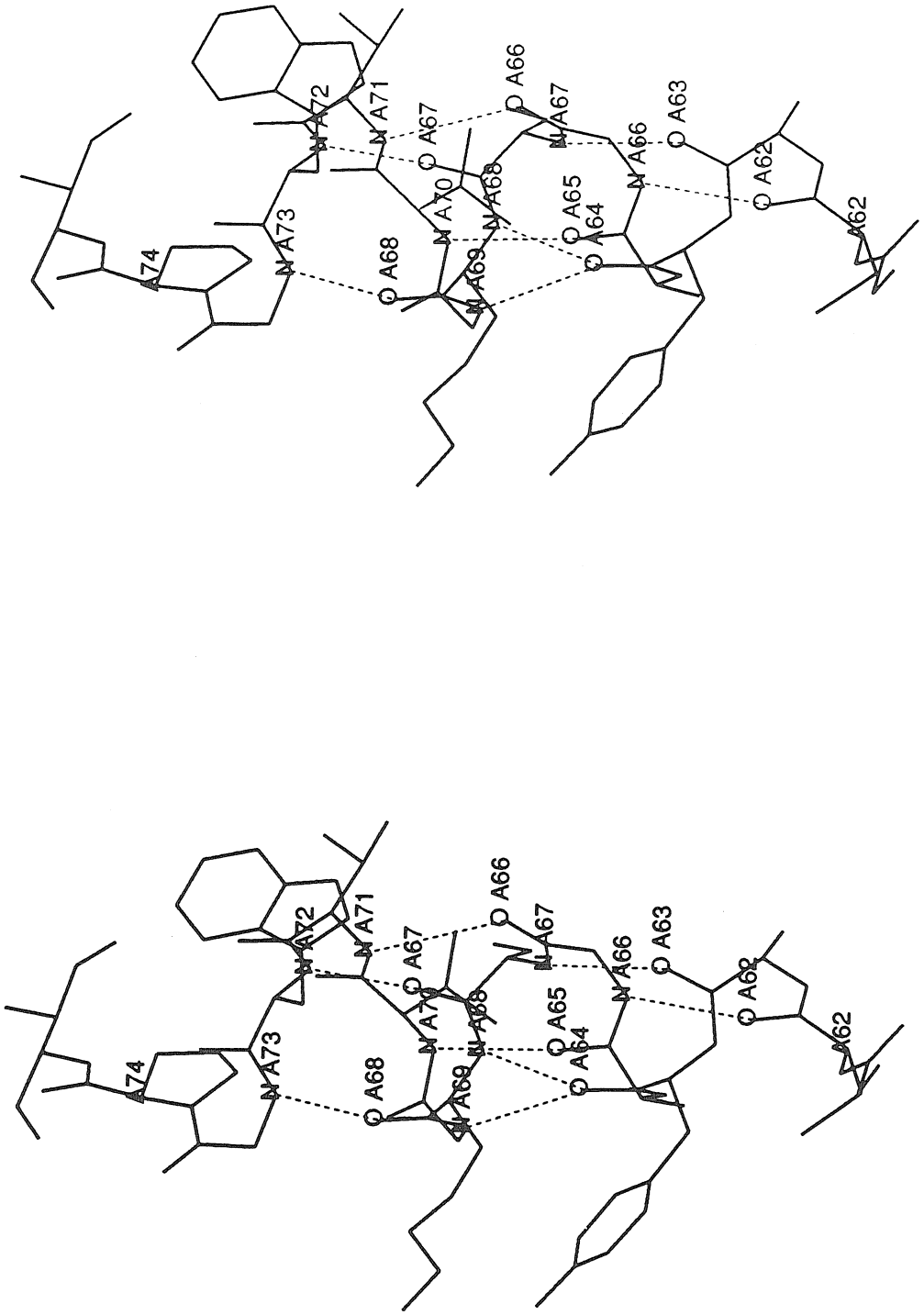


Figure 4-8. Stereoview of a π-helix observed in the Av1 structure. Residues $\alpha 64$-$\alpha 73$ form two turns of a π-helix. The π-helix is conserved in the Cp1 structure.

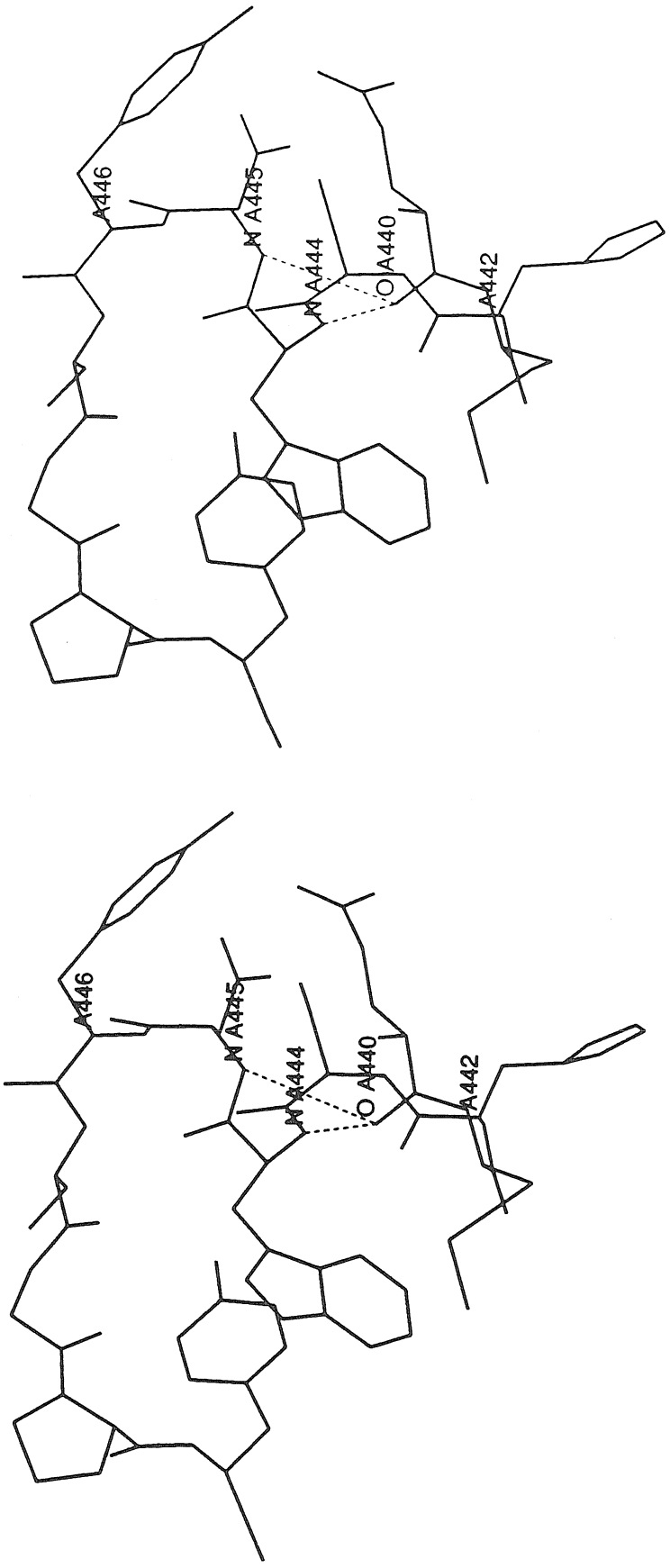


Figure 4-9. Stereoview of a short left-handed α-helix observed in the Av1 structure. Residues α444-α446 form one turn of a left-handed α-helix. This α_L-helix is conserved in the Cp1 structure.

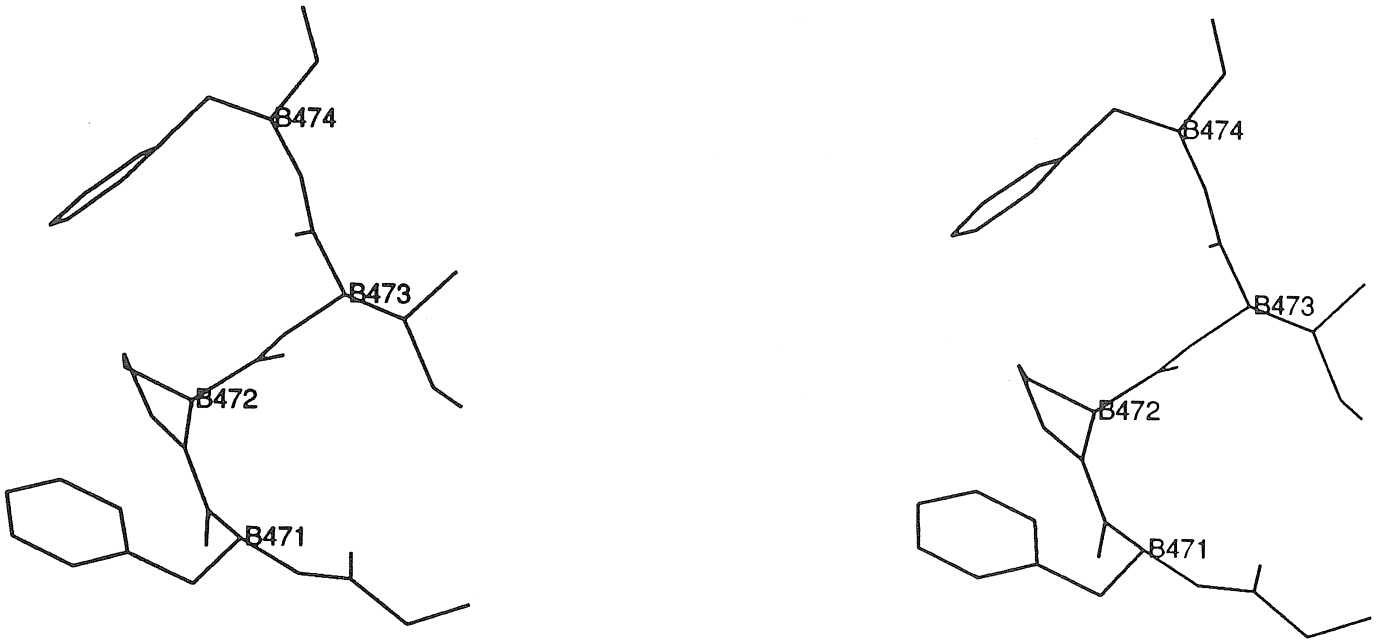


Figure 4-10a. Stereoview of a cis-peptide observed in the Av1 structure. This cis-peptide is also conserved in the Cp1 structure.

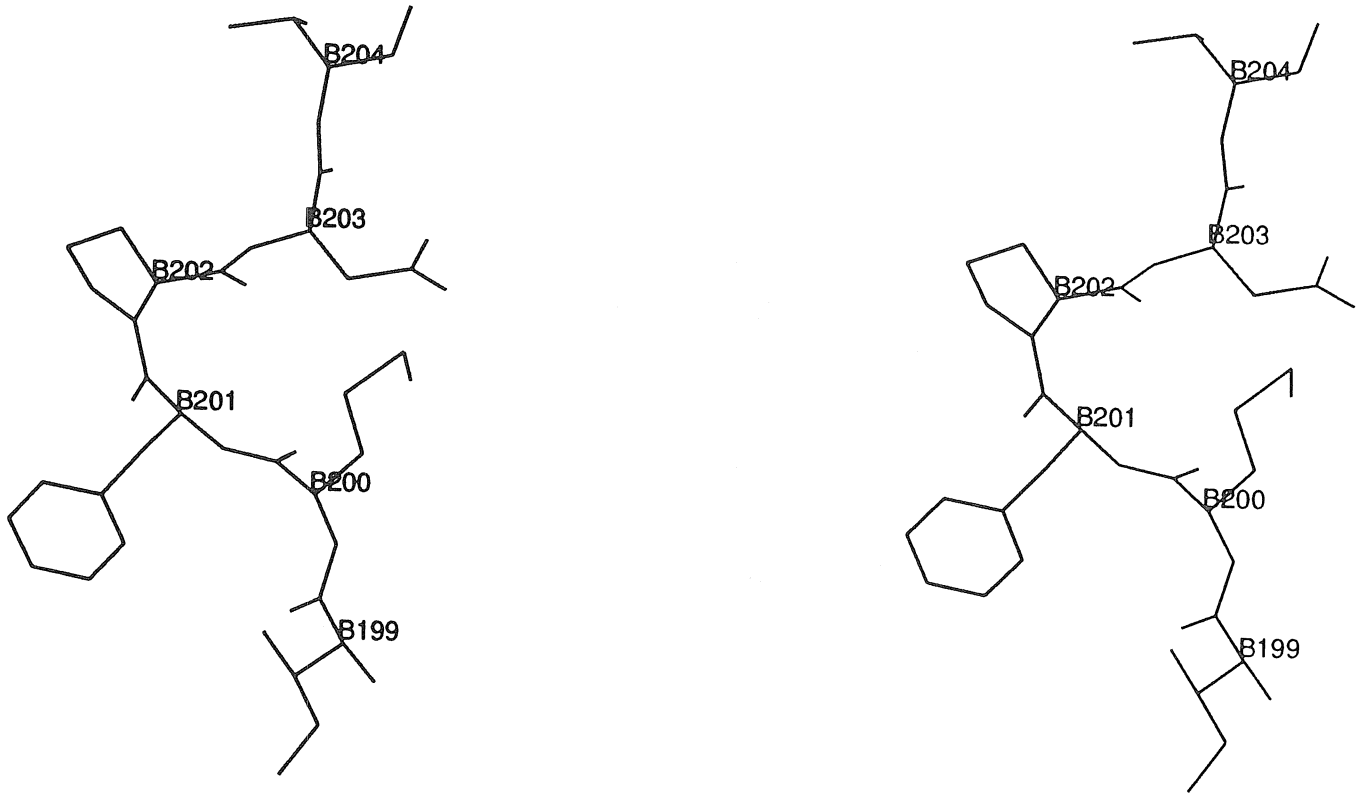


Figure 4-10b. Stereoview of a cis-peptide observed in the Cp1 structure. Pro β202 is not a conserved residue between the two MoFe-proteins.

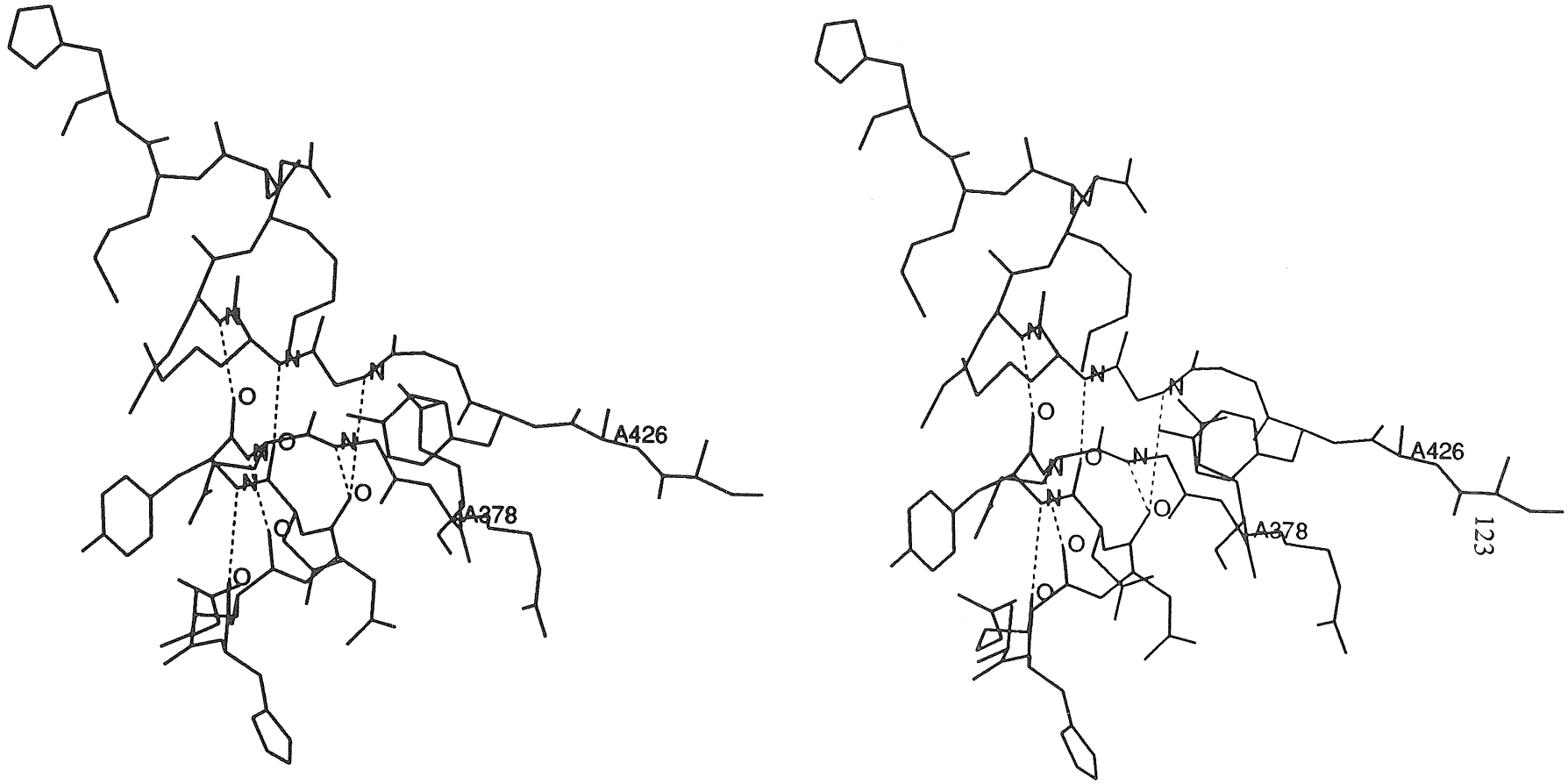


Figure 4-11. Stereoview of an interrupted helix observed in the Cp1 structure. These helical stretches are connected in Av1 forming a normal helix, but are interrupted in Cp1 due to the insertion of the residues α375-α430.

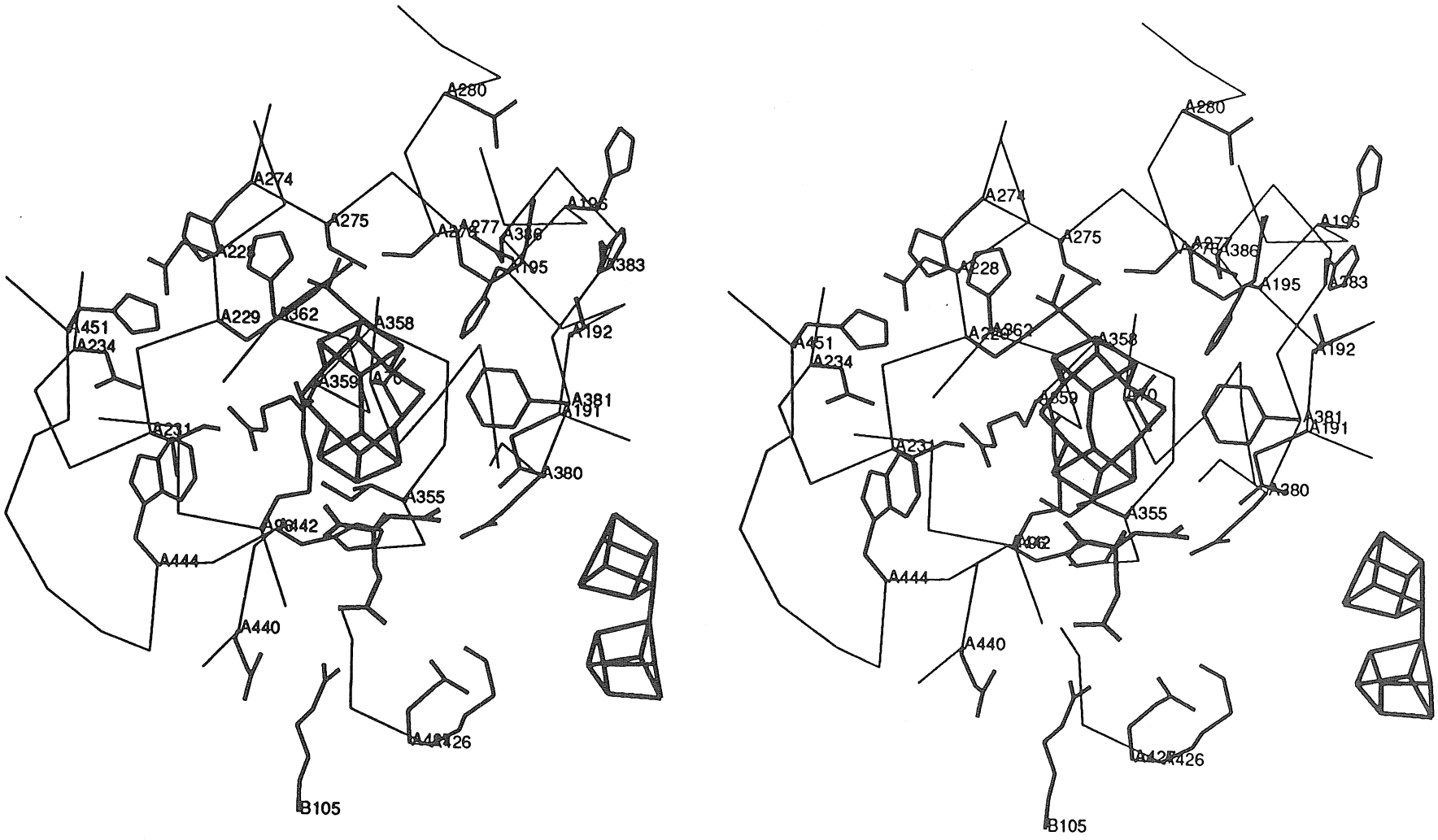


Figure 4-12. Stereoview of the protein environment surrounding the FeMo-cofactor. The FeMo-cofactor is encompassed by the α subunit and buried at least 10Å below the protein surface. The protein environment around the FeMo-cofactor is primarily provided by hydrophilic residues, although there are some hydrophobic residues.

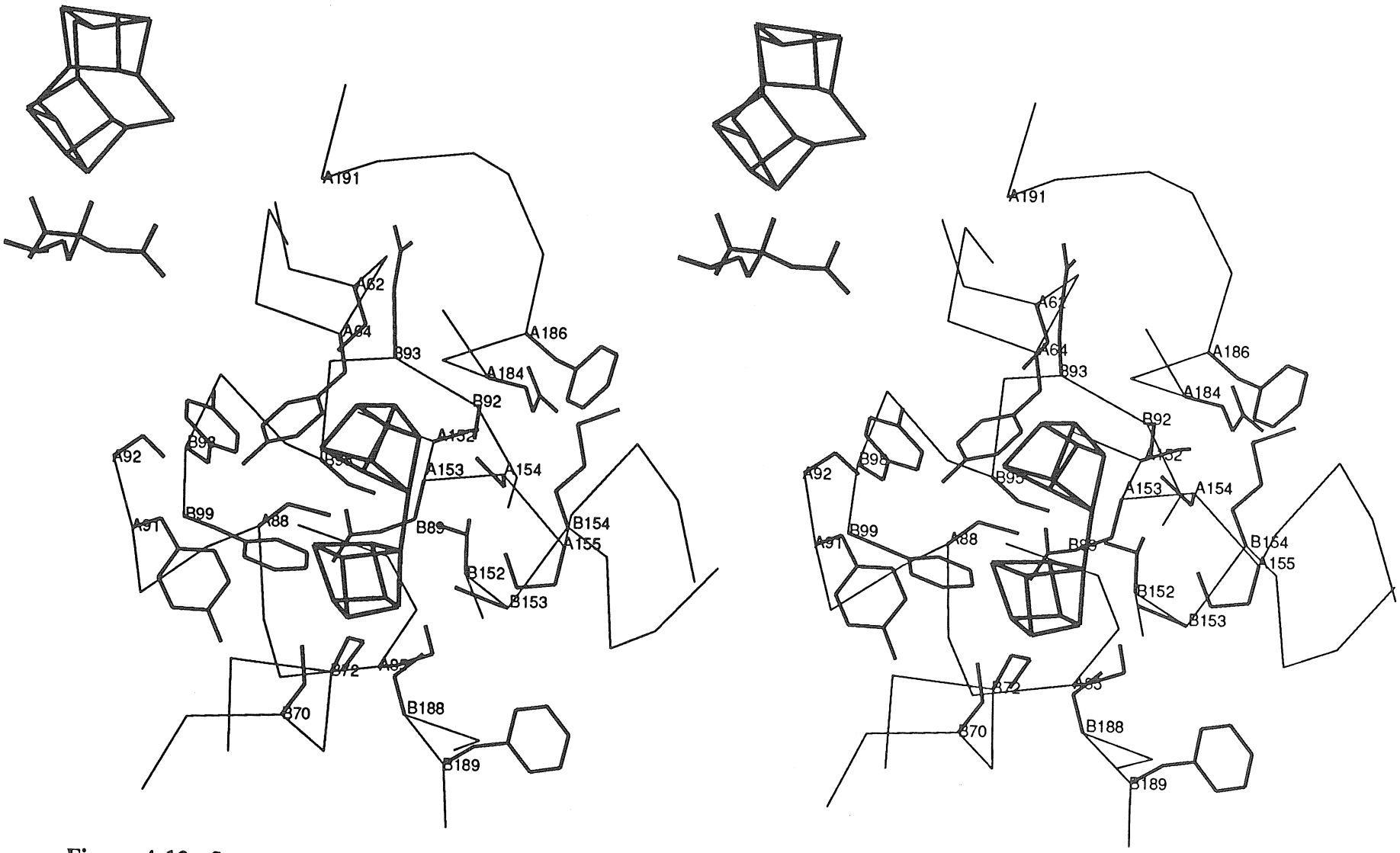


Figure 4-13. Stereoview of the protein environment surrounding the P-cluster pair. The P-cluster pair is located about 12Å below the protein surface, on the 2-fold axis that roughly relates the α and β subunits. The protein environment around the P-cluster pair is mainly provided by hydrophobic residues.

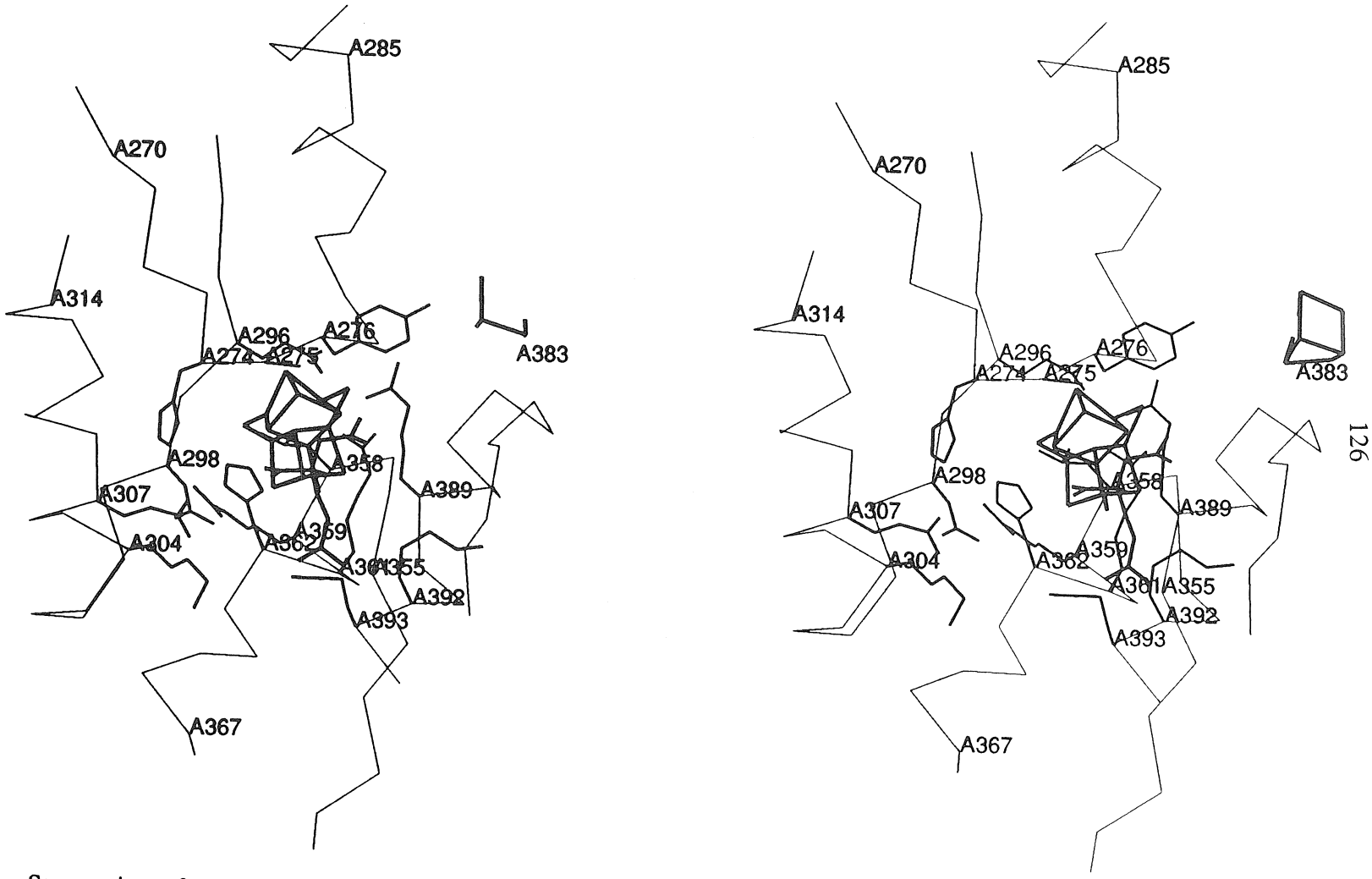


Figure 4-14a. Stereo view of a cleft around the FeMo-cofactor. This cleft could be utilized for substrate entry/product release and/or H_3O^+ transport.

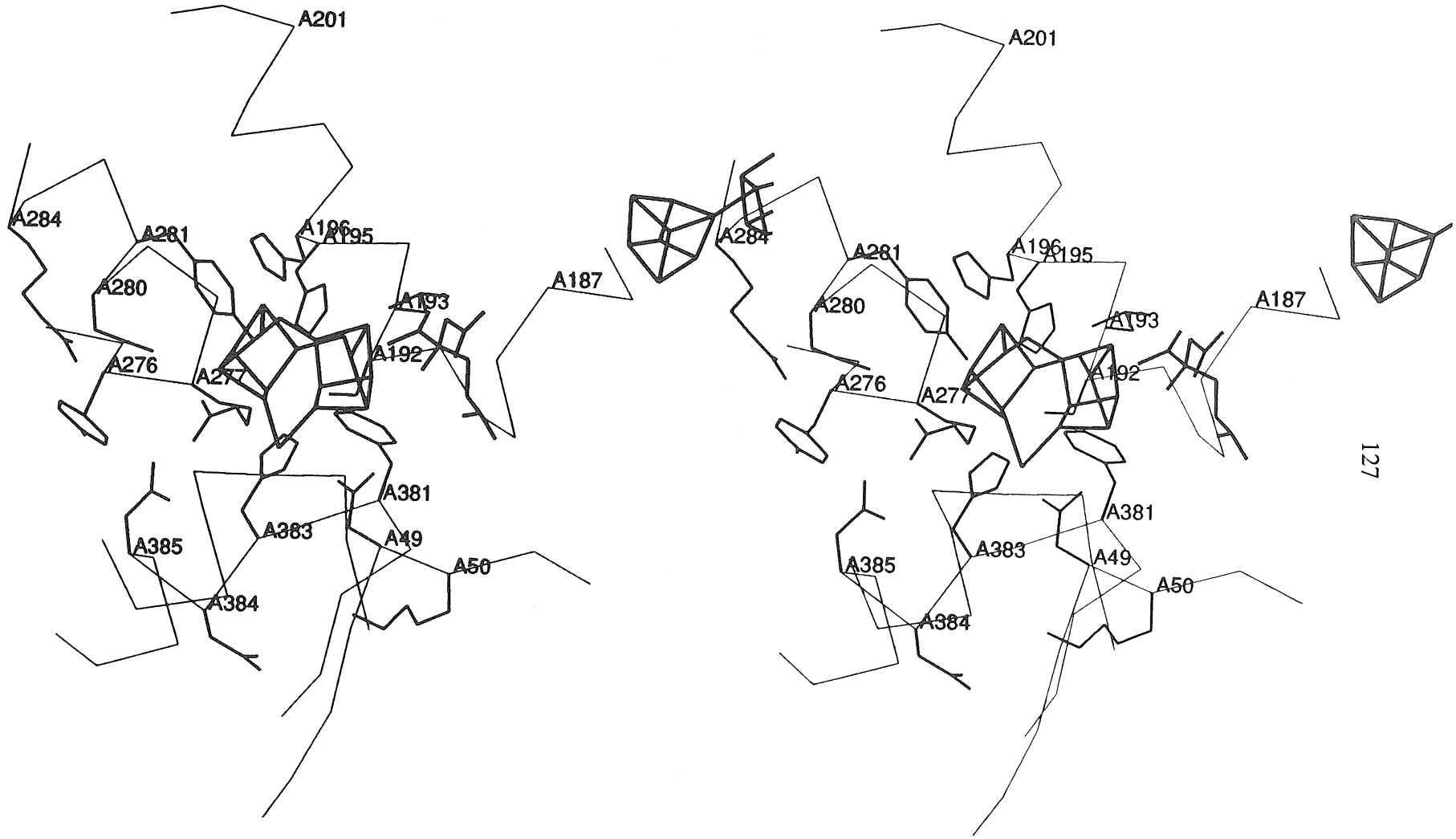


Figure 4-14b. Stereoview of the other cleft around the FeMo-cofactor. This cleft could be also utilized for substrate entry/product release and/or H_3O^+ transport.

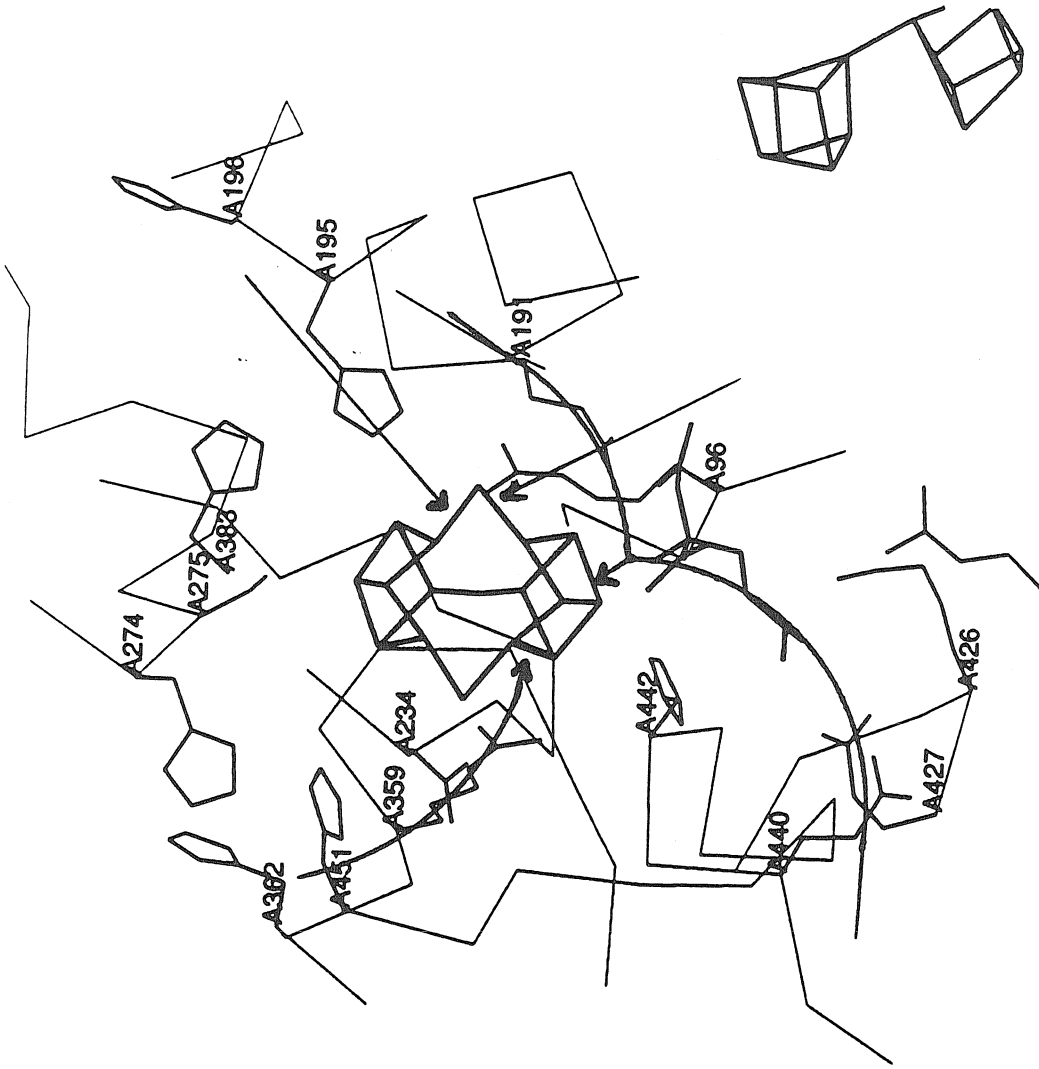


Figure 4-15. Potential proton transfer pathways. The presence of multiple, potential proton transfer routes suggests that there is not a unique pathway by which protons are shuttled from the surface to the active site.

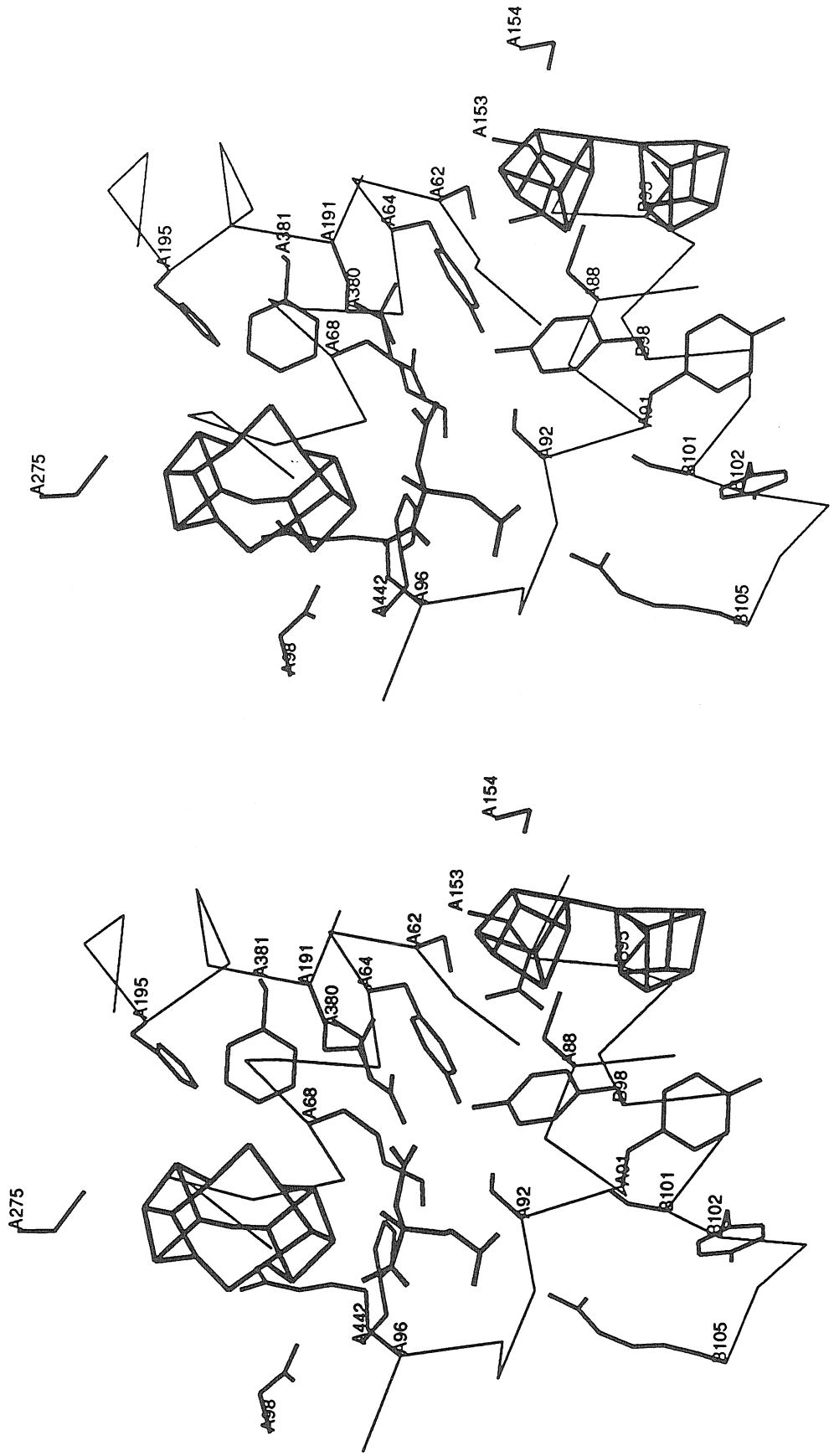


Figure 4-16. Stereoview of the protein environment between the FeMo-cofactor and P-cluster pair. The edge-edge distance of FeMo-cofactor to the P-cluster pair is ~14Å. Four helices are oriented in parallel between the two metal centers and could play a role in electron transfer.

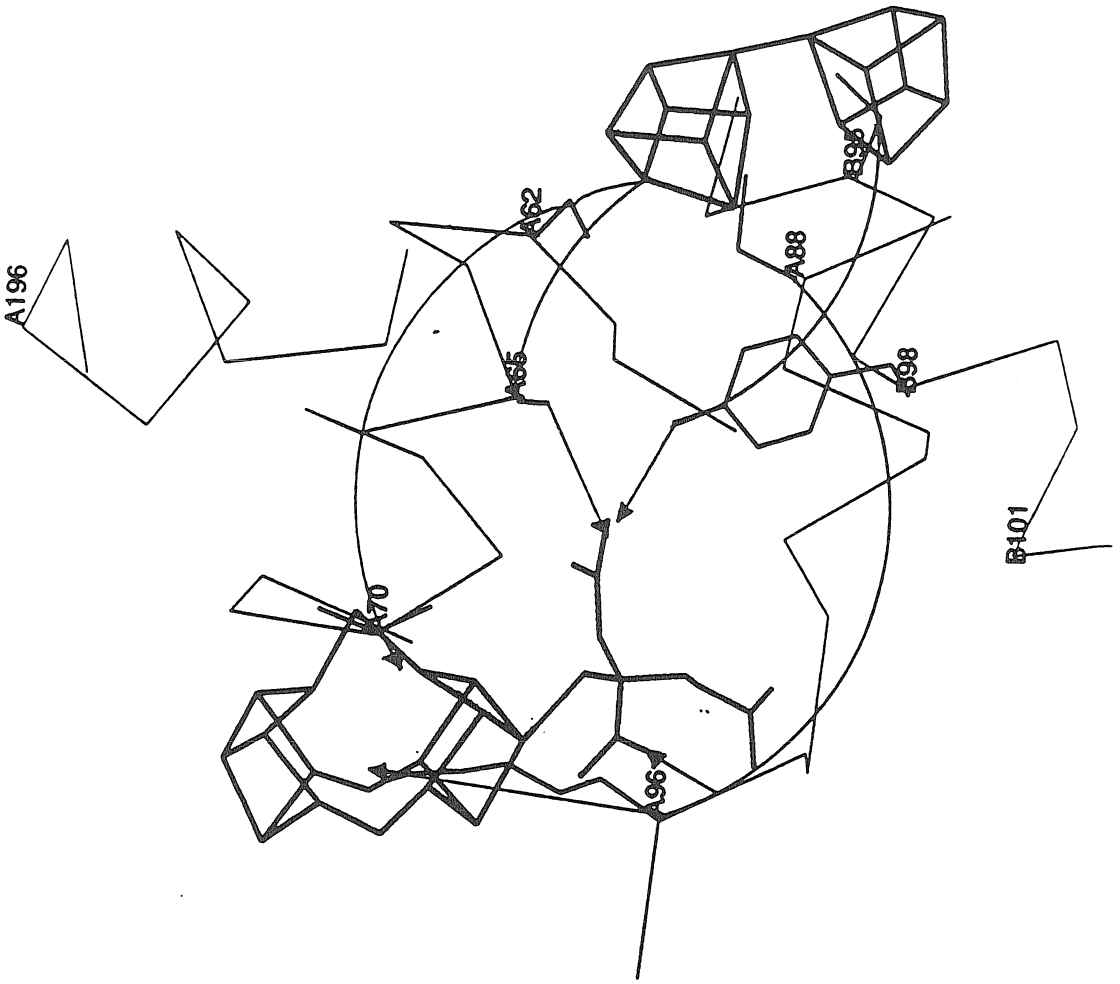


Figure 4-17. Potential electron transfer pathways.

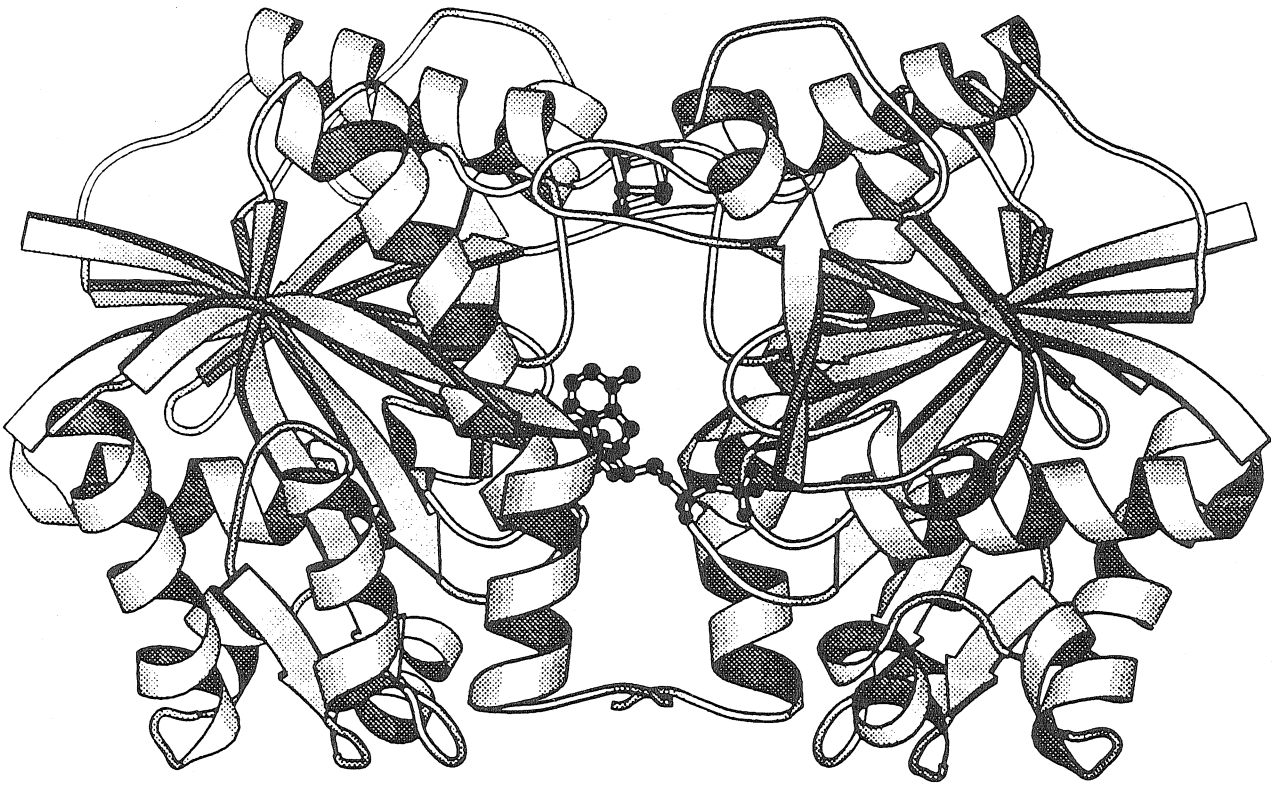


Figure 4-18. Ribbons diagram of the nitrogenase Fe-protein. The 4Fe:4S cluster, which is exposed to the solvent, is coordinated by two identical subunits and the ATP binding site is located at the interface of two subunits.

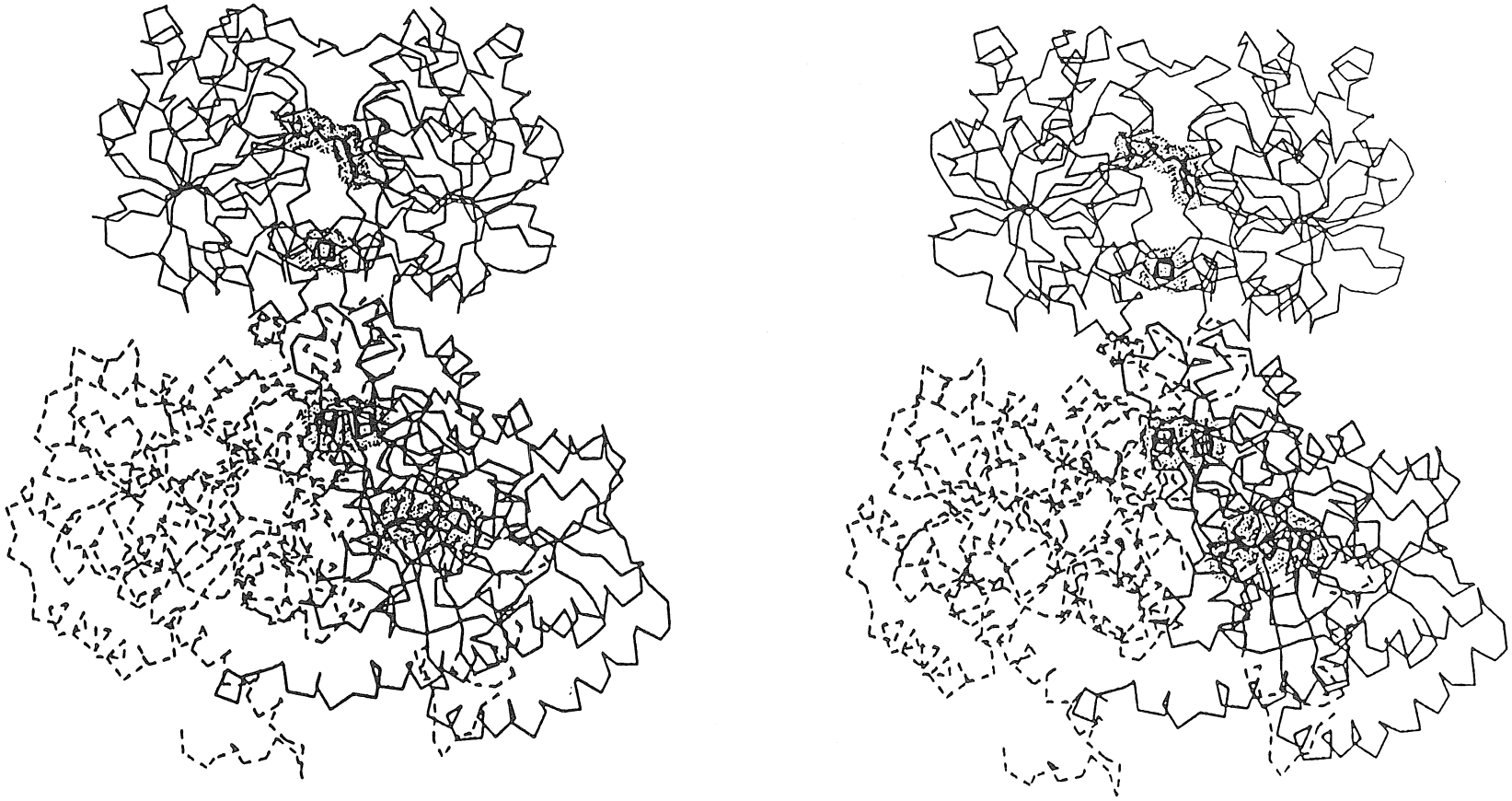


Figure 4-19. Stereoview of a Cα model for a complex formed between Fe-protein and an αβ subunit pair from the MoFe-protein. Fe-protein is represented by the thin lines. The α and β subunits of MoFe-protein are represented by the dark lines and dashed lines, respectively. The metal centers in the two proteins and an ADP molecule at the interface between the two Fe-protein subunits are represented by models surrounded by a dotted van der Waals surface. This model was generated by graphical superposition of the crystal structures of the individual proteins.

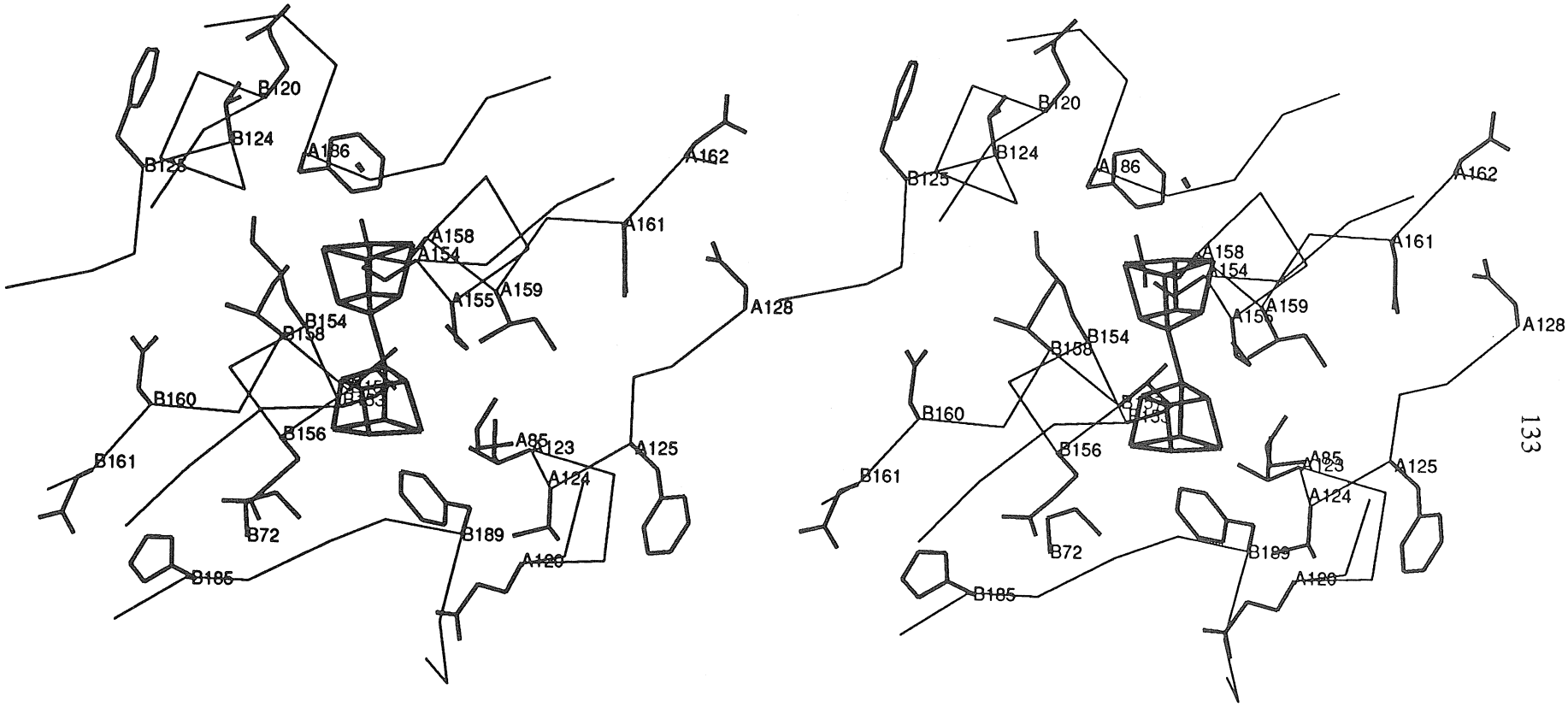


Figure 4-20. Stereoview of the protein environment in the putative Fe-protein binding site. Four short helices are oriented in parallel forming a four-helical bundle, and the 4Fe:4S cluster of Fe-protein could bind to the top surface of these helices.

Chapter 5. Crystallographic Structure of the Nitrogenase Molybdenum-Iron Protein from *Clostridium pasteurianum* at 3.0Å resolution

5-1. Introduction

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 metallo-proteins, the molybdenum iron (MoFe-) protein and the iron (Fe-) protein (reviewed in 1-7), although homologous alternate nitrogenase systems may be induced under the depletion of molybdenum (reviewed in 8,9). 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 under optimal conditions may be represented as follows[10]:

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

The nitrogenase MoFe-protein from *Clostridium pasteurianum* is an $\alpha_2\beta_2$ tetramer with a molecular weight of ~220 kD and the corresponding Fe-protein is a γ_2 dimer with a molecular weight of ~60 kD[11]. The MoFe-protein contains two copies of the FeMo-cofactor (M-center, reviewed in 5,7,12) which are believed to be the substrate binding and reduction sites, and two copies of the P-cluster pair (P-clusters, reviewed in 13) which may serve as the electron mediator 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[14-16].

The physicochemical properties and the primary structures of Fe- and MoFe-proteins are highly conserved among all nitrogen-fixing bacteria so far studied, 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[17,18]. 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[18]. 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[19]. 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)[20]. Since similar studies of the isolated cofactors from each species have shown them to be essentially identical[20], 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[21] and shows a higher specificity for nucleotides[22].

The crystal structure of nitrogenase MoFe-protein from *A.vinelandii,* including the structure of FeMo-cofactor and the P-cluster pair, has been reported[23,24] and is described in Chapter 3 and 4. In this chapter, the tertiary and quaternary structures of the nitrogenase MoFe-protein from *C. pasteurianum* are 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.

5-2. 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[25,26]. The purification, crystallization and heavy atom derivative screening of nitrogenase MoFe-proteins have been described in Chapter 2. 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=70.0Å, b=151.3Å, c=121.9Å, β=110.4°, and a=87.9Å, b=171.4Å, c=73.6Å, 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[27]. Each crystal form was derivatized using ethylmercurithiosalicylate (EMTS) and four common EMTS binding sites which are related by two-fold NCS axis were found for both crystal forms (Table 5-1). The NCS relationships both within and between crystal forms (Av1, Mg1 and Cs1) were determined from rotation functions[28,29] and translation functions[30], 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[26] using the two-fold NCS in both crystal forms. The final R factor of the four-fold averaging was ~22% and that of the subsequent two-fold averaging within the Mg1 crystal form was ~18%. The Mg1 crystal form was used for further crystallographic analysis but 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[31] implemented with the fragment fitting option[32], and subsequently refined using the restrained least square program TNT[33]. The resultant $2F_O$-F_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 X-PLOR[34] 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 (10-3.0Å) with root mean square deviations from ideal bond distances and angles of 0.018Å and 3.9°, respectively (Table 5-2). The correctness of the chain trace is further supported by the location of the heavy atom binding sites (Table 5-1) 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[35] and all the residues have reasonable average 3D-1D scores except for the region, α380-α391 in which electron density is diffused. A Ramachandran plot of the αβ subunit pair in MoFe-protein is presented in Figure 5-1. Residue numbers are prefixed with either α or β to indicate the appropriate subunit number, unless 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 5-1. Heavy atom binding sites in crystals of *C. pasteurianum* MoFe-protein. Locations for the heavy atoms were determined from difference Patterson maps with the aid of molecular replacement phases, and refined with the program HEAVY[65].

Derivative	x	y	z	relative occupancy	location
EMTS (Mg1)	0.356	0.073	0.858	0.23	Cys α302
	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 (Cs1)	0.416	0.231	-0.105	0.17	*
	-0.033	-0.254	0.071	0.16	
	0.737	0.021	0.719	0.14	
	0.019	-0.008	0.924	0.11	

* The binding locations for the Cs1 heavy atom sites presumably composed to the Mg1 sites, but this has not been directly established.

Table 5-2. Refinement statistics of Cp1 structure. The protein model was built with the graphics program TOM/FRODO[31]. The electron density map was improved by iterative cycles of model building, refinement, phase combination and NCS averaging. The program TNT[33] was used for coordinate refinement during the initial stages of modeling, whereas the simulated annealing program X-PLOR[34] was used during the final stages of refinement with the PARAM19X.PRO parameter file.

Model status (Mg1)

# of amino acid residues	1964/1980
Total non-hydrogen atoms	15,273 (99% complete)
# of water molecules	0
Missing residues	α527-α534
# of Ramachandran outliers	18/1960

Refinement statistics

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

5-3. 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[18] (Figure 1-2). 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) (Figure 4-2b). The α subunit of Cp1 consists of three domains of the α/β type with some extra helices (Figures 5-2a and b). The overall shape of the α subunit may be described as a clover leaf. Domain I is composed of 7 helices, 4 parallel β strands, and 1 antiparallel β strand which is provided by the β subunit; domain II is composed of 6 helices, 4 parallel β strands, and 1 antiparallel β strand; and domain III is composed of 10 helices, 5 parallel β strands, and 1 antiparallel β strand which is from N terminus of the α subunit. There is a wide and shallow cleft between the three domains, and the FeMo-cofactor sits under the bottom of this cleft. This cleft has been proposed as a Fe-protein subunit binding site in Av1[24] (Chapter 4). 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 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[18] (Figure 1-2). The structure of the β subunit is also similar to that of Av1 except for the N terminus (~10 residues of Cp1) and loop regions at

the surface where deletions or insertions occur (Figures 5-3a and b). 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[24] (Chapter 4). However, significant differences also exist, especially at the N terminus, in the α375-α430 region, and in surface loop regions (Figure 5-3b). In the β subunit, domain I' is made up of 7 helices and 4 parallel β strands; domain II' is made up of 6 helices and 4 parallel β strands; and domain III' is made up of 6 helices and 5 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[24] (Chapter 4). 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 this region is not as well ordered in Cp1 as in Av1.

The αβ subunit pair consists of six α/β type domains which are arranged like a six petalled flower, with the P-cluster pair located in the center of the dimer like a pistil (Figures 5-4a and b). The α subunit and the β subunit of MoFe-protein are related by an approximate two-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 αβ 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[24] (Chapter 4).

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 ~70Åx80Åx120Å and its general shape is slightly elongated along one axis compared to that of Av1 due to the ~50 residue insertion in the α subunit (Figures 5-5a and b). The two αβ subunit pairs are related by the two-fold NCS rotation axis that was used for initial phase refinement. Even though the α and β subunits in an αβ subunit pair are also related by an approximate two-fold rotation axis, the MoFe-protein tetramer does not exhibit 222 symmetry, which had been proposed by a low resolution rotation function study[36], because the αβ two-fold and the tetramer two-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[37], 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 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 ~8-10Å wide and ~35Å in length (Figure 5-5a). In addition to these helical packing interactions, electrostatic, hydrophobic and hydrogen bonding interactions are important for tetramerization, as is observed in Av1[24] (Chapter 4).

The divalent cation sites found in the Av1 structure also exist in Cp1, and the tertiary structure around these sites is well conserved. Based on the electron density value, temperature factor and the coordination environment, these sites could be assigned as Ca^{2+} (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 (Figure 5-6). Gln D472, 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 αβ dimers and is ~25Å and ~21Å 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[38] and some spherical plant viruses[39] 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, just as in the case of Fe-protein[40], and no ATP/ADP is found in Cp1 and Av1, although some experimental studies have indicated that oxidized MoFe-protein can bind MgADP[41].

5-4. Structures of the metal centers

The structures of the metal centers in Av1 have been described in reference 23 and verified by high resolution X-ray diffraction analysis[42] (Chapter 3). The FeMo-cofactor and P-cluster pair structures of Cp1 appear identical to those of Av1, as had been indicated by various experimental results. The FeMo-cofactor contains 4Fe:3S and 1Mo:3Fe:3S clusters that are bridged by three non-protein ligands (Figure 5-7). 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. X-PLOR refinement results indicate that there may be some iron-iron bonding interactions (average Fe-Fe bond distance = ~2.5Å) that could provide a fourth coordination interaction for the bridging irons. The Mo site has the highest electron density value in the 3.0Å resolution 2Fo-Fc map. Although the other atomic positions are not resolved, the FeMo-cofactor model generally matches the electron density well. Homocitrate, an essential component of the FeMo-cofactor[43], 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[44]. The FeMo-cofactor is attached to the α subunit through two protein ligands, Cys α262 and His α482. Cys α262 coordinates Fe1, and His α482 coordinates Mo along with homocitrate.

A ball and stick model of the P-cluster pair is presented in Figure 5-8. The P-cluster pair consists of two 4Fe:4S clusters, as had been suggested by Mössbauer and extrusion studies[45,46], 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 =~2.2-2.3Å) 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 (Chapter 3). 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.

5-5. 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[47]. In its isolated form, however, it no longer catalyzes dinitrogen reduction[47]. 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 5-9. The FeMo-cofactor is buried at least 10Å 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 5-10b and 5-15. 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) which are important to avoid steric interference with the FeMo-cofactor. Other highly conserved residues near the FeMo-cofactor include: Arg α87 and Arg α347, which can potentially form hydrogen bonds to cluster 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.

5-6. Substrate entry and product release

Even though the FeMo-cofactor is buried at least 10Å below the protein surface and there are no permanent channels between the protein surface and the FeMo-cofactor[24] (Chapter 4), there are two clefts (Figure 5-10a and b) which could be potentially utilized for the substrate entry/product release and/or H_3O^+ transfer to the active site. The first cleft (Figure 5-10a), 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 ~4-10Å. The second cleft (Figure 5-10b), which is 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 also has a funnel shape with an outer diameter of ~3-10Å. 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 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[48].

5-7. Proton transfer

In order to reduce dinitrogen to ammonia, a supply of protons is essential in addition to dinitrogen. In the vicinity of the FeMo-cofactor, several potential pathways occur which could transfer protons to the FeMo-cofactor: Glu α368 and Glu α467 to homocitrate to the FeMo-cofactor, and His α186 to the FeMo-cofactor (Figure 5-9). 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 Arg α87 to the FeMo-cofactor. The two potential channels around the FeMo-cofactor shown in Figure 5-10a and 5-10b could be utilized for the proton transfer to the vicinity of the cofactor. Charged or hydrophilic residues lining these potential 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 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[10].

5-8. 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 5-11. The P-cluster pair is bridged by the α and β subunits and is completely buried at least 12Å 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, Hr α144, α175, Ser β50, His β52 and Gln β24 are generally not conserved, with the exceptions of Gln β46 and Thr β105.

5-9. Mechanism of H_2 evolution

Under optimal conditions, nitrogenase catalyzes the reaction: $N_2 + 8H^+ + 8e^- \rightarrow 2NH_3 + H_2$. H_2 evolution occurs to the same extent in the presence of saturating N_2[49] and is thought to arise from N_2 binding and displacing H_2 from the active site[50]. However, 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[51], or where the pressure of N_2 is low[49]. 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 (Figure 5-8) may provide a site for H_2 evolution[42] (Chapter 3). 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.

5-10. 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 5-12. The edge-edge distance of FeMo-cofactor to the P-cluster pair is ~14Å. 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 structural connection between the P-cluster pair and FeMo-cofactor. However, there are no completely covalent bond or hydrogen bond networks directly linking the two metal centers, suggesting that either some through space jumps or structural fluctuations may occur during substrate reduction that permit the formation of more favorable electron transfer pathways. Additionally, 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 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 FeMo-cofactor, as well as from modulation of the redox properties of the coordinated FeMo-cofactor[23].

5-11. Fe-protein binding sites

The α and β subunits of MoFe-protein are related by an approximate two-fold axis that passes through the center of the P-cluster pair, and there are two wide and shallow clefts related by this pseudo two-fold rotation (Figure 5-13). As the Fe-protein dimer also has a two-fold NCS axis[40], a plausible model for docking the two proteins involves superposition of the Fe-protein two-fold axis with the approximate two-fold axis of MoFe-protein passing through the P-cluster pair[24] (Chapter 4). This docking model, with a stoichiometry of 1 MoFe-protein: 2 Fe-protein, is consistent with both mutagenesis and crosslinking results[52-55]. 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 5-13) and a 4Fe:4S cluster binding site (designated C in Figure 5-13). Interestingly, the long inserted sequences of Cp1 (α375-α430) are 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[19]. 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[19], and this finding indirectly supports the assignment of Fe-protein binding sites in MoFe-protein and the proposed docking model[24] (Chapter 4).

The potential 4Fe:4S cluster binding site (site C in Figure 5-13) and surrounding residues are shown in Figure 5-14. Four short helices (α112-α117, α146-α150, β106-β111 and β73-β78), which are related by the approximate two-fold axis, are oriented in parallel from the P-cluster pair towards the protein surface, forming an approximate four-

helical bundle. The 4Fe:4S cluster of Fe-protein is likely to sit on the top of the helical bundle. The edge-edge distance from the P-cluster pair to the end of these helices is about 12Å, thus the edge-edge distance from the P-cluster pair to the 4Fe:4S cluster of Fe-protein may be about 15Å. 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, Leu β107, 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. Hydrophilic residues 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 two-fold axis, are completely exposed to solvent and are about 18Å away from the P-cluster pair. His β136 is exposed to solvent and is about 16Å 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 5-13), which are related by the pseudo two-fold rotation, and the exposed residues around these sites are shown in Figures 5-15 and 5-16. It is intriguing that the FeMo-cofactor is located under one of the possible Fe-protein subunit 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[56]. 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 interaction 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[57].

5-12. Alternative nitrogenases

The sequences of the structural genes of the vanadium-iron protein (VFe-) subunits (*vnfDK*) and those of the third nitrogenase, iron-only nitrogenase (*anfDK*), show considerable similarity with *nifDK*[7,58]. Of particular significance is the conservation of the Cys residues which ligate the metal centers in these proteins. An additional feature of the operon encoding the structural genes of alternative nitrogenases is the presence of a small additional gene (*vnfG* and *anfG*) between the D and K genes[8,9]. Sequence comparison with the conventional nitrogenase shows that there is no corresponding sequence which is fused within *nifD* or *nifK*[58]. The VFe-proteins of *A. chroococcum* and *A. vinelandii* have been purified and characterized . The protein isolated from *A. chroococcum* has an $\alpha_2\beta_2\delta_2$ subunit structure with the δ subunit being encoded by the *vnfG* gene. Although the δ subunit has not been identified in the protein of *A. vinelandii*, the DNA sequence of the structural gene operon included a homologous gene (*vnfG*) to that found in *A. chroococcum.*

The metal content of current preparations of VFe-proteins is generally lower than those of MoFe-proteins, but V is found in place of Mo[59]. However, EPR and magnetic circular dichroism (MCD) studies on the Vanadium nitrogenase system indicates that the VFe-protein possesses similar redox centers to the MoFe-protein of the conventional nitrogenase system[60,61]. The Vanadium K-edge EXAFS spectrum has been reported for the VFe-proteins of both *A. chroococcum* and *A. vinelandii*[62-64]. These two VFe-proteins appear to be structurally very similar, and the iron-vanadium cofactor (FeVa-cofactor) appears to be apparently analogous to the FeMo-cofactor in MoFe-protein. MCD spectroscopy on the oxidized VFe-protein is also consistent with the presence of the P-cluster pair which is found in MoFe-protein[61].

The α subunit of the VFe-protein is similar in size to that of Av1[58]. However, the β subunit of the VFe-protein has ~50 residue deletion in the N-terminal region compared to that of Av1 as in the case of the Cp1 β subunit. Consequently, the β subunit of VFe-

protein is similar in size to Cp1[58]. Therefore, the three-dimensional structure of VFe-protein could be considered as a hybrid of the Av1 and Cp1 structures; the structure of VFe-protein could be constructed from the Av1 α subunit and Cp1 β subunit. The only remaining question about the overall structure of the VFe-protein is the location of the δ subunit. Since no functional role for the δ subunit has been observed and the δ subunit is missing in some VFe-protein preparations[9], a possible role of the δ subunit may be stabilization of the quaternary structure of VFe-protein. Based on the structure of the Av1 and Cp1, a possible location of the δ subunit is around the N-terminal region of the Av1 β subunit. The N-terminus of the Av1 β subunit, which is absent in Cp1 and VFe-protein, extends from the β subunit and wraps around the α subunit. These residues also interact with other β subunit. Therefore, it is likely that these N-terminal residues may function in stabilization of the quaternary structure of Av1 and may be replaced by the δ subunit in VFe-protein, although there is no sequence homology[58]. Consequently, the structure of the VFe-protein is probably very similar to that of Av1. Based on the amino acid sequence comparisons[8,9,58] the structure of FeFe-protein of iron-only nitrogenase may be very similar to that of VFe-protein .

References

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

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

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

4. Stiefel, E. I., *et al.* (1988)in *Metal Clusters in Proteins* (ed., Que, Jr., L.) ACS Symp. *Ser.* 372, 372-389 (American Chemical Society, Washington, D. C.).

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

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

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

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

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

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

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

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

13. Holm, R. H., Ciurli, S. & Weigel, J. A. (1990) *Prog. Inorg. Chem.* **38**, 1-74.

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

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

16. Hageman, R. V. & Burris, R. H.(1978) *Biochem.* **17**, 4117-4124.

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

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

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

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

21. Guth, J. H. & Burris, R. H.(1983) *Biochem.* **22**, 5111-5122.

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

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

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

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

26. Bricogne, G. (1976) *Acta Cryst.* **A32**, 832-847.

27. Weininger, M.S. & Mortenson, L.E. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 378-380.

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

29. Crowther, R. A. (1972) in *the Molecular Replacement Method* (Rossmann, M. G., ed.) pp. 173-178 (Gordon & Breach, New York).

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

31. Jones, T. A. (1985) *Meth. Enzym.* **115**, 151-171.

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

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

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

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

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

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

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

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

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

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

42. Chan, M. K., Kim, J. & Rees, D. C. (1993)*Science* in press.

43. Hoover, T. R., Imperial, J., Ludden, P. W. & Shah, V. K. (1989)*Biochem.* **27**, 2768-2771.

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

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

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

47. Shah, V. K. & Brill, W. J. (1977)*Proc. Natl. Acad. Sci. USA* **74**, 3249-3253.

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

49. Hadfield, K. L. & Bulen, W. A. (1969) *Biochem.* **8**, 5103-5108.

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

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

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

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

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

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

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

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

58. Pau, R. N. (1991) in *Biology and Biochemistry of Nitrogen Fixation* (Dilworth, M. J. & Glenn, A. R., eds.) pp. 37-57, Elsevier, New York, London.

59. Eady, R. R. (1990) in *Vanadium in Biological Systems* (Chasteen, N. E., ed.) pp. 99-127, Kluwer Academic Publishers, Dordrecht.

60. Morningstar, J. E. & Hales, B. J. (1987) *J. Am. Chem. Soc.* **109**, 6854-6855.

61. Morningstar, J. E., Johnson, M. K., Case, E. E. & Hales, B. J. (1987) *Biochem.* **26**, 1795-1800.

62. Aber, J. M., Dobson, R. R., Eady, R. R., Stevens, P., Hasnain, S. S., Garner, C. D. & Smith, B. E. (1987) *Nature* **325**, 372-374.

63. George, G. N., Coyle, C. L., Hales, B. J. & Cramer, S. P. (1988) *J. Am. Chem. Soc.* **110**, 4057-4059.

64. Harvey, I., Arber, J. M., Eady, R. R., Smith, B. E., Garner, C. D. & Hasnain, S. S. (1990) *Biochem. J.* **266**, 929-931.

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

66. Kraulis, P. J. (1991) *J. Appl. Cryst.* **24**, 946-950.

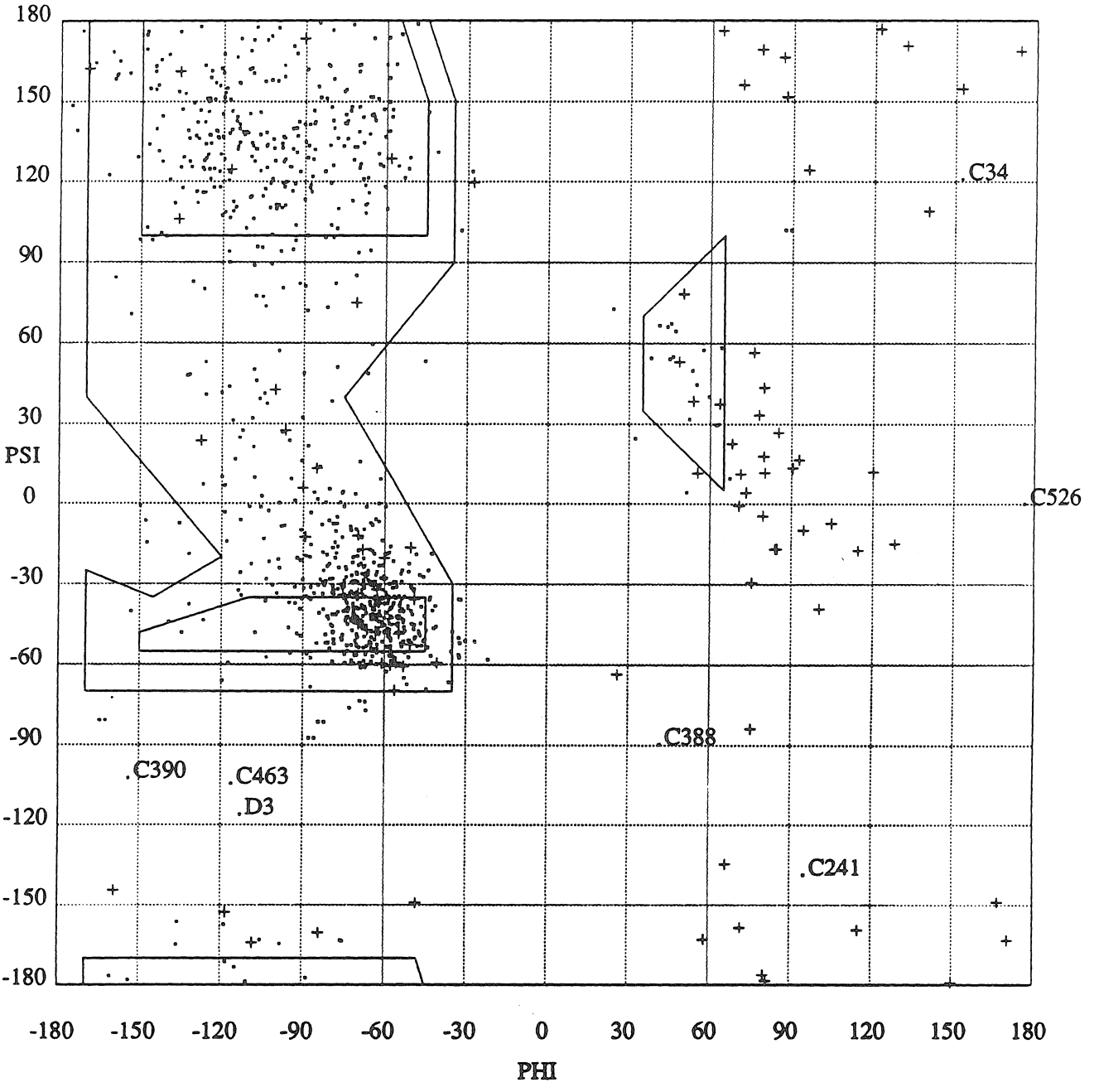


Figure 5-1. Ramachandran plot of the αβ subunit pair inMoFe-protein. Glycine residues are represented as crosses. Ramachandran outliers are labeled.

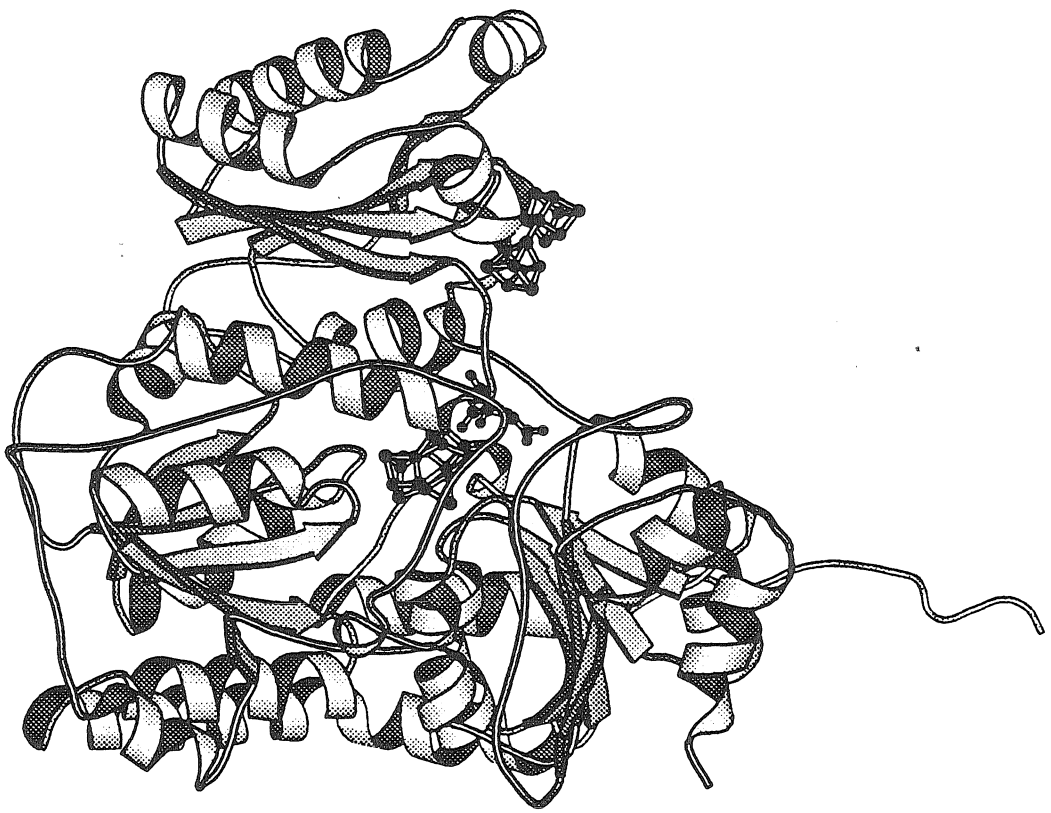


Figure 5-2a. Ribbons[66] diagram of the polypeptide fold of the α subunit of Cp1. The three domains are designated I, II and III, respectively, counterclockwise from the top. The FeMo-cofactor is located in the center of the figure, and the P-cluster pair is on the top right side.

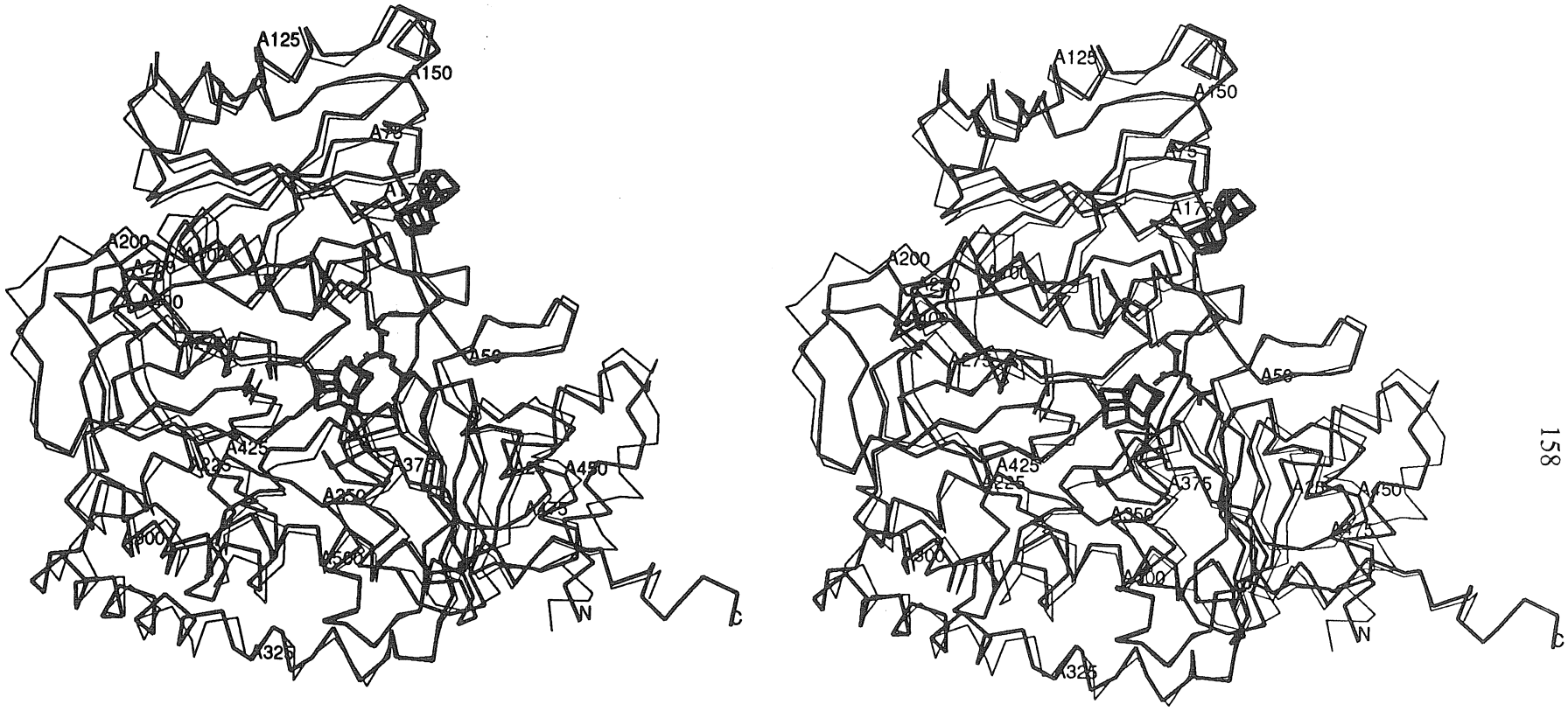


Figure 5-2b. Cα chain trace of the α subunit of Cp1 superimposed on that of Av1. The rms deviation between common Cα positions is 1.57Å. The α subunits of Cp1 and Av1 are represented by the thick and thin lines, respectively. The view is approximately from the same orientation as Figure 5-2a.

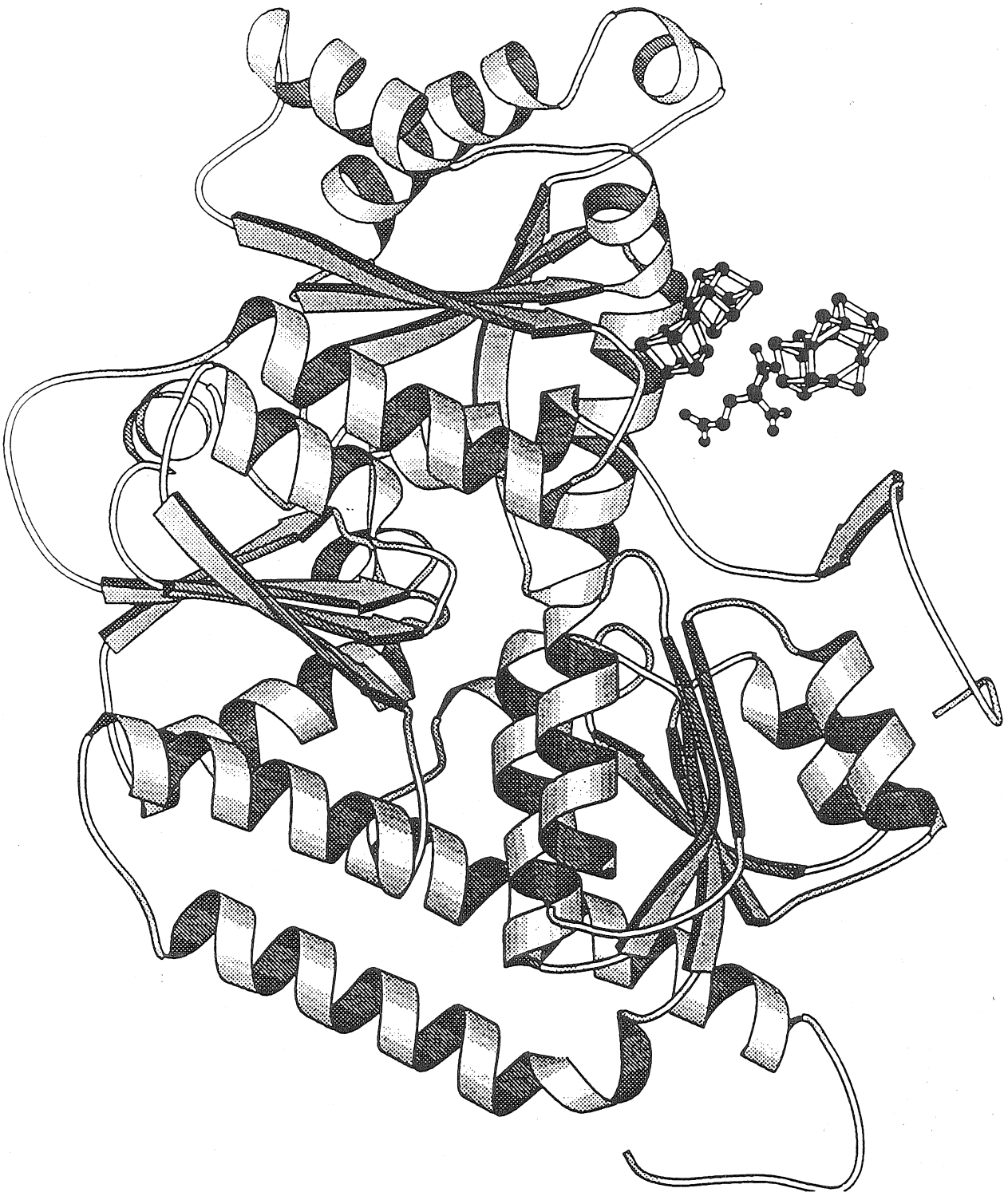


Figure 5-3a. Ribbons diagram of the polypeptide fold of the β subunit of Cp1. The three domains are designated I′, II′ and III′, respectively, counterclockwise from the top. The left-most metal center is the P-cluster pair, while the FeMo-cofactor is to the right of the P-cluster pair.

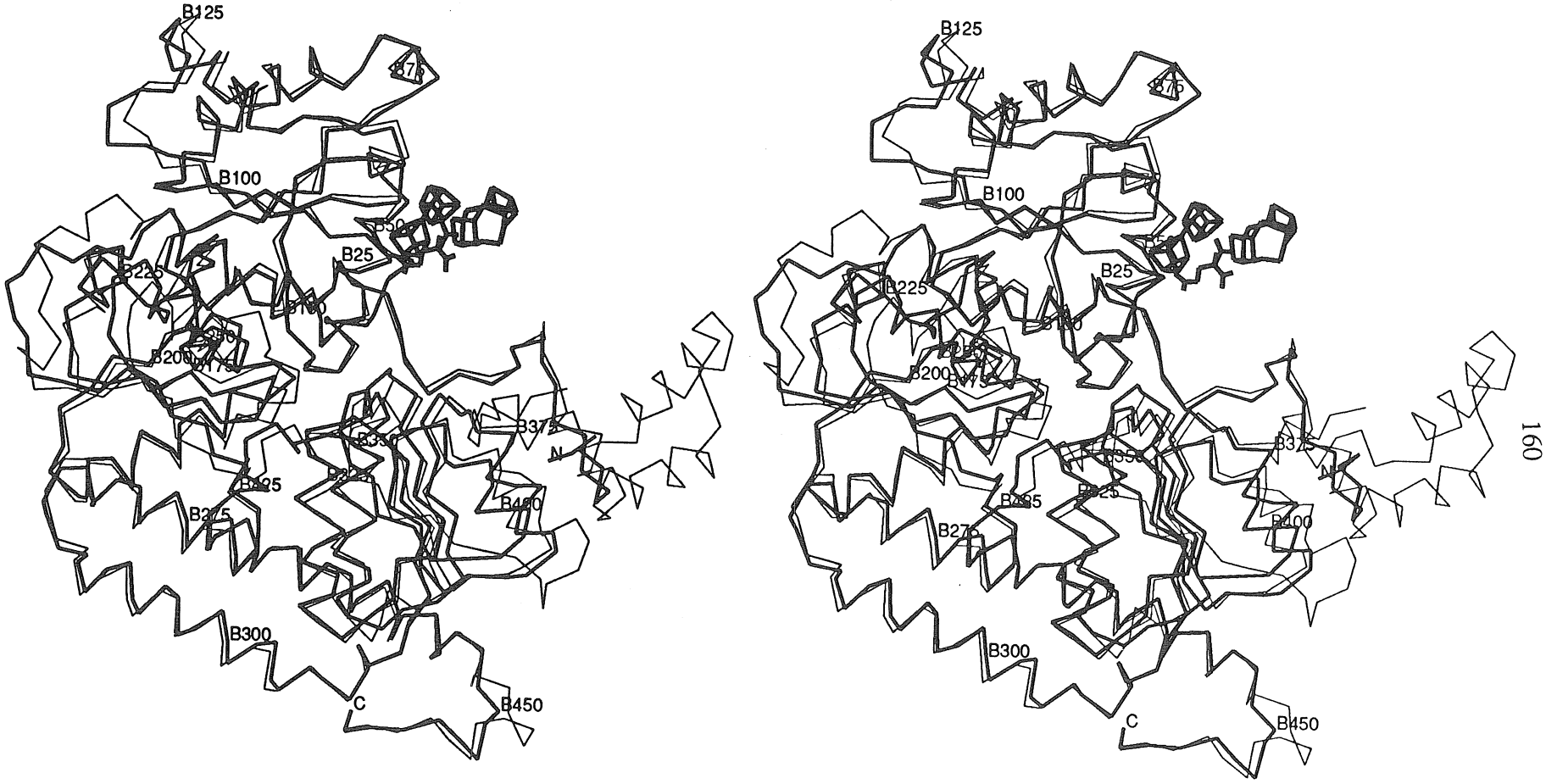


Figure 5-3b. Cα chain trace of the β subunit of Cp1 superimposed on that of Av1. The rms deviation between common Cα positions is 1.61Å. The β subunits of Cp1 and Av1 are represented by the thick and thin lines, respectively. The view is approximately from the same orientation as Figure 5-3a.

Figure 5-4a. Ribbons diagram of the polypeptide fold of an αβ subunit pair. The view is down the two-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.

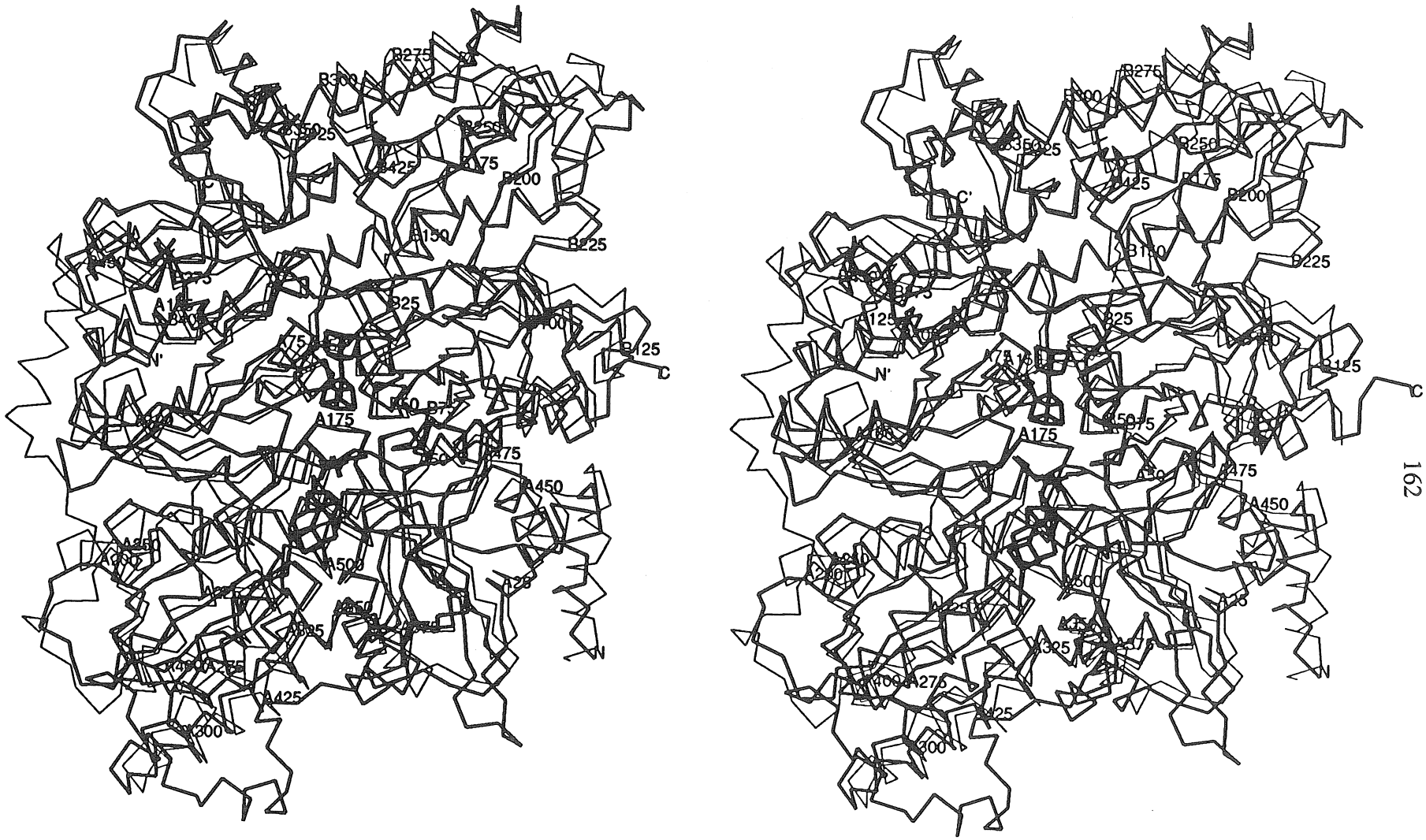


Figure 5-4b. Cα chain trace of the αβ subunit pair of Cp1 superimposed on that of Av1. The rms deviation between 880 common Cα positions is 1.63Å. The αβ subunit pairs of Cp1 and Av1 are represented by the thick and thin lines, respectively. The view is approximately from the same orientation as Figure 5-4a.

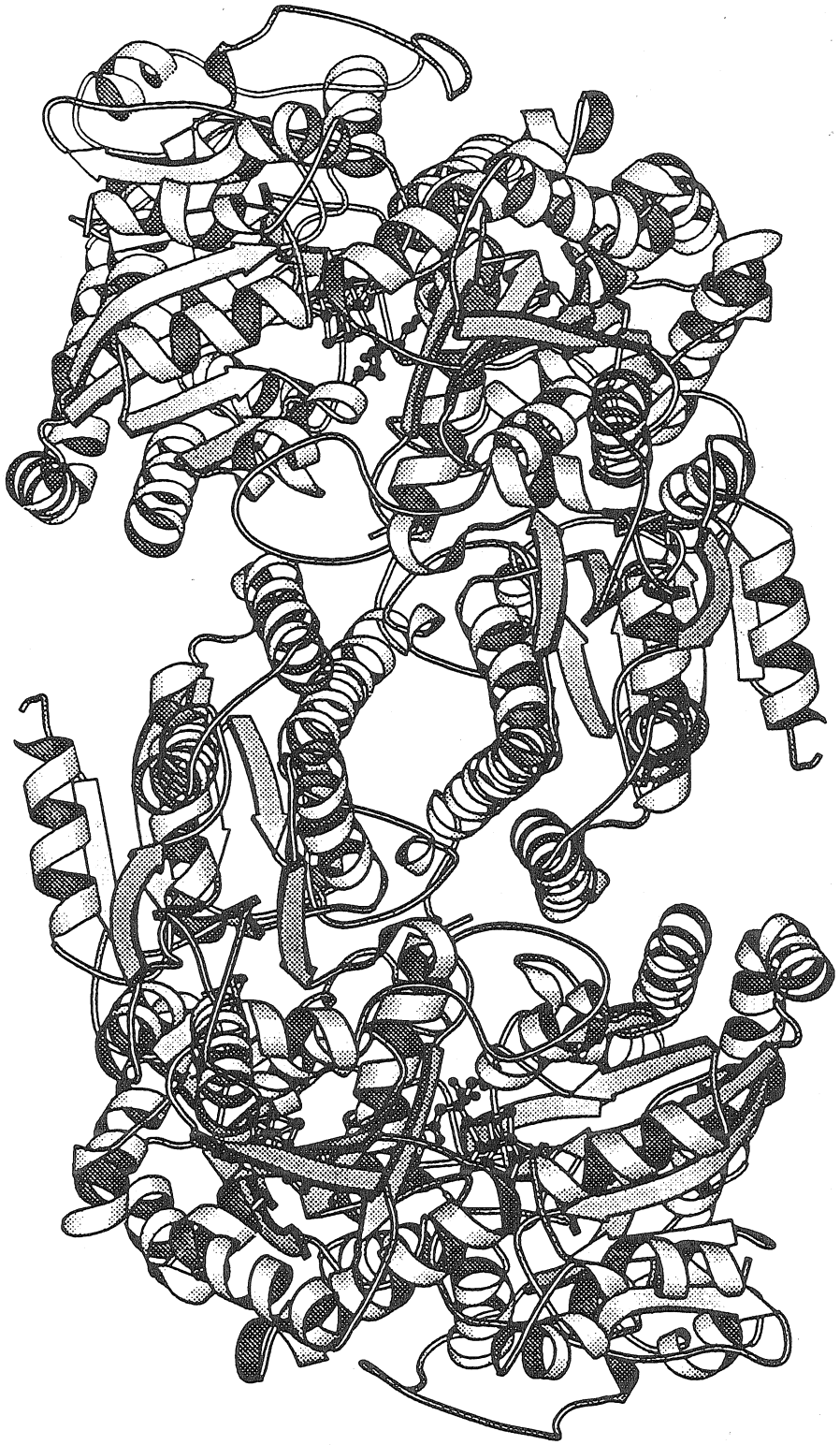


Figure 5-5a. Ribbon diagram of the polypeptide fold of the $\alpha_2\beta_2$ MoFe-protein tetramer I. The view is down the tetramer twofold axis.

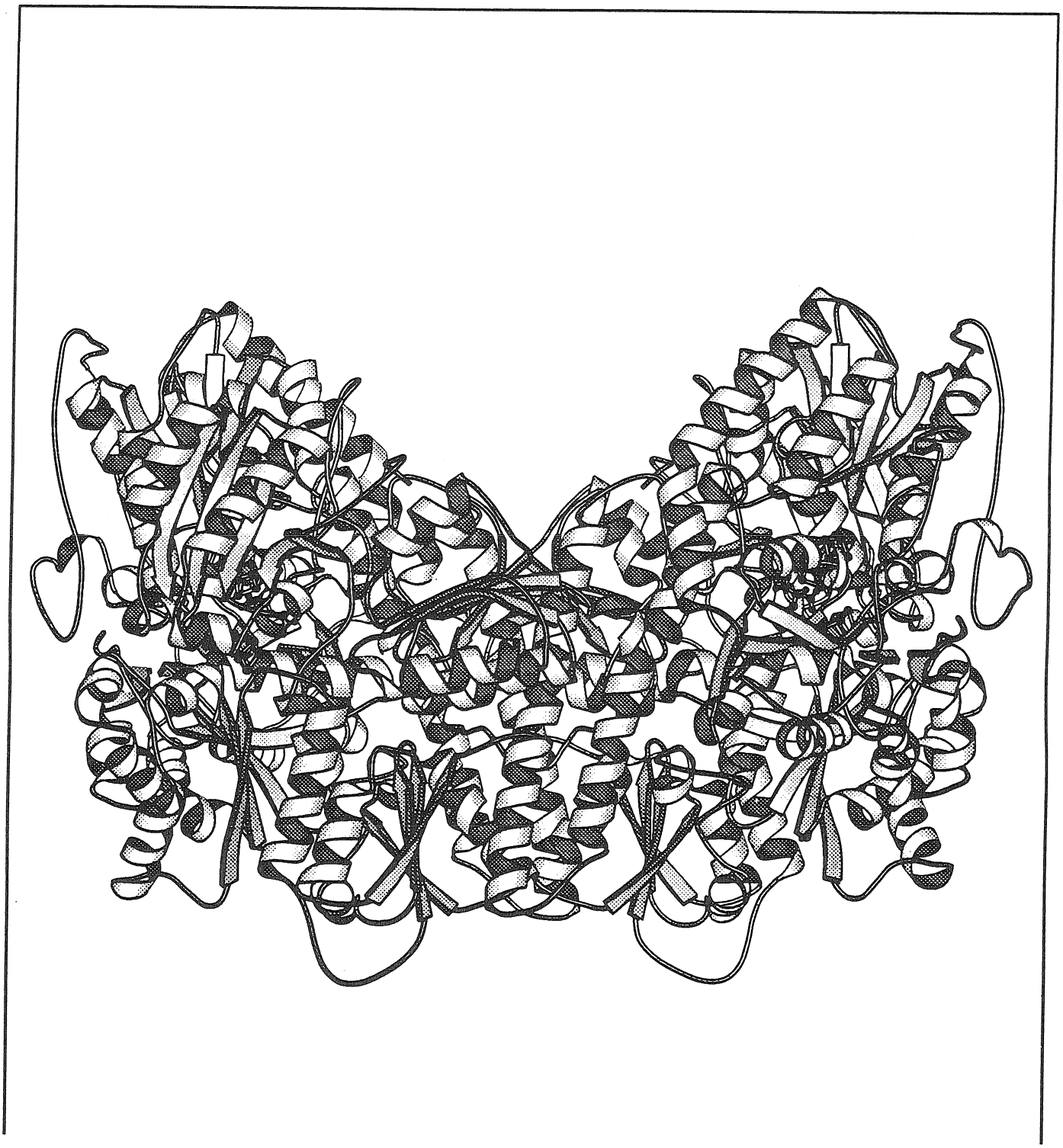


Figure 5-5b. Ribbons diagram of the polypeptide fold of the $\alpha_2\beta_2$ MoFe-protein tetramer II. The view is perpendicular to the tetramer two-fold axis.

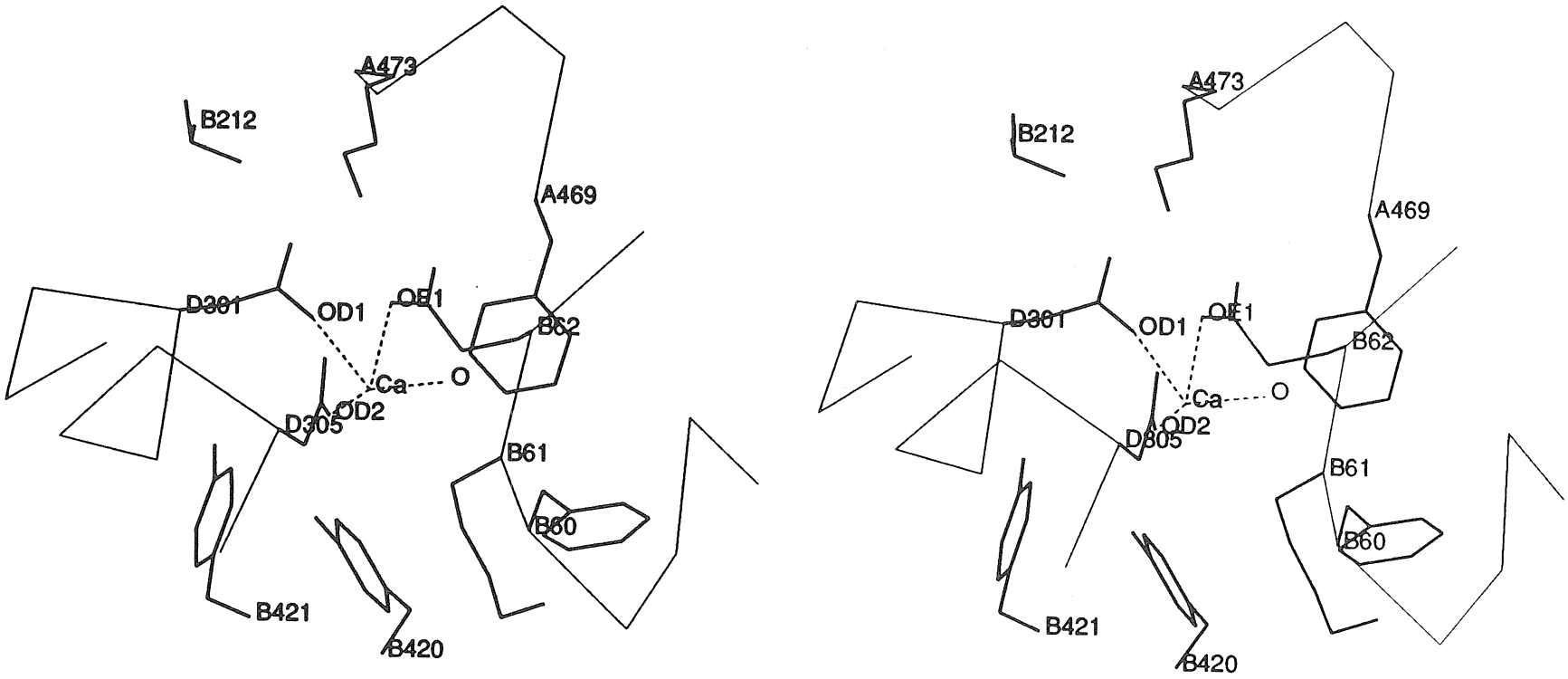


Figure 5-6. A divalent cation binding site and surrounding residues. This ion has an octahedral coordination environment provided the carboxyl oxygens of Glu B62, Asp D301 and Asp D305, carbonyl oxygen of Lys B61 and probably two water molecules (have not yet been modeled).

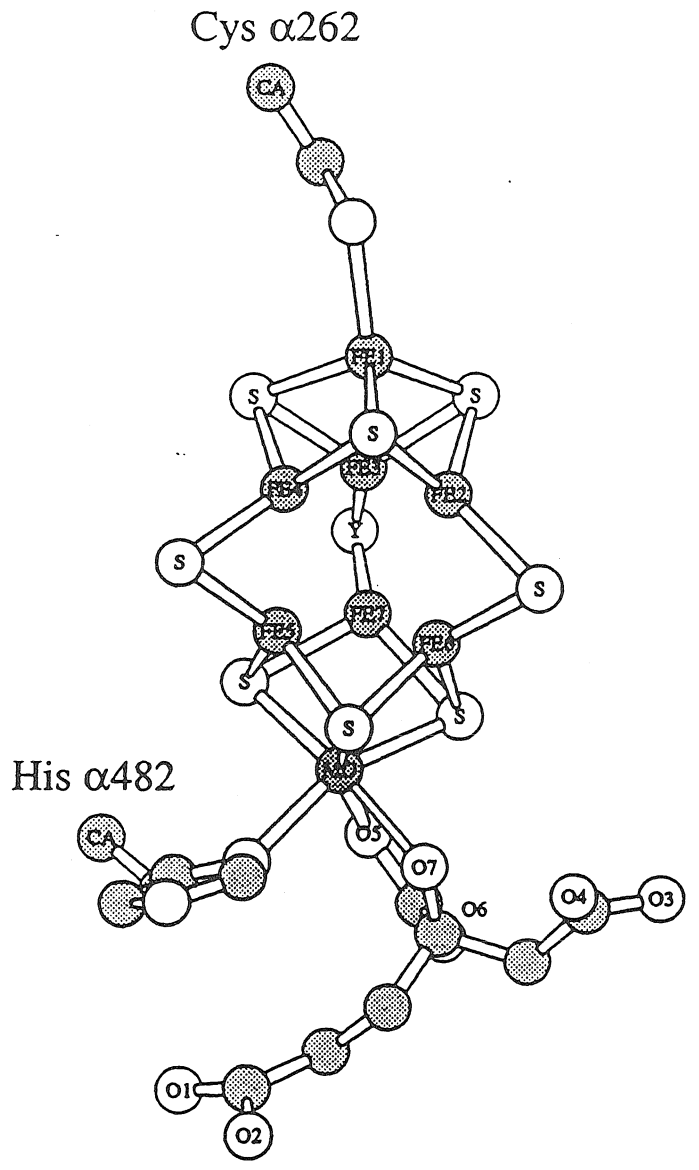


Figure 5-7. Ball and stick model of the FeMo-cofactor with protein ligands. The MOLSCRIPT program[66] was used to draw the model. The "Y" ligand could be S or well-ordered O/N species (less likely).

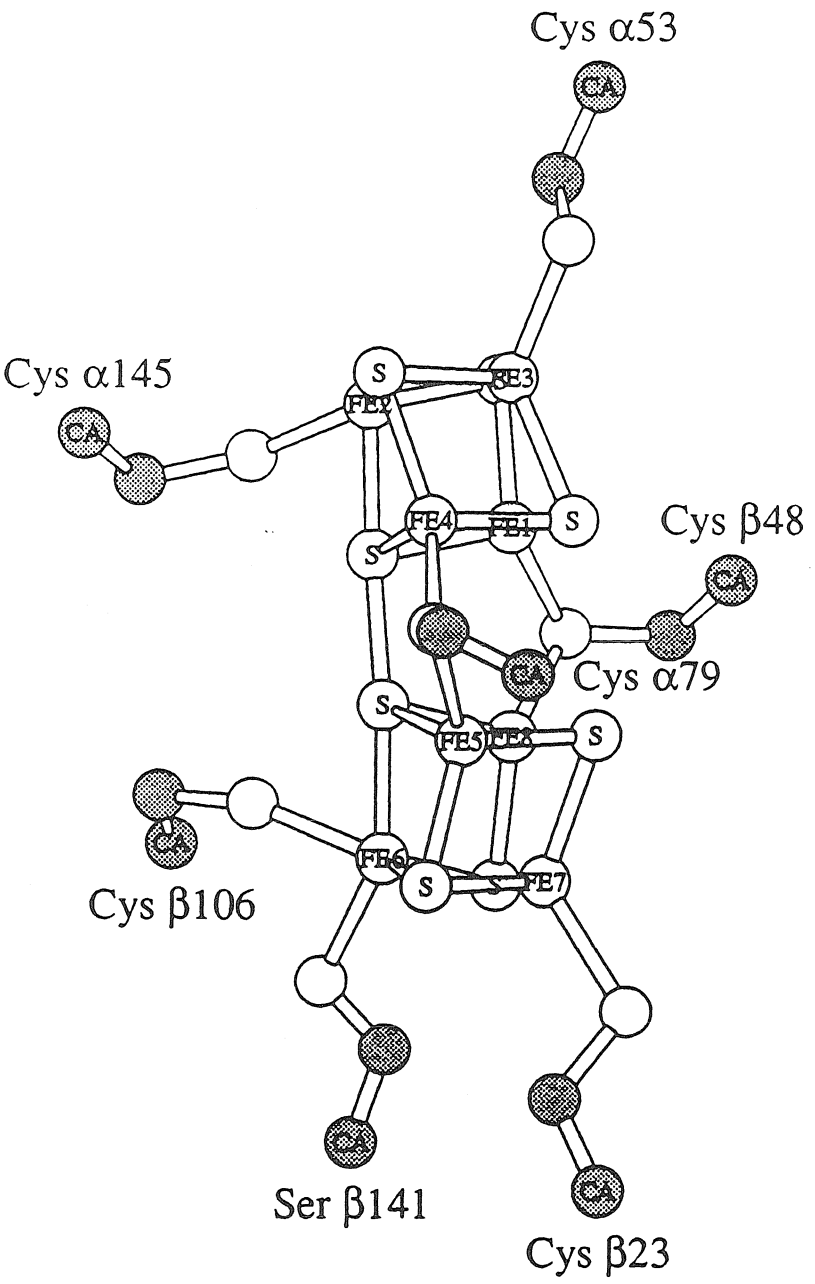


Figure 5-8. Ball and stick model of the P-cluster pair with protein ligands. The MOLSCRIPT program was used to draw the model.

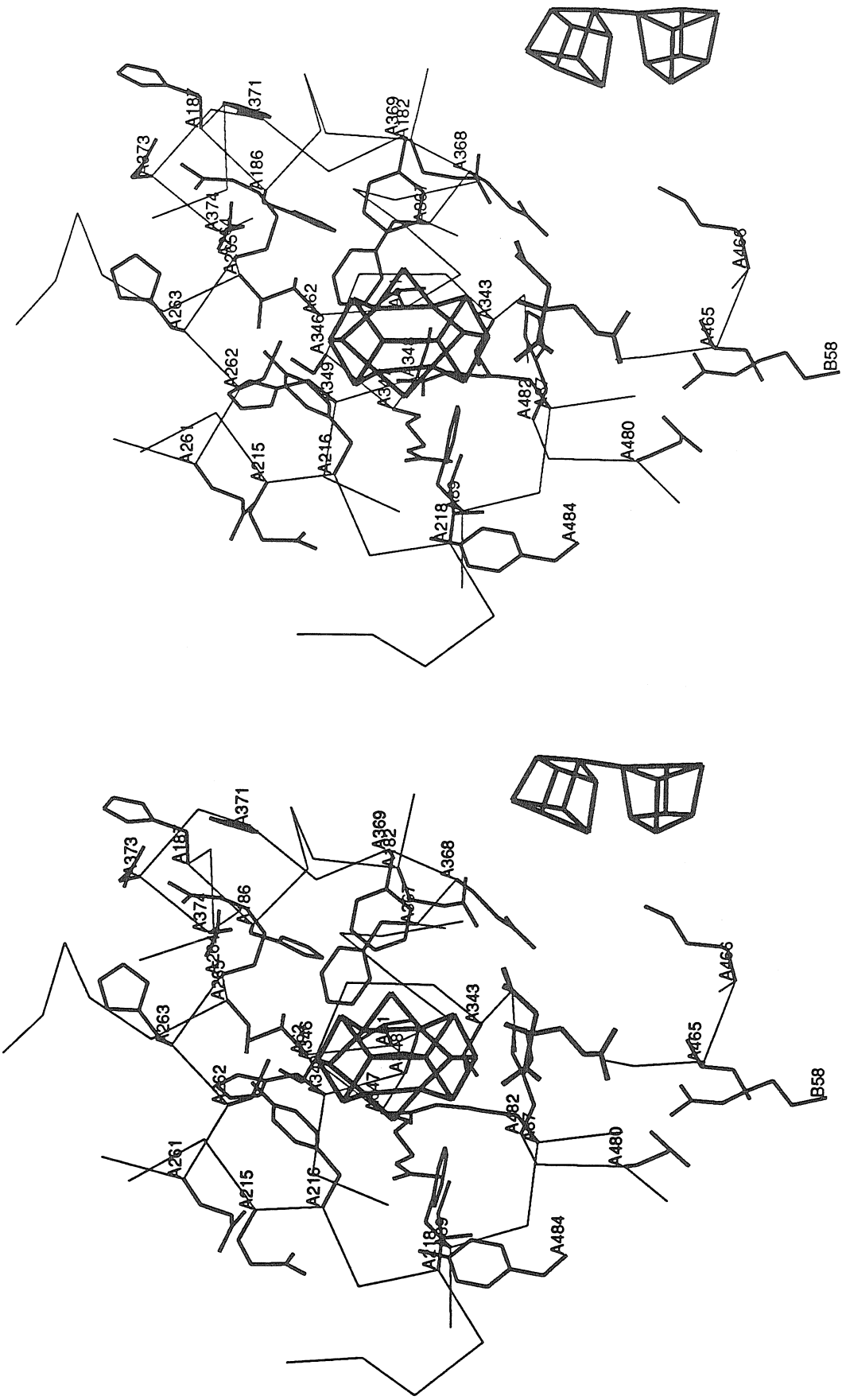


Figure 5-9. Stereoview of the polypeptide environment surrounding the FeMo-cofactor.

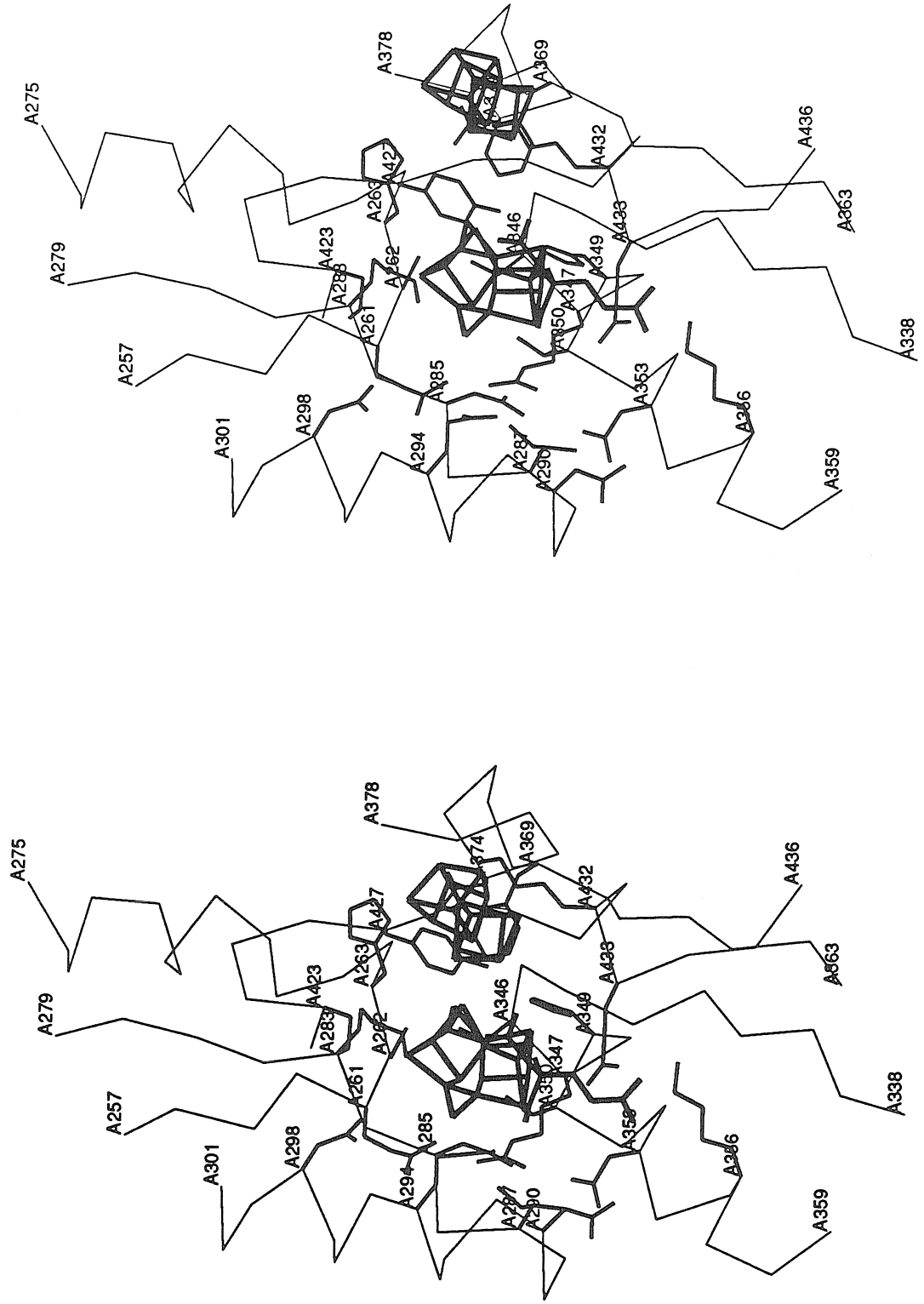


Figure 5-10a. Stereoview of a cleft around the FeMo-cofactor that could be utilized for substrate entry/product release and/or H_3O^+ transport.

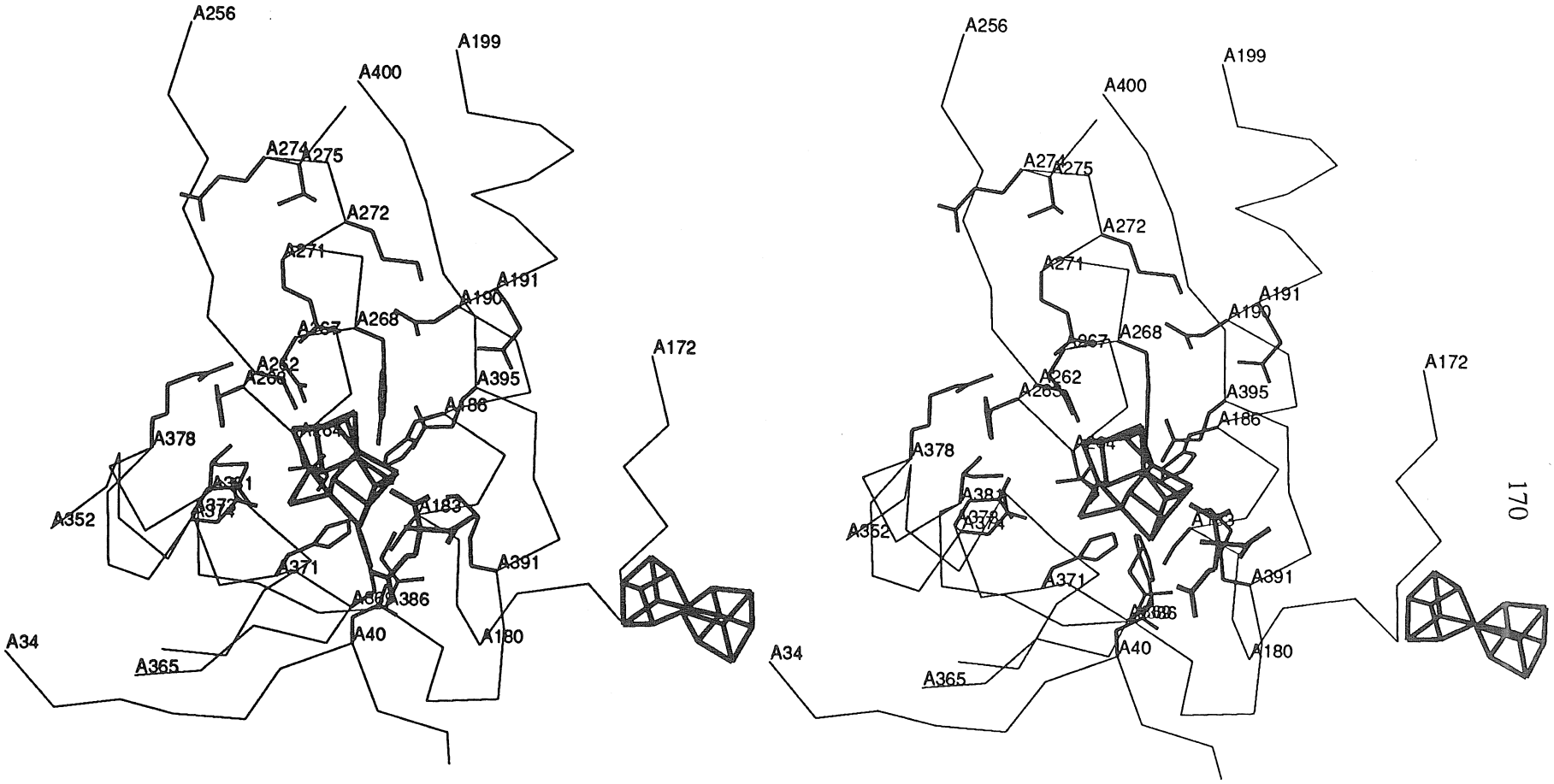


Figure 5-10b. Stereoview of the other cleft around the FeMo-cofactor which could be utilized for substrate entry/product release and/or H_3O^+ transport.

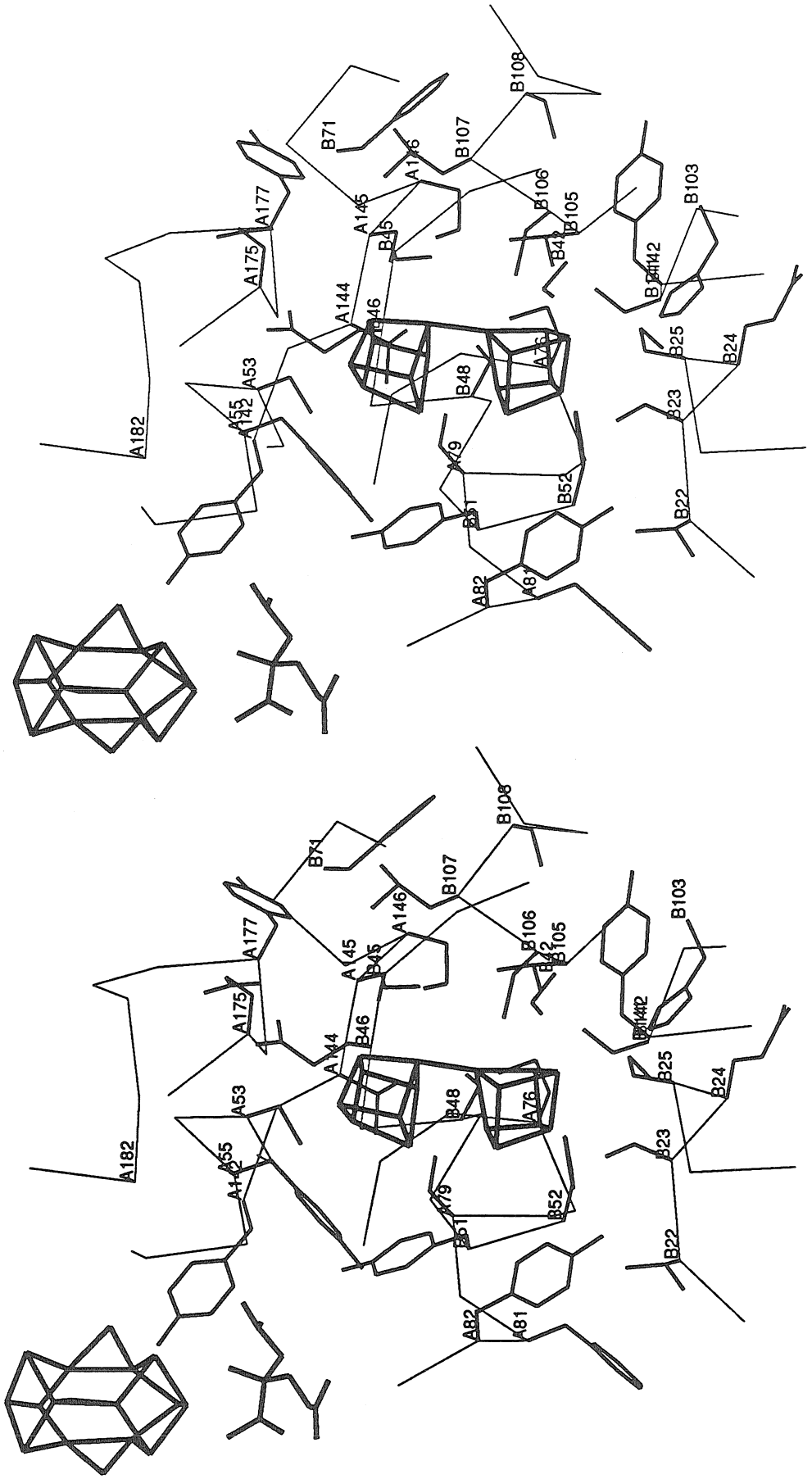


Figure 5-11. Stereoview of the polypeptide environment surrounding the P-cluster pair.

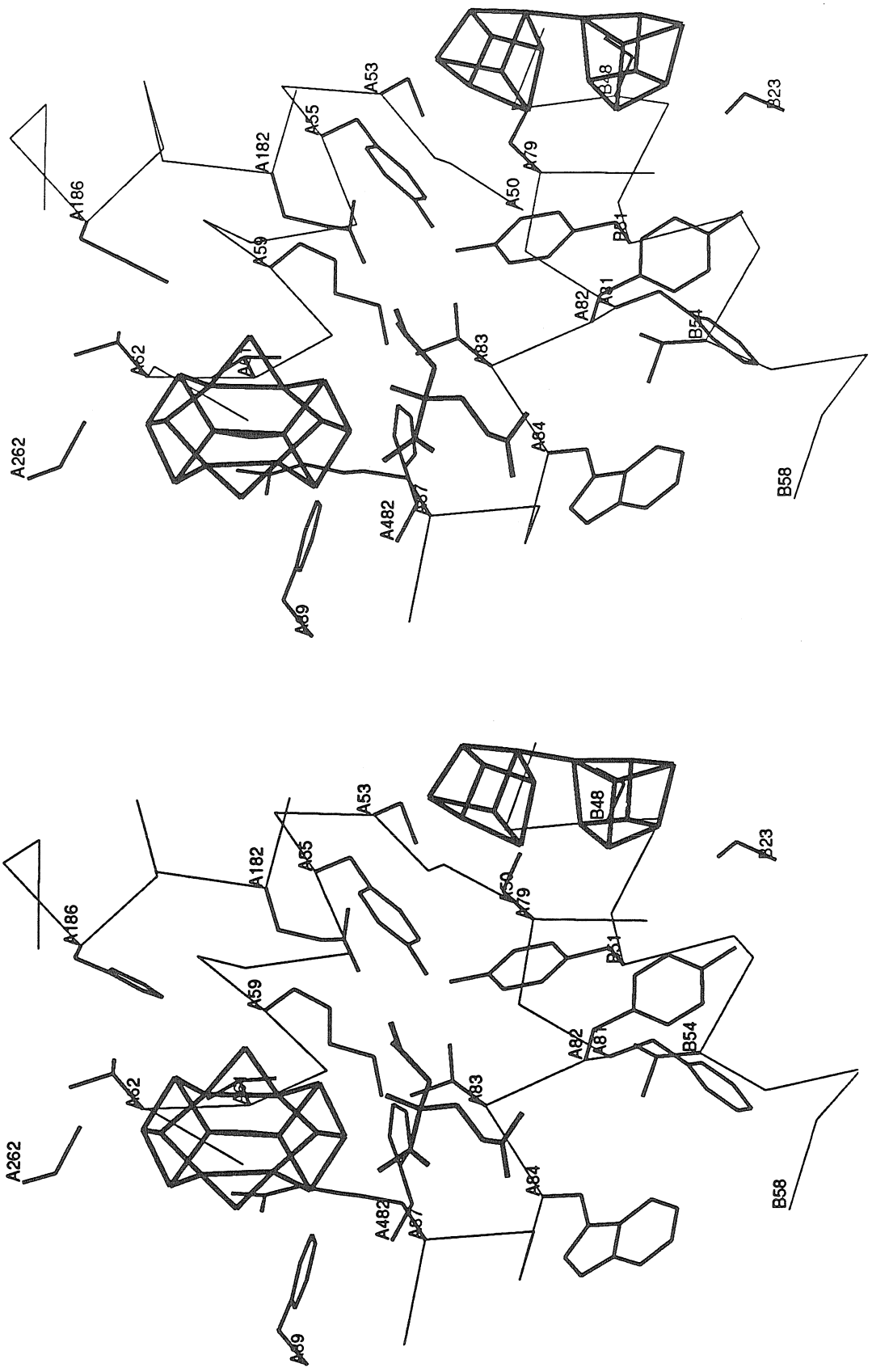


Figure 5-12. Stereoview of the polypeptide environment between the FeMo-cofactor and P-cluster pair.

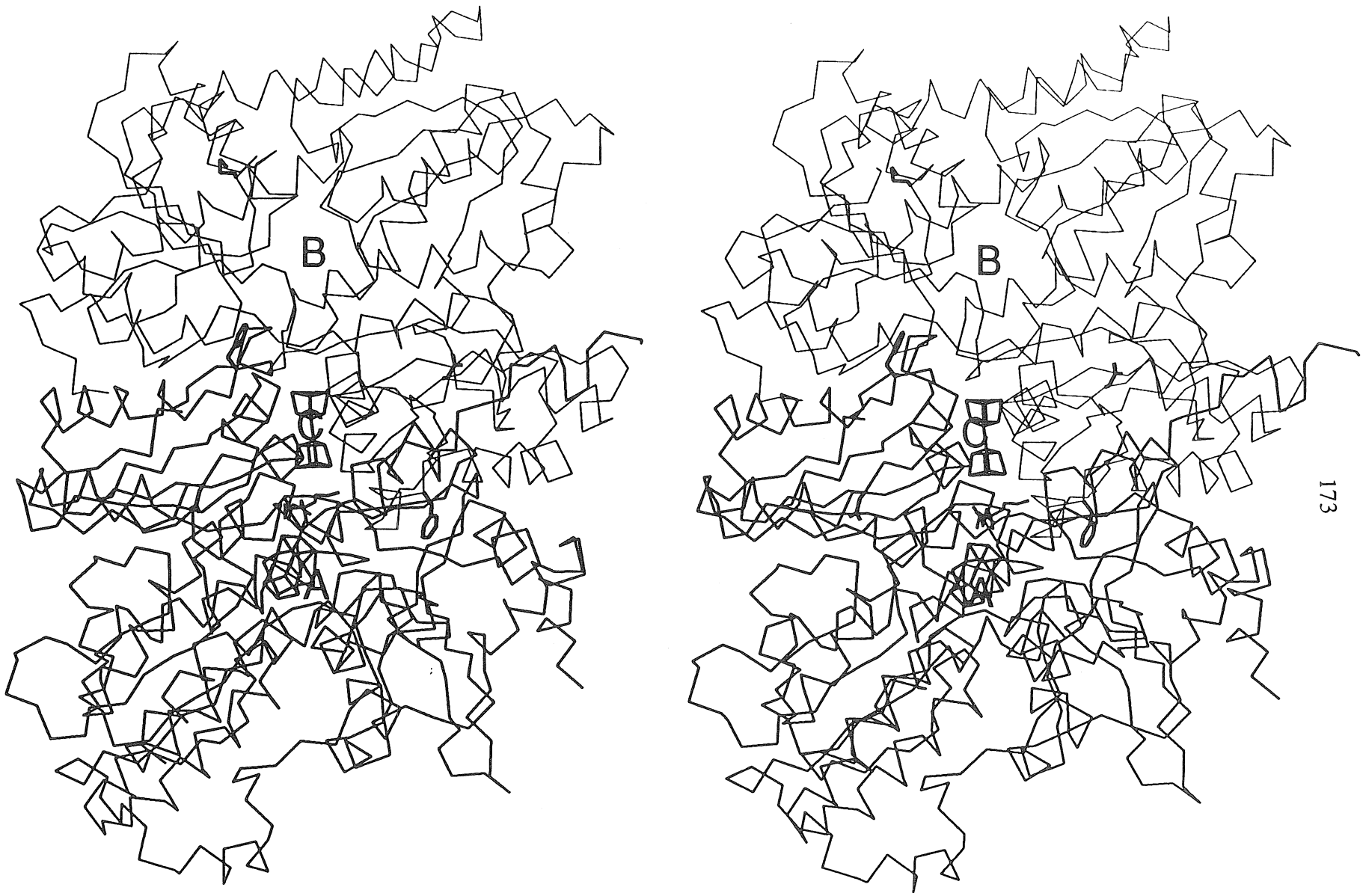


Figure 5-13. Overview of the Fe-protein binding sites. Three potential contact regions are designated A, B and C, and side chains of Phe α117, Phe β78, Asp α153, Asp β114 and Met β348 (Lys β400 in Av1), which correspond to residues implicated in Fe-protein binding in Av1, are also shown. The α and β subunits are represented by the thick and thin lines, respectively.

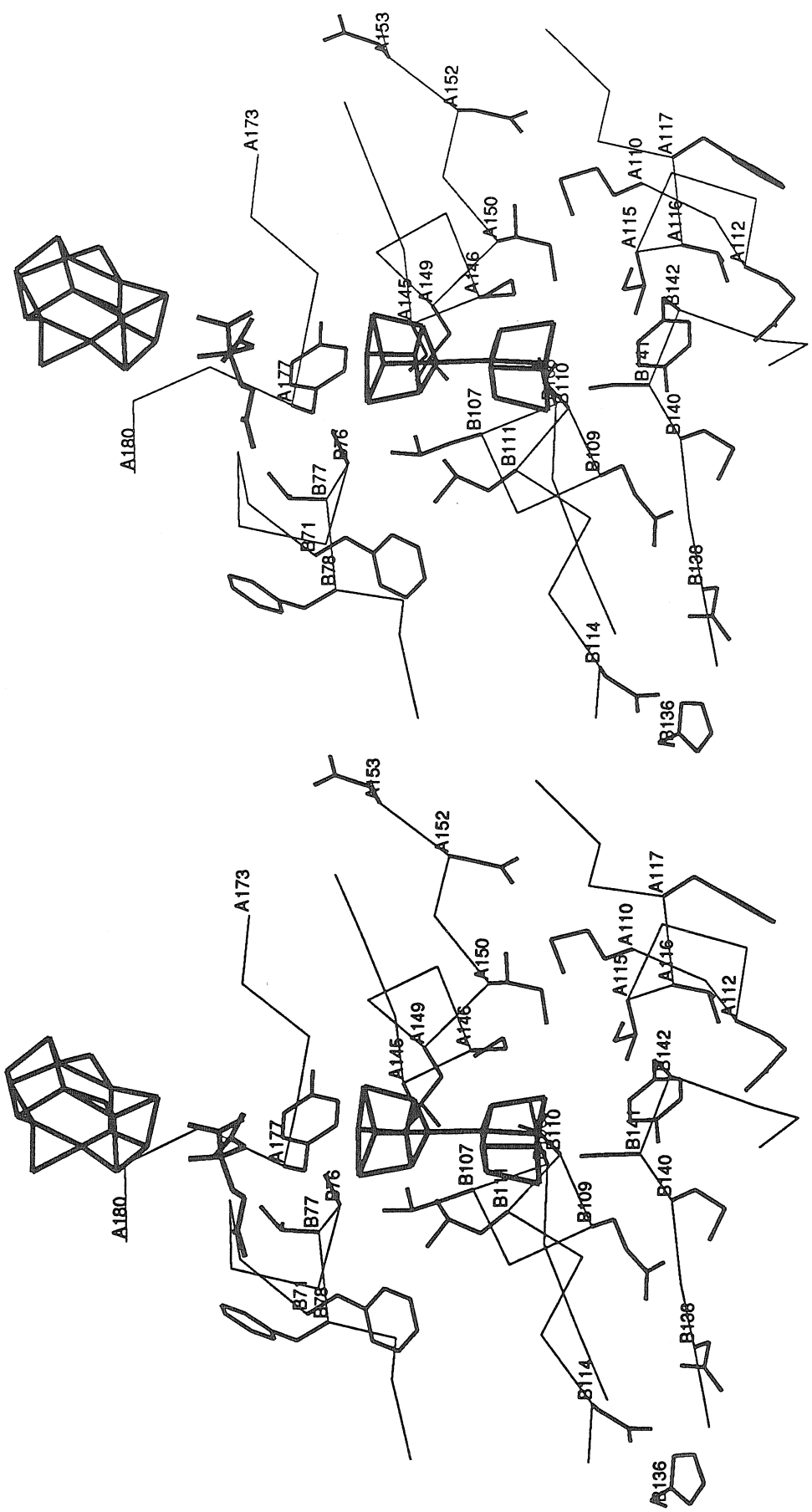


Figure 5-14. Stereoview of the polypeptide environment in the putative Fe-protein cluster binding site (C) near the P-cluster pair.

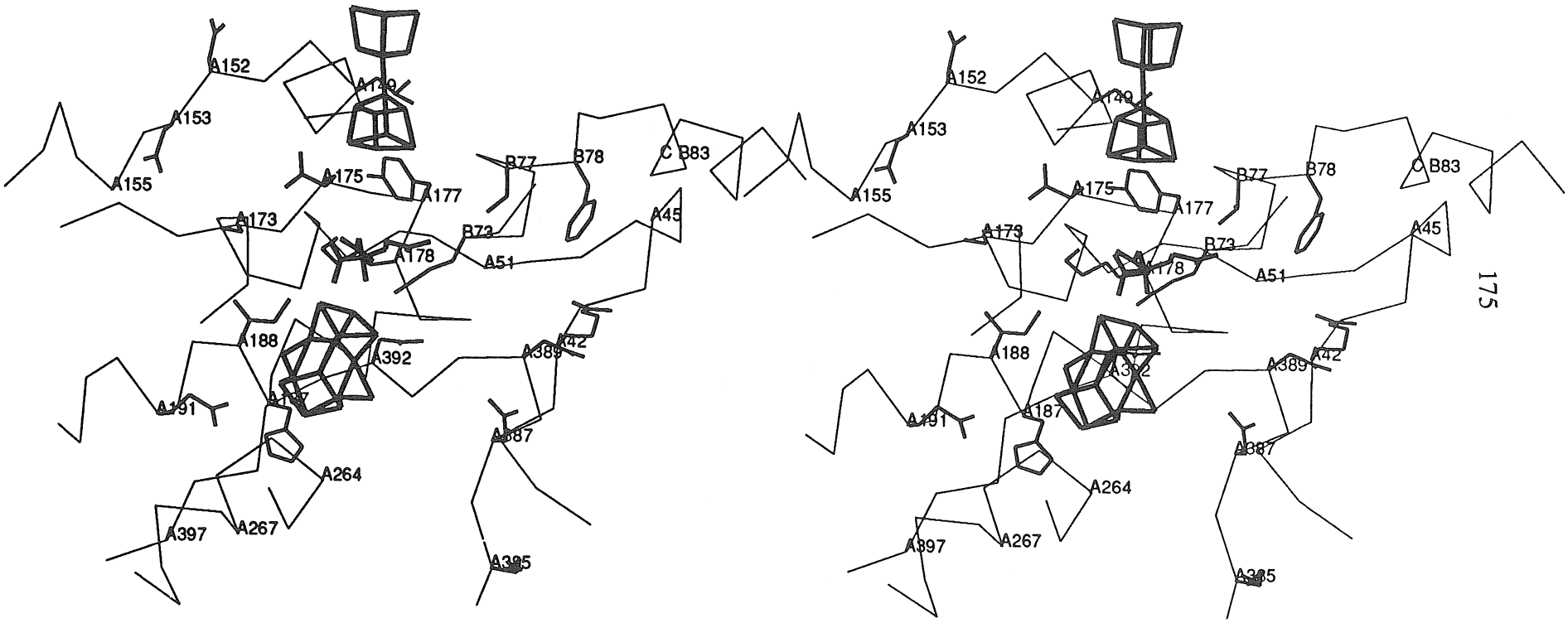


Figure 5-15. Stereoview of the polypeptide environment in the putative Fe-protein subunit binding site (A) on the α subunit, near the FeMo-cofactor.

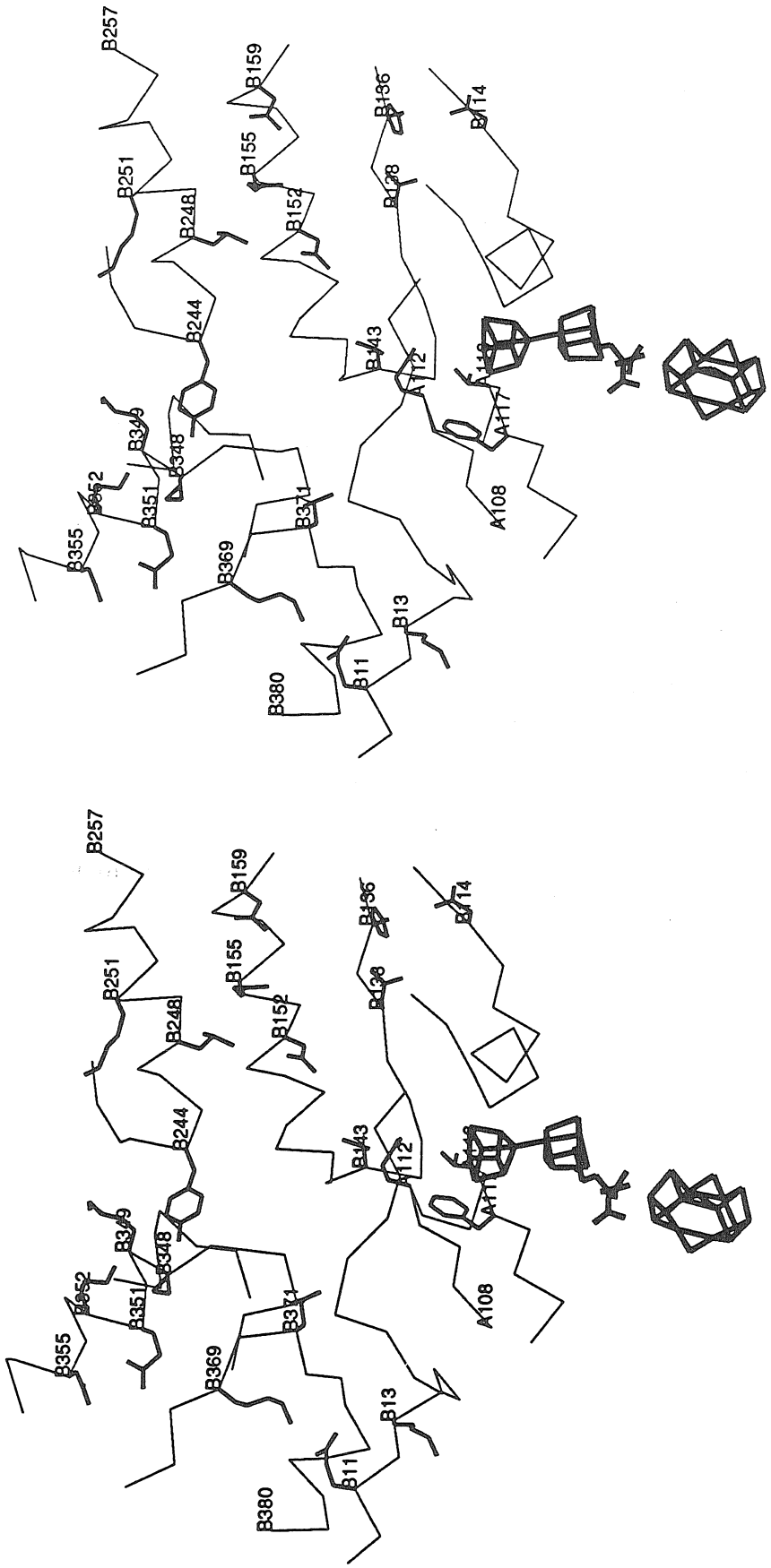


Figure 5-16. Stereoview of the polypeptide environment in the putative Fe-protein subunit binding site (**B**) on the β subunit.

Chapter 6. Similarities with other Electron Transfer Systems

6-1. The H-cluster in Fe-hydrogenases

6-1-1. Introduction

Hydrogenases are a class of enzyme that catalyze the reversible oxidation of molecular hydrogen, $H_2 = H^+ + 2e^-$. The existence of an enzyme capable of using H_2 as a substrate was first demonstrated in *Escherichia coli* by Stephenson and Stickland[1]. Since that time, hydrogenases have been found in a wide variety of bacterial and algal species[2-4]. A comparison of their properties reveals a fairly heterogeneous group of enzymes, that differ in molecular oxidation, electron carrier specificity, cofactor content and sensitivity to inactivation by O_2. However, they have one common feature, all are iron-sulfur proteins, and the majority also contain nickel. Hydrogenases can be divided into two main groups, on the basis of metal content and sequence homology: NiFe-hydrogenases[2,3], which contain nickel and iron, and Fe-hydrogenases[4], which contain only iron.

NiFe-hydrogenases have now been isolated and characterized to varying degrees from well over a dozen different organisms, including photosynthetic bacteria[5], methanogens[6], sulfate-reducing bacteria[7], colon bacteria[8], aerobic hydrogen bacteria[9], aerobic nitrogen fixing bacteria[10], and from an extremely thermophilic archaebacterium[11]. In addition, hydrogenases containing equimolar amounts of selenium and nickel have been purified from various sulfate-reducing bacteria[12]. All NiFe-hydrogenases are comprised of (at least) two dissimilar subunits ($\alpha\beta$) with molecular weights of approximately 60 kD and 30 kD, and the majority are membrane-bound[2,3]. The similarities between the different NiFe-hydrogenases is shown by immunological and DNA sequence analyses, which reveal extensive structural homology between the enzymes of different genera[2,13].

The structure of the Ni center in NiFe-hydrogenases is unknown and may vary in different hydrogenases. X-ray absorption spectroscopy (EXAFS) and EPR studies have suggested that Ni is 5- to 6-coordinate and bound by at least one[14] and more likely three[15] or four[16] sulfur atoms, but not to Fe. Electron spin echo envelope modulation (ESEEM) studies of three hydrogenases[17-19] have all indicated a weak interaction between the Ni center and a ^{14}N nucleus. In one case this was assigned to one N atom of a bound flavin[17], and in the other two to the distal N of a histidine imidazole[18,19]. There is general agreement, however, that the Ni center of the NiFe-hydrogenases is the site of reaction with H_2[20,22]. For example, an Ni(I)-hydride and Ni(I)-CO species, both sensitive to visible light, have been proposed as intermediates in catalysis and the reversible inhibition by CO of *C. vinosum* hydrogenase[22]. The nature of the 4Fe:4S clusters in the NiFe-hydrogenase is not known, since, in contrast to the Ni center, they are not detectable by EPR spectroscopy in most redox states of these enzymes. A recent Mössbauer study of *D. gigas* hydrogenase showed that the two 4Fe-clusters in this enzyme have different redox potentials and have atypical magnetic properties compared to ferredoxin-type 4Fe:4S clusters[23].

In 1984, Peck *et al.*[24] showed that the periplasmic hydrogenase from *D. vulgaris* lacked nickel, and Adams *et al.*[25] demonstrated that the same was true for two hydrogenases (CpI and CpII) from *C. pasteurianum.* All three enzymes appeared to contain approximately equimolar amounts of iron and acid-labile sulfide. The hydrogenase from another strict anaerobes, *Megasphaera elsdenii* (Me)[26] has many properties in common with CpI and CpII[3,25] They are all monomeric, cytoplasmic enzymes with molecular weights of approximately 60 kD, and extremely sensitive to inactivation by O_2, with $t_{1/2}$ values of only minutes in air, and must be purified under strictly anaerobic conditions. In contrast, the hydrogenase from *D. vulgaris* (DvH)[27] is located in the periplasm and can be purified aerobically. However, like the NiFe-hydrogenases, the DvH enzyme in its O_2-stable state is not fully active and requires some form of reductive activation (not just removal of O_2),

otherwise a lag phase is observed in the H_2 oxidation assay[28]. Moreover, the O_2-stable enzyme becomes sensitive to inactivation by O_2 upon reduction by H_2, and the O_2-stable state is regained only by anaerobic oxidation with the dye, 2,6-dichlorophenol, in the presence of Fe and EDTA[29], suggesting that some form of Fe-S cluster interconversion occurs.

Fe-hydrogenases possess two different types of iron clusters, known as "F-" and "H-" clusters. The former are of the ferredoxin 4Fe:4S cluster type[30], whereas the latter, which are thought to be the site of hydrogen interaction, are of unknown structure, and may contain six iron atoms[31]. CpI possesses ~20 Fe atoms per molecule as four F-clusters and one H-cluster, whereas CpII possesses 14 Fe atoms as two different F-clusters (denoted F- and F'-) and one H-cluster[31]. The structure of the H-cluster is of interest not only in connection with the mechanism of the enzyme, but also in that it appears to be spectroscopically unique in bioinorganic chemistry[30,32-35]. A structural model for the H-cluster in Fe-hydrogenases is proposed here, based on the FeMo-cofactor structure of nitrogenase MoFe-protein. The proposed H-cluster model is generally consistent with the experimental results including chemical composition data, amino acid sequence comparison results, ENDOR studies, Mössbauer analysis, resonance Raman spectroscopic studies, and EXAFS data. Two alternative models are also proposed assuming that the chemical composition data , particularly the sulfur composition, is not accurate.

6-1-2. A structural model for the H-cluster in Fe-hydrogenases.

The nitrogenase MoFe-protein has hydrogenase activity. In the absence of other reducible substrates, such as dinitrogen, azides and alkynes, the electron flux through nitrogenase is funneled completely into hydrogen production[36]. Additionally, hydrogen evolution is an obligatory part of nitrogen fixation[37]. Therefore, nitrogenase can be considered as a hydrogenase and it is reasonable to classify hydrogenases into three groups; NiFe-hydrogenases, Fe-hydrogenases, and nitrogenase. It has been postulated

that the nitrogenase MoFe-protein and Fe-hydrogenases may have some common structural motifs in their catalytic metal centers[38,39]. The structures of the metal centers in nitrogenase MoFe-protein have been recently determined[40] and verified by high resolution X-ray diffraction analysis[41]. The FeMo-cofactor, which is believed to be the N_2 binding and reduction site, is composed of 4Fe:3S and 1Mo:3Fe:3S clusters that are bridged by three non-protein ligands (Fig. 6-1). The Fe-Fe distance between bridged iron sites is ~2.5Å suggesting that there may be some iron-iron bonding interactions[41,42], which could contribute the fourth coordination interaction for the bridging irons. Ignoring the partial iron-iron bonding interactions between bridged irons, six of the seven Fe atoms in the FeMo-cofactor have trigonal coordination geometry and are therefore coordinatively unsaturated and potential sites for N_2 activation[40].

From the FeMo-cofactor structure, one can conceive a two-dimensional analog of the FeMo-cofactor model. If the two irons (Fe3 and Fe7), two sulfurs (S2B and S3A) and one bridging ligand ("Y") which protrude from the hypothetical central plane are deleted, a 6Fe:6S cluster model (Figure 6-2a) is obtained. This 6Fe:6S cluster model could be a potential structure of the H-cluster in Fe-hydrogenases, based on the assumption that nitrogenase MoFe-protein and Fe-hydrogenases may have some common structural motifs in their catalytic metal centers[38,39]. The proposed H-cluster model contains two 3Fe:2S clusters that are bridged by two sulfides. There may be some iron-iron bonding interactions that could provide the third and fourth coordination interaction for the four central irons which are bridged by two sulfides. The four central irons form an iron surface which is adequate for the H_2 binding as well as CO (a potent inhibitor of hydrogenase) binding. One possible H_2 binding mode on this cluster is proposed in Figure 6-2b. The central four irons, which are not saturated by ligands, are proposed to be coordinated by H_2, thereby promoting H_2 bond cleavage and formation depending on the physiological electron supply and demand. The weak Fe--Fe bonding interaction in the H-cluster model might be important for H_2 binding to this cluster in a manner analogous to the

N_2 binding to the FeMo-cofactor of nitrogenase MoFe-protein[41,42]. The central four irons are drawn on the same plane in Figure 6-2; however, twisting of four iron atoms out of plane is also conceivable, especially upon oxidation/reduction of the H-cluster as has been indicated by EXAFS studies[39] (see below). The other two irons, the terminal irons of the H-cluster model, are further coordinated by two protein ligands respectively and have a tetrahedral coordination environment.

6-1-3. Experimental evidence.

Chemical composition data. A consensus has not been reached as to the composition of the H-cluster, but comparison of values from different Fe-hydrogenases and the Mössbauer and ENDOR studies of the Fe-hydrogenases have indicated that the H-cluster is probably composed of ~6Fe and ~6 acid-labile sulfur atoms[4,31]. The chemical composition of the non-protein part of the proposed H-cluster model, 6Fe:6S, is within the range of values that have been reported, although this information was used in constructing the model.

ENDOR. ENDOR study of ^{57}Fe enriched CpI[43] has indicated that the H-cluster is comprised of two magnetically distinct types of Fe atoms. The unique nature of this center was shown by its ^{57}Fe hyperfine coupling constants (A1=17 MHz, A2=9.5 MHz), which are substantially smaller than those observed for conventional 2Fe-, 3Fe- or 4Fe-clusters. More detailed ENDOR investigations[34,35] of ^{13}CO treated CpI and CpII, and ^{1}H-ENDOR study show that the H-cluster covalently binds a single CO molecule and this binding causes almost identical changes in the magnetic properties of the two types of Fe atoms in the H-cluster.

The H-cluster model proposed here contains two different types of Fe sites; the terminal irons, Fe1 and Fe6, and the central irons, Fe2, Fe3, Fe4 and Fe6. A carbon monoxide (CO) could bind in an end-on fashion to the center of the cluster. Therefore CO binding does not seem to make any change in the symmetry of the H-cluster, and it may cause

almost identical changes in the magnetic properties of the two types of Fe atoms in the H-cluster.

Mössbauer. The Mössbauer analyses[33] of CpII with and without CO also showed that two types of Fe atoms are present in the H-cluster. In the reduced state, the cluster exhibits, in zero field, spectra with the typical 2:1 quadruple pattern of reduced 3Fe:4S clusters. However, whereas the latter are paramagnetic (S=2), the H-cluster is diamagnetic (S=0). The data obtained for the oxidized H-cluster[33] suggest a novel type of 3Fe cluster and bear little resemblance to those reported for oxidized 3Fe:4S clusters. The data for the oxidized CpII treated with CO also suggests that the H-cluster contains one or more 3Fe motif[33].

The H-cluster model proposed here consists of two novel 3Fe:2S clusters that are bridged by two sulfides. The 6Fe atoms in the H-cluster model can be grouped into two classes with a ratio of 2:1 (Fe2, Fe3, Fe4 and Fe5; and Fe1 and Fe6) in agreement with the Mössbauer observations.

Resonance Raman Spectroscopy. Resonance Raman spectroscopic study[32] on CpI has indicated that the H-cluster might have a 2Fe:2S cluster structural motif. CpI exhibit strong bands near 280 and 390 cm^{-1}, which are seen from 2Fe:2S clusters but from no other known Fe-S cluster type. Analogous bands are extremely weak, if present at all, in CpII. The resonance Raman spectra are dependent on the visible absorption of the cluster, and are very dependent on the number, type, and geometry of the bridging and terminal ligands. That is, the data would be expected to indicate structural features that comprise the cluster rather than cluster itself. An arrangement of Fe and S atoms like those in a 2Fe:2S cluster is therefore a possible component of the H-cluster of CpI[4].

The two Fe atoms in a 2Fe:2S cluster are coordinated by two non-protein sulfide ligands and two cysteine thiol ligands[66]. The Fe1 and Fe6 atoms in the H-cluster model are coordinated by two non-protein sulfide ligands and two cysteine thiol ligands. Therefore it

may be considered that the H-cluster model proposed here contains the 2Fe:2S structural motif.

EXAFS. Only one of the Fe-hydrogenases has been examined by this technique so far, but the unique properties of the H-cluster were apparent[39]. The oxidized CpII exhibited two major peaks, which from simulations were assigned to Fe-S and Fe-Fe interactions at distances of 2.27Å and 2.76Å, respectively, values typical of those of 2Fe-, 3Fe- and 4Fe-clusters. However, reduced CpII gave rise to three major peaks in the transform. Two corresponded to the Fe-S and Fe-Fe interactions seen with the oxidized enzyme, with similar distances, while simulations showed that the additional peak not seen in the oxidized enzyme could be represented by an Fe-Fe interaction at 3.3 Å.

The H-cluster model consists of standard Fe-S and Fe-Fe bonds except for the Fe-Fe bonds for the four central irons. If the bridged Fe-Fe distance (such as Fe2-Fe3, Fe2-Fe5 etc.) is 2.5Å as it is in the FeMo-cofactor[41,42], then the diagonal Fe-Fe distance (such as Fe2-Fe4 and Fe3-Fe5) is ~3.5Å. However, the iron-iron distance may be less than 3.5Å if the central four irons are twisted out of the plane rather than being perfectly planar.

Genetic Studies. Comparisons of the available Fe-hydroganase sequences show that these enzymes constitute a structurally homogeneous family[44]. While they differ in the length of their N-termini and in the number of their 4Fe:4S clusters (F-clusters), they are highly similar in their C-terminal halves, which are postulated to be the H-cluster domain (figure 6-3). Four conserved cysteine residues located in this domain (Cys 300 or Cys 301, Cys 356, Cys 500 and Cys 504 of CpI) are likely ligands of the H-cluster. The sequence comparisons[44] indicate that the H-cluster most probably possesses a common structural framework in all Fe-hydrogenases.

The H-cluster model requires at least four protein ligands for the terminal iron sites (Figure 6-2). The other iron atoms, which form the H_2 binding site, may or may not interact with additional protein ligands.

Model study. A 6Fe:9S cluster which has a similar topology to the proposed H-cluster model has been synthesized and characterized[45] (Figure 6-4). Based on the 6Fe:9S cluster structure and the proposed H-cluster model in Figure 2, two alternative 6Fe:8S and 6Fe:7S H-cluster models are conceivable (Figure 6-5), assuming that the chemical composition data, in particular the sulfur content, of the H-cluster is not accurate. The alternative models are also generally consistent with the experimental data mentioned above.

6-2. Manganese-center in PSII

Photosynthetic reaction centers catalyze the conversion of light into chemical energy. In plants and cyanobacteria, part of this reaction involves the photosynthetic oxidation of water in a four-electron transfer process (equation 1) by the water splitting enzyme in PSII[46].

$$2H_2O \rightarrow O_2 + 4H^+ + 4e^- \qquad (1)$$

The enzyme is located in the thylakoid membrane, and during turnover it donates electrons to PSII, which are then recovered by the oxidation of H_2O to liberate oxygen[47]. The enzyme cycles through five different oxidation levels, known as the S states, S_0 through S_4, which are thought to be reflections of different oxidation levels of a multi-nuclear manganese center[47]. The oxidation of water is not well understood, although it is thought to involve a manganese-center as the active site. Elemental analysis indicates that four manganese atoms are required per reaction center[48], and activation of the enzyme also requires Ca^{2+} and Cl^- ions[49,50]. Despite extensive studies by X-ray[51-53], electron paramagnetic resonance (EPR)[54], and optical spectroscopies[55], the structure of the manganese center in PSII remains unknown.

X-ray absorption studies have provided structural information on the Mn-center. Klein and coworkers have made extensive EXAFS studies on the S_1 and S_2 states of PSII prepared from spinach at manganese concentrations of 500-800 μM. These studies concluded that the average manganese coordination environment was the same for both

states, despite the change in oxidation state of the manganese, and contained a short Mn-O/N distance at 1.75Å, a Mn-O/N shell at 2.0Å and at least one Mn scatterer at 2.7Å[56,57]. Corrie *et al.*[58] found similar results for the S_1 state. This picture has been questioned by George *et al.*[59] and subsequently by Penner-Hahn *et al.*[60], who have obtained S_1 data on more concentrated samples. Both groups confirm the presence of an Mn-Mn scatterer at 2.7Å, but find no evidence for a shell of low-Z scatterers at 1.75-1.85Å. Additionally both groups report the existence of further shells at longer distances. George *et al.*[59], using oriented chloroplasts, resolved a second Mn-Mn shell at 3.3Å whose vector is oriented perpendicular to the membrane plane, and Penner-Hahn *et al.*[60] report a Mn or Ca shell at 3.3Å as well as a possible shell of scatterer at 4.2Å. Recently MacLachlan *et al.*[61] have detected Mn-Mn interaction at 2.7Å and Mn-Mn or Mn-Ca interaction at 3.7Å for the S_1 and S_2 states, but have not observed a Mn-Mn shell at 3.3Å.

It is intriguing that the EXAFS results of the Mn-center in PSII, the H-cluster in Fe-hydrogenase and the FeMo-cofactor in nitrogenase are very similar to each other. In the Mn-center, Mn-Mn interactions at 2.7Å and 3.3Å (or 3.7Å) have been observed by EXAFS studies[56-61]; in the H-cluster, Fe-Fe interactions at 2.7Å and 3.3Å have been observed by EXAFS[39]; and in the FeMo-cofactor, Fe-Fe interactions at 2.7Å and 3.8Å have been observed by EXAFS[62] and the Fe-Fe distances at 2.5Å, 2.7Å and 3.8Å have been determined by x-ray crystallography (Chapter 3). These findings may suggest that those proteins share some common structural motifs in their catalytic metal centers. This idea is also supported by the similarity of reactions that are catalyzed by the three proteins.

$N_2 + 6H^+ + 6e^- \rightarrow 2NH_3$ (Nitrogenase)

$O_2 + 4H^+ + 4e^- \leftarrow 2H_2O$ (PSII)

$H_2 \Leftrightarrow 2H^+ + 2e^-$ (Hydrogenase)

It is proposed that the common structural feature in their catalytic metal centers is a metal surface and the metal surface, which consists of four metal atoms, is adequate for diatomic gas (N_2, O_2 and H_2) binding and activation.

A structural model for the Mn-center is proposed in Figure 6-6a based on the structure of the FeMo-cofactor (Figure 6-1) and the proposed structural model of the H-cluster (Figure 6-2). The proposed Mn-center model contains a 4Mn:4O cluster as the basic structural element, however, additional ligands to the Mn sites could be oxygen as well as protein ligands. This proposed model focuses on the general shape of the Mn-center and the arrangement of four manganese atoms in the Mn-center. There might be some Mn--Mn bonding interactions that could provide the fifth and sixth coordination interactions for the four manganese atoms. The stereochemistry of the Mn-center is not considered in this model; the four manganeses are drawn on the same plane in Figure 6-6, however, twisting of four manganese atoms out of plane is also conceivable. The four manganese atoms form a metal surface which could function in O_2 binding (or binding of two water molecules under physiological conditions). One possible O_2 binding mode on this cluster is proposed in Figure 6-6b. The four manganese atoms, which are not saturated by ligands, could form σ-bonds with 4 p-orbitals of O_2, yielding a μ_4-η_2-O_2 complex, thereby promoting oxygen-oxygen bond breakage and formation.

The proposed Mn-center model is generally consistent with EXAFS results[56-61]. If the adjacent Mn-Mn distance (such as Mn1-Mn2, Mn1-Mn4, etc.) is 2.7Å, then the diagonal Mn-Mn distance (such as Mn1-Mn3 and Mn2-Mn4) can have values ranging between 2.7Å and 3.8Å depending on the torsion angle between the vectors, Mn1-Mn2 and Mn3-Mn4. If the torsion angle is 0°, then the four manganese atoms are on the same plane and the diagonal Mn-Mn distance is ~3.8Å. This model is consistent with the MacLachlan *et al.*'s EXAFS results[61]. If the torsion angle is ~45°, then the diagonal Mn-Mn distance becomes ~3.3Å. This model is consistent with the George *et al.*'s and Penner-Hahn *et al.*'s EXAFS results[59,60] (Figure 6-7). If the torsion angle is 90°, then the four manganese atoms have

tetrahedral geometry and the diagonal Mn-Mn distance becomes ~2.7Å (similar to the cubane model).

6-3. Bacterial photosynthetic reaction center

Although the electron transfer processes catalyzed by nitrogenase and the photosynthetic reaction center (RC, reviewed in 63) are quite different, there are striking similarities in the structural organization of the two systems: (1) Both the MoFe-protein and the RC are composed of two homologous subunits approximately related by a twofold rotation: the α and β subunits of the MoFe-protein, and the L and M subunits of the reaction center[64,65]. The initial electron carriers for both systems - the P-cluster pair in MoFe-protein and the special pair in RC, are buried in the interface between the two subunits. The location of these redox centers at the subunit interface may provide a convenient assembly mechanism for the incorporation of powerful reductants in the protein interior, isolated from contact with solvent. (2) Despite the general twofold symmetry in the protein organization, electron transfer from the initial donor proceeds in only one direction. In the reaction center case, only one of the two electron transfer branches present is actually utilized. In the MoFe-protein, however, the FeMo-cofactor (the final electron acceptor) is present in only one of the two homologous subunits, so that electron transfer necessarily proceeds in this direction. Whether the original MoFe-protein was a homodimer or tetramer containing equivalent branches is an interesting question in molecular evolution. (3) The binding sites for the terminal electron acceptors (the P-cluster pair in MoFe-protein and the special pair in RC) are buried in both the MoFe-protein and RC, so that the protein structures must accommodate both entry and exit of the relevant groups. (4) In the RC, it has been established that separate pathways exist for electron transfer and proton transfer within the protein, and it seems quite likely that the MoFe-protein behaves similarly. (5) A significant fraction of the energy input into these systems (light for the RC and MgATP for nitrogenase) is lost during the overall reaction. This apparently is a consequence of

ensuring that the electron transfer reactions are essentially irreversible, so that short-circuits or futile cycles are minimized.

References

1. Stephenson, M. & Stickland, L. H. (1931) *Biochem. J.* **25**, 205-214.

2. Fauque, G., Peck, H. D., Jr., Moura, J. J. G., Hyunh, B. H., Berlier, Y., Dervartanian, D. V., Texeira, M., Przybyla, A.E., Lespinat, P. A., Moura, I., & Legall, J. (1988) *FEMS Microbiol. Rev.* **54**, 299-344.

3. Voordouw, G. (1992) *Advances in Inorg. Chem.* **38**, 397-422.

4. Adams, M. W. W. (1990) *Biochim. Biophys. Acta* **1020**, 115-145.

5. Colbeau, A. & Vignais, P. M. (1983)*Biochem. Biophys. Acta* **748** 128-138; Albracht, S. P. J., Kalkman, M. L. & Slater, E. C. (1983) *Biochem. Biophys. Acta* **724** 309-316; Gegotov, I. N. (1984) *Arch. Microbiol.* **140**, 86-90.

6. Kojima, N. *et al.* (1983) *Proc. Natl. Acad. Sci. U.S.A* **80**, 378-382; Muth, E., Morshe, E. & Klein, A. (1987) *Eur. J. Biochem.* **169**, 571-577; Jin, S. L. C., Blanchard, D. K. & Chen, J. -S. (1983) *Biochim. Biophys. Acta* **748**, 8-20; Fauque, G., *et al.* (1984) *Eur. J. Biochem.* **142**, 21-28; Fiebig, K. & Friedrich, B. (1989) *Eur. J. Biochem.* **184**, 79-88; Baron, S. F. & ferry, J. G. (1989) *J. Bacteriol.* **171**, 3846-3853.

7. LeGall, J., *et al.* (1982) *Biochem. Biophys. Res. Commum.* **106**, 620-626; Cammack, R., *et al.* (1982) *FEBS Lett.* **142**, 289-292; Kruger, H. J., *et al.* (1982) *J. Biol. Chem.* **257**, 14620-14622; Czechowski, M. H., *et al.* (1985) *Biochem. Biophys. Res. Commum.* **125**, 1025-1032; Lissolo, T., *et al.* (1986) *Biochem. Biophys. Res. Commum.* **139**, 701-708.

8. Unden, G., *et al.* (1982) FEBS Lett. 145, 230-234; Ballantine, S. P. & Boxer, D. H. (1985) *J. Bacteriol.* **163**, 454-459.

9. Friedrich, C. G., Schneider, K. & Friedrich, B. (1982) *J. Bacteriol.* **152**, 42-48; Schneider, K., Patil, D. S. & Cammack, R. (1983) *Biochim. Biophys. Acta* **748**, 353-361; Schneider, K., Cammack, R. & Schlegel, H. G. (1984) *Eur. J. Biochem.* **142**, 75-84; Knuttel, k., Schneider, K., Schlegel, H. G.& Müller, A. (1989) *Eur. J. Biochem.* **179**, 101-108.

10. Harker, A. R., *et al.* (1984) *J. Bacteriol.* **159**, 850-860; Arp, D. (1985) *Arch. Biochem. Biophys.* **237**, 504-512; Seefeldt, L. C. & Arp, D. (1986) *Biochimie* **68**, 25-34; Stults, L. W., O'hara, E. & Maier, R. J. (1984) *J. Bacteriol.* **159**, 153-158.

11. Bryant, F. O. & Adams, M. W. W. (1989) *J. Biol. Chem.* **264**, 5070, 5079.

12. Reider, R., Cammack, R. & Hall, D. O. (1984) *Eur. J. Biochem.* **145**, 637-643; Teixeira, M., *et al.* (1986) *Biochime* **68**, 75-84.

13. Kovacs, K. L., *et al.* (1989) *J. Bacteriol.* **171**, 430-435.

14. Cammack, R., Fernandez, V. M., Schneider, K. (1988) in Bioinorganic Chemistry of Nickel (Lancaster, J. R., Jr., ed.), pp 167-190, VCH publishers, Deerfield Beach.

15. Lindahl, P. A., *et al.* (1984) *J. Amer. Chem. Soc.* **106**, 3062, 3064.

16. Scott, R. A. *et al.* (1984) *J. Amer. Chem. Soc.* **106**, 6864-6865.

17. Tan, S. L., *et al.* (1984) *J. Amer. Chem. Soc.* **106**, 3064-3066.

18. Chapman, A. *et al.* (1988) *FEBS Lett.* **242**, 134-138.

19. Cammack, R., *et al.* (1989) *Eur. J. Biochem.* **182**, 363-366.

20. Teixeira, M., *et al.* (1985) *J. Biol. Chem.* **260**, 8942-8950.

21. Cammack, R., *et al.* (1987) *Biochim. Biophys. Acta.* **912**, 98-109.

22. Van der Zwan, J. W., *et al.* (1986) *Biochim. Biophys. Acta* **872**, 208-215.

23. Teixeira, M., *et al.* (1989) *J. Biol. Chem.* **264**, 16435-16450.

24. Huynh, B. H., *et al.* (1984) *Proc. Natl. Acad. Sci. U.S.A* . **81**, 3728-3732.

25. Adams, M. W. W. & Mortenson, L. E. (1984) *J. Biol. Chem.* **259**, 7045-7055.

26. Grande, H. J., *et al.* (1983) *Eur. J. Biochem.* **136**, 201-207.

27. Van der Westen, H. M., Mayhew, S. G. & Veeger, C. (1978) *FEBS Lett.* **86**, 122-126.

28. Van der Dijk, C., Van Berkel-Arts, A. & Veeger, C. (1983) *FEBS Lett.* **156**, 340-344.

29. Van der Westen, H. M., Mayhew, S. G. & Veeger, C. (1980) *FEMS Microbiol. Lett.* **7**, 35-39

30. Adams, M. W. W., Johnson, M. K., Zambrano, I. C. & Mortenson, L. E. (1986) *Biochemie* **68**, 35-41.

31. Adams, M. W. W., Eccleston, E. & Howard, J. B. (1989) *Proc. Natl. Acad. U.S.A.* **86**, 4932-4936.

32. Macor, K. A., Czernuszewicz, R. S., Adams, M. W. W. & Siro, J. G. (1987) *J. Biol. Chem.* **262**, 9945-9947.

33. Rusnak, F. M., Adams, M. W. W, Mortenson, L. E. & Münck, E. (1987) *J. Biol. Chem.* **262**, 38-41.

34. Telser, J., Benecky, M. J., Adams, M. W. W., Mortenon, L. E. & Hoffman, B. M. (1986) *J. Biol. Chem.* **261**, 13536-13541.

35. Telser, J., Benecky, M. J., Adams, M. W. W., Mortenson, L. E. & Hoffman, B. M. (1987) *J. Biol. Chem.* **262**, 6589-6594.

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

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

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

39. George, G. N., Prince, R. C., Stockley, K. E. & Adams, M. W. W. (1989) *Biochem. J.* **259**, 597-600.

40. Kim, J. & Rees, D. C. (1992) *Science* **257**, 1677-1681.

41. Chan, M. K., Kim, J. & Rees, D. C. (1993) *Science*, in press.

42. Kim, J., Woo, D. & Rees, D. C. (1992) *Biochemistry*, in press.

43. Wang, G., *et al.* (1984) *J. Biol. Chem.* **259**, 14328-14331.

44. Meyer, J. & Gagnon, J. (1991) *Biochemistry* **30**, 9697-9704.

45. Henkel, G., Strasdeit, H. & Krebs, B. (1982) *Angew. Chem. Int. Engl.* **21**, 201-202.

46. Ghanotakis, D. F. & Yocum., C. F. (1990) *Ann. Rev. Plant Physiol. Mo. Biol.* **41**, 255-276.

47. Semenza, G. (1986) *Ann. Rev. Cell Biol.* **2**, 255.

48. Cheniae, G. M. & Martin, I. F. (1984) *Biochim. Biophys. Acta* **197**, 219-239.

49. Coleman, W. J. (1990) *Photosynth. Res.* **23**, 1-27.

50. Rutherford, A. W. (1989) *Trends. Biochem. Sci.* **14**, 227-232.

51. Kirby, J. A., Goodin, D. B., Wydrzynski, T., Robertson, A. S. & Klein, M. P. (1981) *J. Am. Chem. Soc.* **103**, 5537.

52. Guiles, R. D., *et al.* (1987) in *Progress in Photosynthesis Research* (Biggins, J., ed.) Vol. 1, pp. 561-564. Nijhoff, Dordrecht.

53. Yachandra, V. K., *et al.* (1987) *Biochem.* **26**, 5974; McDermott, A. E., *et al.* (1988) *Biochem.* **27**, 4021.

54. Dismukes, G. C. & Siderer, Y. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 274; DePaula, J. C., Beck, W. F., Miller, A. -F., Wilson, R. B. & Brudvig, G. W. (1987) *J. Chem. Soc. Faraday Trans.* 1 **83**, 3635.

55. Lavergne, J. (1987) *Biochim. Biophys. Acta* **894**, 91.

56. Guiles, R. D., Yachandra, V. K., McDermott, A. E., Cole, J. L, Dexheiner, S. L., Britt, R. D., Sauer, K. & Klein, M. P. (1990) *Biochem.* **29**, 486-496.

57. Guiles, R. D., Zimmermann, J. -L., Mcdermott, A. E., Yachandra, A. E., Cole, J. L., Dexheimer, S. L., Britt, R. D., Weighardt, K., Bossek, U., Sauer, K. & Klein, M. P. (1990) *Biochem.* **29**, 471-475.

58. Corrie, A. R., Evans, M. C. W., Hubbard, J. A. M., Strange, R. W. & Hasnain, S. S. (1990) *Curr. Res. Photosynth.* **1**, 793-796.

59. George, G. N., Prince, R. C. & Cramer, S. P. (1989) *Science* **243**, 789-791.

60. Penner-Hahn, J. E., Fronke, R. M., Pecoraro, V. L., Yocum, C. F., Bretts, S. D. & Bowlby, N. R. (1990) *J. Am. Chem. Soc.* **112**, 2549-2557.

61. MacLachlan, J. E., Hallahan, B. J., Ruffle, S. V., Nugent, J. H. A., Evans, M. C. W. Strange, R. W. & Hasnain, S. S. (1992) *Biochem. J.* **285**, 569-576.

62. Arber, J. M., Flood, A. C., Garner, C. D., Gormal, C. A., Hasnain, S. S. & Smith, B. E. (1988) *Biochem. J.* **252**, 421-425.

63. Fehr, G., Allen, J. P., Okamura, M. Y. & Rees, D. C. (1989) *Nature* **339**, 111-116.

64. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. (1985) *Nature* **318**, 618-624

65. Allen, J. P., Fehr, G., Yeates, T. O., Komiya, H. & Rees, D. C. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5730-5734.

66. Fukuama, K., *et al.* (1980) *Nature* **286**, 522-524.

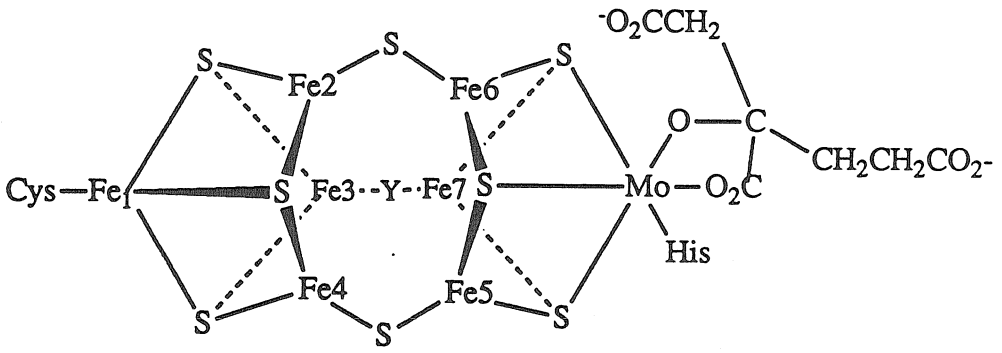


Figure 6-1. Schematic representation of the FeMo-cofactor in nitrogenase MoFe-protein. The FeMo-cofactor contains 4Fe:3S and 1Mo:3Fe:3S clusters that are bridged by three non-protein ligands. There may be some Fe--Fe bonding interactions that could provide a fourth coordination interaction for the six central irons.

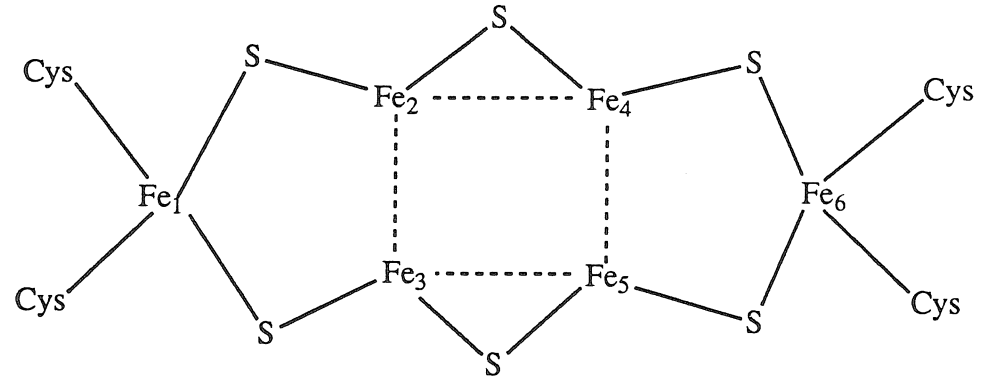


Figure 6-2a. A structural model for the H-cluster in Fe-hydrogenases. The proposed H-cluster model contains two 3Fe:2S clusters that are bridged by two sulfides. There may be some Fe--Fe interactions that could provide the third and fourth coordnation interactions for the four central irons.

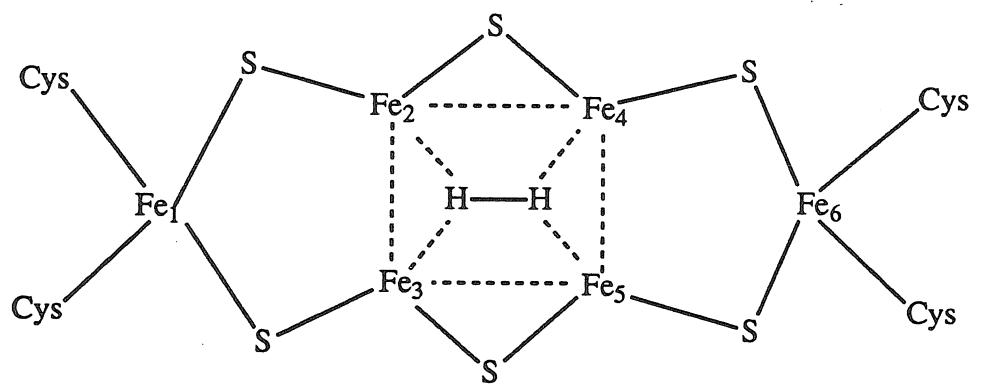


Figure 6-2b. A possible H_2 binding mode on the H-cluster model. The four central irons form an iron surface which is adquate for H_2 binding as well as CO binding. CO may bind in an end-on fashion to the iron surface.

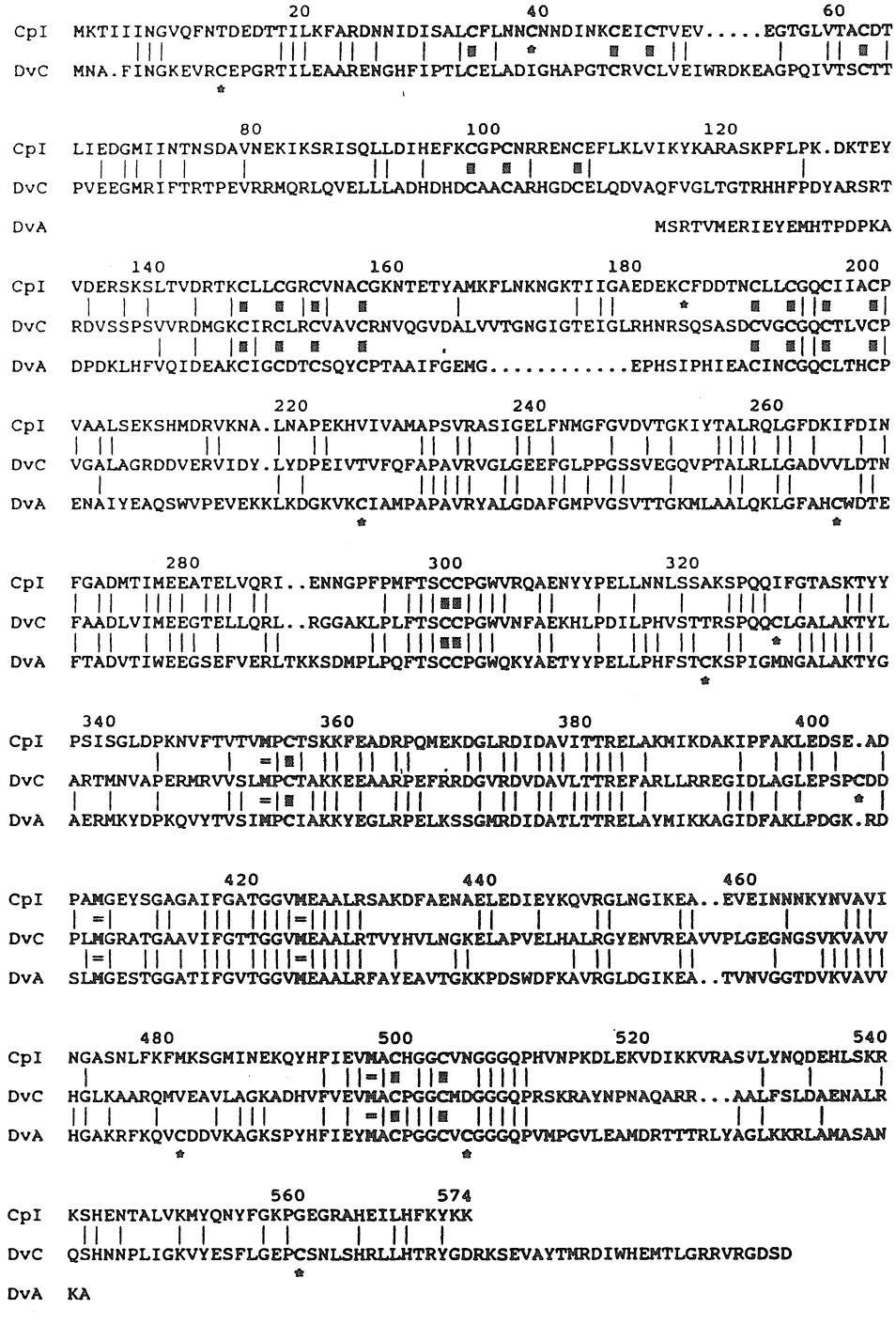


Figure 6-3. Amino acid sequence alignment of Fe-hydrogenases. Comparisons of the sequence of *C. pasteurianum* hydrogenase I (CpI) with those of the translated hydγ gene from *D. vulgaris* (DvC) and of the large hydrogenase subunit from the same strain (DvA). Taken from reference 44.

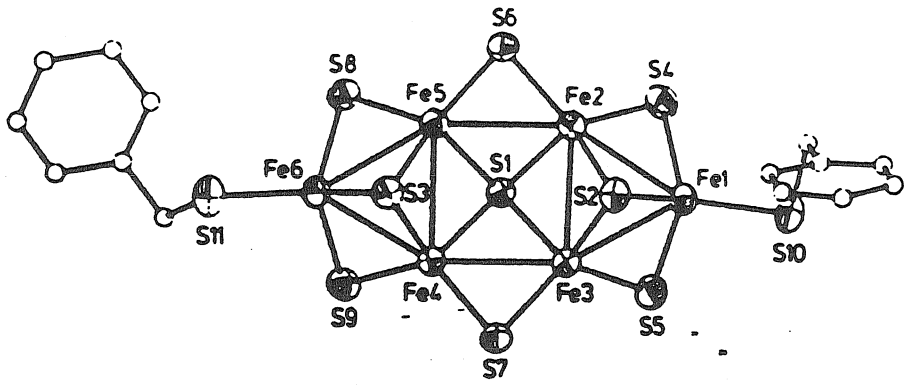
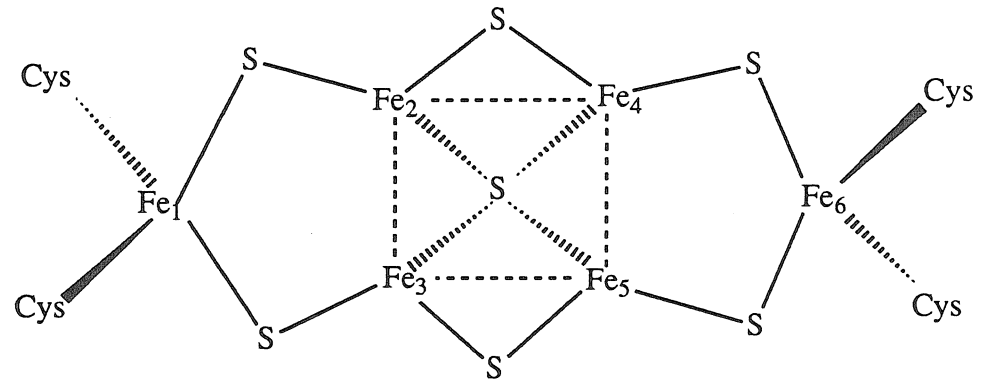
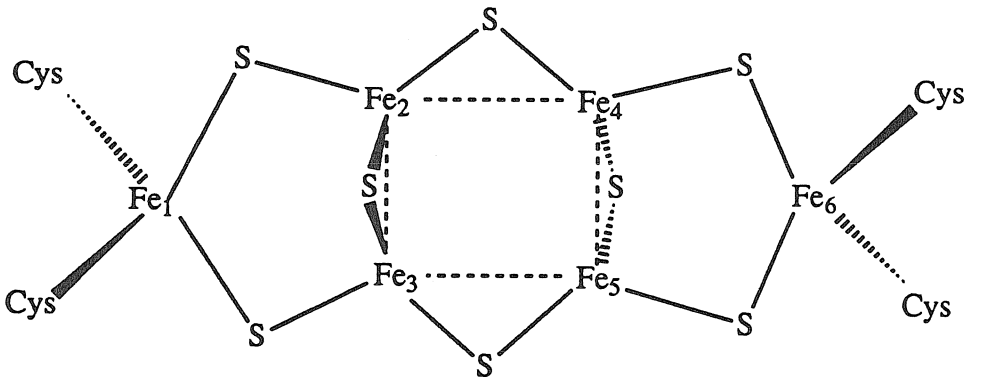


Figure 6-4. Structure of a 6Fe:9S cluster. Taken from reference 45.

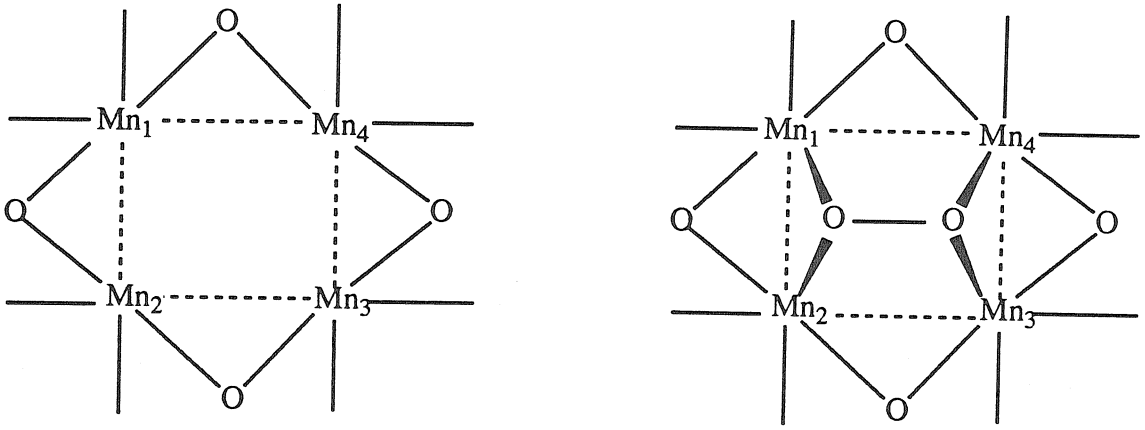


(a)



(b)

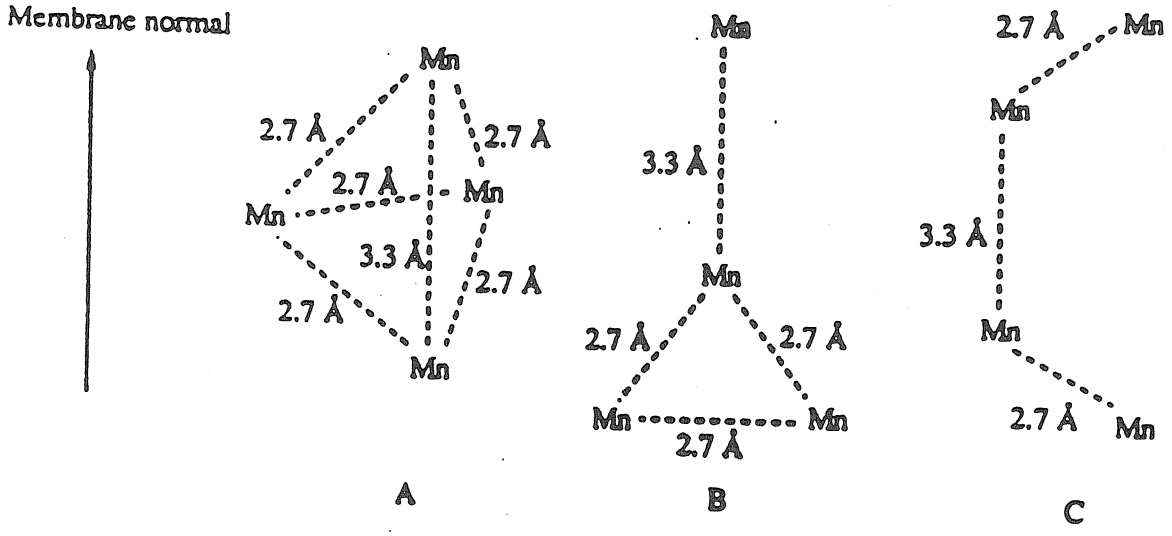
Figure 6-5. Two alternative models for the H-cluster in Fe-hydrogenases: (a) a 6Fe:7S cluster, and (b) a 6Fe:8S cluster.



(a) (b)

Figure 6-6a. A structural model for the Mn-center in PSII. The proposed Mn-center model contains a 4Mn:4O cluster as the basic structural element, however, additional ligands to the Mn sites could be oxygen as well as protein ligands. There might be some Mn--Mn bonding interactions that could prvide the fifth and sixth coordination interactions for the four manganese atoms. The stereochemistry is not considered in this model; the four manganese are drawn on the same plane, however, twisting of four manganese atoms out of plane is also conceivable.

Figure 6-6b. A possible O_2 binding mode on the Mn-center model. The four manganese atoms form a metal surface which is adquate for O_2 binding.



Mn-Mn	2.7Å (N)	3.3Å (N)
Experimental	2.1±0.8	0.8±0.3
A	2.5	0.5
B	1.5	0.5
C	1.0	0.5
4Mn:4O	2.0	1.0

Figure 6-7. Possible geometries of the Mn site based on EXAFS data. Taken from reference 59. The proposed Mn-center model (Figure 6-6a) is a more distorted version of model A containing two 3.3Å Mn--Mn interactions, and best fits the EXAFS data.