CRYSTAL STRUCTURE OF *AZOTOBACTER VINELANDII* NITROGENASE IRON PROTEIN AT 2.2 Å RESOLUTION

Thesis by

Jamie L. Schlessman

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1997

(Submitted June 9, 1997)

ACKNOWLEDGMENTS

I would like to thank my research advisor Doug Rees for the opportunity to work in his laboratory. This research would have been impossible without his support, encouragement, and boundless patience and enthusiasm.

In addition, I am grateful to James Howard, of the University of Minnesota, for excellent and rigorous discussions on nitrogenase. His profound insight into nitrogenase biochemistry has impacted all facets of this work.

I would also like to thank the past and present members of the Rees lab for their friendship and assistance. The contributions of Fe protein structure pioneers John Kornuc, Pinak Chakrabarti, and especially Millie Georgiadis, are much appreciated. I would like to acknowledge the MoFe protein efforts of Jongsun Kim, Mike Chan, and Mike Day, and to thank Hiromi Komiya for sharing her model building expertise. I am particularly grateful to Debbie Woo for teaching me nitrogenase protein purification and crystallization techniques, and engendering a strong attachment to Fe protein. Her meticulous attention to detail remains unparalleled. I thank Art Chirino for keeping the lab current on software developments, for detailed explanations of programming, and for running countless jobs on Av2. Barbara Hsu taught me to collect and process data, and maintained computer and x-ray equipment, day-to-day lab operations, and her composure with exacting standards. I would like to thank Leemor Joshua-Tor for her friendship and for bringing cryocrystallography to the Rees lab. Michael Stowell's brilliance and generosity, particularly in cryocrystallographic hardware advancements, are greatly appreciated. I thank Hermann Schindelin and Caroline Kisker for sharing prep duties and their keen perception of crystallography. I am grateful to John Peters for his encouragement and comraderie, as well as his patient tutelage of *nif* genetics and molecular biology. Tim McPhillips supplied friendship, wit, and tireless computer hardware support. I thank Tricia Takahara for assuming responsibilities for both the anaerobic chambers and lab culinary events, and

Cathy Drennan and Clasa Kielkopf for cheerful assistance in proofreading and thesis assembly. I would also like to acknowledge Geoffrey Chang and Robert Spencer for their support and friendship.

Beyond the Rees lab, I would like to thank the following members of the Caltech community: Tom Dunn, for encouragement and electrical expertise; Guy Duremberg, for superb machining skills, instant service, and wonderful stories; Rick Gerhart, for precision glassblowing; Larry Tutor, for fermentor operation and maintenance; and Phoebe Ray and Dian Buchness, for excellent secretarial support. I am humbly grateful to Larry Henling for emergency room visits, tire changes, and unforgettably delicous berries. I thank Pamela Bjorkman for sharing x-ray equipment and computers, and Dan Vaughn for being an nth year with me. From UCLA, I acknowledge Thomas Sutherland for fermenting *A. vinelandii* cells, and Phoebe Stewart for her patience and support. I am also pleased to acknowledge General Electric and NIH for financial support.

I am indebted to those whose friendship has sustained me. From Albright College, I am grateful to Francis and Lynn Williamson, Brian Pfennig, Jessica Zuber, and Terri Tyson. I thank Barbara Wyslouzil, Christine Nelson, and Elizabeth Burns, the last of the H.S.B.'s, for nurturing during my early graduate years. Missy Lunden, Donnie Cotter, and Chris Tinney have followed me to the ends of the earth (or at least to Dunkeld, Scotland). I am particularly grateful to Kim Tryka for her lasting friendship and relentless self-reliance. Linda Weavers, Donna Womble, Ruth Ann Bertsch, and Gary Mines have all been supportive friends at various stages of the journey. I am extremely grateful to Angela Putney for being a faithful and compassionate friend.

I would like to thank the people who filled the journey with music: the members of Cantores Atri Mortis; Dolores and Bill Bing; Wendy and Don Caldwell; Monica Hubbard; Allen Gross; Frank Brownstead and his choir; James Walker, Timothy Howard, and the members of Coventry Choir; Missy Lunden; Donnie Cotter; Courtney Smith; Toshi Takeuchi; Jerome Claverie; Dan Hall; Tom Lloyd; Heide Li; Deepto Chakrabarti; Anita

Gould; and Ed Lewis. I am extremely grateful for the life of Jordan Kaplan, and the music we shared. I am thankful for the privilege of making music for Nobel Laureates Linus Pauling, Desmond Tutu, and Ed Lewis during my time at Caltech.

I am grateful to my husband, Neill Reid, for his love and seemingly endless patience and tolerance. Thanks go to Ruaraidh and Gwen Reid, for dogged loyalty beyond all reasonable measure. I appreciate the support of Elizabeth Reid, Elspeth Reid, and Lyle Freedman. Finally, I would like to thank my parents, Sharon and Denny, for their unquestioning love and support over many years, and my brother Jason, for his quiet encouragement and constant inspiration.

for Jason

"We achieve, with magnification, a new dimension of strangeness. We enter new worlds. Our own becomes less stable. Imagine seeing this butterfly with no preconceptions about its nature. Or imagine knowing about the way it feeds, about the structure of its wing-tissue, in microscopic detail. Our picture of reality is never fixed but can always be elaborated and made more accurate. And this changes us. **The weight of the butterfly makes an iron bridge bend - in that it redistributes, ever so slightly, some molecules.** *Every new piece of knowledge - in the same way - enlarges our world."*

- A. S. Byatt, "The Game" (1967)*

*Permission granted by Sterling Lord Literistic, Inc.

ABSTRACT

Biological nitrogen fixation by the two-component metalloenzyme nitrogenase provides an elegant solution to the problem of reducing abundant, but relatively inert, dinitrogen to the biologically usable ammonia needed by all organisms. This oxygen sensitive enzyme, consisting of the separately purifiable nitrogenase iron protein and molybdenum iron protein, couples nucleotide hydrolysis to electron transfer to catalyze the ATP-dependent reaction. Iron protein acts as the sole known biological reductant to molybdenum iron protein, which contains the actual site of substrate reduction. MgATP binding to iron protein induces dramatic conformational changes in the protein's structure required for docking with molybdenum iron protein. Complex formation and dissociation are essential for nucleotide hydrolysis, electron transfer, and substrate reduction.

We have determined the crystal structure of *Azotobacter vinelandii* nitrogenase iron protein at 2.2Å resolution in the absence of nucleotide. Crystals grew in space group $P2_12_12_1$, and represent a new crystal form compared with that of the structure previously determined at 2.9 Å resolution. X-ray diffraction data were collected from a single crystal using cryocrystallographic techniques. The structure was solved by molecular replacement, followed by solvent flattening and noncrystallographic averaging. The current model contains 575 of 578 possible amino acid residues and 372 solvent molecules, and has been refined to R-value of 22.3 % (R-free = 29.0 %) for all data to 2.2 Å, with good stereochemistry.

The overall topology of nitrogenase iron protein consists of an Fe_4S_4 cluster symmetrically ligated by two identical subunits of doubly wound α/β structure similar to those of other nucleotide binding proteins. A detailed description is provided of those structural features important for iron protein function, including nucleotide binding regions, the Fe_4S_4 cluster environment, intersubunit interactions, and the molybdenum iron protein binding surface. Comparisons are made between the current model and that of *C.*

pasteurianum iron protein, as well as those of two *A. vinelandii* nitrogenase complexes. Analysis of the various iron protein structures provides a framework for considering the wealth of relevant nitrogenase spectroscopic, biochemical, and genetic information.

TABLE OF CONTENTS

CHAPTER 1 Introduction to Nitrogenase

1. Biological Nitrogen Fixation.....1
2. The Nitrogen Cycle.....2
3. The Haber-Bosch Process.....4
4. Nitrogenase Complex.....5
5. *nif* Gene Products.....8
6. Structural Knowledge of Nitrogenase.....10
7. Research Objective.....16

CHAPTER 2 Preparation of Av2

1. *Azotobacter vinelandii* Cell Culture.....18
    1. Media Requirements.....18
    2. Cell Growth Protocol.....19
2. Purification of Av2.....20
    1. Initial Purification Methods.....21
        1. Pre-Lysis Preparations.....22
        2. Cell Lysis.....24
        3. Chromatographic Separations.....26
        4. Protein Storage and Evaluation.....27
    2. Final Purification Methods.....27
3. Protein Evaluation.....30

CHAPTER 3 Crystallization and X-ray Analysis of Av2

1. Av2 Crystallization..... 32
    1. Preparations of Precipitant Solutions..... 32
    2. Initial Conditions..... 34
    3. Optimization of Conditions..... 36
    4. Crystallization Attempts with Nucleotides..39
2. Crystal Mounting and Data Collection..... 41
    1. Room Temperature Crystal Mounting.....41
    2. Room Temperature Data Collection..... 42
    2. Cryocrystallographic Crystal Preparation...43
    3. Cryocrystallographic Data Collection.....44

CHAPTER 4 Determination of the Av2 Crystal Structure

1. Molecular Replacement.....51
    1. Molecular Replacement using X-PLOR.....52
        1. Self-Rotation Function.....53
        2. Cross-Rotation Function.....55
        3. Translation Function..... 56
    2. Molecular Replacement using AMORE.... 58
2. Anomolous Fourier Function.....59
3. NCS Averaging and Solvent Flattening.....59
4. Crystallographic Refinement.....60
    1. Improved Phasing Techniques..... 62
    2. Final Refinement.....64
5. Current Av2 model.....65

CHAPTER 5 Analysis of the Av2 Crystal Structure

1. Description of the Structure..... 70
2. Similarity to other Iron Protein Structures..... 74
    1. Similarity to 2.9 Å Resolution Av2..... 74
    2. Similarity to Cp2..... 77
    3. Similarity to Complexed Av2..... 78
3. Nucleotide Binding and Hydrolysis..... 82
4. [4Fe-4S] Cluster..... 87
5. MoFe Protein Binding Surface..... 94
6. Intersubunit Interactions..... 95
7. Discussion..... 96

REFERENCES..... 106

LIST OF FIGURES

CHAPTER 1 Introduction to Nitrogenase

1-1. Schematic of the Nitrogen cycle..... 3

1-2. Redox reactions of the nitrogenase complex..... 7

1-3. Organization of the *A. vinelandii nif* genomes..... 9

1-4. Ribbons representation of MoFe protein.....11

1-5. Ribbons representations of Fe proteins.....13

1-6. Ribbons representations of nitrogenase complexes..... 15-16

CHAPTER 2 Preparation of Av2

2-1. PAGE gel of nitrogenase gradient fractions.....30

CHAPTER 3 Crystallization and X-ray Analysis of Av2

3-1. Plot of R_{factor} versus resolution for the $P2_12_12_1$ Av2 data set..... 47

3-2. Plot of <I/sigma> versus resolution for the $P2_12_12_1$ Av2 data set..48

3-3. Plot of completeness versus resolution for the $P2_12_12_1$ Av2 data set..... 48

3-4. Plot of multiplicity versus resolution for the $P2_12_12_1$ Av2 data set..... 49

CHAPTER 4 Determination of the Av2 Crystal Structure

4-1. Amino acid sequence alignment of Av2 and Cp2..... 52

4-2. Plot of self-rotation function for Av2..... 54

4-3. Summary of Av2 molecular replacement search in X-PLOR... 56-57

4-4. Plot of estimated Wilson statistics for the $P2_12_12_1$ Av2 data set.... 67

4-5. Ramachandran plot of each subunit of the current Av2 model..68-69

CHAPTER 5 Analysis of the Av2 Crystal Structure

5-1. Ribbons representations of the current Av2 model.....72

5-2. Superposition of the 2.2Å and 2.9Å resolution Av2 models.....76

5-3. Superposition of the current Av2 and Cp2 models..... 78

5-4. Representation of the current Av2 model with superimposed nucleotide molecules.....80

5-5. Superposition of the current Av2 and Av2 in the (AlF_4^- • Av2)Av1 models..... 81

5-6. Ball-and-stick representation of the P-loop region of Av2..... 83

5-7. Ball-and-stick representation of the Effector 1 region of Av2..... 85

5-8. Ball-and-stick representation of the Effector II region of Av2.....86

5-9. Potential N-H···S bonding networks for Av2, Cp2, and the Av2 in the (AlF_4^- • Av2)Av1 models.....88

5-10. Ball-and-stick representation of lattice packing interactions in the current Av2 model..... 90

5-11. Ball-and-stick representation of the Fe_4S_4 cluster environment of the current Av2 model.....92

5-12. Spatial organization of the Fe protein invariant core..... 98

5-13. Sequence alignment of Av2 and several nifH-like archae sequences..... 101-103

LIST OF TABLES

CHAPTER 2 Preparation of Av2

2-1. Buffers for Av2 purification protocol..... 23

2-2. Chromatographic columns used in Av2 purification protocol..... 24

CHAPTER 3 Crystallization and X-ray Analysis of Av2

3-1. Crystallization conditions for Av2..... 38

3-2. Data collection statistics for $P2_12_12_1$ Av2 SSRL data set I..... 45

3-3. Data collection statistics for $P2_12_12_1$ Av2 SSRL data set II..... 46

3-4. Data collection statistics for $P2_12_12_1$ Av2 CuKα data set.....50

CHAPTER 4 Determination of the Av2 Crystal Structure

4-1. Refinement statistics for the current 2.2 Å Av2 structure..... 66

CHAPTER 5 Analysis of the Av2 Crystal Structure

5-1. Secondary structure elements in $P2_12_12_1$ Av2..... 71

5-2. Potential N-H···S bonding interactions for Av2, Cp2, and the Av2 in the (AlF_4^- • Av2)Av1 models..... 89

5-3. Av2 Fe_4S_4 cluster geometry.....93

5-4. Residues forming the Fe protein invariant core..... 97

LIST OF ABBREVIATIONS

A	alanine
ADP	adenosine-5'-diphosphate
Ala	alanine
Arg	arginine
Asn	asparagine
Asp	aspartic acid
ATP	adenosine-5'-triphosphate
Av1	nitrogenase molybdenum iron protein from *Azotobacter vinelandii*
Av2	nitrogenase iron protein from *Azotobacter vinelandii*
Cp1	nitrogenase molybdenum iron protein from *Clostridium pasteurianum*
Cp2	nitrogenase iron protein from *Clostridium pasteurianum*
C	carbonyl carbon, cysteine
CA	alpha carbon
CD	circular dichroism
Cys	cysteine
D	aspartic acid
DEAE	diethylaminoethyl
Δ	delta
E	glumatic acid
e-	electron
EPR	electron paramagnetic resonance
eq	equilibrium
EXAFS	extended X-ray absorption fine spectroscopy
F	phenylalanine
Fd	ferredoxin

Fe	iron
G	glycine
Gln	glutamine
Glu	glutamic acid
Gly	glycine
H	histidine, hydrogen
Hg	mercury
His	histidine
I	isoleucine
I.D.	inner diameter
Ile	isoleucine
K	lysine
L	leucine
Leu	leucine
Lys	lysine
M	methionine
Met	methionine
Mg	magnesium
M_r	molecular weight
Mo	molybdenum
MoFe	molybdenum iron
MWCO	molecular weight cutoff
N	asparagine, nitrogen
NCS	noncrystallographic symmetry
NMR	nuclear magnetic resonance
O	oxygen
OD	optical density

O.D.	outside diameter
P	proline
PDB	Brookhaven protein data bank
PEG	polyethylene glycol
Phe	phenylalanine
Pro	proline
Q	glutamine
R	arginine
RMS	root-mean-square
rpm	revolution per minute
S	serine, Siemens, sulfur
Ser	serine
T	threonine
Trp	tryptophan
Tyr	tyrosine
V	valine, vanadium
Val	valine
W	tryptophan
Wat	water
w/v	weight-to-volume ratio
Xaa	unspecified amino acid
Y	tyrosine

CHAPTER 1
Introduction to Nitrogenase

1.1. Biological Nitrogen Fixation

Biological nitrogen fixation is the process by which relatively inert atmospheric dinitrogen, N_2, is reduced to metabolically usable ammonia, NH_3, by a small group of diazotrophic microorganisms. These prokaryotes, while physiologically and phylogenetically diverse, all contain the enzyme nitrogenase. Some live symbiotically with either leguminous (*Rhizobium* species) or non-leguminous (*Agrobacterium radiobacter*) plants, while others are free-living (Postgate, 1982). The latter range from heterotrophs such as obligate aerobes (*Azotobacter vinelandii*), facultative anaerobes (*Klebsiella pneumoniae*), and obligate anaerobes (*Clostridium pasteurianum*), to autotrophic purple non-sulfur photosynthetic bacteria (*Rhodospirillum rubrum*) and cyanobacteria (*Anabaena variabilis*) (Postgate, 1982; Burris, 1991; Roberts *et al.*, 1992; Stewart *et al.*, 1977).

Nitrogenase is an oxygen sensitive two-component enzyme which catalyzes nitrogen reduction to ammonia under physiological conditions (recently reviewed in Howard and Rees, 1997; Burgess and Lowe, 1997; Dean *et al.*, 1993). The reaction occurs in an aqueous environment, with temperatures and pressures of 298-310 K and 0.8 atm N_2, respectively. These ambient conditions are accompanied by a huge energetic cost--while thermodynamically favorable ($\Delta G^\circ \approx -15.2$ kcal/mol (Chase *et al.*, 1985)), biological nitrogen fixation by the metalloenzyme nitrogenase hydrolyzes at least 16 ATP for each molecule of N_2 reduced. The overall stoichiometry of the reaction is usually represented as (Simpson and Burris, 1984):

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

where the protons associated with MgATP hydrolysis are not indicated. Under physiological conditions, the energy requirement may approach 20-30 MgATP per N_2 (Hill, 1976). The use of this energy in effecting catalysis represents a fundamental issue in structural bioenergetics.

1.2. The Nitrogen Cycle

The major chemical transformations of the planet involving nitrogen may be summarized in a hypothetical global nitrogen cycle (Svennson and Soderlund, 1976), a schematic of which is shown in Figure 1-1. Within this cycle, plants and microbes synthesize their organic nitrogenous matter, including proteins, from nitrates using nitrites and ammonia as starting materials. Animals, in turn, consume plants and microbes to build their own proteins. This "immobilized" organic nitrogen is inaccessible to plants until eventually released by death, decomposition, and putrefaction as NH_3 in the process of ammonification. Nitrifying bacteria rapidly oxidize this ammonia to nitrate; under conditions of O_2 depletion, nitrate may be reduced to N_2 by denitrifying bacteria and subsequently lost to the biosphere. Additional nitrogen is lost in the form of N_2O as an intermediate in biological denitrification. Nitrous oxide is a greenhouse gas with an energy reflectivity per mole 180-fold higher than that of carbon dioxide (Levi *et al.*, 1991), and may thus contribute to global warming (Hardy, 1994).

The net loss of biologically accessible nitrogen is compensated by the process of nitrogen fixation. While lightning, fires, and atmospheric pollution generate oxides of nitrogen, which can be subsequently washed into the soil as nitrates by rain, the majority of non-commercially prepared fixed nitrogen results from nitrogenase activity in diazotrophic bacteria (Postgate, 1982). Ammonia produced by biological nitrogen fixation is generally converted to nitrate before plant uptake occurs. Nitrification is important on two accounts:

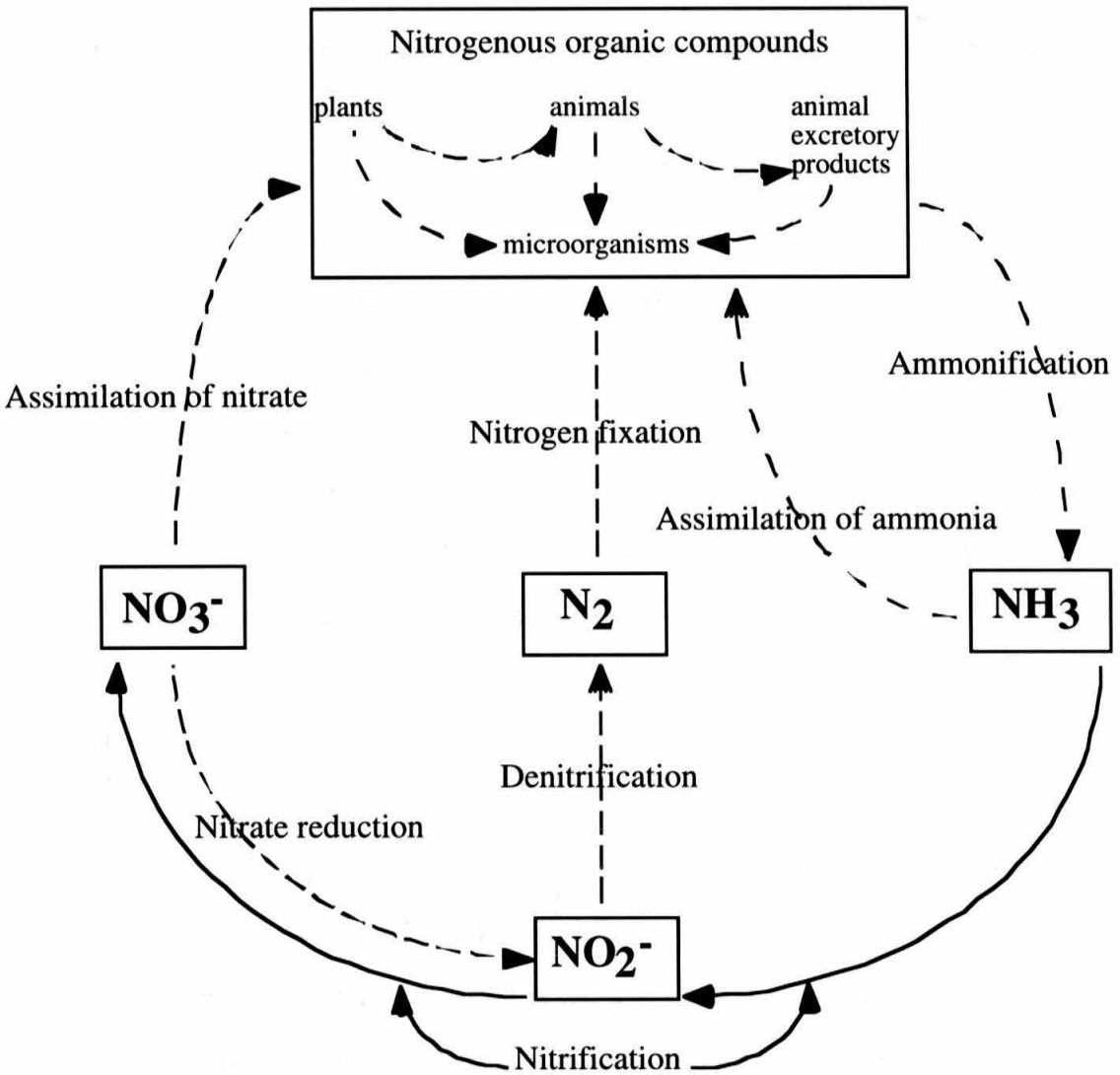


Figure 1-1. A schematic of the reactions involved in the nitrogen cycle. Nitrogen fixation by diazotrophic microorganisms contributes ~60 % of newly fixed nitrogen annually. The figure was adapted from Stanier *et al.* (1976).

(1) plants seems to prefer this form of fixed nitrogen; and (2) after application of commercial fertilizer, nitrate readily washes out of soil, whereas ammonia remains bound (Postgate, 1982). This circulation of fixed nitrogen minimizes the dangerous polluting effects of ammonia leached into water sources.

Globally, the nitrogen cycle involves the turnover of more than 2 x 10^9 tons of nitrogen each year, of which nearly 10 % is lost to the biosphere as N_2 or N_2O (Burns and Hardy, 1975). Biological nitrogen fixation by nitrogenase provides about 60 % of the world's newly fixed nitrogen, and will play a critical role in the attainment of environmentally benign, sustainable farming systems. Increased use of biological nitrogen fixation will mitigate the need for nitrogen from commercial fertilizer, with concomitant benefits accrued in terms of effects on the global nitrogen cycle, global warming, and ground- and surface- water contamination (Hardy, 1994).

1.3. The Haber-Bosch Process

The kinetic stability of the nitrogen-nitrogen triple bond provides the primary energetic barrier to N_2 reduction (Howard and Rees, 1997). The most successful chemical approach to overcoming this obstruction is the use of iron as a catalyst to accelerate the reaction of N_2 and H_2 in the gas phase, in a protocol derived by F. Haber and developed commercially by C. Bosch (Topham, 1985). The mechanism of this Haber-Bosch process involves dissociation of N_2 to atomic nitrogen on the active {111} crystal face of the iron catalyst, with subsequent combination with dissociated hydrogen to produce ammonia (Jennings, 1991). Despite the use of the iron catalyst, extreme temperatures of 600-800 K are required for the reaction. Unfortunately, the exothermic nature of ammonia formation from N_2 and H_2 (ΔH = -10.97 kcal/mol ammonia formed at 298 K (Chase, 1985)) shifts the equilibrium toward reactants with increasing temperature. At atmospheric pressure, this results in a drop in equilibrium ammonia concentration from 96 % at 298 K to ~0.2 % at

723 K (Howard and Rees, 1997). Hence, pressure must be increased to shift favorably the equilibrium towards the products, in accordance with LeChatelier's principle. Typical pressures in the Haber-Bosch process reach 500 atm (Jennings, 1991).

While the Haber-Bosch process allows large-scale economical production of ammonia for use in fertilizers, explosives, and other chemicals (~80 x 10^9 kg ammonia annually (Schlesinger, 1991)), biological nitrogen fixation furnishes even higher annual yields (~170 x 10^9 kg ammonia (Schlesinger, 1991)). Mechanistic details of the nitrogenase-catalyzed reaction remain elusive. Meeting the needs of a burgeoning population upon our resource-limited planet demands provident use of its remaining resources, and a complete mechanistic understanding of the workings of nitrogenase may reap agricultural and industrial benefits. A recent USAID publication reported that synthetic nitrogen use had increased globally from 3 million to 80 million tons over the last 40 years, with an annual worldwide expenditure of US$20 billion for fertilizer nitrogen (Hardy, 1994). Industrial production of these fertilizers by the Haber-Bosch process consumes vast quantities of natural gas, which releases CO_2. Hence, food production may contribute indirectly to global warming. Recently, increased use of biological nitrogen fixation, rather than industrially fixed nitrogen, has been cited independently by USAID and UNESCO as a top priority, with a goal of minimizing perturbations to the natural nitrogen cycle (Hardy, 1994).

1.4. Nitrogenase Complex

The nitrogenase metalloenzyme consists of two separately purifiable components, the iron (Fe) protein and the molybdenum iron (MoFe) protein, each named according its metal center composition (Georgiadis *et al.*, 1992; Kim and Rees, 1992). The proteins associate in a ratio of 1 MoFe protein : 2 Fe protein (Bulen and LeConte, 1966; Burris, 1971). The Fe protein is a γ_2 homodimer of M_r ~60,000 which contains a single Fe_4S_4

cubane cluster ligated symmetrically by two cysteine residues from each protein subunit (Hausinger and Howard, 1982). A single nucleotide binding site is found in each monomer (Georgiadis *et al.*, 1992). The MoFe protein, an $\alpha_2\beta_2$ tetramer of M_r ~240,000, contains two different types of metal centers: (1) a FeMo cofactor, $MoFe_7S_9$, is buried within each α subunit (Chan *et al.*, 1992); and (2) the double cubane P-clusters, Fe_8S_7, that are located at each $\alpha\beta$ interface (Peters *et al.*, 1997). The FeMo cofactor, with its accompanying homocitrate moiety, provides the expected site of substrate reduction (Kim and Rees, 1993), while the P-clusters appear to act as intermediate electron donors between the Fe protein Fe_4S_4 cluster and the FeMo cofactor (Peters *et al.*, 1995; Lanzilotta and Seefeldt, 1996). Fe protein is the sole known biological reductant for MoFe protein, but is also required for numerous other nitrogenase functions, including FeMo cofactor biosynthesis on a protein scaffold (*nifEN* gene products) and insertion into immature MoFe protein (Robinson *et al.*, 1987; Tal *et al.*, 1991).

Both nitrogenase components are required for enzyme activity, including nucleotide hydrolysis, electron transfer, and substrate reduction. The generally accepted mechanism for nitrogenase activity is summarized: (1) Fe protein is reduced *in vivo* by either flavodoxin (in *A. vinelandii*) or ferredoxin (in C. *pasteurianum*), or *in vitro* by dithionite; (2) reduced Fe protein associates with, and transfers a single electron to, MoFe protein in a MgATP-dependent process requiring at least 2 MgATP hydrolyzed per electron transferred; (3) dissociation of the complex of oxidized Fe protein and partially reduced MoFe protein forms the rate-limiting step of the reaction (Lowe and Thorneley, 1983, 1984; Hageman and Burris, 1978); and (4) steps (1)-(3) repeat until eight electrons have accumulated in each $\alpha\beta$ MoFe dimer, resulting in substrate reduction and H_2 evolution. A schematic of the redox reactions of the nitrogenase enzyme is shown in Figure 1-2.

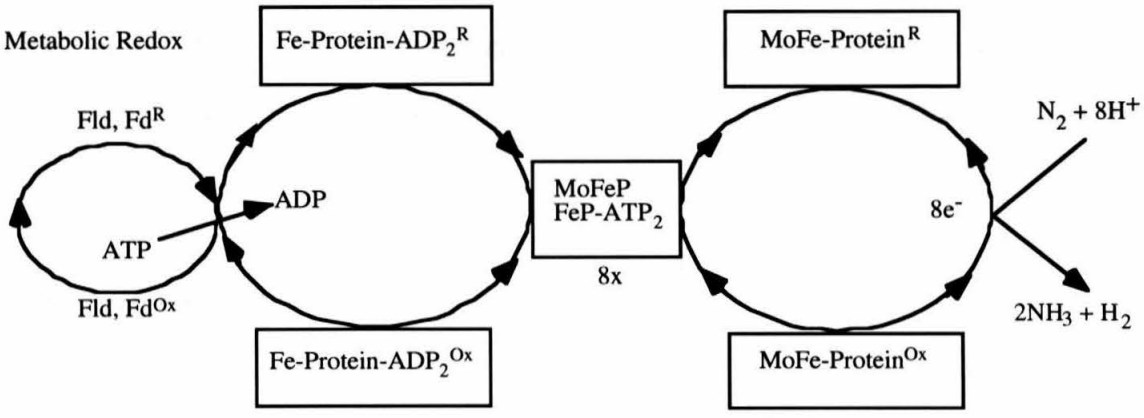


Figure 1-2. Redox reactions of the nitrogenase complex.

Environmental conditions of molybdenum depletion may induce expression of alternative nitrogenase systems homologous to those of the "conventional" molybdenum-containing system. Two such systems are currently known: a vanadium-containing (VFe) nitrogenase (Bishop *et al.*, 1980; Eady *et al.*, 1987; Hales *et al.*, 1986) and, under conditions of both molybdenum and vanadium deficiency, an iron-only nitrogenase (Chisnell *et al.*, 1988). The MoFe cofactor's single molybdenum site is replaced by V or Fe, respectively, in each case. Despite the clear relationship between the three nitrogenases, each alternative system is genetically distinct and does not simply arise from substitution of different metals into the same polypeptide. Each alternative system maintains its own particular Fe protein in *A. vinelandii*. In addition to ammonia formation, nitrogenase catalyzes the reduction of various alternative substrates including acetylene, azide, and cyanide, as well as the reduction of protons to hydrogen (Dilworth, 1966; Schöllhorn and Burris, 1967a, 1967b; Kelly *et al.*, 1967). Hydrogen formation appears to be a concomitant part of dinitrogen reduction (Hadfield and Bulen, 1969).

A convenient nomenclature for nitrogenase designates the MoFe protein and Fe protein from different bacterial sources as components "1" and "2," respectively, preceded by the two-letter abbreviation for the source genus and species (Eady *et al.*, 1972). For example, Av2 represents the Fe protein from *Azotobacter vinelandii*, and Cp1 refers to the

MoFe protein from *Clostridium pasteurianum*. Throughout this manuscript, nitrogenase component protein residues will be numbered according to the *A. vinelandii* sequences, unless otherwise indicated.

1.5. *nif* Gene Products

While the Fe and MoFe proteins provide the catalytic components of nitrogenase, a suite of more than 32 gene products is thought to be expressed under nitrogen fixing conditions (Dean *et al.*, 1993). These "accessory" proteins participate in activation, regulation, maturation, and electron transfer to the structural proteins. A depiction of the organization of *Azotobacter vinelandii* nitrogen fixation, or *nif*, genes is shown in Figure 1-3. *Azotobacter vinelandii* and *Klebsiella pneumoniae* have proven particularly tractable organisms for nitrogenase genetic studies. Studies of alterations engineered into these organisms have yielded extensive details of nitrogenase in the absence of structural information.

Among *nif* genes, only the *nifS* and *nifU* gene products are required for both Fe and MoFe protein production (Jacobson *et al.*, 1989a). Because a common feature between the Fe and MoFe proteins is the presence of Fe:S clusters, NifU and NifS have been proposed to be involved in the mobilization of Fe and S required for Fe:S cluster formation in nitrogenase (Zheng *et al.*, 1993). This hypothesis is supported by observations that nitrogenase structural proteins from *nifU* and *nifS* deletion mutant strains appear pale tan in color, in stark contrast with the rich brown hues of the wild-type equivalents (Jacobson *et al.*, 1989a). NifS, a pyridoxyl-phosphate dependent L-cysteine desulfurase, has been shown, in the presence of ferrous iron and reductant, to catalyze activation of an apo-form of Av2 from which the Fe_4S_4 cluster had been extracted by chelation (Zheng *et al.*, 1993, 1994a, 1994b). Confirmation of NifU's role in Fe:S cluster formation awaits further investigation.

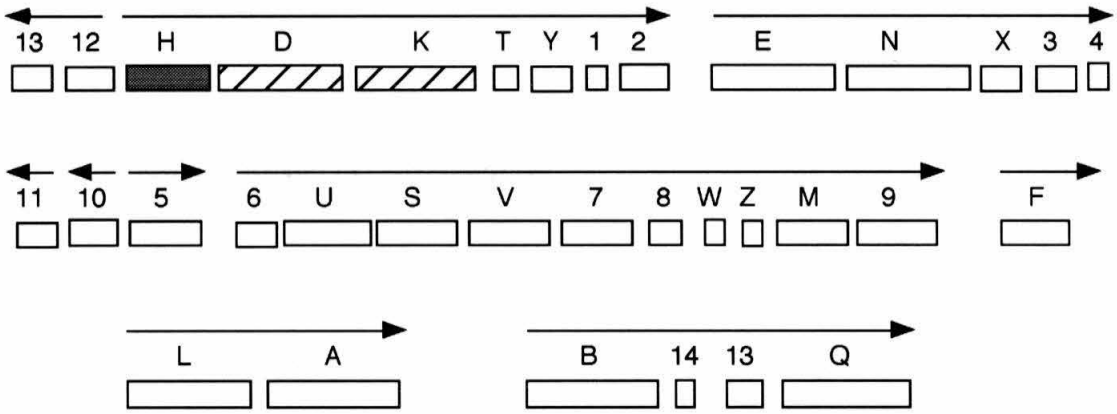


Figure 1-3. Organization of *nif* genes from *Azotobacter vinelandii*. Arrows indicate direction of transcription. Genes encoding for the nitrogenase MoFe protein, *nifD* and *nifK*, are shown crosshatched, while the gene encoding for the nitrogenase Fe protein, *nifH*, is shaded in grey.

Each Fe protein monomer is encoded by the *nifH* gene, and the α and β subunits of MoFe protein are encoded by *nifD* and *nifK*, respectively (Jacobson *et al.*, 1989a). The gene product of *nifM* is required for maturation of competent Fe protein (Robinson *et al.*, 1986), although its specific function remains elusive. The amino acid sequences of NifE and NifN show homology to those of MoFe protein (Brigle *et al.*, 1987), and NifEN is believed to function as a scaffold for the assembly of FeMo cofactor before its insertion into immature MoFe protein (Robinson *et al.*, 1986; Tal *et al.*, 1991). Gene products of *nifW* and *nifZ* are also required for full MoFe protein activity (Jacobson *et al.*, 1989b), although their specific involvement is unknown. MoFe cofactor biosynthesis also requires the Fe protein and gene products of *nifQ*, *nifB*, and *nifV*; Fe protein is also required for cofactor insertion (Robinson *et al.*, 1986; Imperial *et al.*, 1984; Shaw *et al.*, 1994; Allen *et al.*, 1995). The function of Fe protein in MoFe cofactor biosynthesis and insertion is poorly understood. Transcription of the *nif* genes is regulated positively by the *nifA* gene product (Lee *et al.*, 1993; Van de Broek *et al.*, 1992), and negatively by the *nifL* gene

product in response to the presence of NH_4^+ or O_2 (Austin *et al.*, 1994; Blanco *et al.*, 1993).

1.6. Structural Basis of Nitrogenase

The availability of crystal structures for the nitrogenase components has greatly enhanced the scope of nitrogenase understanding, by (1) providing a structural framework for assembling the vast body of relevant genetic, biochemical, and spectroscopic data, and (2) suggesting future experiments, particularly in site-directed mutagenesis, for further elucidating details of the nitrogenase mechanism. To date, the crystal structures of the Fe proteins from *A. vinelandii* and *C. pasteurianum*, (Georgiadis *et al.*, 1992; Woo, 1995) and the MoFe proteins from *A. vinelandii*, *C. pasteurianum*, and *K. pneumoniae* have been determined (Kim and Rees, 1992, 1993; Bolin *et al.*, 1993; Peters et al., 1997; B. E. Smith and coworkers, personal communication), as well as of two *A. vinelandii* complexes (Schindelin *et al.*, 1997, J. Peters, personal communication). Similarities in the structures of each protein from different organisms reflects the high degree of amino acid conservation among proteins of nitrogen fixing organisms.

The MoFe protein adopts a mixed α/β fold, in which the subunits of each $\alpha\beta$ dimer are related by a pseudo two-fold rotation axis passing through each P-cluster (Kim and Rees, 1992). A MoFe cofactor is positioned in each α subunit. However, the real surprise in the structure resulted not from its protein fold, but rather from its metal clusters. In particular, the relatively unassuming position of the molybdenum atom and presence of trigonal iron atoms in the MoFe cofactor were unanticipated from synthetic model chemistry. Controversies over the exact composition of the MoFe cofactor and P-cluster have been resolved recently using high resolution Av1 data collected in different oxidation states (Peters *et al.*, 1997). Figure 1-4 shows a ribbons representation of the present 2.0 Å resolution Av1 model.

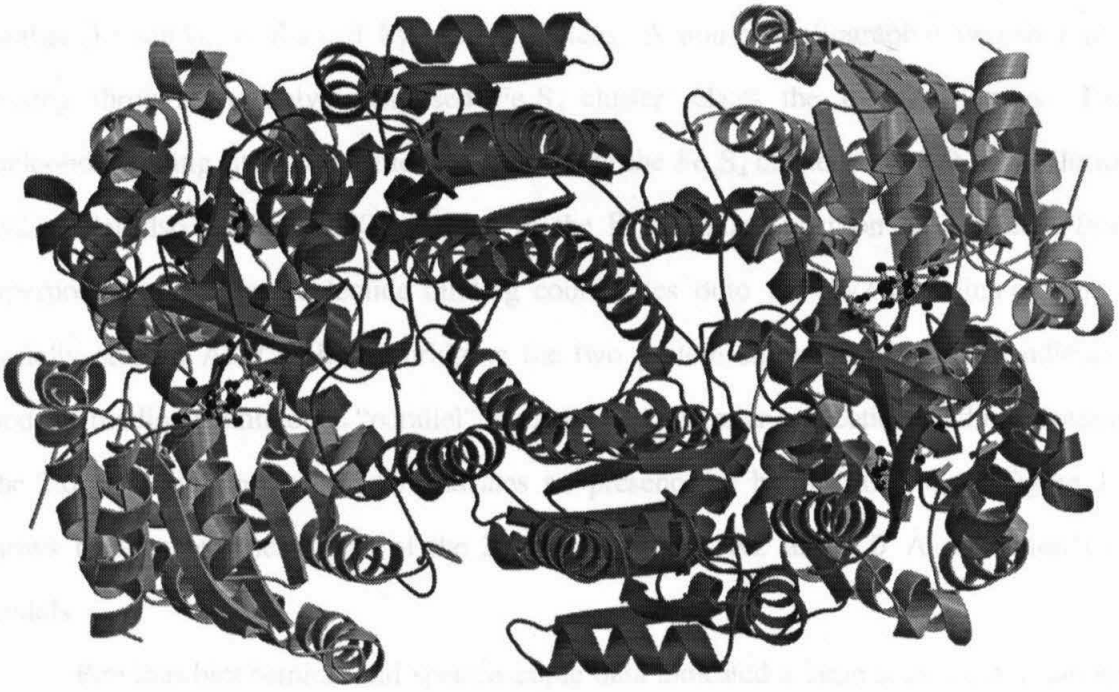


Figure 1-4. Ribbons representation of the MoFe protein from *A. vinelandii*. The FeMo cofactors and P-clusters are depicted as ball-and-stick models. Coils represent α-helices and arrows β-strands. The α subunits are shaded in light grey, while β subunits are depicted in darker grey. The two-fold axis relating the two αβ dimers lies perpendicular to the plane of the page. The pseudo two-fold axis relating the α and β subunits of each αβ dimer lies horizontally in the plane of the page. The figures were generated using MOLSCRIPT/RASTER3D (Kraulis, 1991; Bacon and Anderson, 1988; Merritt *et al.*, 1994), with coordinates courtesy of John Peters.

The 2.9 Å resolution Av2 crystal structure provided an initial glimpse of the Fe protein three-dimensional structure (Georgiadis *et al.*, 1992), which was recently confirmed by the 2.0 Å resolution Cp2 structure (Woo, 1995). The overall topology of nitrogenase iron protein consists of a Fe_4S_4 cluster symmetrically ligated by two identical subunits of doubly-wound α/β structure similar to those of other nucleotide binding

proteins. The eight-stranded twisted β-sheet core, containing seven parallel and one anti-parallel β-strands, is flanked by nine α-helices. A non-crystallographic two-fold axis passing through the solvent-exposed Fe_4S_4 cluster relates the two monomers. Two nucleotide binding sites are positioned ~19 Å from the Fe_4S_4 cluster. The 2.9 Å resolution Av2 model also showed MoO_4^{2-} bound in the β-phosphate position as predicted from superposition of other nucleotide binding coordinates onto the Av2 coordinates, and a partially-ordered ADP molecule bridging the two protein subunits. This "perpendicular" mode of binding is unlike the "parallel" mode found in similar nucleotide binding proteins. The 2.0 Å resolution Cp2 model indicates no presence of bound nucleotide. Figure 1-5 shows ribbons representations of the 2.9 Å resolution Av2 and 2.0 Å resolution Cp2 models.

Previous biochemical and spectroscopic data indicated a large-scale conformational change in the Fe protein upon nucleotide binding. Nucleotide-induced conformational changes in the Fe protein are accompanied by a variety of physiochemical effects on the Fe_4S_4 cluster environment, including a ~120 mV drop in reduction potential (Watt *et al.*, 1986; Zumft *et al.*, 1973). Small-angle X-ray scattering experiments indicate a 2.0 Å reduction in the Fe protein radius of gyration upon addition of MgATP (Chen *et al.*, 1994). The magnetic CD and 1H NMR spectra are altered by nucleotide binding (Stephens *et al.*, 1979; Ryle *et al.*, 1996a; Meyer *et al.*, 1988; Lanzilotta *et al.*, 1995; J. Howard, personal communication), and the EPR signal moves from rhombic to axial (Zumft *et al.*, 1973; Lindahl *et al.*, 1987b). MgATP binding also renders the Fe_4S_4 cluster irons susceptible to chelators such as α,α'-dipyridyl and *o*-bathophenanthroline (Walker and Mortenson, 1974; Ljones and Burris, 1978), and increases the protein's sensitivity to irreversible inactivation by oxygen (Yates, 1992). MgADP is a known inhibitor of Fe protein chelation (Ljones and Burris, 1978; Deits and Howard, 1990).

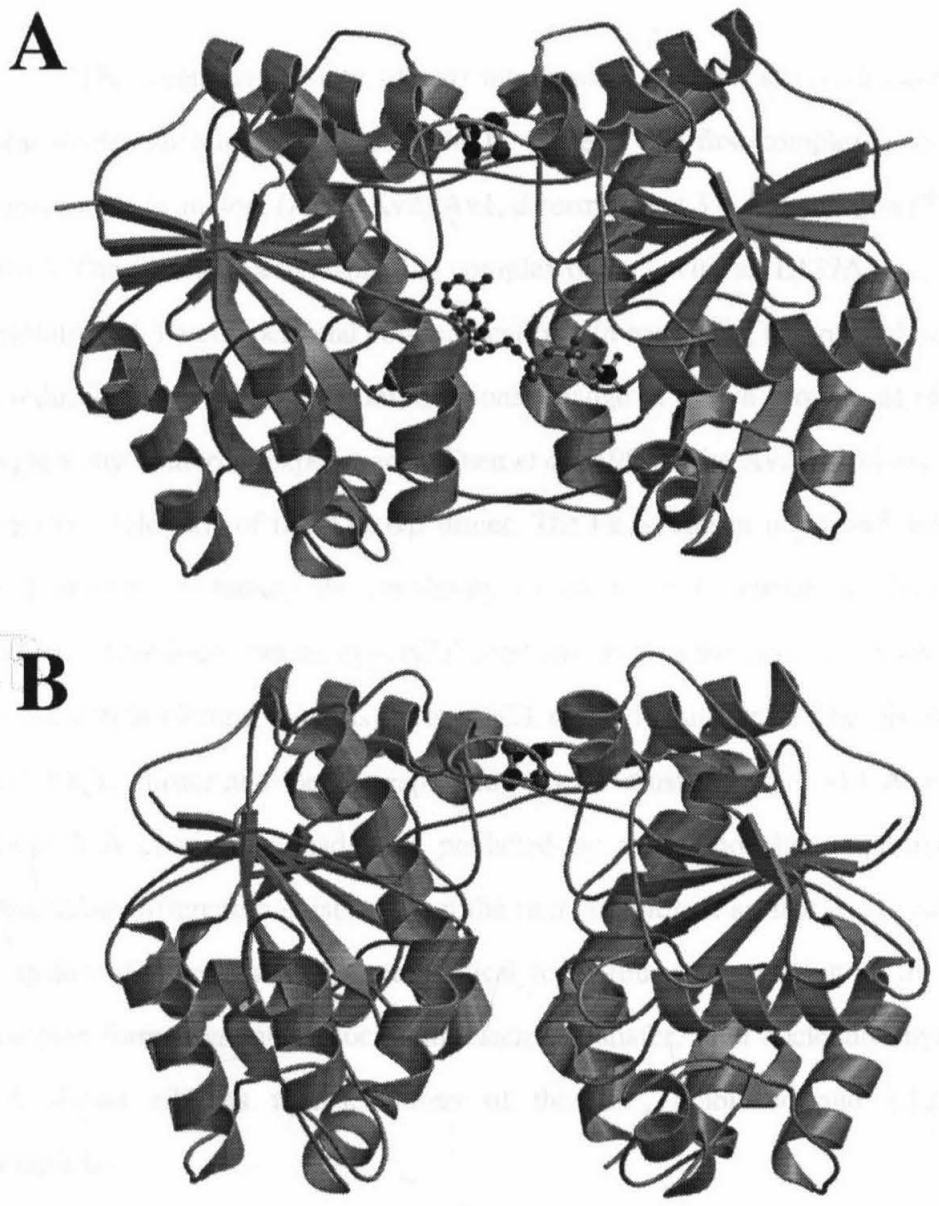


Figure 1-5. Ribbons representations of the Fe protein from (A) *A. vinelandii* and (B) *C. pasteurianum.* The Fe_4S_4 clusters for each model, and the MoO_4^{2-} and ADP molecules of the 2.9 Å resolution Av2 model, are depicted as ball-and-stick models. Coils represent α-helices and arrows β-strands. The two-fold axis relating the two monomers lies vertically in the plane of the page. The figures were generated using MOLSCRIPT/RASTER3D. Cp2 coordinates were generously provided by Debbie Woo, and Av2 coordinates were taken from the PDB (1NIP.pdb).

The recent availability of two nitrogenase complex structures illustrates some of these nucleotide binding effects on the Fe protein. The first complex is an AlF_4^- stabilized transition state analog, $(AlF_4^-•Av2)_2Av1$, determined at 3.0 Å resolution (Schindelin *et al.*, 1997). The second is a tight-binding complex of Av1 with an L127Δ Av2 mutant at 2.2 Å resolution (J. Peters, personal communication). In each case, the intersubunit space of Av2 is reduced due to a dramatic conformational change in the Fe protein, as observed in small angle x-ray scattering experiments (Chen *et al.*, 1994). The Av2 2-fold axis aligns with the pseudo 2-fold axis of the Av1 αβ dimer. The Fe_4S_4 cluster is pushed out away from the Av2 protein, explaining the sensitivity of its irons to chelators. Two $MgADP•AlF_4^-$ moieties were found bound in parallel conformations in the complex, homologous to those of nucleotide binding proteins such as p21 ras and transducin. The distance between the Av2 Fe_4S_4 cluster and the corresponding Av1 P-cluster pair is ~14 Å in each complex, about 5 Å closer than had been predicted by model docking experiments. However, observable differences exist between the two complexes, so that each provides a valuable snapshot of nitrogenase. Both are critical for a full understanding of the mechanisms of complex formation and dissociation, electron transfer, and nucleotide hydrolysis. Figure 1-6 shows ribbons representations of the AlF_4^- stabilized and L127Δ nitrogenase complexes.

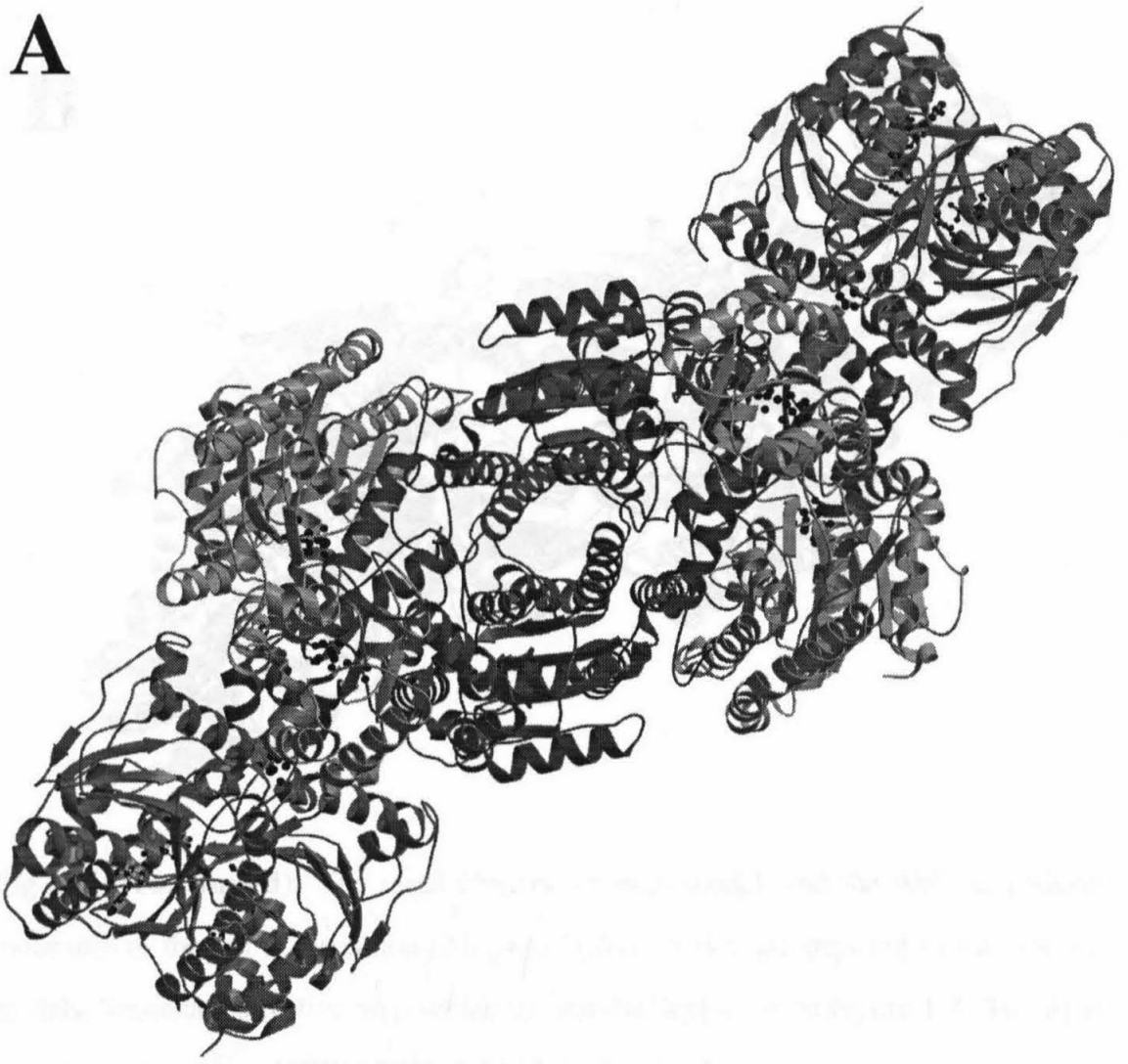


Figure 1-6. Ribbons representations of two *A. vinelandii* nitrogenase complexes. (A) An AlF_4^- stabilized transition state analog at 3.0 Å resolution. (B) A tight-binding complex of Av1 with an L127Δ Av2 mutant at 2.2 Å resolution. In each case, the intersubunit space of Av2 is reduced due to a dramatic conformational change, and the Av2 2-fold axis aligns with the pseudo 2-fold axis of the Av1 αβ dimer. The distance between the Av2 Fe_4S_4 cluster and the corresponding Av1 P-cluster pair is ~14 Å in each complex.

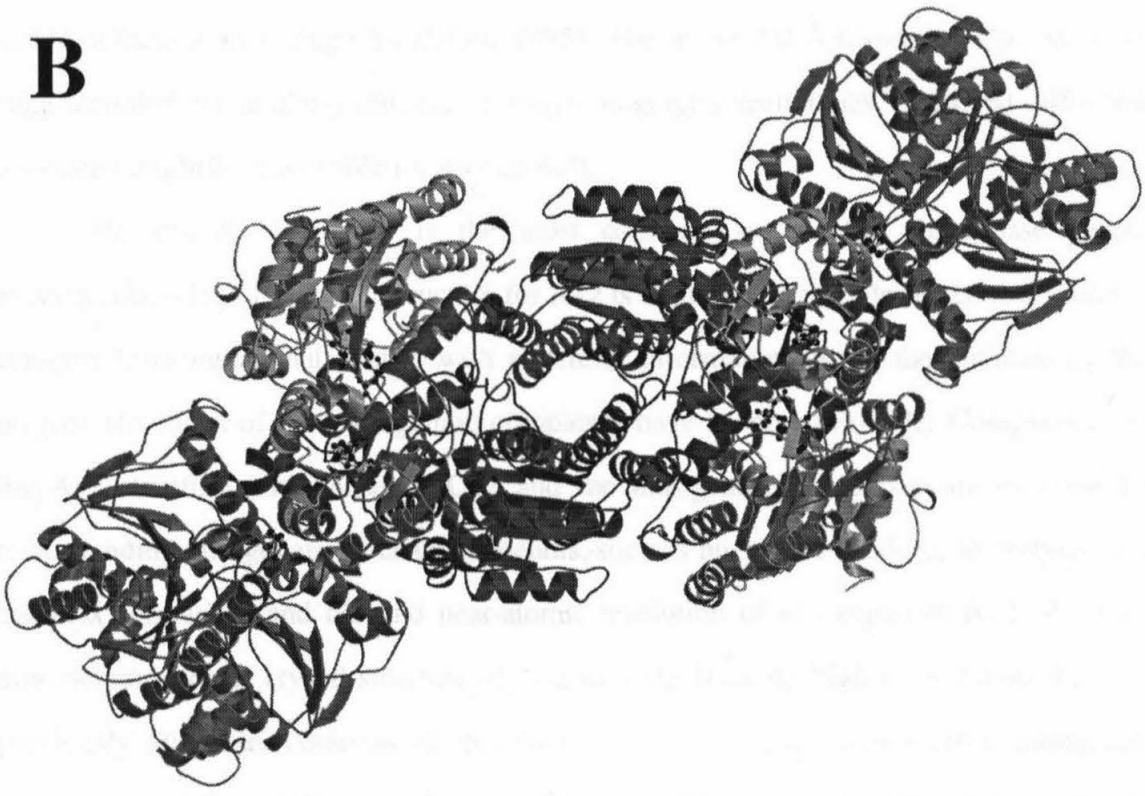


Fig 1-6 (continued). The metal clusters for each model, and the AlF_4^- and MgADP molecules of the 3.0 Å resolution $(AlF_4^-•Av2)_2Av1$ model, are depicted as ball-and-stick models. Secondary structure representations and shading are as in Figure 1-4. The figures were generated using MOLSCRIPT/RASTER3D. Coordinates were generously provided by Hermann Schindelin ($(AlF_4^-•Av2)_2Av1$) and John Peters (L127Δ complex).

1.7. Research Objective

Preliminary crystallographic work revealed the overall topology of Av2 and provided tremendous impetus for mutational studies. However, the 2.9 Å resolution of the previous model was insufficient for detailed mechanistic analysis of the structure which might explain the myriad functions of Fe protein. Av2 crystals are extremely radiation sensitive and subject to thermal diffuse scattering (Georgiadis, 1990). The advent of

cryocrystallography has minimized the problem of crystal radiation decay for a number of macromolecules, including Cp2 (Woo, 1995). The recent 2.0 Å resolution Cp2 structure suggested that, by similarly utilizing cryocrystallographic techniques, improved diffraction resolution might be achievable for Av2 as well.

Because *A. vinelandii* is the most common system for nitrogenase genetic experiments, a high resolution structure for Av2 is required to facilitate direct correlation of findings from mutational studies with structural information. Since the initiation of this project, structures of two nitrogenase complexes have become available. Comparisons of the Av2 structure with those of Cp2 and the nitrogenase complexes are essential for understanding nitrogenase mechanistic details, such as nucleotide binding, hydrolysis, and complex formation, and demand near-atomic resolution of uncomplexed Av2. We have now determined the crystal structure of Av2 to a significantly higher resolution than that previously available. Analysis of this structure and comparisons to other nitrogenase structures place the wealth of available biochemical, genetic, and spectroscopic information on nitrogenase into a structural context to improve our understanding of the structural basis of Fe protein function.

CHAPTER 2
Preparation of Av2

2.1. *Azotobacter vinelandii* Cell Culture

Azotobacter vinelandii is a free-living Gram-negative chemoheterotroph of the family Azotobacteraceae (Brock, 1970). This obligate aerobe was discovered in Vineland, NJ, for which it is named. *A. vinelandii*, generally motile water-dwelling organisms, are characteristically paired and contain peritrichous flagella (Stanier *et al.*, 1976). The cells, growing to lengths of 2 μm or more, are larger than most bacteria and resemble plump rods. *A. vinelandii* may exist as either vegetative cells or cysts; its G+C base pair content is ~66 % (Stanier *et al.*, 1976). The bacteria tend to be sensitive to acid pH values, high phosphate concentrations, and temperatures greater than 35 °C (Postgate, 1982). *A. vinelandii* has the one of the highest respiration rates of any known organism; this is particularly intriguing in light of the extreme oxygen sensitivity of its Fe protein (Brock, 1970). Healthy cells divide with a doubling rate of 3-4 hours under nitrogen fixing conditions in the laboratory.

2.1.1. Media Requirements

Liquid culture of *A. vinelandii* culture utilized a modified Burk's nitrogen-free media containing a mixture of: 1.7 x 10^{-2} M $MgSO_4$, 3.6 x 10^{-5} M $FeSO_4$, 2.1 x 10^{-6} M Na_2MoO_4, 9.0 x 10^{-4} $CaCl_2$, 7.75 x 10^{-3} M K_2HPO_4, 2.5 x 10^{-3} M KH_2PO_4, 3.6 x 10^{-5} M Na citrate, and 2 % w/v sucrose. Media for plates included addition of 2 % w/v agar. Stock solutions A (1.67 M $MgSO_4$), B (0.036 M $FeSO_4$, 0.002 M Na_2MoO_4, and 0.036 M Na citrate), C (0.9 M $CaCl_2$), and D (1.55 M K_2HPO_4, 0.49 M KH_2PO_4) were prepared in

advance, and used to facilitate media preparation. Stock solutions A, B, and C were pipetted into double distilled water in appropriate ratios to create unbuffered media; solid sucrose was added to the mixture. This media and stock solution D were sterilized separately by autoclaving. When the solutions cooled to ~30 °C, an appropriate aliquot of stock D was added to the media; this prevented formation of phosphate precipitates during autoclaving. Flasks of sterile media were placed in an incubator shaker (Series G25; New Brunswick Scientific, Edison, NJ) to bring them to ambient *A. vinelandii* growth temperature before inoculation with cells. Vigorous shaking in the incubator at ~275 rpm provided adequate aeration for the growing cells.

2.1.2. Cell Growth Protocol

Azotobacter vinelandii strain 13705 was obtained in lyophilized form from American Type Culture Collection (ATCC, Rockville, MD). The cell pellet was solubilized in a small volume of liquid culture media and grown at 28 °C. After 12 hours, cells from this stock were streaked onto plates with a sterile loop and stored in the incubator (without shaking). Two to three days were required for colonies to develop on these plates, after which cells were taken from the established growths and streaked onto new plates. Single colonies were obtained from second generation plates and started in 5 mL liquid cultures. Small volume liquid cultures were used to inoculate subsequently larger volumes. A dilution ratio of 1:4 was generally observed. Thriving bacteria were characterized by a fragrant "yeast"-like aroma and yellow-brown color, and a foaming "head" to the culture. Appearance of vivid green siderophores in the culture indicated deficiencies of iron in the media; these cultures, rarely observed, were discarded.

Four flasks of recent culture, each containing 1-2 L cell culture, were maintained until the designated start date for the fermentation run. The initial *A. vinelandii* cell growth involved fermentation in an 180 L fermentor at UCLA by Thomas Sutherland; subsequent

runs utilized a 300 L fermentor at Caltech. Fermentor operation at Caltech was performed by Larry Tudor. Initially, ~1 L cell culture was inoculated into a 10 L seed fermentor; the cells were subsequently transferred to the 300 L fermentor for final growth. Later fermentor runs involved direct transfer to the 300 L fermentor from the laboratory incubator stocks. Temperature ~28 °C and pH 7.0 conditions were maintained throughout the nearly 24 hour run. *A. vinelandii* cells were harvested in mid log phase, with OD typically 3.0 at 595 nm, using a Sharples continuous flow centrifuge. The length of time required to centrifuge the cells, complicated by the large volume of cells, probably prevented uniform bacterial growth phase. Harvested cell paste was frozen in liquid nitrogen, in one gallon plastic freezer bags (Ziploc®, DowBrands L.P., Indianapolis, IN), and stored in a freezer at -80 °C. Cell yields ranged from ~500 g to greater than 2 kg for a given run; typical values were 600-700 g.

2.2. Purification of Av2

The extreme oxygen sensitivity of nitrogenase proteins demands observance of strict anaerobicity throughout the purification and crystallization steps. Fe protein activity shows particular susceptibility to O_2 inactivation, with a half-life of about 45 seconds in air (Smith *et al.*, 1985). Consequently, buffers used during nitrogenase purifications were deoxygenated in the following manner. Ten 10-minute cycles of evacuation by high vacuum followed by flushing with Ar gas which had been passed over heated BTS catalyst (BASF, Paramus, NJ) to remove trace O_2 were repeated for each solution. A Schlenk-type vacuum/gas manifold with two Ar ballasts was used to control gas flow for all manipulations. Reducing conditions were maintained by the addition of solid $Na_2S_2O_4$ to each buffer after the seventh vacuum/Ar cycle to reach a final concentration of 2 mM dithionite. Empty glass containers for collecting and storing protein fractions anaerobically were sealed with Suba-Seal (Aldrich Chemical, Milwaukee, WI) rubber stoppers and

copper wire, and degassed with a minimum of six 3-minute vacuum/Ar cycles. Disposable Luer hub needles and 1 cc syringe tips provided the connections between the vacuum/Ar manifold lines and the rubber stoppers of the bottles and flasks. Positive pressure was maintained on solutions during all transfers, and outgassing was directed to a mercury trap. Protein fractions were stored in sealed containers in an anaerobic chamber under Ar/H_2 atmosphere at room temperature (Coy Laboratory Products, Ann Arbor, Michigan).

Trizma base (tris(hydroxylmethyl)aminomethane), DNase I (Type IV), and RNase I were from Sigma Chemical (Sigma Chemical, St. Louis, MO). Tris was titrated to desired pH using HCl. Ultra-pure (>99 %) $Na_2S_2O_4$ from Noah Technologies (San Antonio, TX) was used as received without further purification. All other chemicals used were of reagent purity or better. Distilled deionized water was used to prepare all solutions. Concentrated dithionite solution was prepared by degassing 5 ml 0.5 M Trizma base and 4.84 g $Na_2S_2O_4$ in separate sealed serum bottles (Wheaton, Millville, NJ). A disposable syringe and needle were used to transfer the Trizma to the dithionite-containing vial. The solution was shaken with additional cycles of degassing.

2.2.1. Initial Purification Methods

Initially, *A. vinelandii* nitrogenase components were purified according to a modification of the method of Burgess *et al.* (1980). The major features of this protocol include: (1) cell lysis by osmotic shock and centrifugation to separate soluble proteins from DNA and membranes; (2) a 56 °C heat step to denature undesired proteins; (3) linear salt gradient elution from a DEAE-agarose column to separate nitrogenase proteins into crude component fractions; (4) gel filtration of the Fe protein to remove trace MoFe protein and flavodoxin; (5) additional gradient elution from a DEAE-agarose column for MoFe protein purification; and (6) precipitation and subsequent resolubilization of MoFe protein. Both Fe

and MoFe proteins were purified in each prep. MoFe protein was used in assays, and for crystallographic studies independent of this manuscript.

2.2.1.1. Pre-Lysis Preparations

All buffers anticipated for the purification were prepared on the day preceding cell lysis, and stored at room temperature. In general, about 24 L Tris-Cl buffers were prepared for each prep; these buffers are summarized in Table 2-1. DEAE Sepharose CL-6B and Sephacryl 200 resins (Pharmacia Biotechnology, Piscataway, NJ) were rinsed with column loading buffer (50 mM Tris-Cl, 0.1 M NaCl, pH 7.75), then degassed with occasional swirling for several hours in stoppered Erlenmeyer flasks attached to the vacuum/Ar manifold. Empty glass Kontes columns (Vineland, NJ) were positioned and leveled near the manifold; they were equipped with column outlet fittings ending in stainless steel needles and checked for leaks. CL-6B resin was loaded into two 1 x 25 cm Kontes columns to resin bed heights of approximately 10 cm (Column 2,4), and into one 4 x 25 cm Kontes column to a resin bed height of 23 cm (Column 1). Additional Sepharose columns were packed for MoFe protein purification. Sephacryl 200 resin was loaded to a bed height of 50 cm in a 2.5 x 50 cm column (Column 3). In later preps, this height was increased to 100 cm. All resin beds were covered with a circle of filter paper (Whatman #2), topped off with buffer, and sealed with three-way stopcocks inserted into rubber stoppers. Parafilm was used to secure the seals. A summary of each column and its function is given in Table 2-2. Buffers required for the initial cell lysis steps were degassed as described in section 2.2. Column 1 was equilibrated with dithionite-reduced loading buffer until a 1 mL volume of column eluate failed to turn blue after addition of one drop of 2 mM methylene blue indicator dye. Column 1 was sealed by inserting the needle tip of its outlet fitting into a rubber stopper; a pair of pliers was used to insert or remove needles, with care taken to avoid crimping the needle.

Table 2-1. Buffers for Av2 purification protocol

Swelling Buffer (1 L)	0.05 M Tris-Cl, pH 8.2 0.1 M NaCl 40 % Glycerol Prepare cells for lysis
Cracking Buffer (1 L)	0.05 M Tris-Cl, pH 8.2 0.1 M NaCl 2 mM $Na_2S_2O_4$ 20 mg each DNase, RNase Lyse cells
Loading Buffer (8 L)	0.05 M Tris-Cl, pH 7.75 0.1 M NaCl 2 mM $Na_2S_2O_4$ Rinse resins; Equilibrate, Load Column 1
No Salt Buffer (4 L)	0.05 M Tris-Cl, pH 8.0 2 mM $Na_2S_2O_4$ Dilute protein fractions, Equilibrate, Load Columns 2,4
High Salt Buffer (8 L)	0.05 M Tris-Cl, pH 8.0 0.45 M NaCl 2 mM $Na_2S_2O_4$ Elute Column 2,4; Equilibrate, Load Column 3; Av2 storage and dialysis
Gradient Salt Buffer I (1 L)	0.05 M Tris-Cl, pH 8.0 2 mM $Na_2S_2O_4$ Elute Column 1 (mixed with Gradient Salt Buffer II)
Gradient Salt Buffer II (1 L)	0.05 M Tris-Cl, pH 8.0 0.6 M NaCl 2 mM $Na_2S_2O_4$ Elute Column 1 (mixed with Gradient Salt Buffer II)

Table 2-2. Chromatographic columns used in gravity-based Av2 purification protocol

Column 1. Anion-Exchange (DEAE-Sepharose CL-6B, 4 x 23 cm)

Equilibrate, Load with Loading (0.1 M NaCl) Buffer
Elute with 0.1 - 0.6 M NaCl Buffer Gradient

Isolates nitrogenase component proteins

Column 2. Anion-Exchange (DEAE-Sepharose CL-6B, 1 x 10 cm)

Equilibrate, Load with No Salt (0.0 M NaCl) Buffer
Elute with High Salt (0.45 M NaCl) Buffer

Concentrates Av2 fraction

Column 3. Gel filtration (Sephacryl S-200, 2.5 x 50(100) cm)

Equilibrate, Load with High Salt (0.45 M NaCl) Buffer

Removes Av1 and Av flavodoxin impurities

Column 4. Anion-Exchange (DEAE-Sepharose CL-6B, 1 x 10 cm)

Equilibrate, Load with No Salt (0.0 M NaCl) Buffer
Elute with High Salt (0.45 M NaCl) Buffer

Concentrates Av2 fraction

2.2.1.2. Cell Lysis

About 250-300 g frozen *A. vinelandii* cell paste was placed in a 1 quart plastic freezer bag (Ziploc®, DowBrands L.P). The bag was placed in a bath of warm water to quick-thaw the cells. About 600 mL swelling buffer (40 % glycerol, 50 mM Tris-Cl, 0.1 M NaCl, pH 8.2) was added to the cells; the bag was kneaded to thoroughly combine the mixture, and stored on ice for 15 minutes. The cells were transferred to three 250 mL polycarbonate centrifuge bottles (Beckman Instruments, Palo Alto, CA) outfitted with Delrin (E. I. du Pont de Nemours, Wilmington, DE) tops modified to accommodate butyl rubber septa, and loaded into a JA-14 rotor in a Beckman Instruments J2-21M centrifuge (Palo Alto, CA). Centrifugation proceeded for 30 minutes at 9000 rpm at 4 °C. At this

time, 20 mg each DNase and RNase, as well as 0.35 g solid $Na_2S_2O_4$, were added to the degassed cracking buffer, and additional cycles of degassing were performed. The supernatant of each bottle was decanted, and three glass marbles were inserted into each bottle (the marbles were used to scrape cells from the centrifuge bottle walls in a later step). The centrifuge bottles were sealed, and a clean disposable needle attached to a manifold line was inserted into each butyl rubber stopper. The *A. vinelandii* cells were degassed thoroughly (six 3-minute vacuum/Ar cycles); needles were changed when they became clogged with cells. Cracking buffer (50 mM Tris-Cl, 0.1 M NaCl, pH 8.2) was rapidly added to the degassed cells using a cannula and tubing; a 0.5 mL portion of 0.1 M $Na_2S_2O_4$ was injected into each bottle. Each centrifuge bottle was shaken vigorously to dissolve all cells in the buffer, resulting in a creamy brown solution; this process was facilitated by the presence of glass marbles in the cell paste. The cells were stored 30 minutes on ice, then centrifuged 40 minutes at 12 000 rpm at 4 °C.

During this time, a 1 L two-necked round-bottom flask containing an egg-shaped stirbar was sealed with one butyl rubber stopper and one thermometer inserted into a rubber stopper. The flask was placed in a heating mantle attached to a Variac and degassed. The centrifuge pellet contained a colorful array of bands corresponding to the cell membranes, nucleic acids, and other insoluble materials. The nearly black supernatant of each bottle was transferred to the degassed flask using a cannula. A 0.5 mL addition of 0.1 M $Na_2S_2O_4$ was made to the solution crude protein extract, and the flask was gradually heated with stirring to 56 °C (the temperature was reduced to 54 °C in later preps to avoid destroying valuable Fe protein). The flask was removed from the heat source and allowed to cool for 5 minutes before the extract was cannulated into clean degassed 250 mL centrifuge bottles and subsequently centrifuged at 9000 rpm for 20 minutes at 4 °C. The brown-black supernatant of each bottle was transferred by cannula to a degassed 1 L round-bottom flask; a grey precipitant remained in the centrifuge bottles.

2.2.1.3. Chromatographic Separations

The crude protein extract was diluted two-fold with no salt buffer (50 mM Tris-Cl, pH 8.0) and loaded onto pre-equilibrated Column 1 using a cannula and tubing. The loaded Sepharose CL-6B column was washed with 4-5 column volumes of loading buffer until the eluate appeared colorless. The MoFe protein was generally observable as a large brown band (with faint red coloring) at the top of the column; the Fe protein was sometimes visible as a small brown band just below the MoFe band. A black-purple color, in proportional to the age of the frozen cells, often covered much of the resin bed and obscured the bands. A set of degassed gradient buffer bottles was attached to the top stopcock port of Column 1, and a 0.1-0.6 M NaCl gradient was begun to elute the nitrogenase proteins. Band detection was performed by eye, and by monitoring the eluate conductivity using a digital conductivity meter (Model 604, Amber Science, Eugene, OR). The MoFe protein eluted as the third band in the series (c_i = ~1.6 x 10^4 µS/cm), while the olive-brown Fe protein followed as the next major band (c_i = ~2.3 x 10^4 µS/cm). The bands representing the MoFe and Fe proteins were often overlapping, as were those of the Fe protein and the flavodoxin which followed it. Several other brown proteins, mainly ferredoxins, eluted towards the end of the gradient. All fractions were collected manually in degassed sealed bottles, and stored in the anaerobic chamber until needed.

Column 2, a 1 x 10 cm Sepharose CL-6B column used for concentrating Av2, was pre-equilibrated with no salt buffer, and Column 3, a 2.5 x 100 cm Sephacryl 200 gel fitration column, with high salt buffer (50 mM Tris-Cl, 0.45 M NaCl, pH 8.0). The Fe protein fraction was diluted three-fold with no salt buffer and loaded onto Column 2. The fraction was washed with no salt buffer, and the headspace over the resin bed with filled with Ar. Step elution with high salt buffer released the concentrated Fe protein fraction, which was collected into a degassed flask containing ~0.9 g sucrose. The Fe protein was carefully loaded directly onto the gel filtration bed by carefully inserting a cannula through

the stopcock assembly of Column 3. Use of the cannula and sucrose assured a tight protein band, to minimize resolution losses from band spreading as the proteins migrated through the column. Despite these precautions, complete separation of the yellow flavodoxin from the Fe protein seldom occurred. The Fe protein fraction was concentrated in Column 4, a 1 x 10 cm Sepharose CL-6B column, by repeating the steps for Column 2.

2.2.1.4. Protein Storage and Evaluation of Protocol

Purified Av2 was stored in the anaerobic chamber, and prepared for crystallization trials by simultaneous dialysis and concentration into fresh high salt buffer using a Micro-Prodicon apparatus outfitted with a degassed 25,000 MWCO cellulose acetate membrane (Spectrum, Houston, TX). Av2 prepared in this way could be stored at room temperature for several days. The quality of crystals grown from this protein rapidly diminished after the first day. Suitable means of freezing Av2 for later crystallizations had not been determined at this time. The inadequacies of the 1 meter Sephacryl 200 column resulted in impure Av2, as detected by denaturing gel electrophoresis. An added shortcoming of this purification scheme was the excessive length of the prep. Maintaining reasonable flow rates with the vacuum/Ar manifold was unpredictable at best, and led to grueling preps of five days and nights. In addition, the reliance on human eyes to discern protein bands by color was fraught with imprecision. Finally, the stopcock fittings of the Kontes gravity columns were unreliable and, despite the use of Parafilm sealant, would unseal due to backpressure within the columns, resulting in lost time and protein.

2.2.2. Final Purification Methods

A more efficient approach to nitrogenase preparation was developed by taking advantage of several fortuitous arrivals at the laboratory. A used Pharmacia FPLC system

(Pharmacia Biosystems, Piscataway, NJ), donated by Amgen Pharmaceuticals (Thousand Oaks, CA), was refurbished and plumbed with PEEK (polyethyletherketone) tubing and Delrin ferrules and fittings (Upchurch Scientific, Seattle, WA). Recommended by Prof. Mike Adams of the University of Georgia and Prof. Jim Howard of the University of Minnesota, PEEK tubing was selected because of its demonstrated superior resistance to O_2. A manual four-port switching valve (Model C4UW; Valco, San Antonio, TX) constructed of dithionite-resistant Halcelloy N stainless steel was placed in line after the buffer mixer. Empty AP (Waters Corporation, Waltham, MA) and XL (Pharmacia) series glass columns replaced the Kontes columns, although Sepharose CL-6B resin was still employed. A pre-packed Superdex 200 HR 26/60 column (Pharmacia) replaced the Sephacryl 200 column and provided excellent Av2 (and Av1) resolution. These columns featured a more rugged construction and allowed columns to be prepared in advance and left unattended with confidence. Two-liter round-bottom flasks and adapters were fashioned by Rick Gerhart in the Caltech Glassblowing Facility. The Rotavisse joints (ChemGlass) of these buffer flasks provided superior seals to those used in previous purifications, and the unique adapter design allowed the flask to be connected as an intermediate between the manifold and FPLC pump. Delrin inserts for the polycarbonate centrifuge bottles were redesigned to improve the seal of the bottles, and constructed by Guy Duremberg of the Caltech Machine Shop. Finally, and perhaps most importantly, Hermann Schindelin arrived as a postdoc in the laboratory and shared nitrogenase purification tasks.

Nitrogenase purification proceeded basically as described in 2.2.1-2.2.3. Jim Howard provided valuable guidelines for converting the protocol from gravity-based to FPLC-driven chromatography. With the exception of the Superdex 200 column, all columns were freshly packed for each prep. The cell lysis and heat step remained essentially unchanged. Loading of crude protein extract onto Column 1 attached to the FPLC was accomplished using a peristaltic pump (Millipore Corporation, Waltham, MA).

The pump's silicon tubing was adapted with a stainless steel cannula to pierce the rubber septum of the protein flask. The FPLC mixer controlled the elution salt gradient. Protein elution was monitored at 405 nm on a UV-vis monitor (Pharmacia); protein fractions, however, were still collected manually into degassed sealed serum bottles. Subsequent columns, analogous to those of the initial protocol, were run in a similar manner. Sucrose was no longer required before loading Av2 onto the gel filtration column. Av1 quality was also improved by passing it over the Superdex 200 column.

One vital improvement in the protocol involved the use of glycerol to stabilize the purified Av2. Glycerol's widespread role in stabilizing protein has long been established, and its benefits in protein crystallization are also well-known. Glycerol had been found to be instrumental in stabilizing purified Cp2 and improving crystal growth. Purified Av2 was subsequently dialyzed and concentrated into a buffer of 20 % glycerol, 50 mM Tris-Cl (pH 8.0), 0.45 M NaCl, and 2 mM $Na_2S_2O_4$. The procedure was performed on a 2 mL sample of concentrated Av2 in the anaerobic chamber over a period of 4 hours, using the Micro-Prodicon apparatus. Av2 in glycerol buffer remained stable for nearly one month at room temperature, and crystallized as described in Chapter 3.

These modifications affected favorably the nitrogenase purification scheme in several ways. Improvements in centrifuge bottle insert, column, and glassware design lowered the incidence of lost protein due to oxygen contamination. The use of an FPLC significantly reduced the amount of time required for each prep, as P-500 FPLC pumps assured constant flow rates. With two people contributing towards the prep effort, purified Av2 and Av1 could be isolated from cells and characterized within 2-3 long days. The availability of a UV-vis monitor allowed quantitative tracking of protein elution. Chromatographic resolution and reproducibility increased dramatically, and protein quality improved. Av2 and Av1 were easily separated with the initial gradient; the superior resolution of the Superdex 200 column completely removed flavodoxin from Av2. A representative PAGE gel of fractions collected from the DEAE-Sepharose gradient column is shown in Figure 2-

1 (Laemmli, 1970). Finally, the inclusion of 20 % glycerol in the final Av2 buffer increased dramatically both protein stability and reproducibility in crystallization.

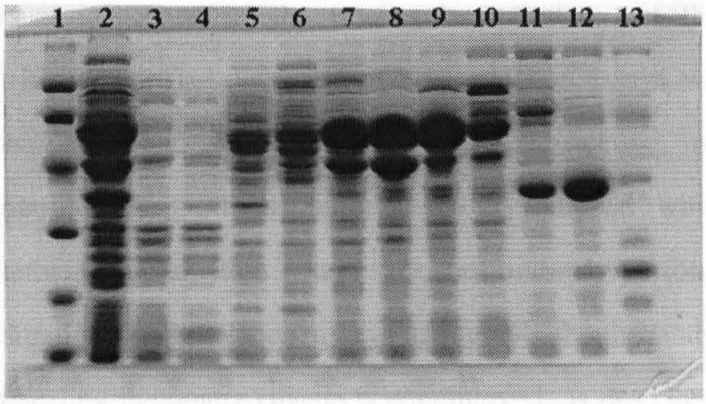


Figure 2-1. PAGE gel of fractions collected from DEAE-Sepharose CL-6B gradient column. Buffer mixing and flow rates were controlled by a Pharmacia FPLC. The 32 kDa Av2 band in Lane 12 was used for subsequent purification and crystallization that led to crystals used in the Av2 structure determination. Lanes 3-6 show bands for the MoFe protein α and β subunits at ~60 kDa. Crude extract as loaded onto the column is shown in Lane 2, and gel markers (Pharmacia, LMW Marker Kit) appear in Lane 1. The marker bands correspond to: phosphorylase b (94 kDa); albumin (67 kDa); ovalbumin (43 kDa); carbonic anhydrase (30 kDa); trypsin inhibitor (20.1 kDa); and α-lactalbumin (14.4 kDa). The gel was stained with Coomassie blue stain.

2.3. Protein Evaluation

Purified nitrogenase proteins were evaluated in a number of ways. Protein concentration was measured by optical absorbance at 420 nm and 430 nm, respectively, for the Fe and MoFe proteins. Optical measurements were made on a Beckman

spectrophotometer. Extinction coefficients are ε_{420}= 7.3 M^{-1} cm^{-1} for Av2 and ε_{430}= 13.8 M^{-1} cm^{-1} for Av1. Typical concentrations for Av2 used in successful crystallization trials were 25-30 mg/mL. SDS PAGE was performed on fractions from each column to monitor protein purity; visual inspection of the final gels indicated that the Av2 used for crystallization was highly purified. Qualitative C_2H_2 reduction assays were performed on cultured *A. vinelandii* cells, and at various stages of the purification, using a Porapak N column in a Varian Model 3700 gas chromatograph equipped with a Model 4270 Integrator. Quantitative C_2H_2 reduction assays were not routinely performed. Av2 specific activity was measured by Fe chelation assays using *o*-bathophenanthroline. A double septum method of sealing the quartz cuvette was employed, as described by Howard and coworkers (Deits and Howard, 1989). The most successful chelation experiment showed 92 % active Fe protein. Lower values in previous assays may have reflected deficiencies in anaerobic technique rather than poor protein quality.

CHAPTER 3
Crystallization and X-Ray Analysis of Av2

3.1. Av2 Crystallization

The original Av2 crystals were obtained more than a decade ago by batch method dialysis using Zeppenzauer cells (Rees and Howard, 1983). These crystals belonged to space group $P2_12_12$ and diffracted to limited resolution. Millie Georgiadis first determined the structure of *Azotobacter vinelandii* Fe protein to 2.9 Å resolution using a second crystal form belonging to space group $P2_1$ with cell constants a = 56.84 Å, b = 90.29 Å, c = 63.50 Å, β=100.02 ° (Georgiadis *et al.*, 1992). The structure was solved by SIRAS methods using a gold myochrysine derivative. The goal of the present work was to extend the resolution of the 2.9 Å resolution model in order to gain additional insights into the structural basis of Fe protein function. Initial crystallization attempts focused on reproducing the $P2_1$ crystals.

3.1.1. Preparation of Precipitant Solutions

Av2 crystallization solutions were made from stock solutions which had been prepared using distilled deionized water and chemicals of reagent purity or better. GC grade PEG 4000 from E. M. Science (Gibbstown, NJ) and 99 % pure Na_2MoO_4 (Fluka Chemical Corp., Ronkonkoma, NY) were used. Stock PEG 4000 was prepared as a 50 % w/w solution. Each stock solution was filtered through a 0.2 μm sterile filter to remove particulates. Nucleotides and nucleotide analogues were purchased from Boehringer Mannheim (Indianapolis, IN). Crystallization solutions were pipetted to a final volume of either 1.0 mL or 2.0 mL, in glass serum bottles (Wheaton, Millville, MA) of 2.0 mL or 3.5

mL capacity, respectively. Glass serum bottles were prepared by first scrubbing with a dilute detergent solution, rinsing, then autoclaving for 1 hr in very dilute detergent solution. Cooled bottles were rinsed 5-6 times with distilled deionized water, then dried in an oven. In initial crystallization trials, serum bottles were plugged with rubber sleeve septa and secured around the neck of each bottle with 22 gauge copper wire to prevent septa from popping off during vacuum transfer to the anaerobic chamber. Later, the use of butyl rubber septa and aluminum ring seals fastened by a manual crimping device (Wheaton) provided a more efficient means of sealing the serum bottles. Sealed bottles were individually held on a tabletop mixer for ~5 s to thoroughly mix the solutions.

The sealed serum bottles were attached to a degassing manifold using disposable 20 gauge needles inserted through the rubber septum of each bottle. The manifold accommodated a maximum of 24 bottles at one time. This manifold was connected to the main Schlenk-type vacuum/Ar manifold used in protein purification. Degassing proceeded through the use of an automatic switching device (Autosuk, v. 1 by Craig Steele) which provided cycles of 45 s vacuum and 15 s Ar to the bottles; typically, 3 sets of 10 cycles were performed on each series of bottles. Effective degassing was detected by bubbling within the crystallization solutions during the vacuum cycles. Additional mixing was supplied with a tabletop mixer to individual bottles during vacuum cycles to promote degassing. Needles frequently clogged during the degassing procedure, and were replaced as needed. Monitoring losses accrued by this technique revealed ~12 % loss of aqueous solution after 30 cycles. A 2.0 mL serum bottle was filled with 1.0 g water and degassed as part of the series of solutions; after the procedure, the bottle contents weighed 0.88 g. Thus, a significant fraction of the crystallization solutions were evaporated by the degassing protocol. Because of the varying viscosities of the solutions, volume losses were expected to differ in each bottle, and crystallization solutions were no longer accurate absolutely or with respect to one another. These inacccuracies may have contributed to inconsistencies in crystal production.

To circumvent the volume losses inherent in degassing small volumes, a new strategy was devised for preparing Av2 crystallization solutions. Large volumes (typically 25-100 mL) of stock solutions were degassed with stirring (using a magnetic stirbar and stirplate) on the vacuum/Ar manifold. Degassed solutions were transferred to the anaerobic chamber through a manual interlock by pumping down the interlock three times to a vacuum of ~27 mm Hg and refilling with a 10 % H_2 / balance Ar gas mixture (Matheson Gas Products). This gas mixture also provided the environment for the anaerobic chamber (residual O_2 was removed from the chamber atmosphere by palladium catalysts, which were regenerated 1-2 times per week by heating to ~160 °C for 90 minutes outside the chamber). Inside the anaerobic chamber, crystallization solutions were pipetted into serum bottles from the stock solutions and sealed as before. The use of solutions which had been pipetted within the anaerobic chamber resulted in more reproducible crystallizations. While the absolute concentration of each solution component may have differed from the original solution due to volume losses from stock solution degassing, the relative concentrations remained constant. Moreover, the availability of large volumes of degassed stock solutions in the anaerobic chamber provided consistency of solution composition between protein preparations. Crystallization solutions containing Na_2MoO_4 were found to be unstable over a period of weeks, requiring that fresh solutions be prepared after each protein purification.

3.1.2. Initial Conditions

The $P2_1$ crystals used to solve the structure of Av2 at 2.9 Å resolution were grown using a liquid-liquid diffusion technique inside the anaerobic chamber. This method, which was also utilized in the present work, involves the layering of precipitant followed by protein solutions into a glass melting point capillary, leaving a small gap between the solutions. Two dedicated 50 mL Hamilton gas tight syringes were used to transfer the solutions. The capillaries were sealed with molten wax; later, each wax seal was coated in

5-minute epoxy and hung inverted to allow the epoxy to polymerize. Epoxy-sealed capillaries were removed from the anaerobic chamber and spun in a clinical centrifuge for ~30 s to begin solution mixing, then returned to the anaerobic chamber and stored upright.

To prepare capillaries for use in crystallization trials, the following protocol was used. Open-ended 1.5-1.8 mm OD melting point capillaries were soaked in dilute detergent solution, rinsed with distilled deionized water, and dried in an oven. The cleaned capillaries were soaked in a silanization solution (Aquasil, Pierce, Rockford, IL) as recommended by the manufacturer, dried in the oven, rinsed a minimum of 5-6 times with distilled deionized water, and dried again in the oven. Finally, capillaries were sealed on one end using a methane-oxygen blowpipe. Silanization of the capillaries was believed to aid in crystal recovery and mounting by reducing the tendency of crystals to adhere to the inner walls of the capillaries.

The use of this liquid-liquid diffusion method for Fe protein crystallizations had several advantages. First, capillaries stored upright required little storage space when compared with the Linbro trays commonly used in hanging or sitting drop methods. This was a strong consideration in light of the limited dimensions of the anaerobic chamber. Also, capillary trials were relatively facile and quick to set up in the anaerobic chamber. The large volumes of solution employed in these capillary trials (~45 μL per trial) kept the crystals bathed in mother liquor during the crystal mounting procedure. This was important to counterbalance the dessicating atmosphere of the anaerobic chamber and the not insignificant challenge of mounting crystals in the anaerobic chamber. However, the large volume of protein required for each capillary (~15 μL) severely limited the number of trials which could be attempted from each protein preparation.

The $P2_1$ crystallization conditions contained a mixture of 30 % PEG 4000 (v/v), 0.1 M Tris-Cl, pH 8.3, and 0.10-0.22 M Na_2MoO_4, in addition to Av2 in 50 mM Tris-Cl, pH 8.0, buffer which also contained 0.45 M NaCl and 2 mM $Na_2S_2O_4$. Crystals grew as brown hexagonal rods. These conditions served as the starting point for the present

studies. Crystals were readily reproduced under these conditions, when protein quality was sufficient. However, diffraction beyond 2.8 Å resolution was not observed for these crystals. PEG 4000 concentration was varied between 20 % and 40 % without improvement in crystal quality. Variation of the ratio and volumes of protein and precipitation, and diameter of the capillary, similarly failed to produce superior crystals.

3.1.3. Optimization of Conditions

Crystals of Av1 had been obtained previously in the laboratory by Jongsun Kim using PEG 4000 and molybdate conditions nearly identical to those for Av2. Isomorphous crystals of Av1 had been grown from PEG 4000 combined with either $MgCl_2$ or CsCl. These latter conditions were tested for Av2. Crystallization attempts with 0.28-0.32 M $MgCl_2$ and 25-30 % PEG 4000 yield occasional needles of Av2, but no three-dimensional crystals. Crystals were obtained under conditions of 0.50-0.60 M CsCl and 20 % PEG 4000. These crystals diffracted only to ~3.5 Å resolution, and were not pursued for structural studies. Attempts to discover additional crystallization conditions with the use of an incomplete factorial were unsuccessful (Jancarik and Kim, 1991).

The availability of sufficiently pure Av2 posed a considerable obstacle to successful crystallization. Improvements in protein purification methods as described in 2.2.2.1, and in solution preparation (3.1.1) greatly improved the reproducibility of Av2 crystals. However, the most substantial crystallization breakthrough involved the stabilization of Av2 by the addition of glycerol in the protein buffer. The role of glycerol in crystal formation may be manifold-- it has been suggested that glycerol not only contributes to protein solvation to avoid phase separation, but also reduces the amount of water available for solvation by interacting directly with water (Sousa and Lafer, 1990). The concentration of glycerol in the precipitant solutions was varied between 0-10 %. Identical trials were made of Av2 which had and had not been dialyzed into glycerol-containing buffer. The

crystals ultimately used in this study contained glycerol in the protein buffer, but not in the precipitant solution; for these capillaries, $[glycerol]_{eq}$ approached 4 %. Apparently, the amount of glycerol in the protein buffer was sufficient to promote growth of crystals of superior diffraction quality. Interestingly, 5-10 % glycerol included in precipitant solutions failed to produce suitable crystals when mixed with Av2 which had not been dialyzed into glycerol-containing buffer. Whether this result reflected insufficient levels of glycerol to sustain crystal growth or poor protein quality is unknown. Glycerol had the added benefit of making Av2 crystals amenable to flash-cooling for cryocrystallographic data collection.

The crystals used in this project were grown from Fe protein of 25-30 mg/mL concentration (as determined by absorbance at 420 nm and Fe chelation assay with OBP), in a buffer of 20 % glycerol, 0.45 M NaCl, 2 mM $Na_2S_2O_4$, and 50 mM Tris-Cl, pH 8.0. The precipitant solutions included 25-30 % PEG 4000, 0.1 M Tris-Cl, pH 8.0, and 0.12-0.20 M Na_2MoO_4. Mature crystals were observed after several months. Crystals had not been observed after one month, at which point the capillaries, sealed in a 50 mL centrifuge tube, were stored lengthwise outside the anaerobic chamber. Whether this unorthodox storage method affected crystal formation and growth is unknown. Crystals generally measured a maximum of 0.2 x 0.2 x 0.5 mm^3, and appeared as brown rectangular prisms. A listing of the conditions used for this protein preparation is shown in Table 3-1.

Table 3-1. Crystallization conditions for Av2, 1-20.

NAME	Jamie Schiessman
DATE	#######
EXPERIMENT	Av2 FPLC-4 prep (1-20)
VOLUME (µL)	1000.0 microliters
STARTING #	1

CONCENTRATIONS				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20				
ID	[STOCK]	U	REAGENT	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	C1	C2	C3	C4	C5	C6	D1	D2	D3	D4	D5	D6
0	100.00	%	Water																								
1	50.00	%	PEG 4000	25.00	25.00	25.00	25.00	25.00	30.00	30.00	30.00	30.00	30.00	25.00	25.00	25.00	25.00	25.00	30.00	30.00	30.00	30.00	30.00				
2	1.00	M	Tris, 8.0	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10				
3																											
4	2.00	M	Na2MoO4	0.12	0.14	0.16	0.18	0.20	0.12	0.14	0.16	0.18	0.20	0.12	0.14	0.16	0.18	0.20	0.12	0.14	0.16	0.18	0.20				
5	2.00	M	CaCl																								
6	100.00	mM	Na2S2O4																								
7	100.00	%	Ethyl. glycol																								
8	100.00	%	Glycerol											5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00				
9																											
10																											
11																											
12																											
13																											
14																											
15																											
16																											
17																											
18																											
19																											
20																											
21																											
22																											
23																											
24																											
25																											
26																											
27																											
28																											
29																											
30																											

VOLUMES				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20				
ID	TOTAL		REAGENT	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	C1	C2	C3	C4	C5	C6	D1	D2	D3	D4	D5	D6
0	4900.0	100.0%	Water	340.0	330.0	320.0	310.0	300.0	240.0	230.0	220.0	210.0	200.0	290.0	280.0	270.0	260.0	250.0	190.0	180.0	170.0	160.0	150.0				
1	11000.0	50.0%	PEG 4000	500.0	500.0	500.0	500.0	500.0	600.0	600.0	600.0	600.0	600.0	500.0	500.0	500.0	500.0	500.0	600.0	600.0	600.0	600.0	600.0				
2	2000.0	1.0M	Tris, 8.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0				
3																											
4	1600.0	2.0M	Na2MoO4	80.0	70.0	80.0	90.0	100.0	80.0	70.0	80.0	90.0	100.0	80.0	70.0	80.0	90.0	100.0	60.0	70.0	80.0	90.0	100.0				
5		2.0M	CaCl																								
6		100.0mM	Na2S2																								
7		100.0%	Ethyl. gly																								
8	500.0	100.0%	Glycerol											50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0				
9																											
10																											
11																											
12																											
13																											
14																											
15																											
16																											
17																											
18																											
19																											
20																											
21																											
22																											
23																											
24																											
25																											
26																											
27																											
28																											
29																											
30																											

3.1.4. Crystallizaton Attempts with Nucleotides

MgATP hydrolysis plays a critical role in nitrogenase function. Sequence and structural comparisons between the 2.9 Å resolution Av2 model and those of nucleotide-bound proteins such as p21 ras and transducin suggested that the nucleotide binding mode for Fe protein would be similar to those of other nucleotide binding proteins. This "parallel" mode of binding was confirmed recently in the 3.0 Å resolution structure of the AlF_4^- stabilized nitrogenase complex. The structure contained two MgADP•AlF_4^- moieties per Av2, one bound to the P-loop region of each monomer in a transducin-like manner. A crystal structure of Fe protein with parallel-bound nucleotide in the absence of MoFe protein, however, remains elusive.

Ironically, a partially-ordered ADP molecule was described in the 2.9 Å resolution structure (Georgiadis *et al.*, 1992). This single endogenous ADP, believed to have copurified with Av2, spanned the intersubunit space. A molybdate anion in each monomer, introduced by the crystallization solution, corresponded to the β-phosphate position of the nucleotide in the AlF_4^- complex structure. The potential physiological significance, if any exists, of this cross-bound ADP remains unproven and no homologous binding has been observed in monomeric nucleotide binding proteins. Previous attempts to produce the parallel-mode Av2-nucleotide complex, either by cocrystallization or by soaking $P2_1$ Av2 crystals in an ADP-containing synthetic mother liquor, failed. Cocrystallization trials produced precipitate but no crystals. Soaking Av2 crystals in nucleotide-containing solutions generally resulted in cracked crystals. The dramatic conformational changes in Fe protein known to accompany nucleotide binding may help to explain these observations.

The present studies included new attempts to produce crystals of Av2 and nucleotide. For cocrystallization experiments, the following protocol was developed. Identical trials were set up using ATP, ADP, ATPγS, AMP-PNP, and AMP-PCP. Solid nucleotide was weighed into a 2.0 mL serum bottle, sealed, and degassed. Inside the

anaerobic chamber, a 0.5 mL aliquot of a degassed solution containing 75 mM $MgCl_2$ and 1.0 M Tris, pH 8.0 was transferred to each bottle; final nucleotide concentration in these solutions was 75 mM. Buffer strength had been selected to maintain pH 8.0 after mixing with the acidic nucleotides. Aliquots of 140 µL Av2 (25-30 mg/ml) were placed in serum bottles with 10 µL of nucleotide solution and allowed to equilibrate for 15 minutes. By mixing a small volume of concentrated nucleotide with the protein, the protein was diluted only minimally, with the expectation that the nucleotide-bound protein would crystallization under conditions used for the native protein. The nucleotide-protein mixture was set up in melting point capillaries as described in 3.1.2. These capillaries produced precipitate, but no crystals.

Failure to crystallize Av2 in the presence of nucleotide may be attributable to a number of factors. First, the number of crystallization conditions which could be tested was severely limited by the quantities of protein available. Also, the experiment utilized Av2 which had not been dialyzed into glycerol buffer, having been performed before the benefits of glycerol for Av2 crystallization were realized. The nucleotide concentration, ~5 mM in the protein-nucleotide mixture, may not have been sufficient to maintain nucleotide binding to Av2 after dilution with the crystallization solution, as nucleotides were not included in the precipitant solution. Final nucleotide concentration in the capillaries was calculated to be ~1.7 mM.

Attempts to soak nucleotides in the $P2_12_12_1$ crystals were also made. A 3 mm length of glass tubing, containing the epoxy and wax seals, was cut from the desired capillary after scoring with a carbide-tip pencil. A gas-tight Hamilton syringe was used to deliver a 2 µL aliquot of concentrated nucleotide to the top level of liquid in the capillary. The capillary was resealed with wax and stored horizontally in the anaerobic chamber. Horizontal storage of the capillaries was expected to promote gradual introduction of nucleotide to the crystal. Solutions containing nucleotide either in the presence or absence of Mg^{+2} were utilized. In most attempts, the crystals dissolved in less than 24 hours; in others the crystals cracked

visibly. In only one attempt, using buffered 75 mM ADP (no Mg^{+2}), did the crystals neither crack nor dissolve. A 3.0 Å resolution data set was collected from this crystal under cryocrystallographic conditions. Analysis of electron density maps calculated from this data set indicated no presence of nucleotide in Av2. However, the data set, collected on an in-house R-AXIS IIC using CuKα radiation at λ = 1.54 Å, was useful for providing anomalous signal for the Fe_4S_4 cluster to confirm the correctness of the molecular replacement solution from the 2.2 Å resolution data set, as described in 4.2. Details of the data collection and statistics are provided in 3.2.4.

3.2. Crystal Mounting and Data Collection

A variety of approaches for mounting Av2 crystals was explored with varying success. Techniques for preparing crystals for room temperature and cryocrystallographic data collection are described below, and a summary of the data processing results is provided.

3.2.1. Room temperature Crystal Mounting

Crystals to be used in room temperature data collection were mounted in the anaerobic chamber, which was equipped with a microscope. Thin quartz x-ray capillaries (Charles Supper Co., Natick, MA) slightly larger in diameter than the crystal were employed; similar glass capillaries, while preferable for minimizing absorption effects during data collection, proved too fragile to handle in the anaerobic chamber. Several methods of transferring a crystal from the melting point capillary in which it was grown to the x-ray capillary were attempted. The traditional method involved expulsion of the contents of the prospective melting point capillary onto a microscope depression slide using a 50 µL syringe equipped with Tygon tubing. Once the desired crystal had been separated

from surrounding crystals with thin glass fibers, it was drawn into the narrow open end of the x-ray capillary with a second syringe and tubing. A second technique was modeled after the Av1 crystal mounting protocol of Jongsun Kim. In this method, the melting point capillary was cut with a carbide-tip pencil on either side of the preferred crystal. The resultant 0.25-0.75 cm length of melting point capillary was suspended in the wide mouth of the x-ray capillary. A 100 μL pipetman outfitted with a pipet tip was used to gently expel the crystal and surrounding liquid into the x-ray capillary. Regardless of the technique employed in crystal transfer, the crystal was positioned in the capillary and excess mother liquor was removed from the crystal vicinity using long paper wicks (Hampton Research, Laguna Hills, CA) and thin glass fibers. A small volume of mother liquor was added to each end of the capillary before sealing with wax and epoxy.

Attempts to mount crystals in the thin capillaries were problematic. Manipulations through the butyl rubber gloves of the anaerobic chamber were often clumsy. Despite silanization of the melting point capillaries, crystals frequently adhered strongly to the capillaries' inner surfaces. Glass fibers were used to gently dislodge crystals from the melting point capillaries, with varying degrees of success. Crystals were fragile, and often broke during transfer to the x-ray capillaries. The pipetman-expulsion method was particularly destructive to potential Av2 crystals, which were smaller and less robust than their Av1 counterparts. X-ray capillaries, themselves fragile, were occasionally crushed during manipulation, particularly while sealing the capillaries with epoxy to maintain an anaerobic environment during data collection.

3.2.2. Room temperature Data Collection

Av2 crystals, once successfully mounted, diffracted to moderate resolution (2.8-3.0 Å) in the x-ray beam. Diffraction quality was impaired by thermal diffuse scattering, which caused a smeared reflection pattern. Crystals were sensitive to radiation damage, and

resolution decreased with beam exposure. Exposure to the intense beam of synchrotron sources severely damaged Av2 crystals, and a data set could not be collected from a single crystal using synchrotron radiation. Despite these limitations, a 2.6 Å resolution data set of $P2_1$ Av2 was collected from a single crystal using an in-house Siemens rotating anode generator and multiwire detector. The data were processed using XENGEN software (Howard *et al.*, 1987), yielding 96.8 % completeness (94.6 % in highest shell), and R_{merge} = 0.058 (R_{merge} = 0.396 in the highest resolution shell). There were 18 252 unique reflections from 51 667 measurements. The crystal was highly mosaic, with cell constants a = 56.84 Å, b = 90.12 Å, c = 58.12 Å, β = 98.52 °. This cell, although similar to that of the published $P2_1$ cell, was sufficiently different to require molecular replacement to solve the structure. This structure was pursued only until higher resolution data were obtained by cryocrystallographic methods, and will not be discussed further.

3.2.3. Cryocrystallographic Crystal Preparation

Cryocrystallography involves the collection of x-ray data from crystals which have been flash-cooled to near liquid nitrogen temperatures (Hope, 1988). The crystal, suspended on a glass fiber or spatula, or glass or rayon loop attached to a magnetic goniometer head, is surrounded by cryoprotectant solution and flash-cooled to low temperature and preserved during data collection in a stream of liquid nitrogen. Radical formation has been implicated in crystal radiation damage at room temperature; cryocrystallography proceeds at temperatures sufficiently low to trap radicals, thereby limiting their propagation and accompanying loss of crystalline order (Hope, 1990; Hope *et al.*, 1989). Cryocrystallography has been successfully employed in a large number of macromolecular structure determinations, including that of Cp2 (Woo, 1995).

Successful cryocrystallographic data collection requires the determination of appropriate cryoprotectant conditions. Initially, Av2 synthetic mother liquor was taken up

in a rayon loop and inserted into the liquid nitrogen stream. The frozen loop formed a glassy surface, suggesting that the Av2 crystals, grown from glycerol and PEG 4000, might be suitable for flash-cooling without modification. Cp2 crystals had been prepared successfully by opening capillaries in the anaerobic chamber and submerging the crystals under degassed Paratone-N oil (Exxon Chemical; Houston, TX) in a petri dish; after the dish had been moved near the x-ray equipment, crystals were removed from the oil with a glass fiber and flash-cooled in the LN_2 stream. Av2 crystals prepared in this manner cracked immediately upon contact with the oil. A more successful approach for Av2 involved opening the capillaries in air near the x-ray equipment. The capillary contents were expelled into a drop of degassed synthetic mother liquor onto a microscope depression slide, and crystals were picked up with a rayon loop and quickly inserted into the liquid nitrogen stream. Rapid technique allowed Av2 crystals to be flash-cooled before irreversible inactivation by O_2 occurred. The best results were obtained by opening a capillary under a microscope and dispelling the contents onto a microscope depression slide without the addition of synthetic mother liquor. The crystal was picked up with a rayon loop and plunged directly into a shallow dewar filled with LN_2; the dewar was moved to the x-ray equipment and carefully transferred to the nitrogen stream for data collection.

3.2.4. Cryocrystallographic Data Collection

Two data sets were collected on beam line 7-1 at the Stanford Synchrotron Radiation Laboratory, equipped with a 18 cm MAR imaging plate detector. Each data set was collected at λ = 1.08 Å wavelength from a single crystal maintained at -160 °C. The detector was set to a distance of 160 cm, which corresponded to a recorded resolution maximum of 2.13 Å. A 1.5 ° oscillation range was used for each data set to produce well-resolved reflections. Oscillation frames were auto-indexed with the program DENZO (Otwinoski, 1993), and crystals were found to belong to space group $P2_12_12_1$.

The first data set was collected using a crystal grown from Av2 in glycerol-containing buffer and a precipitating solution containing 30 % PEG 4000, 0.16 M Na_2MoO_4, and 0.1 M Tris-Cl, pH 8.0 (condition 8 in Table 3-1). The crystal measured 0.2 x 0.2 x 0.4 mm^3, and was flash-cooled from synthetic mother liquor as described in 3.2.3. Its orthorhombic cell constants refined to a = 72.69 Å, b = 90.19 Å, c = 91.22 Å. Data from 75 frames were processed in batch mode using DENZO, and scaled independently using both SCALEPACK (Otwinoski, 1993) and the programs ROTAVATA/AGROVATA from the CCP4 suite (Bailey, 1994). There were 251 789 observations made, yielding 28 078 unique reflections. R_{merge} for the data set was 0.079 to 2.2 Å resolution using SCALEPACK, and 0.267 in the highest resolution shell. Overall completeness was 89.5 %, with 79.9 % completeness in the outermost resolution shell. Multiplicity was 3.7; average I/sigma was 4.8 (2.8 in the highest resolution shell). Mosaicity for this data set refined to 1.1 °. Of the systematically absent reflections, a single 10.2 σ outlier was observed at reflection h = 17, k = 0, l = 0. Data collection statistics for this crystal are summarized in Table 3-2.

Table 3-2. Data collection statistics for 2.2 Å resolution $P2_12_12_1$ Av2 SSRL data set I.

Number of crystals	1
Space group	$P2_12_12_1$
a (Å), *b* (Å), *c* (Å)	72.69, 90.19, 91.22
Resolution limit (Å)	2.13
Mosaicity (°)	1.1
Rmerge	0.079 (0.267)
Number of measurements	251 789
Number of unique reflections	28 078
Completeness (%)	89.5 (79.9)
Multiplicity	3.7 (2.3)
<I/sigma>	4.8 (2.8)

A second data set was collected using a crystal grown from Av2 in glycerol-containing buffer and a precipitating solution composed of 25 % PEG 4000, 0.12 M Na_2MoO_4, and 0.1 M Tris-Cl, pH 8.0 (condition 1 in Table 3-1). The crystal measured 0.2 x 0.2 x 0.4 mm^3, and was flash-cooled without synthetic mother liquor by Michael Stowell. Its orthorhombic cell constants refined to a = 73.12 Å, b = 90.53 Å, c = 91.12 Å. Data from 71 frames were processed in batch mode using DENZO, and scaled independently using SCALEPACK and ROTAVATA/AGROVATA. The data set contained 260 578 independent measurements and 29 060 unique reflections. R_{merge} for the data set was 0.086 to 2.2 Å resolution using SCALEPACK. The R_{merge} in the highest resolution shell was 0.309. Overall completeness was 92.9 %, with 94.6 % completeness in the outermost resolution shell. Redundancy measured 4.0; the average I/sigma was 6.9 (2.4 in the highest resolution shell). Mosaicity for this data set refined to 0.5 °. Of the systematically absent reflections, a single 9.2 σ outlier was observed at reflection h = 0, k = 25, l = 0. A summary of data collection statistics for this crystal is shown in Table 3-3. R_{merge} values as a function of resolution are shown in Figure 3-1; average I/sigma values as a

Table 3-3. Data collection statistics for 2.2 Å resolution $P2_12_12_1$ Av2 SSRL data set II.

Number of crystals	1
Space group	$P2_12_12_1$
a (Å), *b* (Å), *c* (Å)	73.12, 90.53, 91.21
Resolution limit (Å)	2.13
Mosaicity (°)	0.5
Rmerge	0.086 (0.309)
Number of measurements	260 578
Number of unique reflections	29 060
Completeness (%)	92.9 (94.6)
Multiplicity	4.0 (3.8)
<I/sigma>	6.9 (2.4)

function of resolution are shown in Figure 3-2. Completeness as a function of resolution is shown in Figure 3-3, and multiplicity as a function of resolution is shown in Figure 3-4.

Initially, the reflections from both data sets were merged to form a single, more complete (96.7 %), data set, containing 33 131 unique reflections. However, the R_{merge} for the combined data was 0.129 for all data merged to 2.2 Å resolution, and 0.450 in the highest resolution shell. While the merged data set was used in preliminary analyses, it was later abandoned due to poor merging statistics in favor of using the data sets independently. Despite the highest overall completeness and lower R_{merge} for the first data set, the greater completeness at high resolution and lower mosaicity of the second data set deemed the latter a better vehicle for the structural determination. All subsequent work was carried out using only the second data set.

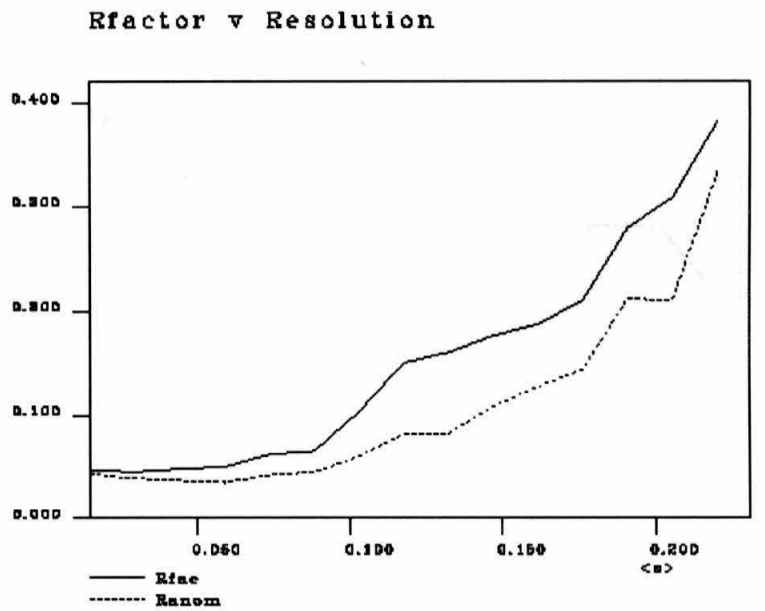


Figure 3-1. R_{merge} values as a function of resolution for 2.2 Å resolution $P2_12_12_1$ Av2 SSRL data set II. The plot was made using XLOGGRAPH in CCP4 (Bailey, 1994).

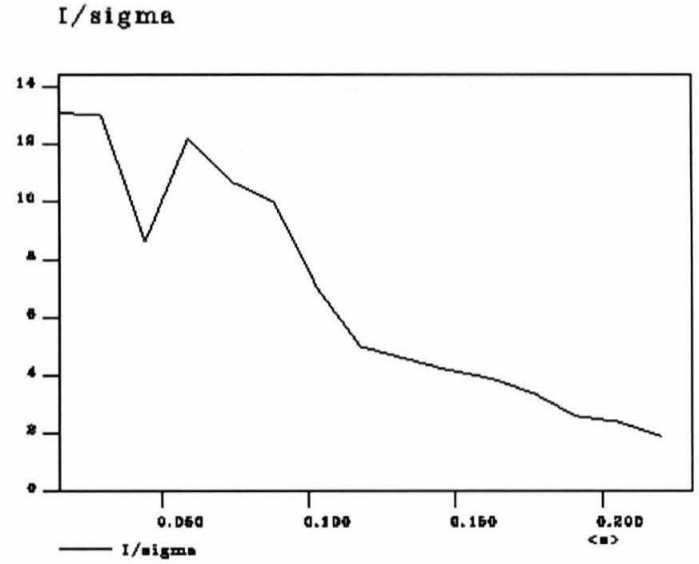


Figure 3-2. Average I/sigma values as a function of resolution for 2.2 Å resolution $P2_12_12_1$ Av2 SSRL data set II. The plot was made using XLOGGRAPH in CCP4 (Bailey, 1994).

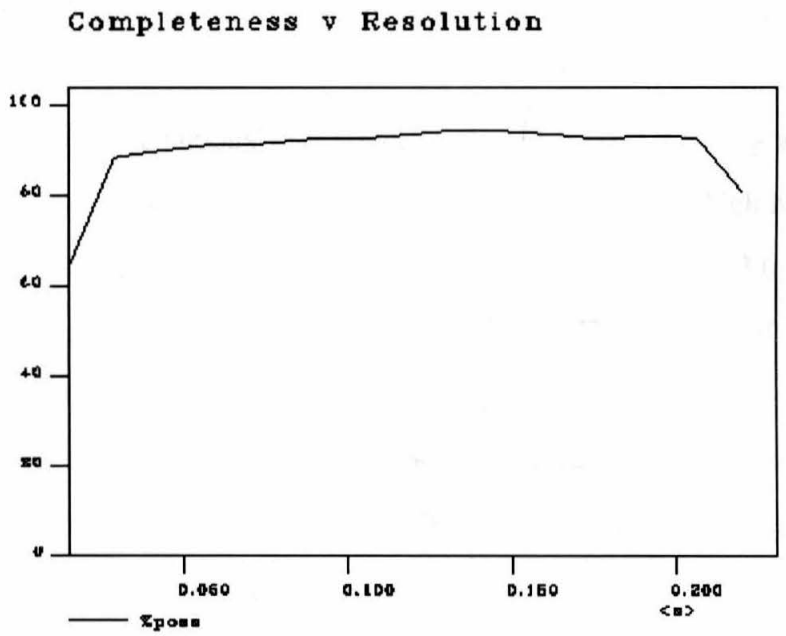


Figure 3-3. Completeness as a function of resolution for 2.2 Å resolution $P2_12_12_1$ Av2 SSRL data set II. The plot was made using XLOGGRAPH in CCP4 (Bailey, 1994).

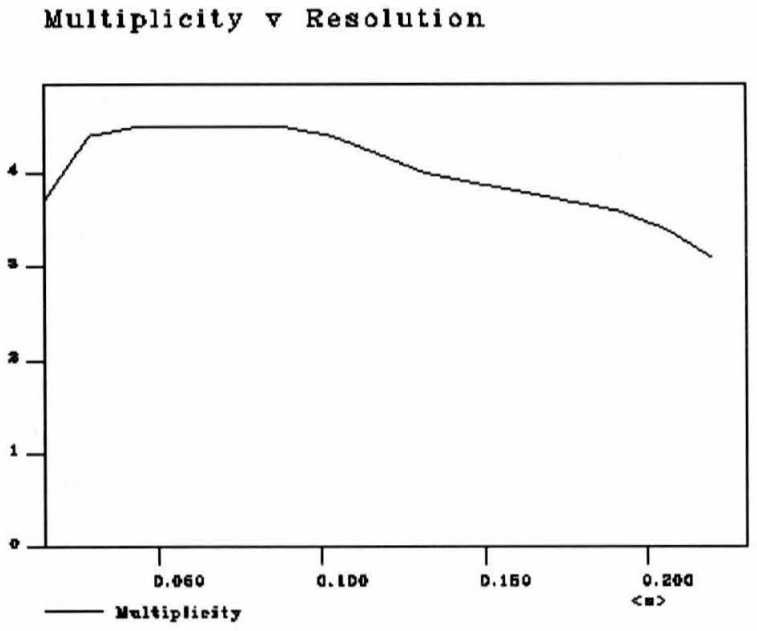


Figure 3-4. Multiplicity as a function of resolution for 2.2 Å resolution $P2_12_12_1$ Av2 SSRL data set II. The plot was made using XLOGGRAPH in CCP4 (Bailey, 1994).

An additional data set was collected at -160 °C using an in-house Rigaku generator and R-AXIS II C imaging plate detector. The crystal, which had been soaked in a solution containing ADP (described in 3.1.4), diffracted to only ~3.0 Å resolution, but a complete data set was collected using CuKα radiation. The refined cell constants were a = 74.05 Å, b = 90.35 Å, c = 90.77 Å. The $P2_12_12_1$ data were processed using DENZO, and scaled and merged using CCP4. A total of 54 654 measurements were made, yielding 12 325 unique reflections to 3.0 Å resolution. The R_{merge} for the data set was 0.095 to 3.0 Å resolution, and 0.338 in the highest resolution shell. Completeness for the data set was 98.2 %, and 99.7 % in the highest resolution shell. Average I/sigma was 5.4, and the mosaicity refined to 0.4 °. The R_{factor} between the CuKα data set and the second 2.2 Å resolution data set was 9.1 %. A summary of data collection statistics is given in Table 3-4.

Table 3-4. Data statistics for the 3.0 Å resolution $P2_12_12_1$ Av2.

Number of crystals	1
Space group	$P2_12_12_1$
a (Å), *b* (Å), *c* (Å)	74.05, 90.35, 90.77
Resolution limit (Å)	3.0
Mosaicity (°)	0.4
Rmerge	0.095 (0.338)
Number of measurements	54 654
Number of unique reflections	12 325
Completeness (%)	98.2 (99.7)
Multiplicity	4.3 (3.9)
<I/sigma>	5.4 (2.3)

Collected at $\lambda = 1.54$ wavelength, this data set carried a strong anomalous signal for the iron atoms of the Av2 Fe_4S_4 cluster, and a much weaker signal for cysteine sulfur atoms in the protein (described in section 4.2). Attempts to gain phasing information from the anomalous signal by combining the low resolution data set with the second SSRL data set were unsuccessful, as the signal was strong only to 5 Å resolution. However, the Fe_4S_4 position for the model in the SSRL data set after molecular replacement corresponded with an anomalous Fourier peak at $x = 0.15$, $y = 0.21$, $z = 0.25$ in the CuKα data set (shown in section 4.2), thereby providing independent confirmation of the correctness of the Fe_4S_4 cluster position in the molecular replacement solution. Analysis of electron density maps failed to indicate the presence of nucleotide in the model.

CHAPTER 4
Determination of the Av2 Crystal Structure

4.1. Molecular Replacement

Isolation of an isomorphous heavy atom derivative in the $P2_1$ 2.9 Å resolution structure proved a difficult task. Despite thorough screening of potential heavy atom derivatives, only a single usable isomorphous derivative was found (Georgiadis *et al.*, 1992). This derivative, myochrysine (gold sodium thiomalate), bound near the Fe_4S_4 cluster. For the $P2_12_12_1$ crystal form, heavy atom efforts concentrated on producing an analogous gold derivative. Attempts to add synthetic mother liquor containing 1 mM myochrysine to capillaries containing $P2_12_12_1$ crystals resulted in crystal dissolution. Molecular replacement methods were sought instead to solve the Av2 structure in the $P2_12_12_1$ crystal form.

Molecular replacement, first explored by Rossmann and Blow as an alternative to isomorphous replacement for macromolecular structures, effectively reduces the crystallographic phase problem to a set of three orientation angles and three translation vectors (Rossman and Blow, 1962). Successful implementation of this technique requires the availability of a search model which bears sufficient structural similarity to the unknown structure. The search model is rotated and translated into position in the unit cell of the unknown structure until correlation is maximized. Phases calculated from the search model are combined with structure factor amplitudes from the observed data to approximate the Fourier synthesis of the unknown structure. The resultant structure is subsequently refined to generate an accurate model of the unknown structure. Molecular replacement techniques using the 2.9 Å resolution Av2 structure as a search model were expected to work extremely well, since the protein was identical in amino acid sequence in both the $P2_1$ and $P2_12_12_1$ crystal forms. Previously, the 2.9 Å resolution Av2 structure had been used

successfully by Debbie Woo as a search model to determine the 2.0 Å resolution Cp2 structure by molecular replacement (Woo, 1995). Av2 and Cp2 share nearly 70 % sequence identity, as shown by the amino acid sequence alignment in Figure 4-1.

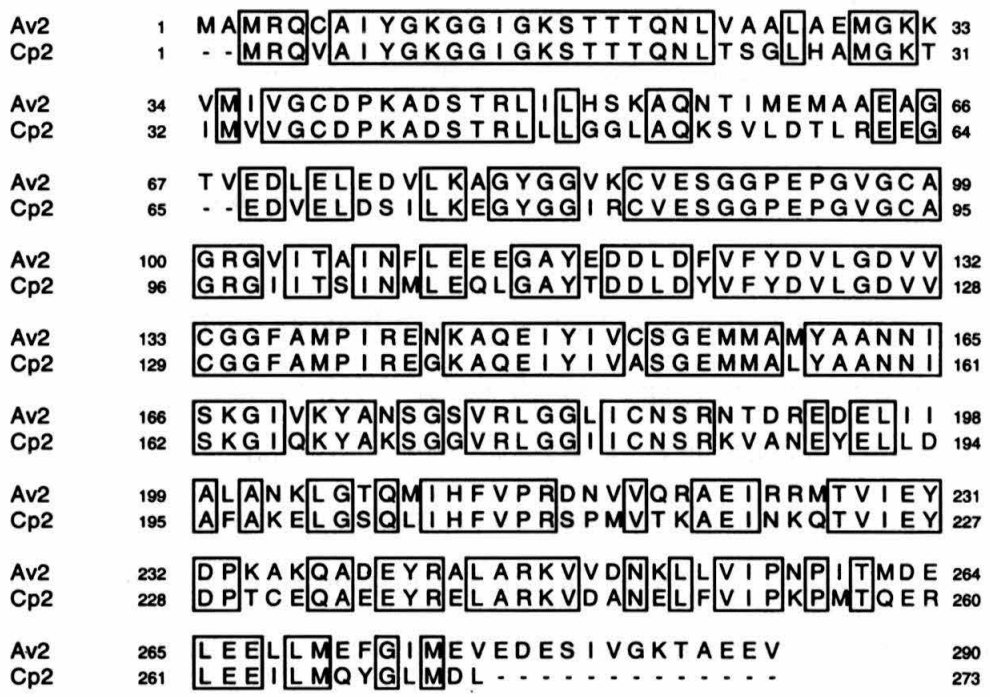


Figure 4-1. Amino acid sequence alignment of Av2 and Cp2. Residues conserved between the two proteins are shown in boxes. Av2 and Cp2 share 69 % sequence identity.

4.1.1. Molecular Replacement using X-PLOR

Initial molecular replacement attempts utilized software from the X-PLOR package (Brünger, 1988). Molecular replacement in X-PLOR involves a Patterson-space search (Brünger, 1991), following the real-space Patterson search method of Huber and Steigemann (Huber, 1985). This approach proved time-consuming on the available

computing facilities, but did produce a unique molecular replacement solution with reasonable packing. All X-PLOR calculations were performed using the parameter and topology files parhcsdx_jpj.pro and tophcsdx_jpj.pro, modified from those of Engh and Huber (Engh and Huber, 1991). The files av2_param.pro and av2_toph.pro were also included to provide information about the geometry and connectivity of the Fe_4S_4 cluster. The target was set to E2E2 for all X-PLOR molecular replacement studies.

4.1.1.1. Self-Rotation Function

Av2 consists of two identical subunits bridged by a single Fe_4S_4 cluster. Two-fold noncrystallographic symmetry (NCS) relates the two monomers; the NCS axis of Av2, known from the 2.9 Å resolution structure, passes between the subunits through the Fe_4S_4 cluster, parallel to the longest dimension of each monomer. The X-PLOR script self_rf.inp was run to calculate the orientation of this axis in the $P2_12_12_1$ cell. In this method, a stationary Patterson map P2 is computed from the observed intensities by fast Fourier transformation on the specified grid (Brünger, 1991). The Patterson map to be rotated in the search, P1, is computed from the observed intensities. The Patterson maps are compared to each other to determine the best superposition of a specified volume of the Patterson function upon itself; the correlations are calculated on a polar coordinate system. The appearance of a transformation which is not part of the crystal symmetry indicates the presence of noncrystallographic symmetry. Hence, a peak in the κ=180 section of the plot was anticipated for dimeric Av2. However, in the case of $P2_12_12_1$ Av2, no obvious NCS peak was observed. In a 15-4 Å resolution search, the strongest peak after the origin peak (ϕ = 0.0, ψ = 0.0, κ = 180.0, σ = 5.72) is located at ϕ = 7.5, ψ = 0.0, κ = 180.0, with σ = 2.53. This peak is located in a shoulder of the crystallographic peak, and is within the range of "noise." Resolution limits and the radius of integration were varied without improvement of signal. It is likely that the self-rotation peak is located near enough to a

crystallographic two-fold axis that the NCS peak is obscured by the shoulder of a crystallographic peak. The κ=180 section of the Av2 self-rotation function, plotted using Mathematica (Wolfram Research, Inc.; Champaign, IL), is shown in Figure 4-2.

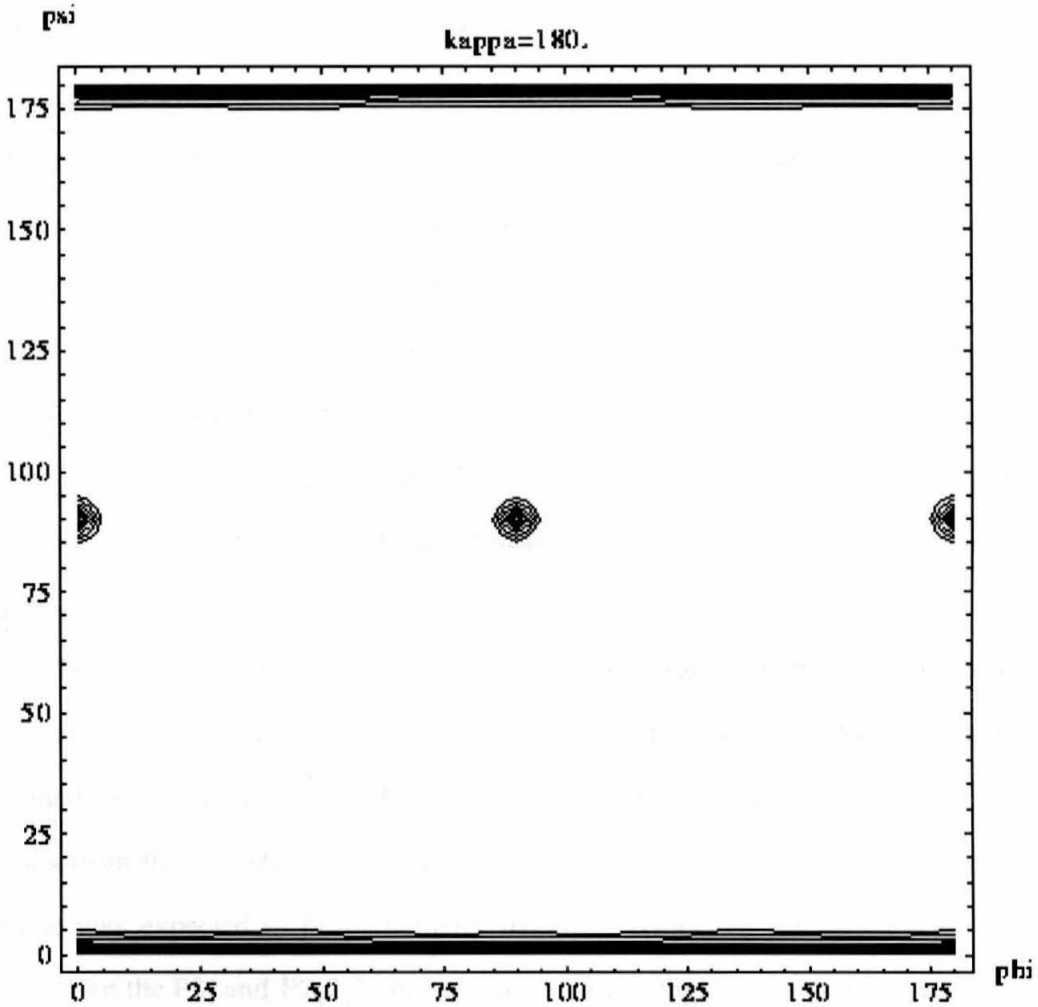


Figure 4-2. Section κ = 180 from a 10-4 Å resolution $P2_12_12_1$ Av2 self-rotation search in X-PLOR. The section was plotted using Mathematica. No obvious noncrystallographic peak is discernible in the section.

4.1.1.2. Cross-Rotation Function

The script rotation.inp was utilized for the initial cross-rotation search in X-PLOR. The stationary Patterson map P2 was computed from the observed intensities as in self_rf.inp; Patterson map P1 was computed from the search model, which consisted of the Av2 dimer and Fe_4S_4 cluster from the 2.9 Å resolution $P2_1$ structure. Searches using only one monomer of Av2 were also performed. Coordinates were taken from the 1NIP.pdb entry of the PDB; solvent molecules and molybdate ions were removed, and amino acid residues for which atoms were missing were truncated to Ala. Using a 40.0 Å integration radius and grid spacing of Δ = 2.5 °, the cross-rotation search using the dimer search model yielded a top solution of θ_1 = 215.0, θ_2 = 80.0, θ_3 = 157.5, over the resolution range of 8-3.5 Å. Correlation for this solution was R_f = 1.5293. The next highest solution, θ_1 = 215.0, θ_2 = 90.0, θ_3 = 15.0 (R_f = 1.5046), was related to the first. The top 10 peaks from this search are listed in Figure 4-3a.

The rotation function peaks were refined in X-PLOR by Patterson Correlation (PC) refinement using the script filter.inp. PC refinement allows macromolecular models to be subdivided into domains, so that each domain may be rotated independently. It also tests solutions from the rotation search which were not among the highest peaks initially. This technique was expected to be particularly useful in correcting for differences that might exist between the $P2_1$ and $P2_12_12_1$ crystal forms in the angle between the Av2 monomers. Three domains were assigned, belonging to subunit A, subunit B, and the Fe_4S_4 cluster. PC refinement yielded four significant solutions: (1) θ_1 = 215.0, θ_2 = 80.0, θ_3 = 157.5 (R_f = 0.12693); (2) θ_1 = 215.0, θ_2 = 90.0, θ_3 = 15.0 (R_f = 0.11097); (3) θ_1 = 215.0, θ_2 = 87.5, θ_3 = 165.0 (R_f = 0.10553); and (4) θ_1 = 205.0, θ_2 = 82.5, θ_3 = 160.0 (R_f = 0.1032). The four peaks were judged to be equivalent solutions. A plot of the PC refinement results, produced in Mathematica, is shown in Figure 4-3b. Analogous tests with the Av2 monomer search model were also performed, with corresponding results. A

direct rotation function search was performed by Art Chirino, which confirmed the results from the Patterson searches.

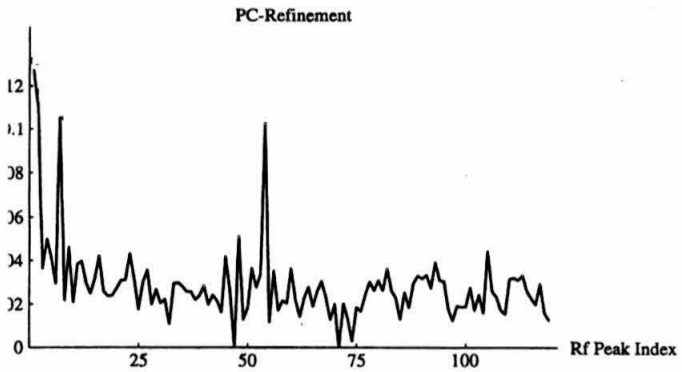
4.1.1.3. Translation Function

The X-PLOR translation search was performed using the script translation.inp. Each prominent PC refinement solution was used as a starting point for the translation search. The dimer translation search over resolution range 8-3.5 Å using orientation angles $\theta_1 = 215.0$, $\theta_2 = 80.0$, $\theta_3 = 157.5$ yielded an 8 σ peak for the top solution 0.350, 0.338, 0.4206 (fractional coordinates given). Correlation for this solution was T = 0.7034, and the error-value was P = 0.4206. A plot of this solution was produced in Mathematica and is shown, contoured at 4 σ, in Figure 4-2c.

A

θ1	θ 2	θ 3	Rf-function
215.000	80.000	157.500	1.5293
215.000	90.000	15.000	1.5046
40.000	90.000	0.000	1.1655
222.500	87.500	150.000	1.1462
220.000	82.500	20.000	1.1444
42.500	82.500	100.000	1.1384
215.000	87.500	165.000	1.1192
40.000	85.000	62.500	1.1159
222.500	85.000	92.500	1.1156
222.500	87.500	90.000	1.1048

B



C

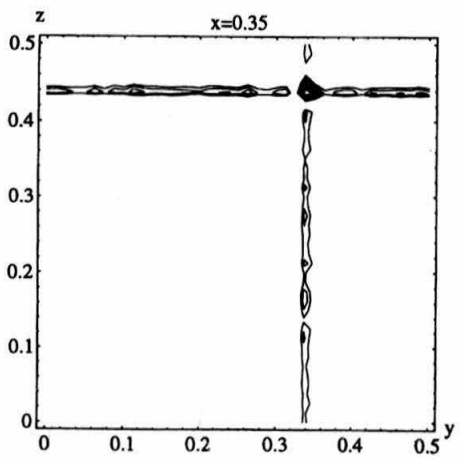


Figure 4-3. Molecular replacement search in X-PLOR for $P2_12_12_1$ Av2, using the $P2_1$ Av2 dimer (1NIP.pdb) as a search model. The resolution range was 8-3.5 Å. (A) The top 10 peaks from a cross-rotation search peak list. (B) Patterson Correlation refinement for the cross-rotation search peaks. (C) Translation search top solution contoured at 4 σ. This solution corresponds to θ_1 = 215.0, θ_2 = 80.0, θ_3 = 157.5, T1 = 0.350, T2 = 0.338, T3 = 0.4206 (given in fractional coordinates). Mathematica was used to produce the plots in (B) and (C).

4.1.2. Molecular Replacement using AMORE

The AMORE (Automated MOlecular REplacement) program by Jorge Navaza provides the possibility of rapid rotation and translation searches in reciprocal space (Navaza, 1997, 1994). A semi-automated version of the program written by Gerard Kleywegt and adapted by Art Chirino was used for self-rotation, cross-rotation, and translation searches and preliminary rigid-body refinement of $P2_12_12_1$ Av2. The 1NIP.pdb coordinates were examined using the graphics program O (Jones, 1991); residues for which atoms had been missing previously were restored to their full length using conservative rotamers, and the occupancy for these atoms was set to q = 0.0; original B-factors were preserved for fully occupied atoms. A variety of search models was employed: (1) Av2 dimer, without Fe_4S_4 cluster; (2) Av2 monomer A; (3) Av2 monomer B; (4) Cp2 dimer, without Fe_4S_4 cluster where non-conserved residues were truncated to Ala; and (5) Av2 poly-Ala (note: here "Av2" refers to the 2.9 Å resolution Av2 coordinates with conservatively restored side chains).

The best molecular replacement results were produced using the Av2 subunit A. The program was run twice. In the first run, a single monomer solution was sought. With the top solution from this run fixed, the program was rerun to identify the second monomer. These two solutions were: (1) α = 53.71, β = 18.63, γ = 318.59, T1 = 0.07977, T2 = 0.36014, T3 = 0.12941; and (2) α = 84.36, β = 13.19, γ = 324.25, T1 = 0.67128, T2 = 0.81230, T3 = 0.14011. AMORE optimizes its search by first orienting the search model according to its center of mass; this rotated search model is transformed by the solutions to produce the true coordinates. Following rigid-body refinement, correlation for this solution was 0.74, and the error-factor was 0.38. When the molecular replacement solutions were examined graphically in O, the individual subunits did not appear to form homodimers; this was an artifact of AMORE's search model orientation feature. However, generation of symmetry-related mates of each solution with the sym_sphere command in O

revealed the correct position of each dimer partner. The Fe_4S_4 cluster from 1NIP.pdb was superimposed onto the dimer in O, using the lsq_exp command. Finally, the full dimer and cluster coordinates were written out to produce starting coordinates for the $P2_12_12_1$ Av2 refinement.

4.2. Anomalous Patterson and Fourier Functions

The 2.2 Å resolution Av2 data sets were collected using radiation of wavelength λ = 1.08 Å at SSRL. This wavelength is not optimal for detection of anomalous iron signal. However, anomalous Fourier maps, calculated with the CCP4 suite of programs, of the 2.2 Å resolution data sets revealed small peaks to 5 Å resolution for the Fe_4S_4 cluster position as determined by molecular replacement. An analogous map of the 3.0 Å resolution $P2_12_12_1$ Av2 data set, collected using CuKα radiation, revealed much stronger peaks in the corresponding position at 5 Å resolution. A single peak was seen at x = 0.15, y = 0.21, z = 0.25 in this map; discrete peaks for the individuals atoms of the cluster were not visible at this, or higher resolutions. However, the strong anomalous Fourier signal confirmed the correctness of the Fe_4S_4 cluster position in the molecular replacement solution. In addition, native anomalous Patterson maps of the CuKα data set displayed a strong peak at x = 0.16, y = 0.21, z = 0.25. This 6 σ peak corresponded to the cluster position in the molecular replacement solution. No anomalous signals corresponding to bound molybdate ions were detected in the maps.

4.3. NCS Averaging and Solvent Flattening

To produce interpretable maps for model building, the Av2 molecular replacement solution underwent preliminary rounds of rigid-body and positional refinement in X-PLOR, followed by solvent flattening and two-fold noncrystallographic averaging with the

program SOLOMON (Abrahams and Leslie, 1996). X-PLOR rigid-body refinement before solvent flattening and averaging yielded a model with R_{cryst} = 0.428 and R_{free} = 0.435 for all data between 10 - 3.5 Å resolution. Hendrickson-Lattman AB coefficients (Hendrickson and Lattman, 1970) from the model were calculated using the PHASES package (Furey and Swaminathan, 1990). A molecular envelope for the Av2 monomer was created using the program MAMA (Klegweyt and Jones, 1993, 1994a), with a 3.5 Å van der Waals radius; NCS operators were calculated using the program LSQMAN (Kleywegt and Jones, 1994b; Kleywegt, 1996). The map grid in SOLOMON was set to h = 104, k = 128, l = 128. Ten cycles of SOLOMON solvent flattening were performed at 3.5 Å maximum resolution. The output .MTZ file from the solvent flattening, mask, and preliminary NCS operators were input into the program trimmasks.com (Abrahams and Leslie, 1996) to calculate SOLOMON input NCS operator files for NCS averaging in SOLOMON. Finally, 110 cycles of two-fold NCS averaging and phase extension in SOLOMON were performed, from 3.5 Å to 2.6 Å maximum resolution. The starting and final FOM values for the averaging run were 0.24 and 0.88, respectively, while the starting and final error values for the run were 0.38 and 0.19, respectively. The final statistics are somewhat inflated by a feature in SOLOMON known as the "flipping factor." Resultant electron density maps revealed largely connected main chain density and density for many side chains.

4.4. Crystallographic Refinement

Refinement of the Av2 model involved two iterative processes: (1) computational refinement of atom positions and temperature factors to reduce the discrepancy between the observed ($|F_{obs}|$) and calculated ($|F_{calc}|$) structure factor amplitudes; and (2) manual rebuilding of the model using a graphics system to fit atoms to electron density. Progress in crystallographic refinement is generally assessed by the residual index, R_{cryst}, and cross-

validated by the use of R_{free}. The R-factor measures the error in the agreement between F_{obs} and F_{calc}; R_{free}, in particular, estimates the mean phase error of the model. Inclusion of an R_{free} term requires that a subset of the observed reflection, 6 % was used in the case of Av2, must be partitioned from the data into a "test set" of reflections. The test set is not included in the refinement process, and therefore remains a more objective means of determining model quality. The balance of the reflections, known as the "working set," is used in refinement and model building, and in the calculation of R_{cryst}.

Av2 refinement was performed in X-PLOR, using both energy minimization and molecular dynamics. The program XPLORPDBPREP, written by Art Chirino, converted standard coordinate files to X-PLOR format. Coordinate and molecular structure files were generated using generate.inp. Rigid-body refinement was performed on the Av2 dimer using rigid.inp, first refined as a single rigid body, then as three units (subunit A, subunit B, and the Fe_4S_4 cluster). Values for R_{cryst} and R_{free} for all data 10-3.5 Å resolution were 0.428 and 0.435, respectively, at this stage in the refinement. An overall temperature factor of 12.63 $Å^2$ was applied to the model using baoverall.inp. NCS restraints were employed to maintain similarity between the protein subunits. A weight of 100 was used for each main chain atom, and 50 for each side chain atom. Weights for the x-ray data, WA, were estimated by running the script check.inp; actual values used in refinement were determined by monitoring the value of R_{free}, but generally approximated the recommended check WA value. Positional refinement and simulated annealing techniques were utilized throughout this portion of the refinement. Simulated annealing runs utilized an initial temperature of 4000 K and timestep of 1.0 fsec.

Following solvent flattening and averaging with SOLOMON, model building and refinement on the Av2 model reduced R_{cryst} and R_{free} to 0.365 and 0.415, respectively. Correctness of fit in model building was monitored using the rs_fit command in O; a real-space correlation coefficient (RSCC) was plotted to identify regions of the model that fit the map poorly. More than 20 additional iterations of model building and refinement were

performed. Individual temperature factors were refined and solvent molecules conservatively incorporated at resolutions at or beyond 2.6 Å. A bulk solvent correction was applied to the data set, which reduced both R_{cryst} and R_{free} by 1.5 %. Despite these efforts, the refinement could not be improved beyond $R_{cryst} = 0.25$ and $R_{free} = 0.32$ for all data to 3.0 Å resolution; R-factors rose precipitously with increasing resolution. Average RSCC for the model was 0.758. Geometry for the model was good, with the following RMS deviations: bond lengths, 0.014 Å; bond angles, 1.9 °; dihedral angles, 1.6 °; and improper angles, 24.74 °. The model appeared in general to conform to the electron density of $(2|F_{obs}| - |F_{calc}|)$ maps; however, some residues, including several cluster-ligating cysteines, and many charged side chains, remained poorly defined. Efforts to slowly increase the resolution of the model in 0.1 Å increments by iterations of simulated annealing, positional and temperature factor refinement were unsuccessful. Attempts with other software packages, including RSREF (Chapman, 1995) and TNT (Tronrud, 1992; Tronrud *et al.*, 1987), also failed to improve the statistics.

4.4.1. Improved Phasing Techniques

Failure to improve the R-factors in the Av2 refinement was a source of chronic consternation. The data were tested for twinning and other anomalies by Doug Rees, but none were identified. Reprocessing the data, including attempts to process the data in other orthorhombic space groups, failed to remedy the problem. The accuracy of the molecular replacement solution was questioned, and the procedure was repeated using various search models and programs. No new solutions were found, and no improvement in the refinement was made. Molecular replacement in AMORE using a poly-alanine Av2 dimer as the search model was followed by solvent flattening and NCS averaging in SOLOMON. Inspection of the averaged maps revealed reasonable electron density for many protein side chains, and fostered confidence in the original molecular replacement solution. With

soundness of the data and correctness of the molecular replacement solution confirmed, focus on completing the refinement turned to improvements in the model itself.

At the suggestion of Doug Rees, a "composite" iron protein was created by superimposing the Cp2 model onto the Av2 molecular replacement solution. Each dimer was rotated by 180 ° along its 2-fold NCS axis and overlaid on the coordinates. Hence, a coordinate file containing four complete dimers was produced. The composite iron protein was used as the starting model for phasing in PHASES. A mask with 4.0 Å van der Waals radius was created about one "monomer" and Fe_4S_4 cluster using MAMA. Solvent flattening and NCS averaging in SOLOMON were performed at 3.5 Å resolution. The averaged maps showed excellent electron density, and the averaging process was repeated at 3.0 Å, 2.8 Å, 2.6 Å, 2.4 Å, and 2.2 Å resolution. The 2.2 Å resolution maps displayed exceptionally beautiful electron density. Numerous solvent molecules could be clearly identified.

Comparisons of the Av2 molecular replacement solution (which had been based on the 1NIP.pdb coordinates) with the map using rs_fit in O produced a correlation value RSCC = 0.58. The correlation value for the Cp2 model to this map was RSCC = 0.61. In contrast, RSCC for the partially-refined Av2 model in this maps was 0.71. Hence, the partially refined Av2 model was a better match for the composite averaged maps. Particularly convincing were the clear electron density for residues in the loop A39-A43 and the C-terminus of subunit B. The A39-A43 loop region had been rebuilt in the $P2_12_12_1$ Av2 molecule, and differed markedly from those of Cp2 and the $P2_1$ Av2. Average RMS deviation for CA atoms in the loop was 3.4 Å between the two Av2 models. Yet, the 2.2 Å resolution averaged maps displayed contiguous density consistent with the new loop conformation, including side chain density for Lys[41]. Also, the Av2 C-terminus of subunit B, which had been incorporated into the $P2_12_12_1$ model during previous model building efforts, had not been present in the $P2_1$ Av2 and was not included in the initial phasing model. Therefore, the strong electron density for this region in the averaged maps was not

attributable to model bias (Cp2 is shorter than Av2 by 13 residues at its C-terminus, and could not contribute model bias either). The 2.2 Å resolution averaged maps created from the composite phasing model were regarded as objective, and were used to rebuild the $P2_12_12_1$ Av2 model.

4.4.2. Final Refinement

The Av2 model rebuilt from the composite averaged maps refined to R_{cryst} = 0.32 and R_{free} = 0.38 to 2.2 Å resolution. The averaged maps had identified regions of the model which did not obey noncrystallographic symmetry, primarily due to packing interactions, so NCS restraints were removed from the refinement parameters. Among these regions were the Fe_4S_4 cluster ligands and the loop residues 60-75, 116-118, and 172-176. After refinement, all atoms with poor density or temperature factors greater than 60.0 $Å^2$ were identified. These atoms were found primarily on charged side chains and in loop regions. The occupancy of each of these atoms was set to q = 0.0, and its temperature factor to B = 20.0. More than 15 additional iterations of model building and refinement were performed; simulated annealing methods were abandoned, and only positional and temperature refinements were used. In each round of model building, small changes in the atomic positions were made, and maps improved iteratively. Side chain atoms were restored as electron density became unambiguous. Solvent molecules were added as waters according to the following criteria: (1) potential solvent molecules must display 1.0 σ ($2|F_{obs}| - |F_{calc}|$) and 3.0 σ ($|F_{obs}| - |F_{calc}|$) minimal electron density; (2) potential solvent molecules must make at least one protein contact with hydrogen bond length less than 3.5 Å; and (3) after refinement, solvent molecules must possess temperature factors less than b = 70.0 $Å^2$. Model building and refinement were ended when R_{cryst} = 0.223 and R_{free} = 0.290 were achieved. The mean RMS deviation for the final ($|F_{obs}| - |F_{calc}|$) electron density map was 0.001 σ, and the average RSCC was 0.863 for all atoms.

4.5. Current Av2 Model

The present Av2 model contains 4342 non-hydrogen protein atoms in residues A1-A286 and B1-B289, an Fe_4S_4 cluster, and 372 water molecules. Interpretable electron density was not observed for the C-terminal three residues of subunit A (A287-A289), and these residues have not been included in the model. Despite the inclusion of Na_2MoO_4 in the crystallization conditions, examination of electron density maps identified no molybdate ions bound in the P-loop region of the protein as been described for the 2.9 Å resolution Av2 structure. As in the Cp2 model, nucleotide was not observed bound to the protein.

Using all reflections to 2.2 Å resolution, R_{cryst} is 0.223 and R_{free} is 0.290 for the Av2 model. Exclusion of weak reflections below 3 σ reduces the values to 0.200 and 0.264, respectively. A summary of R_{cryst} and R_{free} statistics is given in Table 4-1. The model observes good stereochemistry, with RMS deviations from ideality for bond lengths, and bond, dihedral, and improper angles of 0.017 Å, 1.93 °, 23.99 °, and 1.64 °, respectively.

In general, electron density contoured at 1.5 σ in ($2|F_{obs}| - |F_{calc}|$) maps was continuous along the protein backbone. Most protein side chains had clear electron density visible at 1.0 σ in ($2|F_{obs}| - |F_{calc}|$) maps. Some regions of the model, however, remain conformationally ambiguous. In particular, the loop of residues B60-B75 eluded model building efforts. Strong density (> 5 σ ($|F_{obs}| - |F_{calc}|$) electron density) in the region suggests the presence of residues, but these residues appear to be arranged in at least two different conformations. The loop residues have been modeled conservatively, with occupancies set to q = 0.5 for main chain atoms. Some surface loop regions, including residues 110-112, 116-118, and 171-176, contain side chains with poor or no ($2|F_{obs}| - |F_{calc}|$) electron density. Where appropriate, occupancies have been set to q = 0.0.

Table 4-1. Refinement Statistics for the current 2.2 Å resolution $P2_12_12_1$ Av2 structure

Resolution	Working Set				Test Set			
Range	Reflections		R_{cryst}		Reflections		R_{free}	
(Å)	Shell	% Complete	Shell	Cumulative	Shell	% Complete	Shell	Cumulative
4.74-50.00	2737	82.02	0.184	0.184	192	5.75	0.230	0.230
3.76-4.74	2713	85.23	0.154	0.169	155	4.87	0.226	0.228
3.29-3.76	2737	86.67	0.187	0.174	167	5.29	0.275	0.241
2.99-3.29	2724	87.25	0.219	0.182	172	5.51	0.323	0.254
2.77-2.99	2747	88.07	0.256	0.191	172	5.51	0.361	0.266
2.61-2.77	2760	88.92	0.292	0.197	174	5.61	0.346	0.273
2.48-2.61	2774	89.05	0.296	0.207	159	5.10	0.376	0.280
2.37-2.48	2759	89.26	0.307	0.213	179	5.79	0.350	0.284
2.28-2.37	2745	88.98	0.327	0.219	182	5.90	0.341	0.288
2.20-2.28	2747	88.87	0.296	**0.223**	185	5.99	0.336	**0.290**

Packing interactions for the $P2_12_12_1$ Av2 cell are manifested in the arrangement of several loop regions of the model. These loop regions are not uniform between subunits with respect to atomic position and temperature factor. The regions of greatest difference in atomic positions are residues 96-100 and 131-134, which contain the Fe_4S_4 cluster ligands, 60-75, and the C-termini. In general, electron density for subunit A atoms is superior to corresponding atoms of subunit B; this difference is reflected in the average temperature factors for each subunit. Subunit A has an average temperature factor of 27.73 $Å^2$ for non-hydrogen atoms; the corresponding value in subunit B is 33.22 $Å^2$. The overall temperature factor for non-hydrogen atoms (including solvent molecules) is 31.72 $Å^2$. The average temperature factors for the Fe_4S_4 cluster and solvent molecules are 20.11 $Å^2$ and 46.67 $Å^2$, respectively. These values are in agreement with the average temperature factor estimated for the data set by a Wilson plot, 35.13 $Å^2$, shown in Figure 4-4.

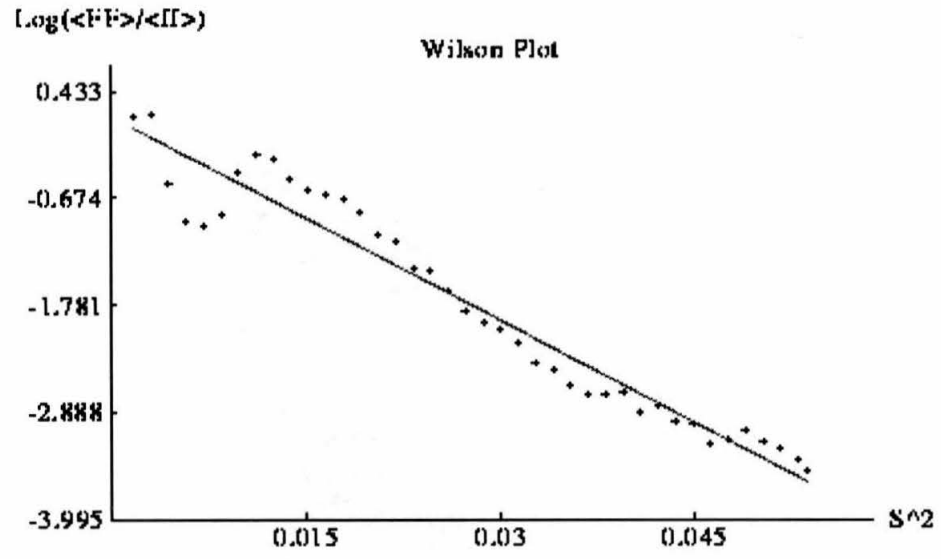


Figure 4-4. Wilson plot for the 2.2 Å resolution Av2 data set. The plot was generated in Mathematica, after a calculation in X-PLOR. The average temperature factor estimated by the plot is 35.13 $Å^2$.

The stereochemistry of the $P2_12_12_1$ Av2 model was analyzed using the programs PROCHECK and WHATIF. Ramachandran plots of each subunit calculated in PROCHECK show all residues are located within regions of sterically-allowed conformations, with 88.6 % in subunit A and 83.3 % in subunit B in the most favored regions. The remaining residues are found entirely in additionally allowed regions; no residues are found in generously allowed or disallowed regions of the Ramachandran plot. Ramachandran plots of each subunit are shown in Figure 4-5. The main chain and side parameters of Av2 also compare favorably in PROCHECK with well-refined structures of the same resolution. Stereochemical analysis with WHATIF is favorable as well, indicating normal variation of rotamers, planarities, bond lengths, and bond angles.

A

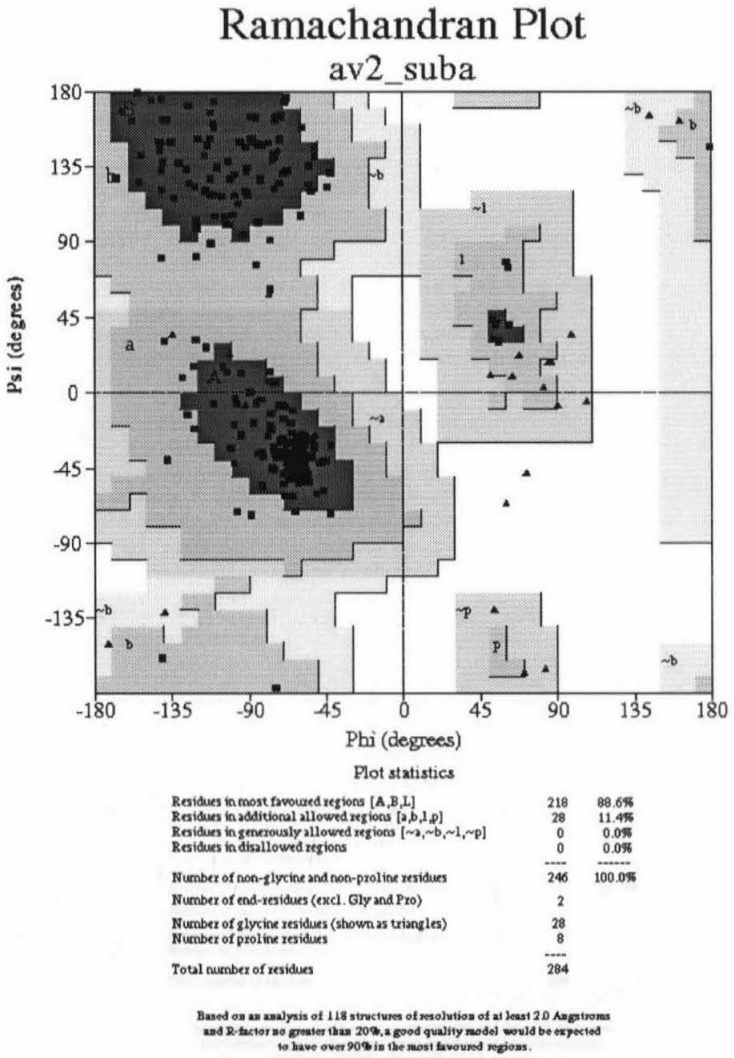
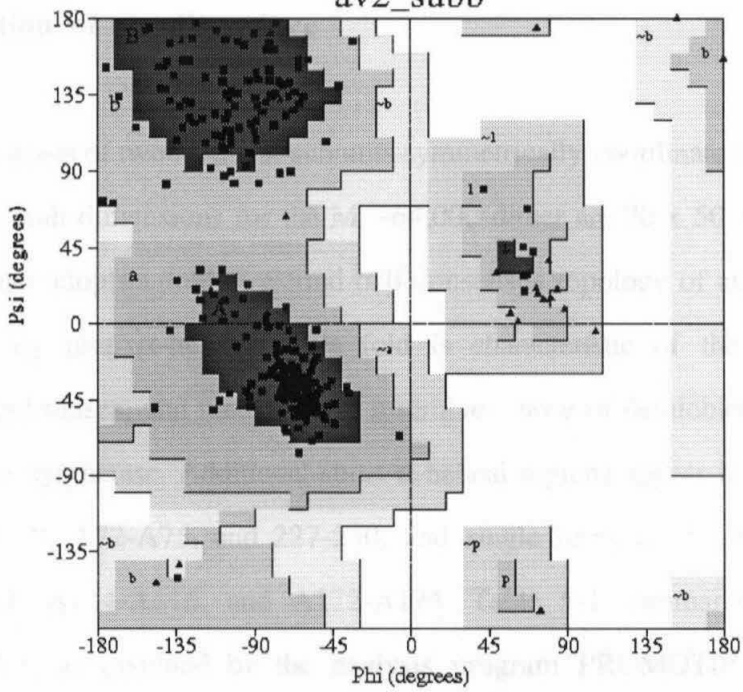


Figure 4-5. Ramachandran plots of main chain (ϕ,ψ) torsion angles for the current Av2 model. (A) Subunit A. (B) Subunit B. All residues are located within regions of sterically-allowed conformations, with 88.6 % in subunit A and 83.3 % in subunit B in the most favored regions. The remaining residues are found entirely in additionally allowed regions; no residues are found in generously allowed or disallowed regions of the plot. The figures were generated using PROCHECK.

B

Ramachandran Plot

av2_subb



Plot statistics

Residues in most favoured regions [A,B,L]	209	83.3%
Residues in additional allowed regions [a,b,l,p]	42	16.7%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	251	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	28	
Number of proline residues	8	
Total number of residues	289	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

CHAPTER 5
Analysis of the Av2 Crystal Structure

5.1. Description of the Structure

Av2 consists of two identical subunits symmetrically coordinated by a single Fe_4S_4 cluster. The overall dimensions for the M_r ~64,000 dimer are 70 x 50 x 55 $Å^3$. Each Fe protein monomer adopts a doubly-wound α/β consensus topology of an eight-stranded β-sheet flanked by nine α-helices. This fold is characteristic of the P-loop class of triphosphate hydrolases, and most closely resembles those of dethiobiotin synthetase and adenylsuccinate synthetase. Additional short α-helical regions appear in the current model at residues 46-49, A72-A75, and 227-230, and single turns of 3_{10} helix are found at residues 80-82, A114-A116, and A172-A174. Table 5-1 summarizes the secondary structure of Av2 as assigned by the analysis program PROMOTIF (Hutchinson and Thornton, 1996). Nomenclature for secondary structure elements follows the precedent of the 2.9 Å resolution Av2 structure (Georgiadis *et al.*, 1992). A noncrystallographic two-fold axis passing through the solvent-exposed Fe_4S_4 cluster relates the two monomers. The RMS deviation for all superimposed Cα atoms between subunits is 0.7 Å for Av2; this value is 0.6 Å for the 2.9 Å resolution Av2 structure, and 0.7 Å for Cp2. Cys^{97} and Cys^{132} from each subunit ligate the Fe_4S_4 cluster. The current Av2 model is depicted in Figure 5-1.

The twisted β-sheet core of each monomer contains seven parallel and one short antiparallel β-strands. The single antiparallel strand, β3 (residues 76-78), is located at one edge of the β-sheet. The strand order of the β-sheet, 3-4-2-5-1-6-7-8, reverses between strands β1 and β4, forming topological "switches" in the resultant crevices in the loops carboxyl to the affected strands. These switch points correspond to substrate and cofactor binding sites in numerous α/β proteins. In the case of Fe protein, loops following β1 and

Table 5-1. Secondary Structure Elements in $P2_12_12_1$ Av2

2° Structure Element	Residue Range	Number of Residues
β1	3-9	7
α1	15-28	14
β2	33-37	5
α1'	46-49	4
α2	57-61	5
α2'	A72-A75	4
β3	76-78	3
3_{10}	80-82	3
β4	84-87	4
α3	101-111	12
3_{10}	A114-A116	3
β5	121-127	7
α4	137-140	4
β6	146-151	6
α5	155-170	16
3_{10}	A172-A174	3
β7	178-185	8
α6	192-203	12
β8	207-211	5
α7	215-221	7
α8	227-230	4
α8'	235-249	15
α9	261-270	10

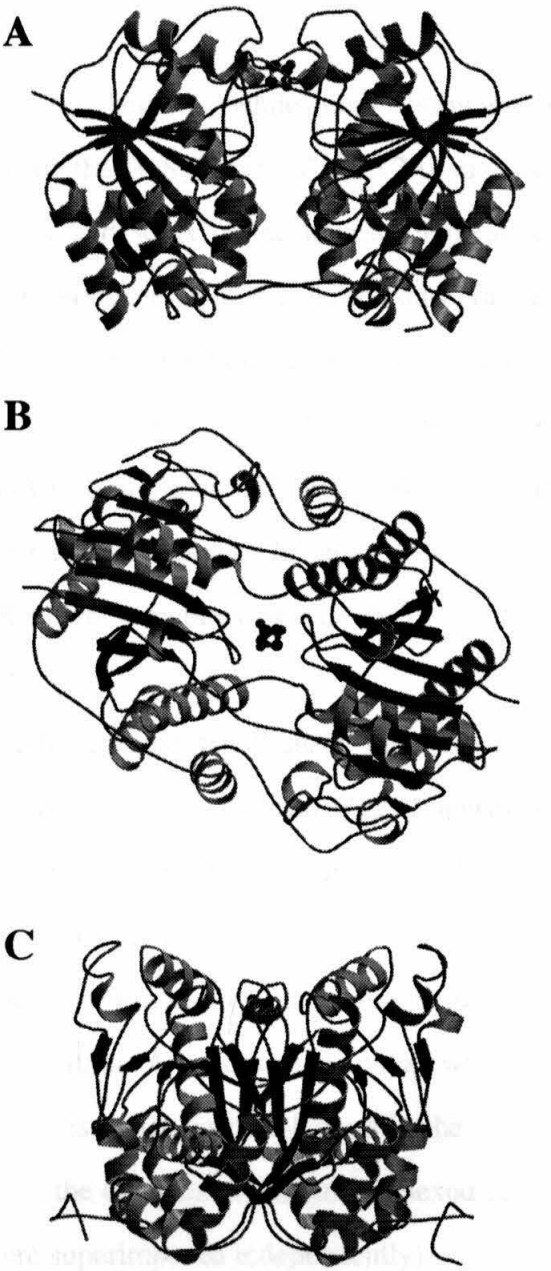


Figure 5-1. Ribbon representation of the 2.2 Å resolution Av2 structure. The model is shown from the "front" (A), "top" (MoFe binding face) (B), and "side" (C). The two-fold noncrystallographic axis lies vertically in the plane of the page in (A), perpendicular to the plane of the page in (B), and horizontally in the plane of the page in (C). The Fe_4S_4 cluster is depicted as a ball-and-stick model. Coils represent α-helices and arrows represent β-strands. The figure was generated using MOLSCRIPT/RASTER3D.

β2 form the nucleotide binding regions, while loops following β4 and β5 contain the cysteine residues which ligate the Fe_4S_4 cluster. The Av2 β-strands, marked by excellent 2 $|F_{obs}| - |F_{calc}|$ electron density and some of the lowest temperature factors in the model, represent areas of strong sequence and structural conservation. The RMS deviation between superimposed β-sheet Cα atoms of the two Av2 subunits is 0.2 Å. When the β-sheet Cα atoms of Cp2 are superimposed onto those of Av2, the RMS deviation is 0.3 Å for each subunit and 0.6 Å for the dimer. The Av2 β-sheet appears to remain constant in the AlF_4^- stabilized complex structure, as the Cα atom RMS deviation for β-sheet residues superimposed between the two structures is 0.5 Å for each subunit (subunits were superimposed independently).

The α-helices which surround the β-sheet form are less structurally conserved between Av2 monomers, and between Av2 and Cp2. Superimposing the α-helical Cα atoms of each Av2 subunit yields an RMS deviation of 0.4 Å. Between Cp2 and Av2, the value rises to 0.8 Å for each subunit and 1.1 Å for the dimer. Comparison of the Av2 and the AlF_4^- stabilized complex structures shows large reorganization not only between Av2 subunits, but also of the position of helices (and loops) within each subunit. Based on superposition of those residues which form α-helices in the uncomplexed Av2 structure, the RMS deviation between the complexed and uncomplexed Av2 Cα atoms is 1.1 Å for each subunit (subunits were superimposed independently).

The loop regions in the present model varied in quality. While many loops, such as those of residues 9-15 and 39-43 were extremely well-defined, others, such as those of residues 61-75, 116-118, and 173-175, were less clear. The positions of most loop atoms were confidently assigned by the termination of the refinement. However, one particular loop, residues B61-B75, eluded model building attempts. This long loop extends into the MoFe binding region, and while reasonable electron density was available for subunit A, that for subunit B remained ambiguous. Packing interactions prevented modeling of the B

loop to match that of subunit A. A conservative loop was modeled instead for the B subunit, based on the strong but uninterpretable electron density in the region. The occupancy for main chain atoms in residues B67-B74 was set to q = 0.0.

5.2. Similarity to other Iron Protein Structures

Comparisons between the present model and those of other Fe protein structures are important for identifying regions of structural conservation and variability. Each structure provides a valuable snapshot of the protein. Taken together, the various models form a "composite" view of Fe protein, a vehicle for understanding the structural basis of its functions.

5.2.1 Similarity to the 2.9 Å Resolution Structure

While the 2.2 Å resolution Av2 structure generally resembles the 2.9 Å resolution structure, several notable differences were identified. The present structure was determined to a superior resolution that allowed confident modeling of most of the atoms. Also, no nucleotide was observed in the present structure, while a partially-ordered, endogenous ADP molecule was detected bound between the Av2 subunits in the 2.9 Å resolution structure. In the current structure, additional single helical turns have been identified at residues 46-49, A72-A75, 80-82, 114-116, A172-A174, and 227-230. While the RMS deviation for all Cα atoms superimposed between the two structures is 1.3 Å (1.1 Å for each monomer), larger differences are found in several regions. For example, the loop at residues 40-43 adopts a different conformation in the 2.2 Å resolution structure. This loop, located adjacent to the nucleotide-binding loop, points toward the nucleotide-binding domain in the 2.9 Å resolution structure, but towards the Fe_4S_4 cluster in the present model. This difference may reflect a conformational flexibility for the region which is

involved in Av2 function. Alternatively, the loop may have been poorly defined in electron density maps for the lower resolution structure. The RMS deviation for Cα atoms in this loop is 3.1 Å.

The C-terminus of each subunit also differs considerably between the two Av2 structures. Crystal packing interactions have resulted in rearrangement of the C-terminus of subunit B in the present model, so that the C-terminus of subunit B of one molecule packs tightly against the MoFe protein binding face of a neighboring molecule. In the $P2_12_12_1$ cell, lattice contacts include hydrogen bonds from A97 N to B'289 OT1 (d = 2.83 Å), A97 N to B'289 OT2 (d = 2.67 Å), and B171 OH to B'287 Oε1 (d = 3.36 Å). Additional water-mediated interactions exist between A140 Nζ and B'278 Oδ2 and B'280 Oγ. In the $P2_1$ cell, the C-terminus of subunit B approaches the neighboring molecule less closely. While several potential contacts have been identified (A100 Nη2 - A'238 Oδ2 (d = 2.72 Å), A140 Nη2 - A'231 Oδ2 (d = 3.49 Å), B95 Cγ2 - B'285 O (d = 3.12 Å), A140 Nη1 - B'285 Cβ (d = 2.91 Å), the MoFe binding face of the $P2_1$ model does not appear to be altered by the packing arrangement. In contrast, the $P2_12_12_1$ model displays significant distortion by the neighboring molecule, particularly in the Fe_4S_4 ligating regions. The RMS deviation between Av2 models for the packing-affected residues is 2.7 Å. Other loop residues of Av2 which show high RMS deviations are 66-73, 110-113, 116-118, and 173-175. Changes in the $P2_12_12_1$ Av2 structure at the MoFe binding surface are intriguing because they may hold some physiological relevance. A superposition of the 2.9 Å and 2.2 Å resolution Av2 structures is shown in Figure 5-2.

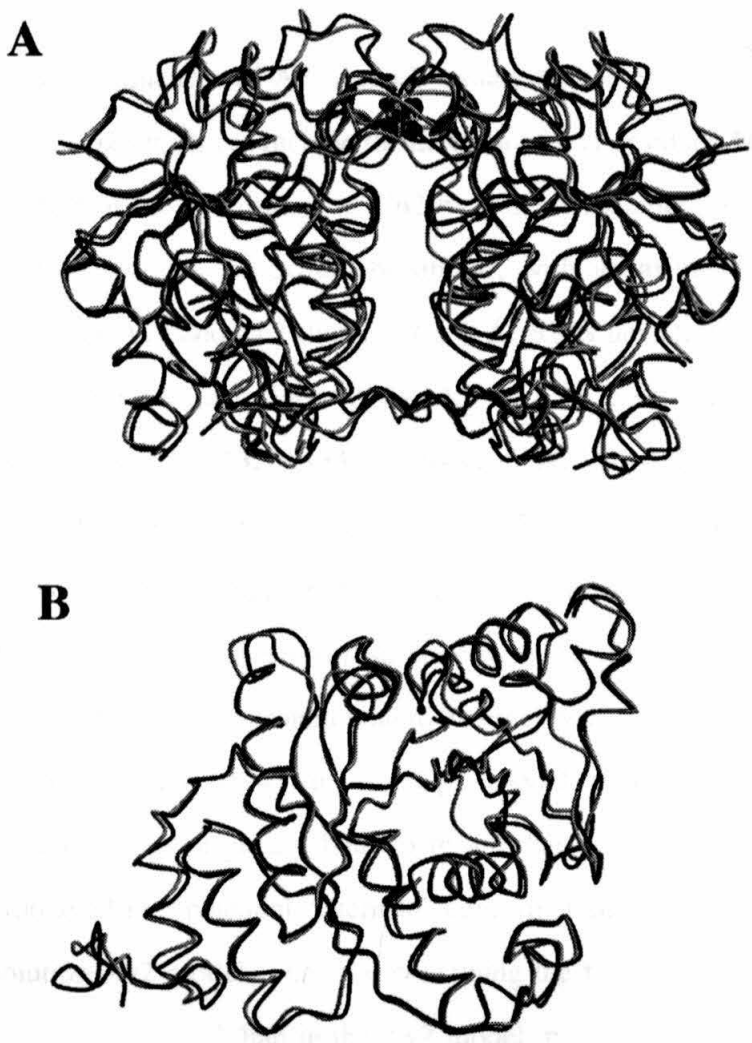


Figure 5-2. Superposition of the present 2.2 Å resolution $P2_12_12_1$ Av2 model and the previously determined 2.9 Å resolution $P2_1$ Av2 model. Dimers are shown superimposed in (A), and subunit A from each model is superimposed in (B). The 2.2 Å resolution $P2_12_12_1$ Av2 model is shown in dark grey, and the 2.9 Å resolution $P2_1$ Av2 model is in light grey. The figure was generated using MOLSCRIPT/RASTER3D.

5.2.2. Similarity to Cp2

The amino acid sequences of Av2 and Cp2 share 69% identity. Major differences between the Av2 and Cp2 sequences include the deletion of Av2 residues Thr^{66} and Val^{67} in Cp2, and the truncation of the C-terminus of Cp2 by thirteen residues. As anticipated, the Av2 and Cp2 crystal structures are extremely similar, with an average RMS deviation between Cα atoms of the Av2 and Cp2 dimers of 1.3 Å. RMS deviations greater than 2.0 Å occur between the Cα atoms of residues A40-A42, A51-A52, A64-A68, A90-91, A98-A100, A112-A113, A116-117, A132-A134, A189-A195, A266-A272, B40-42, B51-B52, B64-B71, B91-B96, B116-B18, B189-B194, and B259-B272. These large deviations occur mainly in loop regions, and may reflect differences due to crystal packing effects or model uncertainty, as well as genuine differences between the structures which reflect the conformational variability of the Fe protein vital to its functions.

In particular, residues 40-43 in the Cp2 model appeared "conformationally challenged." This loop, while originally built to match that of the 2.9 Å resolution Av2 model, was also modeled in a potential "alternate" conformation which was similar to that of the 2.2 Å resolution Av2 model. The loop containing the two-residue deletion in Cp2 (residues 61-75) is better defined than in the Av2 model, presumably due, in part, to the fact that the loop is shorter in the Cp2 sequence. Large differences (RMS deviation > 2.0 Å) in the loop at residues 187-194 appear both between the Cp2 subunits, and between Av2 and Cp2. These differences are suspected to be attributable to poor density in the Cp2 model. A superposition of the 2.2 Å resolution Av2 and 2.0 Å resolution Cp2 structures is shown in Figure 5-3.

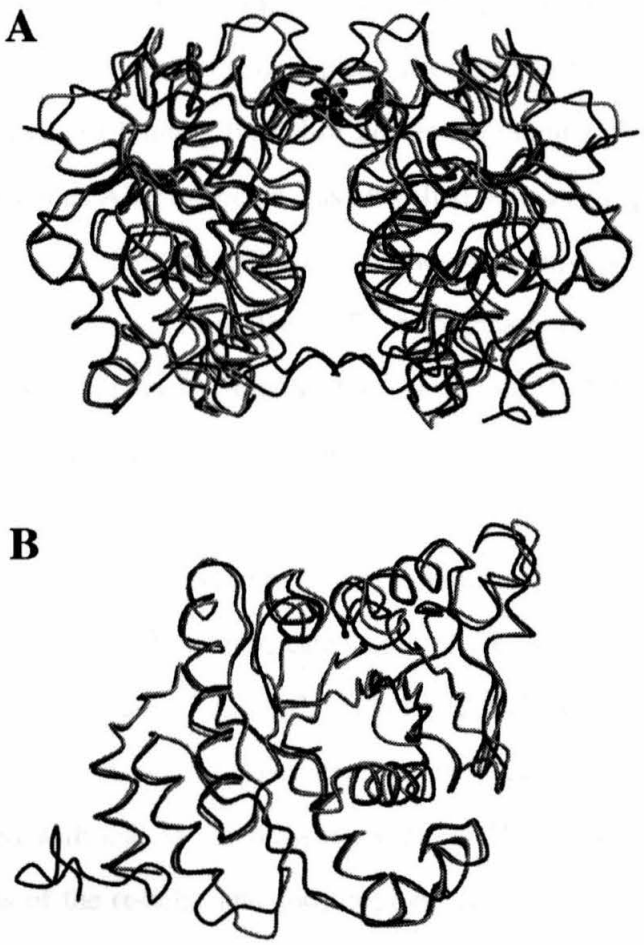


Figure 5-3. Superposition of the present 2.2 Å resolution $P2_12_12_1$ Av2 model and the recently determined 2.0 Å resolution Cp2 model. Dimers are shown superimposed in (A), and subunit A from each model is superimposed in (B). The 2.2 Å resolution $P2_12_12_1$ Av2 model is shown in dark grey, and the 2.0 Å resolution Cp2 model is in light grey. The figure was generated using MOLSCRIPT/RASTER3D.

5.2.3. Similarity to Av2 in the $(AlF_4^-•Av2)_2Av1$ Complex

Comparisons of the nucleotide-free Fe protein models with those of the AlF_4^- nitrogenase complex and numerous related nucleotide-bound proteins suggest the site of

MgATP binding in uncomplexed Fe protein, in the P-loop (phosphate-binding loop), lies 19 Å from the Fe_4S_4 cluster. MgATP binding in uncomplexed Fe protein is expected to adopt this "parallel" mode of binding typical of nucleotide-binding proteins. Previously, a low-occupancy endogenous ADP molecule was identified bound in a "perpendicular" mode which spans the subunit interface in the 2.9 Å Av2 model. A superposition of the parallel-bound MgATPs (in light grey) from the AlF_4^- nitrogenase complex, as well as the perpendicular-bound ADP (in black) from the 2.9 Å resolution Av2 model, onto the current 2.2 Å resolution Av2 model, is shown in Figure 5-4.

The most striking difference between the complexed and uncomplexed Av2 structures is the change in the orientation between the protein subunits. If the Fe_4S_4 cluster is considered to be a hinge for the dimer, calculation of the hinge bend angle between the subunits suggests a 13 ° closure of the AlF_4^- complexed Av2 with respect to the present model. Within each subunit, the β-sheet core of the AlF_4^- complexed Av2 remains structurally conserved with that of the present structure. However, the relative positions and/or conformations of the α-helix and loop regions change between the structures. In particular, helix α5 appears to shift dramatically downward in the complexed Av2 structure. Assessment of changes in the MoFe binding region is complicated by the crystal packing interactions in the $P2_12_12_1$ Av2 structure. A superposition of the AlF_4^- complexed Av2 and the present model is shown in Figure 5-5.

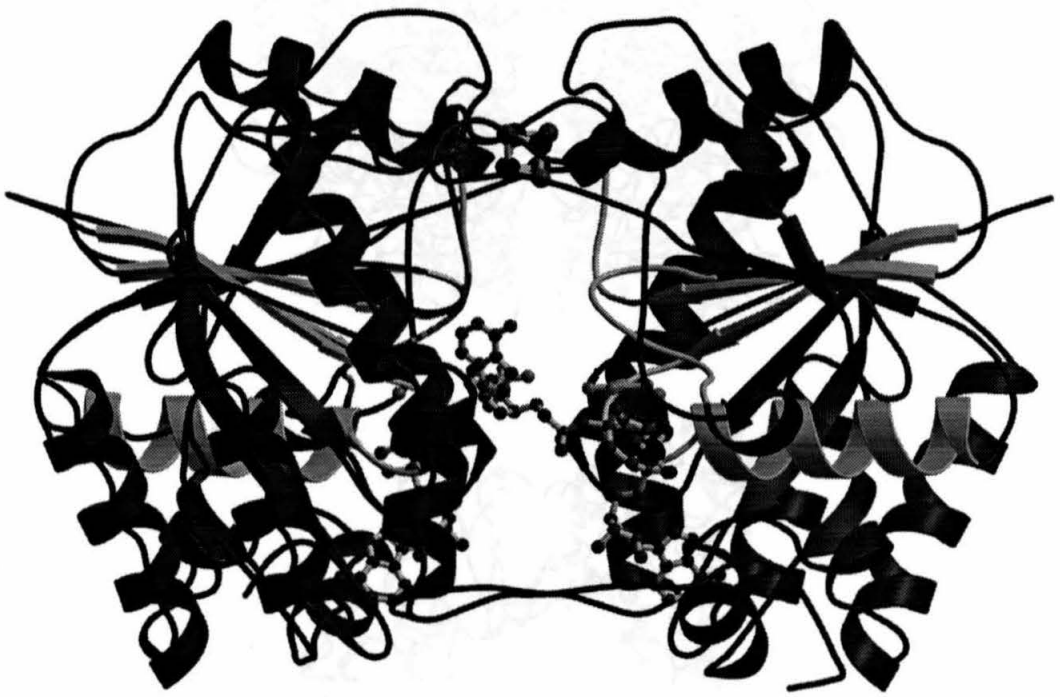


Figure 5-4. Ribbon representation of the 2.2 Å resolution Av2 structure, with nucleotides from the 2.9 Å resolution Av2 structure and AlF_4^- complex structure superimposed to indicate "perpendicular" and "parallel" modes of nucleotide binding. Regions of the protein involved in nucleotide hydrolysis (the P-loop, Effector I and II regions) are shown in white. The two-fold noncrystallographic axis lies vertically in the plane of the page. The Fe_4S_4 cluster, MoO_4^{2-} and AlF_4^- anions, and MgADP molecules are depicted as ball-and-stick models. Coils represent α-helices and arrows represent β-strands. The figure was generated using MOLSCRIPT/RASTER3D.

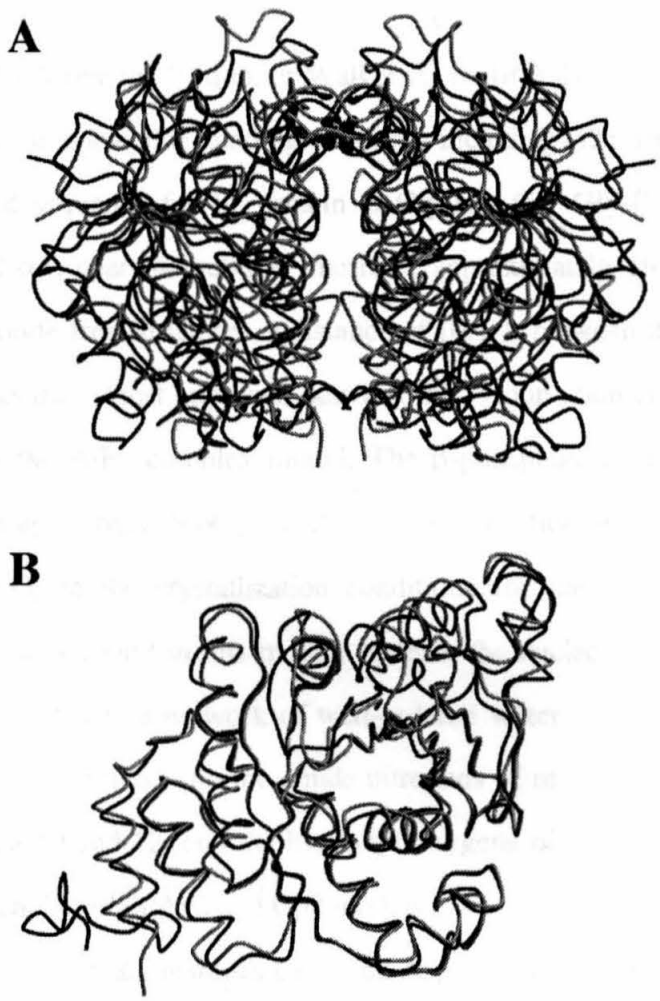


Figure 5-5. Superposition of the present 2.2 Å resolution $P2_12_12_1$ Av2 model and the Av2 model from the recently determined 3.0 Å resolution AlF_4^- stabilized complex structure. Dimers are shown superimposed in (A), and subunit A from each model is superimposed in (B). The Fe_4S_4 cluster, AlF_4^- anions, and MgADP molecules are depicted as ball-and-stick models. The 2.2 Å resolution $P2_12_12_1$ Av2 model is shown in dark grey, and the 3.0 Å resolution complex model is in light grey. Coils represent α-helices and arrows represent β-strands. The figure was generated using MOLSCRIPT/RASTER3D.

5.3. Nucleotide Binding and Hydrolysis

Residues 9-15 define the P-loop, or Walker A, motif, Gly-Xaa-Xaa-Xaa-Xaa-Gly-Lys-Ser/Thr, typical of nucleotide binding proteins, including p21 ras, transducin, and recA. The conserved sequence for Fe protein is Gly-Lys-Gly-Gly-Ile-Gly-Lys-Ser, and lies between β1 and α1, near the subunit interface, approximately 19 Å from the Fe_4S_4 cluster. In the nucleotide-free structures, a distance of 10.1 Å between the Gly^{12} Cα atoms of each subunit marks the closest approach between the P-loop main chains. This distance reduces to 4.6 Å in the AlF_4^- complex model. The β-phosphate position of ATP in the AlF_4^- complex is occupied by a MoO_4^{-2} in the 2.9 Å resolution Av2 model. Despite the inclusion of Na_2MoO_4 in the crystallization conditions for the 2.2 Å resolution Av2 crystals, no MoO_4^{-2} was found in this model. Instead, the nucleotide binding regions of Av2 (and Cp2) are lined with a network of well-ordered water molecules. These solvent molecules form hydrogen bonds with the amide nitrogens of residues 10-17, the carbonyl oxygens of residues 8-10 and 12, and the hydroxyl oxygens of Ser^{16} and Thr^{17}; the water channel is not identical between Av2 and Cp2 models.

A salt bridge connecting residues Lys^{15} and Asp^{125} is critical to MgATP hydrolysis. Lys^{15} has been implicated in γ-phosphate binding of MgATP, and the ionic bond provides a clear link between the nucleotide binding and Fe_4S_4 cluster regions of the Fe protein. Mutations in either residue abolish nitrogenase activity . This salt bridge is clearly defined in the model, with a representative distance of 2.72 Å between Av2 B Lys^{15} Nζ and B Asp^{125} Oδ1. Lys^{15} Nζ makes additional potential hydrogen bonds with Ile^7 O and Val^{126} O in Av2, and a salt bridge with Asp^{39} Oδ2 in Cp2. Ser16 Oγ lies within hydrogen bonding distance to Asp^{43} Oδ2 and O in Av2 subunit A and Asp^{125} Oδ2 in Av2 subunit B. In Cp2, Ser^{16} Oγ can hydrogen bond with Asp^{125} Oδ2, and B Ser^{16} Oγ may bond with B Asp^{39} Oδ2 in addition to those contacts observed for Av2. Lys^{10} Nζ hydrogen bonds to the carbonyl O

of Gly^{11}, while Gly^{11} N is bridged to Asp^{129} Oδ1. Hydrogen bonds between P-loop backbone atoms and neighboring amide nitrogens of residues Val^{150}, Gly^{128}, Thr^{18} help to stabilize the region. Mean RMS deviation for eight superimposed Av2 and Cp2 Cα atoms in the P-loop is 0.24 Å. Superposition of residues 10-17 of p21 ras onto Av2 A9-A15 yields an Cα RMS deviation of 0.71 Å. Figure 5-6 shows 2 $|F_{obs}|$ - $|F_{calc}|$ electron density for the Av2 P-loop region.

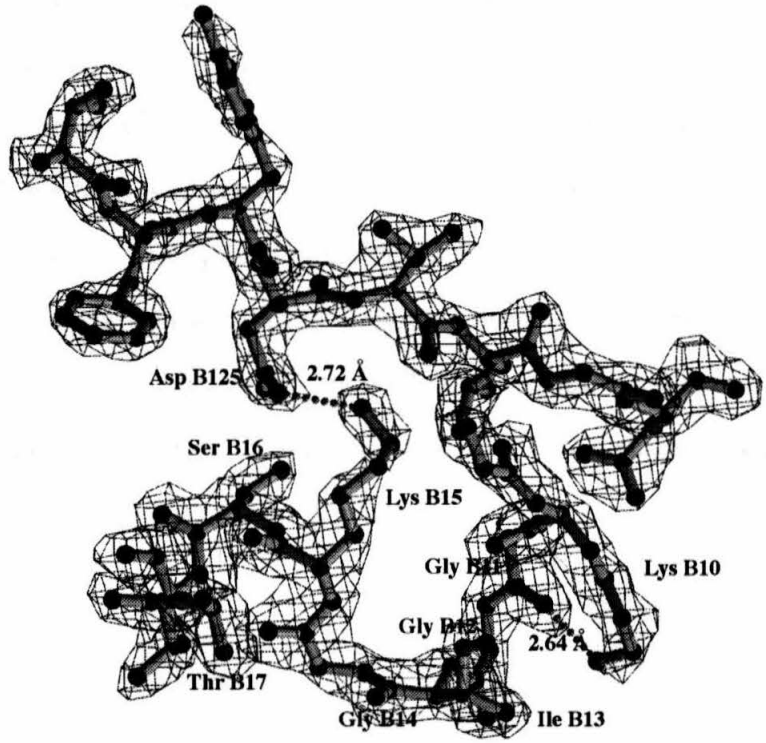


Figure 5-6. Ball-and-stick model of the P-loop of Av2. Subunit B, with 1.5 σ 2 $|F_{obs}|$ - $|F_{calc}|$ electron density, is shown. A 2.72 Å salt bridge connecting B Lys^{15} Nζ and B Asp^{125} Oδ1 is critical for nucleotide hydrolysis. B Lys^{10} also forms a 2.83 Å hydrogen bond to B Gly^{11} O. The figure was generated using FRODO/TOM (Jones, 1988), FRODOMOL (M. Ludwig), and MOLSCRIPT.

Adjacent to the P-loop, between β2 and α2, residues 38-44 comprise the "Effector I" region of Fe protein. This loop, with conserved sequence Cys-Asp-Pro-Lys-Ala-Asp-Ser, is marked by conformational flexibility. Av2 shows a significant (RMS 4.0 Å) change compared with the previous 2.9Å Av2 model, with the main chain and side chain atoms of residues 39, 41, and 43 exchanged from the earlier model. A potential salt bridge connects Lys^{41} Nζ and Asp^{43} Oδ1 in subunit A. The AlF_4- complex model matches the 2.2 Å resolution Av2 model in this region. The Cp2 model favors that of the 2.9 Å resolution model, although an alternative conformation similar to that of the present structure was also modeled. In the current Av2 model, the loop points towards the Fe_4S_4 cluster, while in the previous model, the loop is directed towards the P-loop. Such flexibility is likely to have mechanistic implications. Analysis of mutations in this region, including Cys^{38}, Asp^{39}, and Asp^{43}, has begun to identify roles for the individual residues . In particular, Asp^{39} has been suggested to function as a molecular "trigger" for complex dissociation and nucleotide release following electron transfer (Lanzilotta *et al.*, 1997). An electron density map of the loop for the current structure is shown in Figure 5-7.

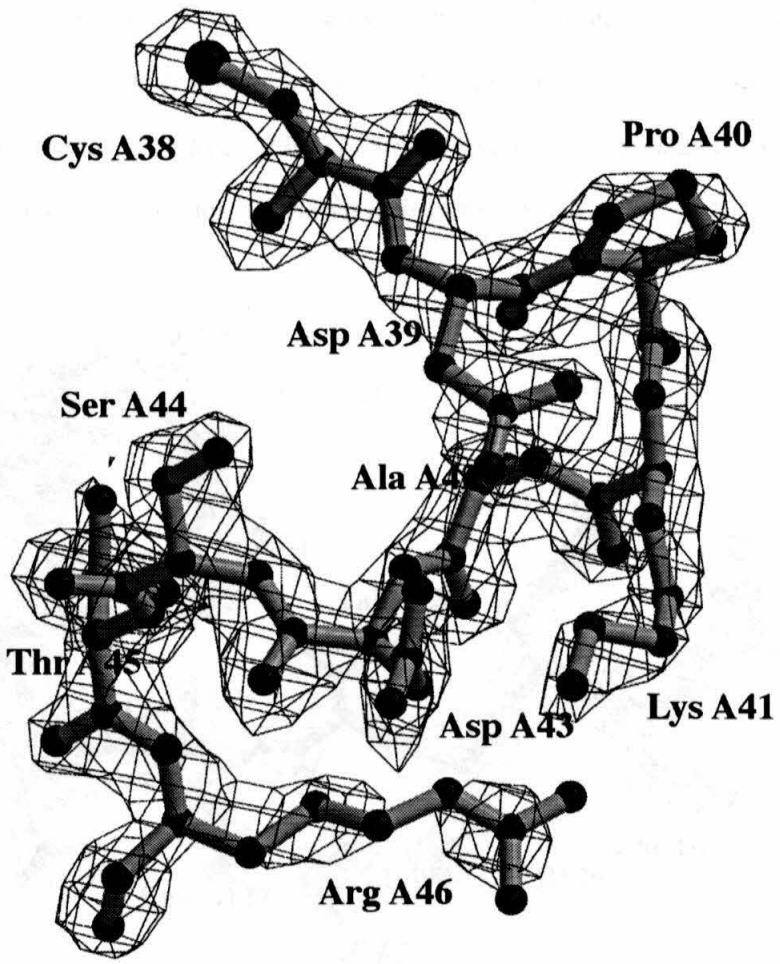


Figure 5-7. Ball-and-stick model of the Effector I loop of Av2. Subunit A, with 1.2 σ 2 $|F_{obs}| - |F_{calc}|$ electron density, is shown. This region shows considerable change compared with that of the previous 2.9 Å resolution model. Side chain atoms of Asp^{39} and Asp^{43} both point in the direction of Lys^{41} Nζ, although the distances between them (3.69 Å) are slightly long for strong interaction. The figure was generated using FRODO/TOM, FRODOMOL, and MOLSCRIPT.

Residues 125-132, the so-called "Effector II" region or "Walker B motif," complete the link between the nucleotide binding region of Fe protein and the Fe_4S_4 cluster. Asp^{125}, salt bridged to Lys^{15} of the P-loop, is located on strand β5, preceding the loop which

contains cluster ligand Cys[132]. This pathway has been confirmed by the Av2 deletion mutant L127Δ, which has been shown to exhibit spectroscopic characteristics of MgATP-bound Av2 (Ryle and Seefeldt, 1996a). Well-defined electron density was observed for this region in the current model, as shown in Figure 5-8.

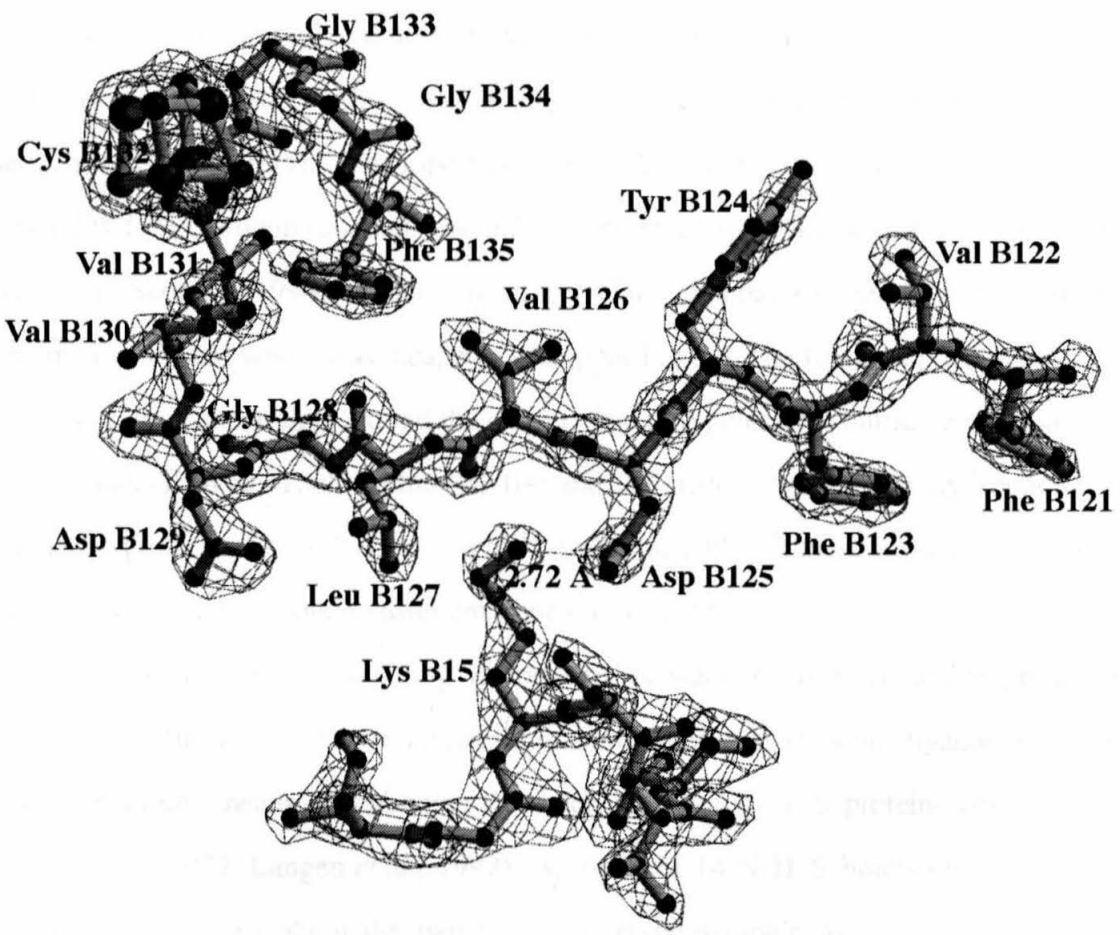


Figure 5-8. Ball-and-stick model of the Effector-II loop of Av2. The salt bridge between Lys15 Nζ and Asp125 Oδ1 is again shown, as well as the loop following β5 which includes the Cys[132] Fe_4S_4 ligand. Subunit B, with 1.5 σ 2 $|F_{obs}|$ - $|F_{calc}|$ electron density, is shown. The figure was generated using FRODO/TOM, FRODOMOL, and MOLSCRIPT.

5.4. [4Fe-4S] Cluster

The thiol ligands of Cys^{97} and Cys^{132} from each Fe protein subunit coordinate the cubane Fe_4S_4 cluster. The cluster's solvent-exposed position on one face of the dimer makes it exquisitely poised for electron donation to the MoFe protein. Amino acids immediately surrounding each ligand, Val^{95}, Gly^{96}, Ala^{98}, Gly^{99}, Val^{130}, Val^{131}, Gly^{133}, and Gly^{134}, are small, hydrophobic residues, conserved among all known Fe protein sequences (Woo, 1995). The importance of small residues in these loops in conferring flexibility for Fe protein function was underscored recently by studies of a mutant of Ala^{98} (Ryle and Seefeldt, 1996b). Replacement by valine rendered a protein which bound only one nucleotide and which was incapable of supporting nitrogen fixation, presumably due to increased steric constraints around the cluster. The Fe_4S_4 cluster is buttressed on three sides by the main chain atoms of residues 96-100 and 126-135, as well as the hydrophobic side chains of Pro^{40}, Pro^{91}, Ala^{98}, Leu^{127}, Val^{130}, Val^{131}, and Phe^{135}; these invariant residues are likely to be crucial in proper cluster environment formation.

Potential hydrogen bonding interactions between main chain amide groups and inorganic sulfur atoms of the cluster and thiol groups of cysteine ligands have been proposed as one means of regulating the redox potential in Fe-S proteins (Adman *et al.*, 1975; Carter, 1977; Langen *et al.*, 1992). A crown of 14 N-H···S bonds encircles the Cp2 cluster symmetrically about the two-fold non-crystallographic axis passing through the cluster center. Nine of these bonds obey a maximum bond distance criterion of 3.7 Å and minimum angle criterion of 120° (Backes *et al.*, 1991); the remaining five are incipient. Cys^{97} SG is situated at the amino end of helix α3 leading away from the cluster, allowing formation of a hydrogen bond with Arg^{100} N. Additional electrostatic interactions between the cluster and the helix dipole, which is in line with the cluster, further stabilize binding (Dao-Pin *et al.*, 1991).

Crystal packing forces distort the cluster region of Av2, resulting in differences in the N-H···S bond arrangement. The C-terminus of a neighboring molecule spirals into hydrogen bonding distance (<3.5 Å) of the cluster face, with loss of amide hydrogen bonds to Cys^{132} SG and the inorganic S1. Helix α3 is also affected by the packing, and begins at residue Gly^{101} rather than Cys^{97}. Mean RMS deviation for Cα atoms in the cluster environment between Av2 and Cp2 is 1.6 Å. However, deviations for the B subunit Cys^{132} ligand region, which is not affected by crystal packing forces in Av2, are significantly lower (< 0.4 Å). In the AlF_4^- complex structure, the amide hydrogen bonding pattern also shifts, with loss of hydrogen bonding to S2 and S3 and increased contacts to S1. Taken together, these three snapshots of the Fe protein cluster confirm its dynamic nature. Table 5-2 summarizes the N-H···S interactions for all three structures; Figure 5-9 indicates the N-H···S bonding network for each molecule. The packing interaction between Av2 molecules in the $P2_12_12_1$ cell is shown in Figure 5-10.

Figure 5-9. Potential N-H···S hydrogen bonding networks for Cp2 (A), Av2 (B), and the Av2 in $(AlF_4^-•Av2)_2Av1$ (C). Dark dashed lines indicate hydrogen bond interactions which meet the requirements of N-H···S > 3.7Å or N-H···S < 120°. Light dashed lines indicate incipient hydrogen bonds.

Table 5-2. Potential N-H···S Hydrogen Bond Interactions for Cp2, Av2, and the Av2 in $(AlF_4^-•Av2)_2Av1$

Fep	Donor	Acceptor	N···S Distance (Å)	N-H···S Angle (°)
Cp2	A96 N	A94 SG	3.62	108.2§
	A97 N	A94 SG	3.67	152.0
	A131 N	A129 SG	3.47	122.0
	A132 N	A129 SG	3.55	158.1
	B96 N	B94 SG	3.40	147.7
	B97 N	B94 SG	4.09§	150.4
	B131 N	B129 SG	3.65	111.3§
	B132 N	B129 SG	3.62	155.2
	A95 N	S4	3.70	135.4
	A96 N	S4	4.12§	173.3
	A131 N	S1	3.36	126.1
	B95 N	S2	3.44	143.5
	B96 N	S2	4.52§	138.8
	B131 N	S3	3.29	141.4
Av2	A99 N	A97 SG	3.70	153.7
	A134 N	A97 SG	3.58	171.4
	B99 N	B97 SG	3.38	161.8
	B134 N	B132 SG	3.65	122.4
	B135 N	B132 SG	3.67	153.1
	A98 N	S4	4.06§	133.7
	A99 N	S4	4.57§	147.5
	B98 N	S2	3.43	129.1
	B99 N	S2	3.69	126.5
	B134 N	S3	3.65	122.4
AlF_4^- Av2	E100 N	E97 SG	3.65	158.2
	E134 N	E132 SG	4.67§	129.9
	E135 N	E132 SG	4.83§	156.4
	F100 N	F97 SG	3.63	156.2
	F134 N	F132 SG	4.67§	129.8
	F135 N	F132 SG	4.83§	157.2
	E98 N	S4	3.52	92.5§
	E99 N	S4	3.01	158.1
	F98 N	S2	3.50	92.6§
	F99 N	S2	3.01	158.0

§ Considered to be an incipient hydrogen bond (N-H···S > 3.7Å or N-H···S < 120°).

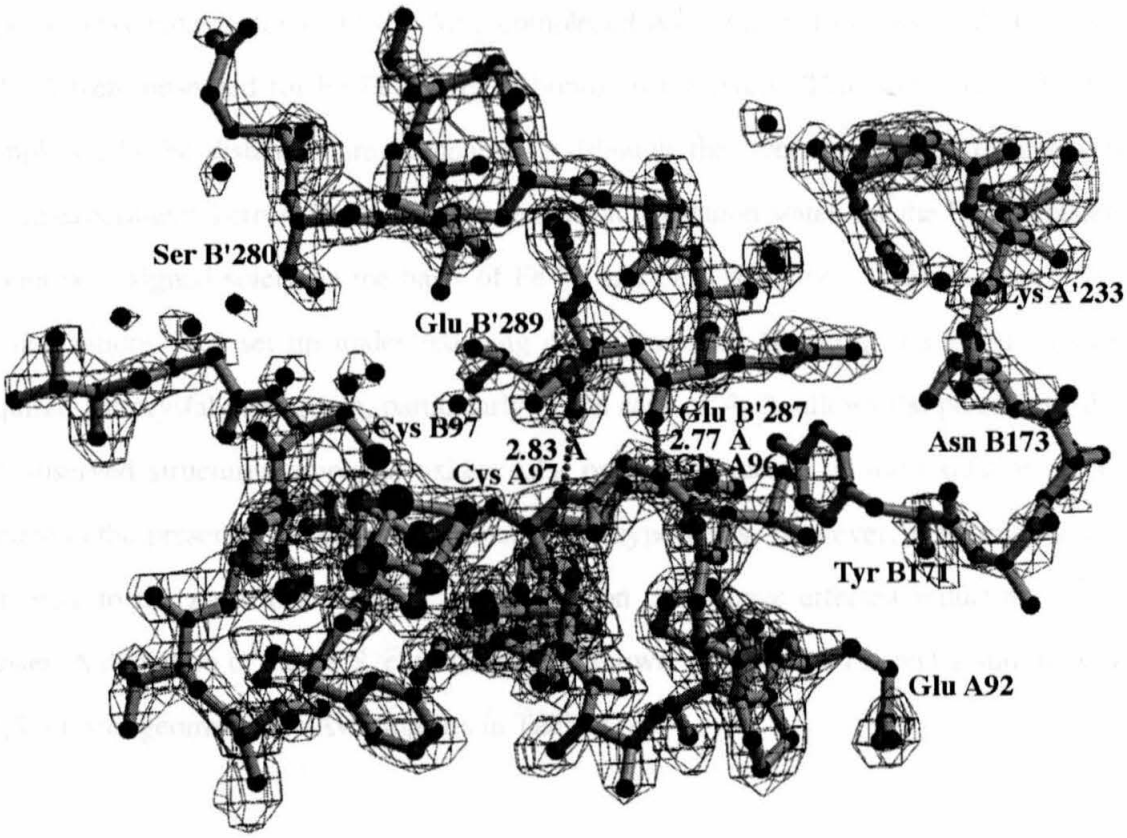


Figure 5-10. Ball-and-stick model of the lattice packing interaction between molecules in the $P2_12_12_1$ Av2 cell. The MoFe binding face of one molecule is impacted by the C-terminus of subunit B of a neighboring molecule. The models are shown with 1.5 σ 2 $|F_{obs}|$ - $|F_{calc}|$ electron density. The closest contacts involve hydrogen bonds between A96 N and B'287 O (2.83 Å), and A97 N and B'289 OT (2.77 Å). The figure was generated using FRODO/TOM, FRODOMOL, and MOLSCRIPT.

Mean Fe-Fe and Fe-S bond distances are 2.63 Å and 2.29 Å in Cp2, respectively. In the Av2 cluster, these values shorten to 2.49 Å for Fe-Fe and 2.19 Å for Fe-S distances. Peculiarities in the Av2 distances may reflect crystal packing interactions. Target geometries in the parameter files used in the X-PLOR refinements of Cp2 and Av2 were identical (2.7 Å for Fe-Fe, 2.3 Å for Fe-S), although the weight of geometric factors in each refinement

protocol may have differed. In the AlF_4^- complexed Av2, mean distances of 2.50 Å and 2.28 Å were observed for Fe-Fe and Fe-S bonds, respectively. Thus, the Av2 and AlF_4^- complexed Fe-Fe distances are comparable, although the overall range of distances is within experimental error. At the present resolution, oxidation states for the Fe_4S_4 clusters cannot be assigned solely on the basis of Fe-Fe and Fe-S distances. While Av2 and Cp2 crystallizations were set up under reducing conditions with $Na_2S_2O_4$, the length of time required for crystal emergence, particularly in the case of Av2, allows the possibility that the observed structures represent oxidized Fe protein. Evidence of self-oxidation of Fe protein in the presence of dithionite supports this hypothesis. However, it is possible that exposure to the x-ray beam during data collection could have effected reduction of the cluster. A depiction of the Fe_4S_4 environment is shown in Figure 5-11, and a summary of Fe_4S_4 cluster geometry for Av2 appears in Table 5-3.

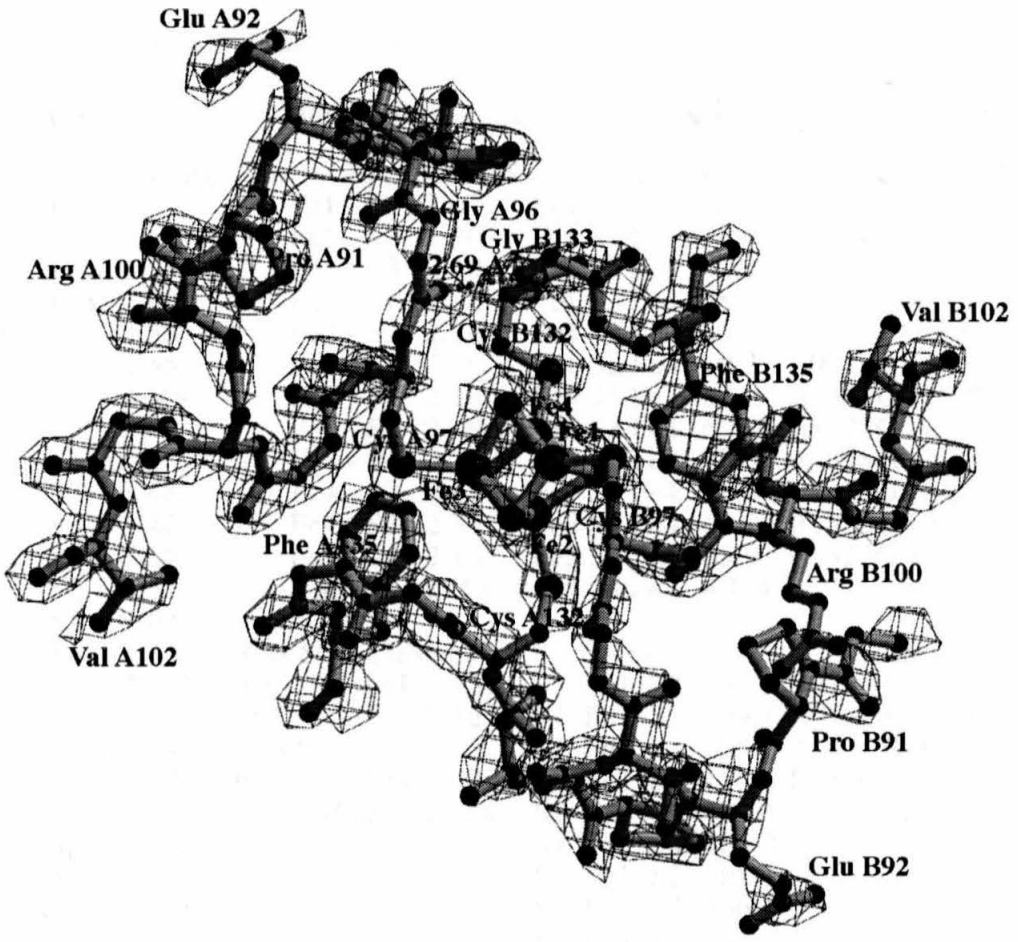


Figure 5-11. Ball-and-stick model of the local Fe_4S_4 environment in Av2. The model is shown with 1.5 σ 2 $|F_{obs}|$ - $|F_{calc}|$ electron density. The figure was generated using FRODO/TOM, FRODOMOL, and MOLSCRIPT.

Table 5-3. Av2 Fe_4S_4 Cluster Geometry

Interatomic Distances (Å)			Bond Distances (Å)			Angles (°) Cluster-Cluster			Angles (°) Cluster-Ligand		
Cluster Fe - Cluster Fe			*Cluster Fe - Cluster S*			*Fe - S - Fe*			*CysSG - Fe - CysSG*		
Fe1	Fe2	2.52	Fe1	S1	2.07	Fe1	S1 Fe3	72.5	B97 SG	Fe1 S1	118.9
Fe1	Fe3	2.52	Fe1	S2	2.20	Fe1	S1 Fe2	70.6	B97 SG	Fe1 S3	122.8
Fe1	Fe4	2.53	Fe1	S3	2.19	Fe1	S3 Fe3	70.2	B97 SG	Fe1 S2	87.6
Fe2	Fe3	2.49	Fe2	S1	2.28	Fe1	S3 Fe4	68.4	A97 SG	Fe3 S1	120.9
Fe2	Fe4	2.33	Fe2	S2	2.22	Fe1	S2 Fe2	69.6	A97 SG	Fe3 S3	114.5
Fe3	Fe4	2.55	Fe2	S4	2.16	Fe1	S2 Fe4	70.3	A97 SG	Fe3 S4	101.0
Mean		**2.49**	Fe3	S1	2.18	Fe3	S1 Fe2	67.9	A132SG	Fe2 S1	101.4
			Fe3	S3	2.19	Fe3	S3 Fe4	69.0	A132SG	Fe2 S2	111.9
Cluster S - Cluster S			Fe3	S4	2.14	Fe3	S4 Fe2	70.8	A132SG	Fe2 S4	120.9
S1	S2	3.48	Fe4	S2	2.19	Fe3	S4 Fe4	73.4	B132SG	Fe4 S3	101.4
S1	S3	3.42	Fe4	S3	2.31	Fe2	S2 Fe4	63.9	B132SG	Fe4 S2	113.0
S1	S4	3.55	Fe4	S4	2.13	Fe2	S4 Fe4	65.9	B132SG	Fe4 S4	117.8
S2	S3	3.57	*Mean*		**2.19**	*Mean*		**69.4**	*Mean*		**111.0**
S2	S4	3.64									
S3	S4	3.46	*Cys SG - Cluster Fe*			*S - Fe - S*			*CysCB - SG - Fe*		
Mean		**3.52**	B97 SG	Fe1	2.15	S1	Fe1 S3	106.9	B97 CB SG	Fe1	113.0
			A97 SG	Fe3	2.12	S1	Fe1 S2	109.1	A97 CB SG	Fe3	106.6
Cluster Fe - Cluster S			A132SG	Fe2	2.35	S1	Fe3 S3	103.1	A132CBSG	Fe2	117.7
Fe1	S4	3.69	B132 SG	Fe4	2.34	S1	Fe3 S4	110.6	B132CBSG	Fe4	101.3
Fe2	S3	3.65	*Mean*		**2.24**	S1	Fe2 S2	101.4	*Mean*		**109.7**
Fe3	S2	3.76				S1	Fe2 S4	106.2			
Fe4	S1	3.63				S3	Fe1 S2	108.7			
Mean		**3.68**				S3	Fe3 S4	106.1			
						S3	Fe4 S2	104.7			
Cys SG - Cys SG						S3	Fe4 S4	102.5			
A97	A132	5.99				S2	Fe2 S4	112.2			
A97	B97	6.54				S2	Fe4 S4	114.8			
A97	B132	5.97				*Mean*		**107.2**			
A132	B132	6.56									
B97	A132	5.65									
B97	B132	5.72									
Mean		**6.07**									

5.5. MoFe Protein Binding Surface

The recent AlF_4- nitrogenase complex structure confirmed the role of Arg^{100} in providing a key Fe protein contact with the MoFe protein, revealing a direct electron pathway from the Fe protein Fe_4S_4 to the P-clusters of the MoFe protein (Schindelin *et al.*, 1997). Arg^{100} extends from the Fe_4S_4 cluster region of Fe protein to bind to glutamate residues α120 and β120 of the MoFe protein. ADP-ribosylation of Arg^{100} in some photosynthetic nitrogen-fixing bacteria has been shown to function in nitrogenase regulation by blocking complex formation (Pope *et al.*, 1985; Ludden *et al.*, 1989ab). Additional residues along the MoFe binding surface include a long loop from residues 64 to 75. This loop of primarily acidic residues contains a two residue deletion in Cp2 compared with Av2 and has been speculated to contribute towards the formation of a tight-binding inactive complex of Cp2 and Av1 (Peters *et al.*, 1995). Side chain atoms of Glu^{68}, Asp^{69}, and Asp^{73} in the Cp2 loop lack convincing density, although main chain atoms are well-ordered. The longer Av2 loop is reasonably well-defined in Av2 subunit A, but only partially ordered in subunit B. These loops are ordered in the nitrogenase complex structures, presumably due to stabilization by the MoFe protein.

Cross-linking experiments revealed a specific interaction between Av2 Glu^{112} and Lys^{399} of the Av1 β subunit (Willing *et al.*, 1989, 1990). While residue Glu^{112} is located on the MoFe protein interface, it probably has no universal role in nitrogenase complex formation, as it is not conserved in Fe protein sequences-- glutamine is the preferred amino acid. Additional charged residues along the MoFe protein interface of Fe protein include Glu^{92}, Glu^{110}, Glu^{111}, Asp^{117}, Arg^{140}, Glu^{141}, Lys^{143}, Lys^{166}, Lys^{170}, and Arg^{178}. In the unbound Av2 structure, many of these residues form internal salt bridges, such as those between Arg^{140} and Glu^{141}, and between Glu^{110} and Lys^{143}. However, of the charged MoFe interface residues, only Arg^{100} is invariant among known Fe protein sequences, so the roles of the remaining residues in nitrogenase activity are probably secondary.

5.6. Intersubunit Interactions

Extraction of the Fe_4S_4 cluster does not result in immediate loss of Fe protein dimerization, presumably due to extensive intersubunit contacts (Georgiadis *et al.*, 1992). In the immediate vicinity of the Fe_4S_4 cluster, residues 88-97 form a loop which overlaps the opposing subunit at the MoFe binding surface of the protein and includes cluster ligand Cys^{97}. Three pairs of cross-subunit hydrogen bonds form between Gly^{96} O and Gly^{133} N, Gly^{94} N and Val^{131} O, and Gly^{94} O and Tyr^{171} OH in Cp2 (Woo, 1995). Differences in the Fe_4S_4 region of the Av2 model due to packing interactions prohibit formation of cross-subunit hydrogen bonds between Tyr^{171} OH and Gly^{94} O, and between B Gly^{96} O and A Gly^{133} N. However, a unique 3.09 Å hydrogen bond can form between B Gly^{94} O and A Lys^{170} Nζ in Av2. Lys^{170} Nζ had been observed in a salt bridge with Glu^{92} Oε atoms from the opposing subunit in the 2.9 Å resolution Av2 structure. The Glu^{92}-Lys^{170} interaction had been considered significant, because the residues are conserved in all but one known Fe protein sequence (Woo, 1995). In Cp2, the distances between the corresponding residues are too long for ionic interactions (4.2 Å, 4.4Å), although the side chains seem poised to interact. In Av2, Glu^{92} seems more likely to interact with Lys^{166} Nζ of the opposing subunit, although distances are still slightly long (3.5 Å, 3.8 Å). Residue 166 is a positively charged residue (Arg or Lys) in all known Fe protein sequences. The cyclic arrangement of salt bridges in the 2.9 Å resolution Av2 structure, A Lys^{41} - B Asp^{129} - B Lys^{41} - A Asp^{129} (Georgiadis *et al.*, 1992), could not be confidently assigned in the Cp2 model (Woo, 1995), and is not present in the 2.2 Å resolution Av2 model. However, a more extensive network of intersubunit interactions was identified in the AlF_4^- complexed protein (Schindelin *et al.*, 1997).

Hydrophobic interactions, such as those between opposing Met^{155} and Met^{156} residues, may also help to stabilize the dimer. This particular region is additionally

interesting because A157S substitution results in a mutant protein able to bind nucleotide but incapable of undergoing any conformational change (Gavini *et al.*, 1992). The Fe protein C-terminus provides an additional degree of intersubunit interaction in Av2, as the C-terminal residues of each monomer wrap around the body of the opposing subunit. In particular, cross-subunit salt bridges form between Lys^{224} and Glu^{277}, and Lys^{233} and Glu^{287}. Hydrogen bonds connect Arg^{223} of the A subunit with Lys^{284} O and Val^{289} OT1 of the B subunit in the Av2 model. Although the C-terminus has not been implicated in Fe protein function, such interactions may contribute to the overall stabilization of Av2. The shorter C-terminus of Cp2 does not extend far enough onto the opposing subunit to create substantial interactions. By calculation with DSSP (Kabsch and Sander, 1972), the total amount of buried surface area for each monomer resulting from dimer formation is 1786 $Å^2$ in Av2 and 985 $Å^2$ in Cp2; once the effects of the Av2 C-terminus are taken into account, the Av2 value drops to 965 $Å^2$.

5.7. Discussion

The three-dimensional structures of Av2 and Cp2 display extensive similarity which reflects their high degree of sequence conservation. Despite their 69 % amino acid sequence identity, Av2 and Cp2 represent two of the more disparate organisms among nitrogen fixing organisms. In an attempt to identify which features of the Av2 and Cp2 structures might be generally applicable to all nitrogen fixing organisms, Debbie Woo prepared an alignment of the amino acids of 39 Fe protein sequences (Woo, 1995). At the time of the alignment, this represented the full range of sequences available. A 64 residue invariant Fe protein core was identified, which is listed in Table 5-4. Not surprisingly, the majority of the core residues are located in the nucleotide binding and cluster regions. Several, such as Arg^{100}, Ile^{103}, Pro^{212}, Leu^{243}, Pro^{256}, and Pro^{258}, are located on the surface of the protein. Others, such as Thr^{226}, do not appear to serve any obvious role. A spatial representation of

the invariant residues in the Av2 structure is shown in Figure 5-12; conserved residues are indicated by spheres at the Cα position.

Table 5-4. Residues Forming the Invariant Core of Fe Protein (Adapted from Woo, 1995)

Number	Type	Number	Type	Number	Type
2	M	93	P	150	V
8	Y	96	G	156	M
9	G	97	C	157	A
10	K	98	A	159	Y
11	G	99	G	160	A
12	G	100	R	162	N
13	I	101	G	163	N
15	K	103	I	167	G
16	S	125	D	179	L
17	T	127	L	180	G
37	G	128	G	181	G
38	C	129	D	185	N
39	D	130	V	187	R
40	P	131	V	212	P
43	D	132	C	221	E
45	T	133	G	226	T
79	G	134	G	243	L
85	C	135	F	250	N
87	E	136	A	256	P
89	G	143	K	258	P
90	G	144	A		
91	P	148	Y		

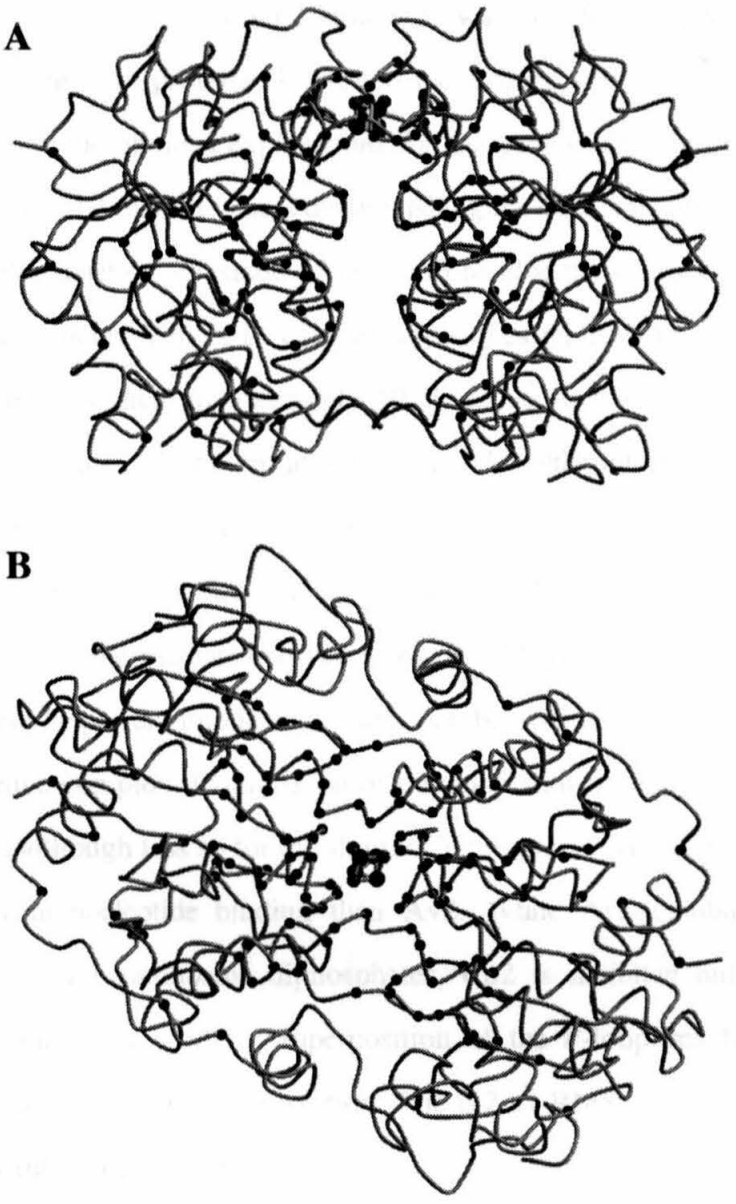


Figure 5-12. Spatial organization of the 64 residue Fe protein invariant core, shown from the "front" (A) and the MoFe binding face (B). Each of the 64 residues conserved among 39 known nitrogenase Fe protein sequences is depicted by a sphere at its Cα atom position. Many, but not all, of the residues are located near the nucleotide-binding and Fe_4S_4 cluster regions. The figure was generated using MOLSCRIPT/RASTER3D.

Comparison of the Av2 and Cp2 structures was sought to explain the functional differences between the two proteins. However, the striking similarities between the two structures do not provide obvious explanations for the genetic and functional differences between the proteins. The major features distinguishing the Cp2 structure from that of Av2 include a two residue deletion at residue 65 and a 13-residue truncation of the C-terminus. An Av2 mutant constructed with its 18 C-terminal residues replaced by the five C-terminal residues of Cp2 demonstrated approximately 50 % native Av2 activity (Jacobson *et al.*, 1990), indicating that the Av2 C-terminus plays a role, although non-essential, in Av2 function. A second construction, in which Av2 residues 59-67 were replaced by their Cp2 counterparts, including the Cp2 two residue deletion, produced a hybrid Fe protein with Av1 binding properties intermediate between those of wild-type Av2 and Cp2 (Peters *et al.*, 1995b). This hybrid Fe protein implicated residues in the Cp2 deletion loop in formation of the Cp2-Av1 inactive complex. Examination of this loop shows conformational variability in both structures, although less so for the shorter Cp2 sequence. In addition, Cp2 displays greater specificity in nucleotide binding than Av2. While Av2 exhibits inhibition by adenosine, inosine, and xanthosine diphosphates, Cp2 is inhibited only by adenosine diphosphate (Weston *et al.*, 1983). Superposition of the P-loop residues of the two proteins shows strong structural conservation. The 0.2 Å RMS deviation for the region suggests that any differences in the P-loops themselves must be subtle. Residues Arg^{213} and Val^{217} were implicated in binding the adenosine purine ring atoms in the AlF_4^- structure (Schindelin *et al.*, 1997). Comparison of these residues in the uncomplexed Av2 and Cp2 structures shows strong three-dimensional similarity. However, two neighboring residues, Asn^{215} and Gln^{218} in Av2 are substituted to Pro and Thr, respectively, in Cp2. Any potential role of these residues in mediating nucleotide specificity must await future mutational studies. The availability of crystal structures of both Av2 and Cp2 complexed with nucleotides may provide insights into nucleotide specificity. One particularly

interesting future effort might involve the generation of an Fe protein which contained one Av2 monomer and one Cp2 monomer.

Of the 39 Fe protein sequences previously aligned, only 5 were taken from organisms classified as Archae: two sequences each from *Methanobacterium ivanovii* and *Methanococcus thermolithotrophicus*, and one from *Methanococcus voltae* (Souillard *et al.*, 1986, 1988, 1990). The Archae include organisms which grow optimally under extreme conditions of temperature, pH, and salinity, and may represent some of the earliest life forms on this planet. Since the previous alignment, *nifH*-like sequences from several additional Archae have become available. Among those are *Methanococcus jannaschii* (Bult *et al.*, 1996), *Methanococcus maripaludis*, and *Methanosarcina barkeri* (Sibold *et al.*, 1991). An additional *M. thermolithotrophicus nifH*-like sequence was also identified. The *M. jannaschii* sequence shares 50.4 % identity (and 73.6 % similarity) with that of Av2. An alignment of Av2, Cp2, and the Archae sequences is shown in Figure 5-13. An understanding of nitrogenase activity in Archae is important. For example, thermophilic nitrogen-fixing organisms might be beneficial in developing improved symbiotic relationships with plants for agricultural goods grown at elevated temperatures. This might allow previously unused land to become arable, to meet the increased food needs of the planet's growing population. Also, since the Archae are presumed to be ancient organisms, their nitrogenase proteins may provide a glimpse of a nitrogenase progenitor. Not all Archae can be expected to carry nitrogen fixation genes, of course, as underscored by a BLAST search of the sequence of *Pyrobaculum aerophilum* which failed to identify a *nifH*-like sequence (U.-J. Kim, personal communication)

BLAST searches of Archae sequences have revealed that while *nifH*-like sequences may be present, their *nifD* and *nifK* counterparts are often lacking, leading to speculation as to what role Fe protein might play in an organism which does not encode MoFe protein. Fe protein has been implicated in ATPase activity, electron donation to MoFe protein, and FeMo cofactor assembly and insertion. Perhaps additional roles for Fe protein await

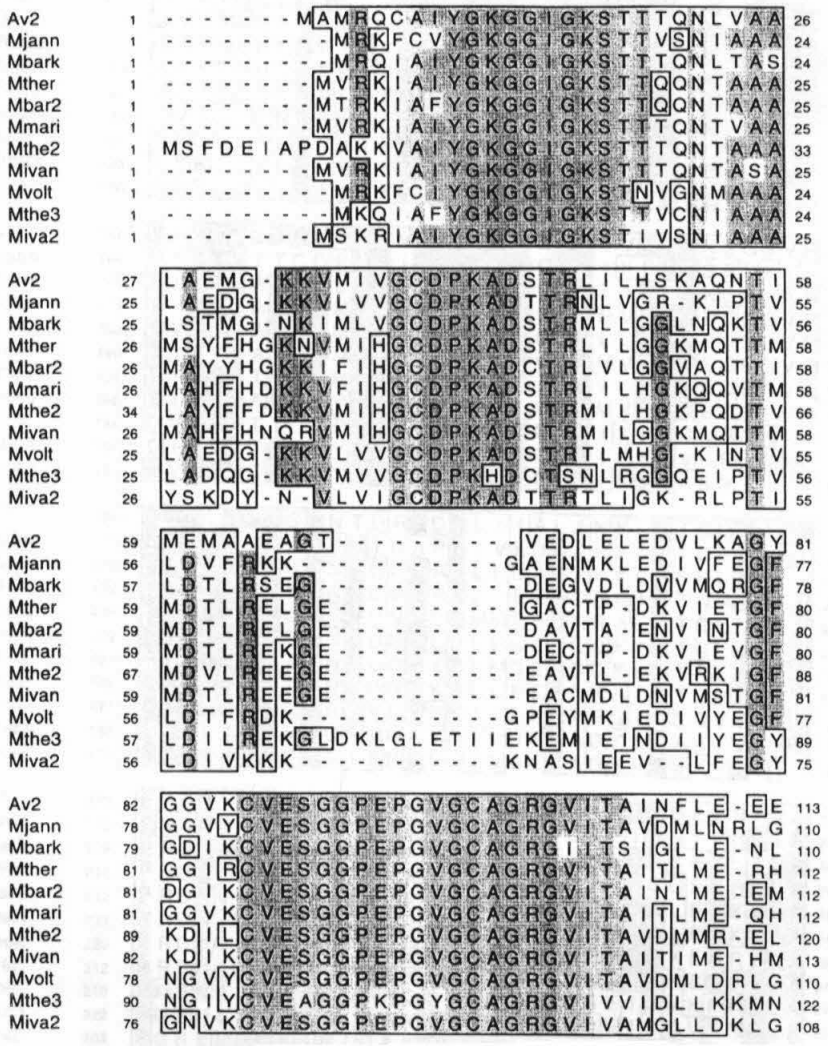
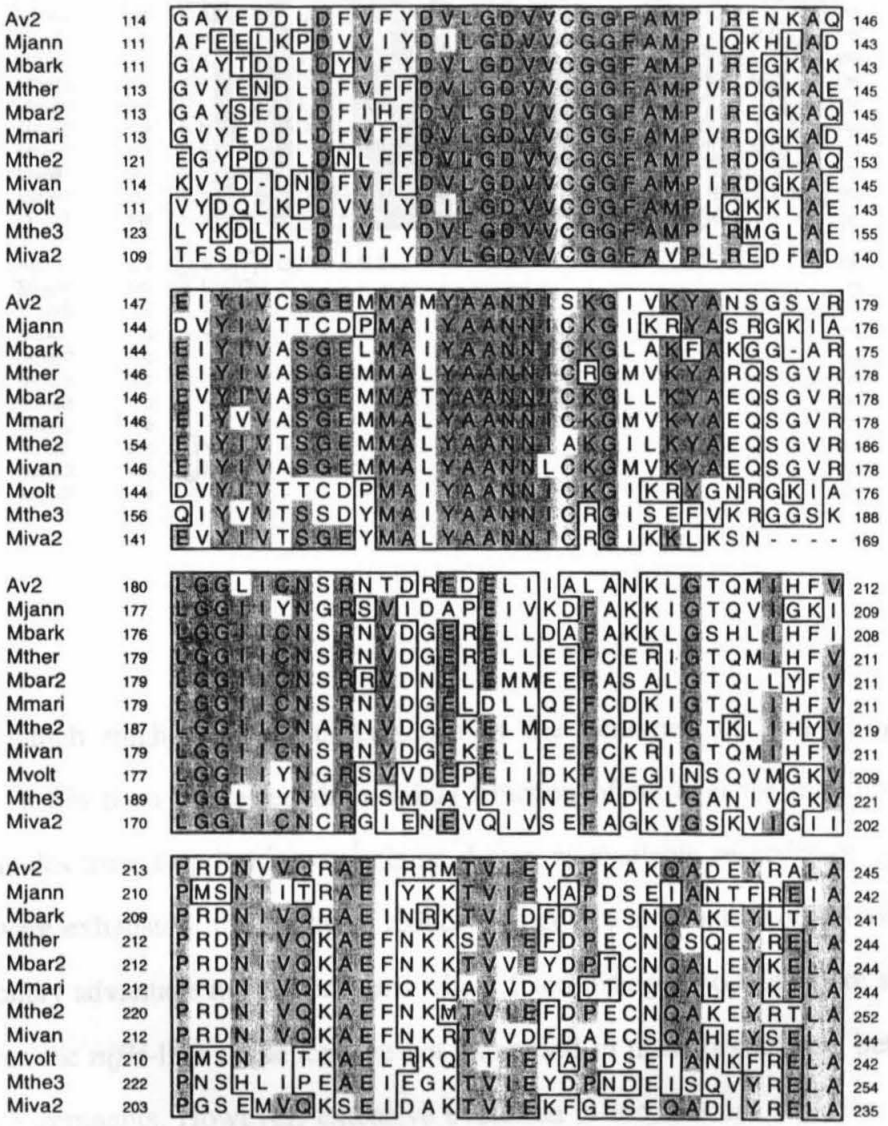
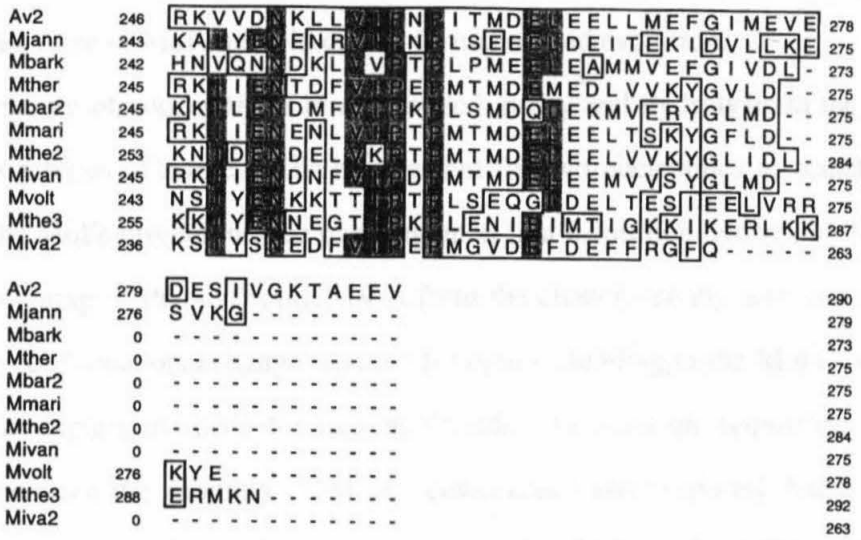


Figure 5-13. Sequence alignment of Av2 and several *nifH*-like sequences from methanobacteria. The sequences were aligned using ClustalW.


Av2 114 GAYEDDLDFVFYDVLGDVVCGGFAMPIRENKAQ 146
Mjann 111 AFEELKPDVVIYDILGDVVCGGFAMPLQKHLAD 143
Mbark 111 GAYTDDLDYVFYDVLGDVVCGGFAMPIREGKAK 143
Mther 113 GVYENDLDFVFFDVLGDVVCGGFAMPVRDGKAE 145
Mbar2 113 GAYSEDLDFIHFDVLGDVVCGGFAMPIREGKAQ 145
Mmari 113 GVYEDDLDFVFFDVLGDVVCGGFAMPVRDGKAD 145
Mthe2 121 EGYPDDLDNLFFDVLGDVVCGGFAMPLRDGLAQ 153
Mivan 114 KVYD-DNDFVFFDVLGDVVCGGFAMPIRDGKAE 145
Mvolt 111 VYDQLKPDVVIYDILGDVVCGGFAMPLQKKLAE 143
Mthe3 123 LYKDLKLDIVLYDVLGDVVCGGFAMPLRMGLAE 155
Miva2 109 TFSDD-IDIIIYDVLGDVVCGGFAVPLREDFAD 140
Av2 147 EIYIVCSGEMMAMYAANNISKGIVKYANSGSVR 179
Mjann 144 DVYIVTTCDPMAIYAANNICKGIKRYASRGKIA 176
Mbark 144 EIYIVASGELMAIYAANNICKGLAKFAKGG-AR 175
Mther 146 EIYIVASGEMMALYAANLICRGMVKYARQSGVR 178
Mbar2 146 EVYIVASGEMMATYAANNICKGLLKYAEQSGVR 178
Mmari 146 EIYVVASGEMMALYAANNICKGMVKYAEQSGVR 178
Mthe2 154 EIYIVTSGEMMALYAANNIAKGILKYAEQSGVR 186
Mivan 146 EIYIVASGEMMALYAANNLCKGMVKYAEQSGVR 178
Mvolt 144 DVYIVTTCDPMAIYAANNICKGIKRYGNRGKIA 176
Mthe3 156 QIYVVTSSDYMAIYAANNICRGISEFVKRGGSK 188
Miva2 141 EVYIVTSGEYMALYAANNICRGIKKLKSN---- 169
Av2 180 LGGLICNSRNTDREDELIIALANKLGTQMIHFV 212
Mjann 177 LGGIIYNGRSVIDAPEIVKDFAKKIGTQVIGKI 209
Mbark 176 LGGIICNSRNVDGERELLDAFAKKLGSHLIHFI 208
Mther 179 LGGIICNSRNVDGERLLEEFCERIGTQMIHFV 211
Mbar2 179 LGGIICNSRRVDNELEMMEERASALGTQLLYFV 211
Mmari 179 LGGIICNSRNVDGELDLLQEFCDKIGTQLIHFV 211
Mthe2 187 LGGIICNARNVDGEKELMDEFCDKLGTKLIHYV 219
Mivan 179 LGGIICNSRNVDGEKELLEEFCKRIGTQMIHFV 211
Mvolt 177 LGGIIYNGRSVVDEPEIIDKFVEGINSQVMGKV 209
Mthe3 189 LGGLIYNVRGSMDAYDIINEFADKLGANIVGKV 221
Miva2 170 LGGIICNCRGIENEVQIVSEFAGKVGSKVIGII 202
Av2 213 PRDNVVQRAEIRRMTVIEYDPKAKQADEYRALA 245
Mjann 210 PMSNIITRAEIYKKTVIEYAPDSEIANTFREIA 242
Mbark 209 PRDNIVQRAEINRKTVIDFDPESNQAKEYLTLA 241
Mther 212 PRDNIVQKAEFNKKSVIEFDPECNQSQEYRELA 244
Mbar2 212 PRDNIVQKAEFNKKTVVEYDPTCNQALEYKELA 244
Mmari 212 PRDNIVQKAEFQKKAVVDYDDTCNQALEYKELA 244
Mthe2 220 PRDNIVQKAEFNKMTVIEFDPECNQAKEYRTLA 252
Mivan 212 PRDNIVQKAEFNKRTVVDFDAECSQAHEYSELA 244
Mvolt 210 PMSNIITKAELRKQTTIEYAPDSEIANKFRELA 242
Mthe3 222 PNSHLIPEAEIEGKTVIEYDPNDEISQVYRELA 254
Miva2 203 PGSEMVQKSEIDAKTVIEKFGESEQADLYRELA 235



discovery which might explain its presence in the genomes of non-nitrogen fixing organisms. It has been suggested the original function of nitrogenase may have been to reduce cyanides toxic to primitive organisms. Later, as available reserves of cyanide and ammonia were exhausted, the ability to reduce nitrogen to ammonia would have provided an evolutionary advantage in environments where fixed nitrogen sources were scarce. It is possible that the *nifH*-like sequences are not expressed in these organisms, but rather are evolutionary remnants. However, extensive evolution is inconsistent with the hypothesis that Archae are truly ancient organisms.

The dimeric nature of Fe protein provides another evolutionary puzzle. Many P-loop nucleotide-binding proteins are functional monomers, which suggests that Fe protein also evolved from a monomeric species. However, most of these proteins do not contain a metallocenter. Perhaps the monomer contained a ferredoxin-like Fe_2S_2 cluster, but found additional stability (and a lower reduction potential) in dimerization to form the cubane Fe_4S_4 cluster. Alternatively, the presence of the Fe:S cluster may have evolved after

dimerization. Understanding the nature of Fe protein evolution may provide insight into the more complex issue of MoFe protein cluster formation and evolution.

The nature of cooperativity in nucleotide binding and hydrolysis in the Fe protein also remains a focus of interest. The requirement for two nucleotides to transfer a single electron to the MoFe protein may be puzzling upon first inspection. However, it is possible that the positioning of the nucleotide, 19 Å from the cluster, merely acts as a fulcrum for effecting the conformational change required for optimal binding to the MoFe protein. This might help to explain the need for two nucleotides, to maintain symmetric movements within and between the subunits. Watt and coworkers have proposed that a $2e^-$ transfer from the Fe protein to the MoFe protein is possible (Watt *et al.*, 1994), which would almost certainly require the presence of two nucleotides. Further investigation will be necessary to elucidate the details of cooperativity in nitrogenase.

The emerging picture of nitrogenase indicates that the Fe protein is a molecule of multiple functions and, indeed, multiple structures. Fe protein is involved not only in nucleotide hydrolysis and electron transfer, but also in FeMo cofactor biosynthesis and assembly. Additional functions, and perhaps additional conformational changes, may await discovery, as details emerge about the interactions between Fe protein and its electron donor (flavodoxin or ferredoxin), or between Fe protein and the NifEN scaffold protein. The available structures of Fe protein, alone and in complex with the MoFe protein, reveal at least two dramatically different Fe protein conformations. Comparisons between the $(AlF_4^-\bullet Av2)_2Av1$ and $(L127\Delta)_2Av1$ nitrogenase complex structures show discernible differences even within the complexed Fe proteins. These variations may account for their functional differences: the inactive $(AlF_4^-\bullet Av2)_2Av1$ complex is a transition state analog (Renner and Howard, 1996; Duyvis *et al.*, 1996), whereas L127Δ has been shown capable of transferring a single electron to the MoFe protein P-clusters in the absence of nucleotide (Lanzilotta and Seefeldt, 1996). The tight-binding nature of the L127Δ complex prevents its dissociation and, hence, substrate reduction. Similarly, the three available uncomplexed

Fe protein structures, Cp2 at 2.0 Å resolution, $P2_1$ Av2 at 2.9 Å resolution, and $P2_12_12_1$ Av2 at 2.2 Å resolution, are not identical. For instance, changes in the protein structure caused by crystal packing interactions in the 2.2 Å resolution Av2 structure reveal a plasticity in the MoFe binding surface which may be functional, perhaps participating in the initial steps of complex formation. Taken together, the various structures suggest that a continuum of iron protein structure exists, reinforcing our perception of nitrogenase as a dynamic enzyme.

Integration of biochemical, genetic, spectroscopic, synthetic, and structural information about nitrogenase has vastly enhanced our understanding of the general processes in biological nitrogen fixation. This multidisciplinary approach perhaps reflects the very symbiosis in which many nitrogen-fixing organisms live. The availability of a structural framework allows for informed interpretation of the results of genetic or spectroscopic results. In turn, easily manipulated nitrogenase genetic systems allow the functions of individual protein residues to be probed. Taken together, the range of research approaches forges a molecular picture of biological nitrogen fixation. Application of a fully detailed understanding of nitrogenase brings the possibility of enhanced agricultural productivity, reduced reliance on commercially produced fertilizers, and a more balanced ecosystem. However, the benefits of nitrogenase knowledge extend beyond potential applications to agriculture and industry. Similarities between nitrogenase and other signal transduction proteins such as p21 ras, transducin and F1 ATPase suggest a general mechanism in molecular gating pathways, impacting such diverse fields as regulation and DNA recombination. Details of the nitrogenase system may be relevant to studies of other electron transfer systems, particularly to those with novel metallocenters. Finally, nitrogenase knowledge may be extended to general areas of protein-protein interaction and to other large, dynamic macromolecular assemblies.

REFERENCES

Abrahams, JP, Leslie, AGW (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr.* D52: 30-54.

Adman, E, Watenpaugh, KD, Jensen, LH (1975). NH···S bonds in *Peptococcus aerogenes* ferredoxin, *Clostridium pasteurianum* rubredoxin, and *Chromatium* high potential iron protein. *PNAS USA* 72: 4854-4858.

Allen, RM, Homer, MJ, Chatterjee, R, Luden, PW, Roberts, GP, Shah, VK (1993). Dinitrogenase reductase- and MgATP-dependent maturation of apodinitrogenase from *Azotobacter vinelandii. J Biological Chemistry* 268: 23670-23674.

Austin, S, Buck, M, Cannon, W, Eydmann, T, Dixon, R (1994). Purification and *in vitro* activities of the native nitrogen fixation control proteins NifA and NifL. *J Bacteriology* 176: 3460-3465.

Bacon, D and Anderson, WF (1988). A fast algorithm for rendering space-filling molecular pictures. *J Molecular Graphics* 6:219-220.

Backes, G, Mino, Y, Loehr, TM, Meyer, TE, Cusanovich, MA, Sweeney, WV, Adman, ET, Sanders-Loehr, J (1991). The environment of Fe_4S_4 clusters in ferredoxins and high-potential iron protein: New information fromx-ray crystallography and resonance Raman spectroscopy. *JACS* 113: 2055-2064.

Bailey, S. (1994). The CCP4 Suite - Programs for protein crystallography. *Acta Crystallogr.* D50: 760-763.

Bishop, PE, Jarlenski, DML, Hetherington, DR (1980). Evidence for an alternative nitrogen fixaton system in *Azotobacter vinelandii. PNAS USA* 77:7342-7346.

Blanco, G, Drummond, M, Woodley, P, Kennedy, C (1993). Sequence and molecular analysis of the nifL gene of *Azotobacter vinelandii. Mol Microbiology* 9: 869-879.

Bolin JT, Campobasso N, Muchmore SW, Minor W, Morgan TV, Mortenson LE (1993a). Structure of the nitrogenase MoFe protein: spatial distribution of the intrinsic metal atoms determined by x-ray anomalous scattering. See Palacios *et al.,* 1993, pp. 89-94.

Bolin JT, Campobasso N, Muchmore SW, Morgan TV, Mortenson (1993b). Structure and environment of the metal clusters in the nitrogenase molybdenum-iron protein from *Clostridium pasteurianum.* See Stiefel *et al.,* 1993, pp. 186-95.

Brigle KE, Newton WE, Dean DR (1985). Complete nucleotide sequence of the *Azotobacter vinelandii* nitrogenase structural gene cluster. *Gene* 37:37-44.

Brigle KE, Setterquist RA, Dean DR, Cantwell JS, Weiss MC, Newton WE (1987a). Site–directed mutagenesis of the nitrogenase MoFe protein of *Azotobacter vinelandii. PNAS USA* 84:7066-69.

Brigle KE, Weiss MC, Newton WE, Dean DR (1987b). Products of the iron-molybdenum cofactor-specific biosynthetic genes, *nifE* and *nifN*, are structurally homologous to the products of the nitrogenase molybdenum-iron protein genes, *nifD* and *nifK*. *J. Bacteriology* 169:1547-53.

Brock, TD (1979). Biology of Microorganisms. Prentice-Hall, Englewood Hills, New Jersey.

Brünger, AT (1991). X-PLOR version 3.1 - A system for x-ray crystallography. Yale University Press: New Haven.

Bulen WA, LeComte, JR (1966). The nitrogenase system from Azotobacter: two enzyme requirements for N2 reduction, ATP dependent H2 evolution anf ATP hydrolysis. *PNAS USA* 56:979-986.

Bult, CJ and many others (1996). Complete genome sequence of the methanogenic archeon *Methanococcus jannaschii. Science* 273: 1058-1073.

Burgess, BK and Lowe, DJ (1996). Mechanism of molecular nitrogenase. *Chem Reviews* 96:2983-3011.

Burgess BK, Jacobs DB, Stiefel EI (1980a). Large-scale purification of high activity *Azotobacter vinelandii* nitrogenase. *Biochim. et Biophys. Acta* 614:196-209.

Burns, RC, Hardy, RWF (1975). Nitrogen fixation in bacteria and higher plants. New York: Springer.

Burris RH (1991). Nitrogenases. *J. Biol. Chem.* 266:9339-42.

Burris, RH (1971). Fixation by free-living micro-organisms: enzymology. In The chemistry and biochemistry of nitrogen fixation. ed. J R Postgate, pp 105-160. London and New York: Plenum Press.

Carter, CW, Jr (1977). New sterochemical analogies between iron-sulfur electorn transport proteins. *J Biological Chemistry* 252: 7802-7811.

Chan MK, Kim J, Rees DC (1993). The nitrogenase FeMo-cofactor and P-cluster pair: 2.2 resolution structures. *Science* 260:792-94.

Chapman, MS (1995). Restrained real-space macromolecular atomic refinement using a new resolution-dependent electron-density function. *Acta Crystallogr.* A51: 69-80.

Chase, MW, Jr, Davies, CA, Downe, JR, Jr, Frurip, DJ, McDonald, RA, Syverue, AN (1985). *Phys. Chem. Ref. Data* 14, Suppl. No. 1.

Chen, L, Gavini, N, Tsuruta, H, Eliezer, D, Burgess, BK, Doniak, S, Hodgson, KO (1994). MgATP-induced conformational changes in the iron protein from *Azotobacter vinelandii,* as studied by small-angle x-ray scattering. *J Biological Chemistry* 269: 3290-3294.

Chisnell, JR, Premakumar, R, Bishop, PE (1988). Purification of a second alternative nitrogenase from a *nifHDK* deletion strain of *Azotobacter vinelandii. J Bacteriology* 170: 27-33.

Cordewener, J, Haaker, H, Van Ewijk, P, Veeger, C (1985). Properties of the MgATP and MgADP binding-sites on the Fe protein of nitrogenase from *Azotobacter vinelandii. Eur J Biochem* 148:499-508.

Dao-pin, S, Nicholson, H, Baase, WA, Zhang, X-J, Wozniak, JA, Matthews, BW (1991). Structural and genetic analysis of electrostatic and other interactions in bacteriophage T4 lysozyme. *In* Protein Conformation, Chadwick, DJ, Widdows, K (Eds.), Chichester, England. John Wiley and Sons, 52-62.

Dean DR, Bolin JT, Zheng L (1993). Nitrogenase metalloclusters: Structures, organization, and synthesis. *J. Bacteriol.* 175:6737-44.

Dean DR, Jacobson MR (1992). Biochemical genetics of nitrogenase. See Stacey *et al.,* 1992, pp. 763-834.

Dean, DR, Bolin, JT, Zheng, L (1993). Nitrogenase metalloclusters: Structures, organization, and synthesis. *J Bacteriology* 1756:6737-6744.

Deits TL, Howard JB (1990). Effect of salts on *Azotobacter vinelandii* nitrogenase activities. *J. Biol. Chem.* 265:3859–67.

Dilworth MJ (1966). Acetylene reduction by nitrogen-fixing preparations from *Clostridium pasteurianum. Biochim. Biophys. Acta* 127:285-94.

Dutton, PL (1978). Redox potentiometry: Determination of midpoint potentials of oxidation-reduction components of biological electron-transfer systems. *Methods Enzymol* 54:411-435.

Duyvis, MG, Wassink, H, Haaker, H (1996). Formation and characterization of a transition-state complex of *Azotobacter vinelandii* nitrogenase. FEBS Lett 380: 233-236.

Eady, RR, Robson, RL, Richardson, TH, Miller, RW, Hawkins, M (1987). The vanadium nitrogenase of *Azotobacter chroococcum* - Purification and properties of the Vfe protein. Biochem J 244: 197-207.

Eady RR, Smith BE, Cook KA, Postgate JR (1972). Nitrogenase of *Klebsiella pneumoniae*: Purification and properties of the component proteins. *Biochem. J.* 128:655-75.

Emerich DW, Burris RH (1976). Interactions of heterologous nitrogenase components that generate catalytically inactive complexes. *Proc. Natl. Acad. Sci. USA* 73:4369-73.

Emerich DW, Ljones T, Burris RH (1978). Nitrogenase: Properties of the catalytically inactive complex between *Azotobacter vinelandii* MoFe protein and the *Clostridium pasteurianum* Fe protein. *Biochim. Biophys. Acta* 527:359-69.

Engh, RA, Huber, R (1991). Accurate bond and angle parameters for x-ray protein structure refinement. Acta Crystallogr. A47: 392-400.

Ennor AH (1957). Determination and preparation of N-phosphates of biological origin. *Methods in Enzymol.* 3:850-56.

Fisher K, Lowe DJ, Thorneley RNF (1991). *Klebsiella pneumoniae* nitrogenase: The pre-steady-state kinetics of MoFe-protein reduction and hydrogen evolution under conditions of limiting electron flux show that the rates of association with the Fe-protein and electron transfer are independent of the oxidation level of the MoFe-protein. *Biochem. J.* 279:81-85.

Gavini, N Burgess, BK (1992). FeMo cofactor synthesis by a *nifH* mutant with altered MgATP reactivity. *J Biological Chemistry* 267:21179-21186.

Georgiadis MM, Komiya H, Chakrabarti P, Woo D, Kornuc JJ, Rees DC (1992). Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii. Science* 257:1653-59.

Georgiadis, MM (1990). Crystal structure of the nitrogenase iron protein from *Azotobacter vinelandii.* Ph.D. Dissertation, UCLA.

Hageman RV, Burris RH (1978a). Kinetic studies on electron transfer and interaction between nitrogenase components from *Azotobacter vinelandii. Biochemistry* 17:4117-24.

Hageman RV, Burris RH (1978b). Nitrogenase and nitrogenase reductase associate and dissociate with each catalytic cycle. *PNAS USA* 75:2699-2702.

Hageman RV, Burris RH (1979). Changes in the EPR signal of dinitrogenase from *Azotobacter vinelandii* during the lag period before hydrogen evolution begins. *J. Biol. Chem.* 254:11189-92.

Hageman RV, Burris RH (1980). Electron allocation to alternative substrates of *Azotobacter vinelandii* nitrogenase is controlled by the electron flux through dinitrogenase. *Biochim. Biophys. Acta* 591:63-75.

Hales, BJ, Case, EE, Morningstar, JE, Dzeda, MF, Mauterer, LA (1986). Isolation of a new vanadium-containing nitrgoenase from *Azotobacter vinelandii. Biochemistry* 25: 7251-7255.

Hallenbeck PC (1983). Nitrogenase reduction by electron carriers: Influence of redox potential on activity and the ATP/$2e^-$ ratio. *Arch. Biochem. Biophys.* 220:657-60.

Hardy, RWF (1994). Biological Nitrogen Fixation: Research Challenges - A review of research grants funded by the US Agency for International Development. Washington, D.C.: National Academy Press.

Hausinger RP, Howard JB (1982). The amino acid sequences of the nitrogenase-iron protein from *Azotobacter vinelandii. J. Biol. Chem.* 257:2483-90.

Hausinger RP, Howard JB (1983). Thiol reactivity of the nitrogenase Fe-protein from *Azotobacter vinelandii. J. Biol. Chem.* 258:13486-92.

Hawkes TR, McLean PA, Smith BE (1984). Nitrogenase from *nifV* mutants of *Klebsiella pneumoniae* contains an altered form of the iron-molybdenum cofactor. *Biochem. J.* 217:317-21.

Hill, S (1976). The apparent ATP requirement for nitrogen fixation in growing *Klebsiella pneumoniae. J. Gen. Microbiol.* 95:297-312.

Hope, H (1988). Cryocrystallography of biological macromolecules: A generally applicable method. *Acta Crystallogr.* B44, 22-26.

Hope, H (1990). Crystallography of biological macromolecules at ultra-low temperature. *Ann Rev. Biophys. Biophys. Chem.* 19: 107-126.

Hope, H, Frolow, R, von Böhlen, K, Makowski, L, Kratky, C, Halfon, Y, Dantz, H, Webster, P, Bartels, KS, Wittmann, HG, Yonath, A (1989). Cryocrystallography of ribosomal particles. *Acta Crystallogr.* B45: 190-199.

Howard, AJ, Gilliland, GL, Finzel, BC, Poulos, TL, Ohlendorf, DH, Salemme, FR (1987). The use of an imaging proportional counter in macromolecular crystallography. *J. Applied Crystallography* 20: 383-387.

Howard JB (1993). Protein component complex formation and adenosine triphosphate hydrolysis in nitrogenase. See Stiefel *et al.,* 1993, pp. 271-89.

Howard JB, Davis R, Moldenhauer B, Cash VL, Dean DR (1989). Fe-S cluster ligands are the only cysteines required for nitrogenase Fe-protein activities. *J. Biol. Chem.* 264:11270-74.

Howard JB, Rees DC (1994). Nitrogenase: A nucleotide–dependent molecular switch. *Annu. Rev. Biochem.* 63:235–64.

Howard, JB and Rees, DC (1996). Structural basis for biological nitrogen fixation. *Chem Reviews* 96:2965-2982.

Huber, R (1985). Experience with the application of Patterson search techniques. *In* Molecular Replacement, Proceedings of the Daresbury Study Weekend, Science and Engineering Council, The Librarian, Daresbury Laboratory, Daresbury, United Kingdom, pp 58-61.

Hutchinson, EG, Thornton, JM (1996). PROMOTIF - A program to identify and analyze structural motifs in proteins. *Protein Science* 5: 212-220.

Imam S, Eady RR (1980). Nitrogenase of *Klebsiella pneumoniae*: Reductant-independent ATP hydrolysis and the effect of pH on the efficiency of coupling of ATP hydrolysis to substrate reduction. *FEBS Lett.* 110:35-38.

Imperial, J, Ugalde, RA, Shah, VK, Brill, WL (1984). Role of the *nifQ* gene product in the incorporation of molybdenum into nitrogenase in *Klebsiella pneumoniae. J Bacteriology* 158: 187-194.

Jacobson MR, Brigle KE, Bennett LT, Setterquist RA, Wilson MS, Cash, VL, Beynon, J, Newton, WE, Dean, DR (1989a). Physical and genetic map of the major *nif* gene cluster from *Azotobacter vinelandii. J. Bacteriol.* 171:1017-27.

Jacobson MR, Cash VL, Weiss MC, Laird NF, Newton WE, Dean DR (1989b). Biochemical and genetic analysis of the *nifUSVWZM* cluster from *Azotobacter vinelandii*. *Molec. Gen. Genet.* 219:49-57.

Jacobson MR, Cantwell JS, Dean DR (1990). A hybrid *Azotobacter vinelandii-Clostridium pasteurianum* nitrogenase iron protein that has *in-vivo* and *in-vitro* catalytic activity. *J. Biol. Chem.* 265:19429-33.

Jancarik, J, Kim, S-H (1991). Sparse Matrix sampling - A screening method for crystallization of proteins. *J Applied Crystallography* 24: 409-411.

Jennings, JR, Ed. Catalytic Ammonia Synthesis: Fundamentals and Practice. New York: Plenum Press. 1991.

Johnson MK (1988). Variable-temperature magnetic circular dichroism studies of metalloproteins. In *Metal clusters in proteins*, ed. L Que, pp. 326-42. Washington DC: ACS Books Symposium Series.

Jones, TA, Zou, JY, Cowan, SW, Kjeldgaard, M (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* A47: 110-119.

Jones, TA, Thirup, S (1986). Using known substructures in model building and crystallography. *EMBO J* 5: 819-822.

Jurnack F (1985). Structure of the GDP domain of EF-Tu and localization of the amino acids homologous to *ras* oncogene proteins. *Science* 230:32-36.

Kabsch, W, Sander, C (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22, 893-899.

Kelly, M, Postgate, JR, Richards, RL (1967). Reduction of cyanide and isocyanide by nitrogenase of *Azotobacter chroococcum. Biochem J* 102: 1c-3c.

Kim J, Rees DC (1992a). Structural models for the metal centers in the nitrogenase molybdenum-iron protein. *Science* 257:1677-82.

Kim J, Rees DC (1992b). Crystallographic structure and functional implications of the nitrogenase molybdenum-iron protein from *Azotobacter vinelandii. Nature* 360:553-60

Kim J, Rees DC (1994). Nitrogenase and biological nitrogen fixation. *Biochemistry* 33:389-97.

Kim J, Woo D, Rees DC (1993). X-ray crystal structure of the nitrogenase molybdenum-iron protein from *Clostridium pasteurianum* at 3.0 Å resolution. *Biochemistry* 32:7104-15.

Kjeldgaard M, Nyborg J (1992). Refined structure of elongation factor EF-Tu from *Escherichia coli. J. Mol. Biol.* 223:721-42.

Koonin EV (1993). A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. *J. Mol. Biol.* 229:1165-74.

Kraulis PJ. (1991). Molscript: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystall.* 24:946-50.

Laemmli UK (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. *Nature* 227:680-85.

Langen, R, Jensen, GM, Jacob, U, Stephens, PJ, Warshel, A (1992). Protein control of iron-sulfur cluster redox potentials. *J Biological Chemistry* 267: 25625-25627.

Lanzilotta, WN, Fisher, K, Seefeldt, LC (1997). Evidence for electron transfer-dependent formation of a nitrogenase iron protein molybdenum iron protein tight complex - the role of Aspartate-39. *J Biological Chemistry* 272: 4157-4165.

Lanzilotta, WN, Seefeldt, LC (1996) Electron transfer from the nitrogenase iron protein to the [8Fe-(7/8)S] clusters of the molybdenum-iron protein. *Biochemistry* 35:16770-16776.

Lanzilotta, WN, Fisher, K, Seefeldt, LC (1996). Evidence for electron transfer from the nitrogenase iron protein to the molybdenum-iron protein without MgATP hydrolysis - Characterization of a tight protein-protein complex. *Biochemistry* 35:7188-7196.

Lanzilotta, WN, Holz, RC, Seefeldt, LC (1995a). Proton NMR investigation of the [4Fe-4S](1+) cluster environment of nitrogenase iron protein from *Azotobacter*

vinelandii - Defining nucleotide-induced conformational changes. *Biochemistry* 34:15646-15653.

Laskowski, RA, Macarthur, MW, Moss, DS, Thornton, JM. (1993). PROCHECK - A program to check the stereochemical quality of protein structures. *J Applied Crystallogr.* 26: 283-291.

Lee, H-S, Berger, DK, Kustu, S (1993). Activity of purified NIFA, a transcriptional activator of nitrogen fixing genes. *PNAS USA* 90: 2266-2270.

Levi, BG, Hafemeister, D, Scribner, R, Eds. (1992). Global Warming: Physics and Facts. NY: American Institute of Physics. Washington, DC, 1991.

Lindahl PA, Day EP, Kent TA, Orme-Johnson WH, Münck E (1985). Mössbauer, EPR, and magnetization studies of the *Azotobacter vinelandii* Fe protein. *J. Biol. Chem.* 260:11160-73.

Lindahl, PA, Gorelick, NJ, Münck, E, Orme-Jonson, WH (1987a). Electron-paramagnetic-resonance and Mössbauer studies of nucleotide-bound nitrogenase iron protein from *Azotobacter vinelandii. J Biological Chemistry* 262:14945-14953.

Lindahl, PA, Teo, BK, Orme-Johnson, WH (1987b). EXAFS studies of the nitrogenase iron protein from *Azotobacter vinelandii. Inorganic Chemistry* 26:3912-3916.

Ljones T, Burris RH (1972). ATP hydrolysis and electron transfer in the nitrogenase reaction with different combinations of iron protein and molybdenum-iron protein. *Biochim. Biophys. Acta* 275:93-101.

Ljones, T and Burris, RH (1978). Nitrogenase: The reaction between the Fe protein and bathophenanthrolinedisulfonate as a probe for interactions with MgATP. *Biochemistry* 17:1866-1872.

Lowe DJ, Fisher K, Thorneley RNF (1990). *Klebsiella pneumoniae* nitrogenase: Mechanism of acetylene reduction and its inhibition by carbon monoxide. *Biochem. J.* 272:621-25.

Lowe DJ, Fisher K, Thorneley RNF (1993). *Klebsiella pneumoniae* nitrogenase: Pre-steady state absorbance changes show that redox changes occur in the MoFe protein that depend on substrate and component protein ratio; a role for P-centres in reducing nitrogen. *Biochem. J.* 292:93-98.

Lowe DJ, Thorneley, RNF (1983). Nitrogenase of Klebsiella pneumoniae - Kinetics of the dissociation of oxidized iron protein from molybdenum iron protein - identification of the rate-limiting step for substrate reduction. *Biochem J* 215:393-403.

Lowe, DJ, Thorneley, RGNF (1984). The mechanism of Klebsiella pneumoniae nitrogenase action - simulation of the dependences of H_2 evolution rate on component-protein concentration and ratio and sodium dithionite concentration. *Biochem J.* 224: 895-901.

Ludden PW, Shah VK, Roberts GP, Homer M, Allen R, *et al.* (1993). Biosynthesis of the iron–molybdenum cofactor of nitrogenase. See Stiefel *et al.*, 1993, pp. 196–215.

Ludden, PW, Lehman, L, Roberts, GP (1989a). Reversible ADP-ribosylation of dinitrogenase reductasein a *nifD-* mutant of *Rhodospirillum rubrum. J Bacteriology* 171: 5210-5211.

Ludden, PW, Roberts, GP (1989b). Regulation of nitrogenase activity by reversible ADP-ribosylation. *Curr Top Cell Regulation* 30: 23-55.

Masepohl B, Angermuller S, Hennecke S, Hubner P, Moreno–Vivian C, Klipp W. (1993). Nucleotide sequence and genetic analysis of the *Rhodobacter capsulatus* ORF6–*nifU1SVW* gene region: possible role of *NifW* in homocitrate processing. *Mol. Gen. Genet.* 238:369–82.

Merritt, EA, Murphy, MEP (1994). RASTER-3D version 2.0 - A program for photorealistic molecular graphics. *Acta Crystallographica* D50:869-873.

Meyer, J, Gaillard, J, Moulis, J-M (1988).Hydrogen-1 nuclear magnetic resonance of the nitrogenase iron protein (Cp2) fromClostridium pasteurianum. Biochemistry 27: 6150-6156.

Münck E, Rhodes H, Orme-Johnson WH, Davis LC, Brill WJ, Shah VK (1975). Nitrogenase. VIII. Mössbauer and EPR spectroscopy. The MoFe protein component from *Azotobacter vinelandii* OP. *Biochim. Biophys. Acta* 400:32-53.

Navaza, J (1994). AMORE - An automated package for molecular replacement. *Acta Crystallogr.* A50: 157-163.

Navaza, J (1997). AMORE - An automated molecular replacement program package. *Methods Enzymol.* 276: 581-594.

Otwinosky, Z (1993). *In* Data Collection and Processing, L Sawyer, L Isaacs, S Bailey, Eds., (SERC Daresbury Laboratory, Daresbury, UK). pp 56-62.

Pai EF, Krengel U, Petsko GA, Goody RS, Kabsch W, Wittinghofer A (1990). Refined crystal structure of the triphosphate conformation of H-*ras* p21 at 1.35 resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* 9:2351-59.

Palacios R, Mora J, Newton WE, eds. 1993. *New Horizons in Nitrogen Fixation.* Dordrecht: Kluwer Academic.

Paustian TD, Shah VK, Roberts GP (1989). Purification and characterization of the *nifN* and *nifE* gene products from *Azotobacter vinelandii* mutant UW45. *Proc. Natl. Acad. Sci. USA* 86:6082-86.

Peters JW, Fisher K, Dean DR (1994). Identification of a nitrogenase protein-protein interaction site defined by residues 59 through 67 within the *Azotobacter vinelandii* Fe protein. *J. Biol. Chem.* 269:28076-83.

Peters, JW, Fisher, K, Newton, WE, Dean, DR (1995). Involvement of the P-cluster in intramolecular electron transfer within the nitrogenase MoFe protein. *J Biological Chemistry* 270: 27007-27013.

Peters, JW, Stowell, MHB, Soltis, MS, Finnegan, MG, Johnson, MK, Rees, DC (1997). Redox-dependent structural changes in he nitrogenase P-cluster. *Biochemistry* 36:1181-1187.

Pope MR, Murrell SA, Ludden PW (1985). Covalent modification of the iron protein of nitrogenase from *Rhodospirillum rubrum* by adenosine diphosphoribosylation of a specific arginine residue. *PNAS USA* 82:3173:77.

Postgate, JR. 1982. The Fundamentals of Nitrogen Fixation. New York: Cambridge University.

Rees, DC, Howard, JB (1983). Crystallization of the *Azotobacter vinelandii* nitrogenase iron protein. *J Biological Chemistry* 258: 2733-2734.

Renner, KA, Howard, JB (1996). Aluminum fluoride inhibition of nitrogenase: Stabilization of a nucleotide Fe protein MoFe protein complex. *Biochemistry* 35: 5353-5358.

Roberts, GP and Ludden, PW. 1992. Nitrogen fixation by photosynthetic bacteria. In Biological Nitrogen Fixation, Stacey, G, Burris, RH, Evans, HJ, eds. New York: Chapman and Hall. pp. 135-165.

Robinson AC, Burgess BK, Dean DR (1986). Activity, reconstitution, and accumulation of nitrogenase components in *Azotobacter vinelandii* mutant strains containing defined deletions within the nitrogenase structural gene cluster. *J. Bacteriol.* 166:180–86.

Robinson, AC, Dean, DR, Burgess, BK (1987). Iron-molybdenum cofactor biosynthesis in *Azotobacter vinelandii* requires the iron protein of nitrogenase. *J Biological Chemistry* 262:14327-14332.

Robson RL (1984). Identification of possible adenine nucleotide-binding sites in nitrogenase Fe- and MoFe-proteins by amino acid sequence comparison. *FEBS Lett.* 173:394-98.

Rossman, MG, Blow DM (1962). The detection of subunits within the crystallographic asymmetric unit. *Acta Crystallogr.* 15: 24-31.

Ryle, MJ and Seefeldt, LC (1996a). The [Fe4-S4] cluster domain of the nitrogenase iron protein facilitates conformational changes required for the cooperative binding of two nucleotides. *Biochemistry* 35:15654-15662.

Ryle, MJ, Seefeldt, LC (1996b). Elucidation of a MgATP signal-transduction pathway in the nitrogenase iron protein - Formation of a conformation resembling the MgATP-bound state by protein engineering. *Biochemistry* 35:4766-4775.

Ryle, MJ, Lanzilotta, WN, Seefeldt, LC, Scarrow, RC, Jensen, GM (1996a) Circular-dichroism and x-ray spectroscopies of *Azotobacter vinelandii* nitrogenase iron protein - MgATP and MgADP induced protein conformational changes

affecting the [4Fe-4S] cluster and characterization of a [2Fe-2S] form. *J Biological Chemistry* 271:1551-1557.

Ryle, MJ, Lanzilotta, WN, Seefeldt, LC (1996b). Elucidationg the mechanism of nucleotide-dependent changes in the redox potential of the [4Fe-4s] cluster in nitrogenase iron protein - The role of Phenylalanine 135. *Biochemistry* 35:9424-9434.

Ryle, MJ, Lanzilotta, WN, Mortenson, LE, Watt, GD, Seefeldt, LC (1995). Evidence for a central role of Lysine 15 of *Azotobacter vinelandii* nitrogenase iron protein in nucleotide binding and protein conformational changes. *J Biological Chemistry* 270:13112-13117.

Schindelin, H, Kisker, C, Schlessman, JL, Howard, JB, Rees, DC (1997). Structure of the ADP·AlF4⁻ stabilized nitrogenase complex: Implications for nucleotide dependent signal transduction mechanisms. *Nature,* 387: 370-376.

Schlesinger, WH (1991) Biogeochemistry: An Analysis of Global Change. San Diego: Academic Press.

Schöllhorn R, Burris RH (1967). Acetylene as a competitive inhibitor of N_2 fixation. *PNAS USA* 58:213-16.

Schulz GE (1992). Binding of nucleotides by proteins. *Curr. Opin. Struct. Biol.* 2:61-67.

Seefeldt LC (1994). Docking of nitrogenase iron– and molybdenum–iron proteins for electron transfer and MgATP hydrolysis: The role of arginine 140 and lysine 143 of the *Azotobacter vinelandii* iron protein. *Protein Sci.* 3:2073-81.

Seefeldt LC, Morgan TV, Dean DR, Mortenson LE (1992). Mapping the sites of MgATP and MgADP interaction with the nitrogenase of *Azotobacter vinelandii:* Lysine 15 of the Fe protein plays a major role in MgATP interaction. *J. Biol. Chem.* 267:6680-88.

Seefeldt LC, Mortenson, LE (1993). Increasing nitrogenase catalytic efficiency for MgATP by changing serine 16 of its Fe protein to threonine: Use of Mn^{++} to show interaction of serine 16 with Mg^{++}. *Protein Sci.* 2:93-102.

Shah VK, Brill WJ (1977). Isolation of an iron–molybdenum cofactor from nitrogenase. *PNAS USA*. 74:3249–53.

Shah VK, Davis LC, Brill WJ (1972). Repression and derepression of the iron-molybdenum and iron proteins of nitrogenase in *Azotobacter vinelandii. Biochim. Biophys. Acta* 256:498-511.

Shah, VK, Allen, JR, Spangler, NJ, Ludden, PW (1994). In vitro synthesis of the iron-molybdenum cofactor of nitrogenase: Purification and characterization of NifB cofactor, the product of NIFB protein. *J Biological Chemistry* 269: 1154-1158.

Sibold, L, Henriquet, M, Possot, O, Aubert, J-P (1991). Nucleotide sequence of nifH genes from *Methanobacterium ivanovii* and *Methanosarcina barkeri* 227 and characterization of *glnB*-like genes. *Res Microbiol* 142: 5-12.

Simpson, FB, Burris, RH (1984). A nitrogen pressure of 50 atmospheres does not prevent evolution of hydrogen by nitrogenase. *Science* 224:1095-1097.

Smith BE, Eady RR, Lowe DJ, Gormal C (1988). The vanadium-iron protein of vanadium nitrogenase from *Azotobacter chroococcum* contains an iron-vanadium cofactor. *Biochem. J.* 250:299-302.

Smith BE, Lang G. (1974). Mössbauer spectroscopy of the nitrogenase proteins from *Klebsiella pneumoniae*. *Biochem. J.* 137:169-80.

Souillard, N, Sibold, L (1986). Primary structure and expression of a gene homologous to nifH from the archaeobacterium *Methanococcus voltae*. *Mol Gen Genet* 203: 21-28.

Souillard, N, Sibold, L (1989). Primary structure, functional organization and expression of nitrogenase structural genes of the thermophilic archaebacterium *Methanococcus thermolithotrophicus*. *Mol Microbiol* 3: 541-551.

Souillard, N, Magot, M, Possot, O, Sibold, L (1988). Nucleotide sequence of regions homologous to nifH from the nitrogen-fixing archaebacteria *Methanococcus thermolithotrophicus* and *Methanococcus ivanovii:* Evolutionary implications. *J Mol Evol* 27: 65-76.

Sousa, R and Lafer, EM (1990). The use of glycerol in crystallization of T7 RNA polymerase: Implications for use of cosolvents in crystallizing flexible proteins. *Methods Companion Methods Enzym.* 1, 50-56 (1990).

Spiro TG, ed. 1985. *Molybdenum Enzymes*, New York: John Wiley & Sons.

Stacey G, Burris RH, Evans HJ, eds. 1992. *Biological Nitrogen Fixation,* New York: Chapman and Hall.

Stanier, RY, Ingraham, JL, Wheelis, ML, Painter, PR (1976). The Microbial World. 4th Ed. New York (englewood Cliffs, NJ): Prentice-Hall.

Stephens, PJ, McKenna, CE, Smith, BE, Nguyen, HT, McKenna, MC, Thomson, AJ, Devlin F, Jones, JB (1979). Circular dichroism and magnetic circular dichroism of nitrogenase proteins. *PNAS USA* 76:2585-2589.

Stephens PJ. 1985. The structures of the iron-molybdenum and the iron proteins of the nitrogenase enzyme. See Spiro, 1985, pp. 117-160.

Stewart, WDP, Rowell, P, Apte, SK (1977). Cellular physiology and the ecology of N_2-fixing blue-green algae. In Recent developments in nitrogen fixation. Newton, W, Postgate, JR, Rodriguez-Barrueco, Eds. London: Academic Press. pp. 287-307.

Stiefel EI, Coucouvanis D, Newton WE, eds. 1993. *Molybdenum Enzymes, Cofactors and Model Systems.* Washington DC: Am. Chem. Soc.

Strandberg GW, Wilson PW (1968). Formation of the nitrogen-fixing enzyme system in *Azotobacter vinelandii. Can. J. Microbiol.* 14:25-31.

Story RM, Steitz TA (1992). Structure of the recA protein-ADP complex. *Nature* 355:374-76.

Surerus KK, Hendrich MP, Christie PD, Rottgardt D, Orme-Johnson WH, Münck E (1992). Mössbauer and integer-spin EPR of the oxidized P-clusters of nitrogenase: P^{ox} is a non-kramers system with a nearly degenerate ground doublet. *J. Am. Chem. Soc.* 114:8579-90.

Svennson, BH and Soderland, R (1976). Nitrogen, phosphorus and sulphur - global cycles. Scope Report 7, *Ecol. Bull.* 22:23-73.

Tal, S, Chun TW, Gavini, N, Burgess, BK (1991).The ΔNifB (or ΔNifE) FeMo cofactor-deficient MoFe protein is different from the ΔNifH protein. *J Biological Chemistry* 266: 10654-10657.

Thorneley RNF, Deistung J (1988). Electron-transfer studies involving flavodoxin and a natural redox partner, the iron protein of nitrogenase: Conformational constraints on protein-protein interactions and the kinetics of electron transfer with the protein complex. *Biochem. J.* 253:587-95.

Thorneley RNF, Lowe DJ (1983). Nitrogenase of *Klebsiella pneumoniae*: Kinetics of the dissociation of oxidized iron protein from molybdenum-iron protein: Identification of the rate-limiting step for substrate reduction. *Biochem. J.* 215:393-403.

Thorneley RNF, Lowe DJ (1984a). The mechanism of *Klebsiella pneumoniae* nitrogenase action: Pre-steady-state kinetics of an enzyme-bound intermediate in N_2 reduction and NH_3 formation. *Biochem. J.* 224:887-94.

Thorneley RNF, Lowe DJ (1984b). The mechanism of *Klebsiella pneumoniae* nitrogenase action: Simulation of the dependence of H_2-evolution rate on component-protein concentration and ratio and sodium dithionite concentration. *Biochem. J.* 224:903-09.

Thorneley RNF, Lowe DJ (1985). Kinetics and mechanism of the nitrogenase enzyme system.See Spiro, 1985, pp. 222-284.

Tong L, de Vos AM, Milburn MV, Kim S-H (1991). Crystal structures at 2.2 Å resolution of the catalytic domains of normal *ras* protein and an oncogenic mutant complexed with GDP. *J. Mol. Biol.* 217:503-16.

Topham, SA. 1985. *In* Catalysis: Science and Technology, Anderson, JR, Boudart, M, Eds. Berlin: Springer-Verlag. p 1.

Tronrud, DE (1992). Conjugate-directed minimization - an improved method for the refinement of macromolecules. *Acta Crystallogr.* A48: 912-916.

Van de Broek, A, Michiels, J, de Faria, SM, Milcamps, A, Vanderleyden, J (1992). Transcription of the Azospirillum brasilience nifH gene is positively regulated by NifA and NtrA and is negatively controlled by the cellular nitrogen status. *Mol Gen Genet.* 232: 279-283.

Walker GA, Mortenson LE (1974). Effect of magnesium adenosine 5'-triphosphate on the accessibility of the iron of clostridial azoferredoxin, a component of nitrogenase. *Biochemistry* 13:2382-88.

Walker, JE, Saraste, M, Runswick, MJ, Gay, NJ (1982). Distantly related sequences in the alpha-subunits and beta-subunits of ATP synthase, myosin, kinases, and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* 8:945-951.

Watt GD, Wang Z-C, Knotts RR (1986). Redox reactions of and nucleotide binding to the iron protein of *Azotobacter vinelandii*. *Biochemistry* 25:8156-62.

Watt GD, Reddy, KRN (1994). Formation of an all-ferrous Fe_4S_4 cluster in iron protein component of *Azotobacter vinelandii* nitrogenase. *J. Inorganic Biochemistry* 53:281-294.

Weston, MF, Kotake, S, Davis, L (1983). Interaction of nitrogenase with nucleotide analogs of ATP and ADP and the effect of metal ions on ADP inhibition. *Arch. Biochem Biophys* 225: 809-817.

Willing AH, Georgiadis MM, Rees DC, Howard JB (1989). Cross-linking of Nitrogenase Components. *J. Biol. Chem.* 264:8499-8503

Willing A, Howard JB (1990). Cross-linking site in *Azotobacter vinelandii* complex. *J. Biol. Chem.* 265:6596-99.

Wolle D, Dean DR, Howard JB (1992a). Nucleotide iron-sulfur cluster signal transduction in the nitrogenase iron-protein: The role of Asp^{125}. *Science* 258:992-95.

Wolle D, Kim C-H, Dean DR, Howard JB (1992b). Ionic interactions in the nitrogenase complex: Properties of Fe-protein containing substitutions for Arg-100. *J. Biol. Chem.* 267:3667-73.

Woo, D (1995). Crystal Structure of the Nitrogenase Iron Protein from *Clostridium pasteurianum.* Ph.D. dissertation, UCLA.

Yates, MG (1972). The effect of ATP upon the oxygen sensitivity of nitrogenase from *Azotobacter chroococcum. Eur J Biochem* 29:386-392.

Zheng, L, White, RH, Cash, VL, Jack, RF, Dean, DR (1993). Cysteine desulfurase activity indicates a role for NifS in metallocluster biosynthesis. *PNAS USA* 90: 2754-2758.

Zheng, L, White, RH, Cash, VL, Dean, DR (1994a). Mechanism for the desulfurization of L-cysteine catalyzed by the *nifS* gene product. *Biochemistry* 33: 4714-4720.

Zheng, L, White, RH, Cash, VL, Dean, DR (1994b). Catalytic formation of a nitrogenase iron sulfur cluster. *J Biological Chemistry* 269: 18723-18726.

Zumft, WG, Palmer, G, Mortenson, LE (1973). Electron paramagnetic resonance studies on nitrogenase. II. Interactions of adenosine 5'-triphosphate with azoferredoxin. *Biochim Biophys Acta* 292:413-420.