STRUCTURAL AND FUNCTIONAL STUDIES OF JAMM DOMAIN PROTEINS AND THEIR ROLES IN THE UBIQUITIN SYSTEM

Thesis by

Xavier I. Ambroggio

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena California

2005

(Defended September 22, 2004)

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisors for their support, guidance, and patience. In particular, I am immensely grateful to Doug for allowing me to pursue a project that was outside of the agenda of his laboratory. I am continually amazed by his benevolence towards his students and hope that I can follow his example in my interactions with others. I am grateful to Ray for sharing his enthusiasm for his work. His excitement in our experiments was a great source of motivation. I have often told people that Ray and Doug combined make the perfect advisor.

I would like to thank all of the members of the Rees lab and especially: Kaspar Locher for helping me troubleshoot experiments and teaching me the importance of objectively analyzing the big picture; Andrew Yeh for introducing me to crystallography and local microbreweries; Pavel Strop for making the lab a fun place to work; Oliver Einsle for having a trick to solve any computational problem; Benedikt Schmid for always having the exact thing I needed when the stockroom was closed; Akif Tezcan for talks on the stairs and over coffee and both together; Pam Lum for introducing the lab to dim sum and group activities; Liz Borths and Ivana Hughes for gossip and parenting advice; Jens Kaiser for proving that it is possible to live off of alcohol and cigarettes; Alex Gagnon for replacing Pavel; Leonard Thomas and Welison Floriano for keeping everything going; and Phoebe Ray for so much that it would be an entire thesis if I wrote it out.

Likewise, I would like to thank everyone in the Deshaies lab and especially: Greg Cope and Rati Verma for inviting me to participate in their research of the JAMM domain; Liz Ottesen for being a great rotation student and introducing me to anime; Robert Oania for great screensavers; Johannes Graumann for helping me and everyone else with the mass spectrometer; Matt Petroski for enzymes; and Thibault Mayor for bringing down the house and humoring me on the courts.

Last but not least, I would like to thank my wonderful family for all of their support. I could not have accomplished anything without the constant support of my wife, Jennifer. The smiles of my daughter, Ella, keep me smiling and put everything into perspective. My parents (both sets) have been wonderful in helping us settle down and start our own family.

ABSTRACT

The JAMM (JAB1/MPN/Mov34 metalloenzyme) motif is a conserved amino acid sequence, EX(n)HS/THX(7)SXXD, found in proteins from all domains of life. Eukaryotic proteins possesing a JAMM motif are responsible for the selective hydrolysis of iso-peptide linkages involving ubiquitin and ubiquitin-like proteins and often exist as subunits of large complexes. The iso-peptidase activity of JAMM proteins plays a major role in key points of regulation in the ubiquitin system. In particular, the JAMM motif of CSN5 of the COP9 signalosome is responsible for the cleavage of the ubiquitin-like Nedd8 from SCF ubiquitin ligases. A homolog of CSN5 in the lid subcomplex of the 19S proteasome regulatory particle, Rpn11, cleaves ubiquitin from proteasome substrates as they are processed by the proteasome. In order to understand the mechanism underlying iso-peptide bond hydrolysis by the JAMM motif, we have solved the crystal structure of a JAMM domain protein from *Archaeoglobus fulgidus*, AfJAMM. The JAMM motif forms a thermolysin-like active site on a cytidine deaminase fold. We have demonstrated through biochemical analysis of mutations in the JAMM motif of Csn5 that the mechanism of hydrolysis is similar to that of thermolysin. To achieve an integrated understanding of a JAMM domain protein within its cognate complex, we have purified and crystallized the lid subcomplex of the 19S proteasome regulatory particle for structural studies.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Chapter 1: Introduction to JAMM Domain Proteins

Abbreviations: ATP, adenosine triphosphate; Da, Dalton; COP, constitutively photomorphogenic ; CSN, COP9 signalosome ; E1, ubiquitin activating enzyme; E2, ubiquitin carrier protein; E3, ubiquitin ligase; lys, lysine; HECT, homologous to E6-AP carboxyl terminus; JAMM, Jab1/MPN metalloenzyme; Mpn, Mpr1p Pad1p N-terminal domain; nedd8, neural precursor cell expressed, developmentally downregulated 8; PCI, proteasome, COP9 complex, initiation factor 3; PINT, Proteasome, Int-6, Nip-1 and TRIP-15; RING, really interesting new gene; Rpn, regulatory particle non-ATPase; SCF, Skp1/Cullin/F-box

The Ubiquitin System

In eukaryotic cells, there are two major routes by which proteins may be degraded, through lysosomes or proteasomes. Proteins encapsulated by vesicles, through processes such as endocytosis and autophagy, are normally degraded in lysosomes. Lysosomes are membrane bound organelles harboring acid hydrolases. The bulk of cytosolic and nuclear proteins are degraded by the proteasome as a result of targeting by the ubiquitin system. Ubiquitin is a small protein (8 kDa) which is covalently attached in chains to proteins to signal their destruction. By a mechanism that is not well understood,

these ubiquitinated proteins are ultimately recognized and degraded by the 26S proteasome, a ~2 MDa protein complex.

Ubiquitin is conjugated to proteins through a series of enzymes in a process requiring the input of energy in the form of ATP (Figure 1.1). Hershko and coworkers found that the ability to conjugate ubiquitin to substrates required factors from three eluates (E1-E3) of a ubiquitin affinity column loaded with rabbit reticulocyte lysates (1983). The first factor, E1, catalyzes the formation of a ubiquitin c-terminal adenylate which serves as the substrate for attack by a specific cysteine residue. This results in a thioester bond between the terminal glycine in ubiquitin and a cysteine in the E1. The ubiquitin is then conjugated in a similar manner to a cysteine of an E2. The third factor, E3, may either assist the E2 in ligating the ubiquitin to the substrate by a cooperative allosteric type mechanism or the ubiquitin may be conjugated to a cysteine in the E3 and subsequently ligated to the substrate. The c-terminus of ubiquitin is ligated to the amino side chain of a lysine in the substrate. Ubiquitin itself may serve as a substrate for the formation of ubiquitin chains on the initial substrate. Ubiquitin chains are predominantly formed through linkages of Lys48 of ubiquitin, but may also occur through Lys63, Lys29, or Lys11 linkages. A Lys48 linked tetra-ubiquitin chain is the minimum signal required for degradation by the 26S proteasome.

Regulation of Ubiquitin Ligases

E3 ubiquitin ligases are multi-subunit complexes or single chain proteins that are classified by their E2 binding module. These modules are divided into two major classes,

the HECT type and the RING-finger type. A proposed third class, the U-box type, is predicted to be structurally similar to the RING-finger type (Aravind and Koonin 2000). HECT type ubiquitin ligases form thioester intermediates with ubiquitin and ultimately ligate ubiquitin to substrates, whereas RING-finger ubiquitin ligases mediate transfer of ubiquitin non-enzymatically by recruiting and orienting substrates and E2 enzymes. The largest family of E3 ubiquitin ligases is made up of SCF complexes belonging to the RING-ringer class.

SCF complexes are heteroligomeric complexes made up of a Skp1 subunit, a Cullin subunit, an F-box protein, and a RING-finger protein (Feldman *et al* . 1997, Skowyra *et al.* 1997). The F-box protein binds the substrate targeted for ubiquitination after it has been phosphorylated. The Cullin subunit serves as a scaffold upon which all of the subunits are organized (Deshaies 1999, Zheng *et al.* 2002). In addition, reversible modification of the Cullin subunit with the ubiquitin-like protein, Nedd8, modulates SCF activity (Podust *et al.* 2000, Read *et al.* 2000, Wu *et al.* 2000, Morimoto *et al.* 2000). Cullins are the only proteins known to be modified by Nedd8 and this modification, known as neddylation, occurs in a manner similar to ubiquitination with cognate E1 and E2 enzymes (Figure 1.1 inset).

In an effort to characterize SCF subunit composition, Lypania and coworkers, purified a *myc9* tagged and truncated form of a Cullin subunit and found that all of the subunits of the COP9 signalosome co-purified (2001). The COP9 signalosome (CSN) is a complex of eight subunits that was discovered through genetic analysis of constitutively photomorphogenic *Arabidopsis* (COP) mutants. Following up on this lead, they were able to demonstrate that in the absence of functional CSN, Cullin subunits accumulated *in vivo*

in the neddylated form and that adding purified CSN to these extracts was sufficient to restore Cullin to the deneddylated state. In addition, budding yeast cells lacking *Rri1*, the only well conserved CSN subunit homolog corresponding to CSN5, resulted in the accumulation of neddylated Cullin. These data suggested that the CSN was responsible for deneddylation of Cullin subunits of SCF and that the activity was tightly associated with the CSN5 subunit.

The COP9 Signalosome and the 26S Proteasome

The CSN and a portion of the 26S proteasome descended from a common ancestral complex and have retained certain functional properties (Glickman *et al.* 1998). The 26S proteasome is made up of two sub-particles, the 20S core particle and 19S regulatory particles. The core particle is a chamber composed of two α and β heptameric rings joined by the β rings that harbor the proteolytic active sites. The full 26S proteasome has two 19S regulatory particles, one at each opening of the chamber. The 19S regulatory particle is itself composed of two subparticles, known as the base and lid. The lid particle has eight subunits, each of which is homologous to one of the eight subunits of the CSN. Five of the eight subunits of the CSN and the 19S lid have a C-terminal PINT/PCI domain and of the other three, two have N-terminal MPN domains (Hofmann and Bucher, 1998). One of the subunits containing an MPN domain, CSN5 in the CSN and Rpn11 in the 19S lid, is the most highly conserved subunit between the two complexes.

The observation that ubiquitin is released from substrates when they are degraded by the 26S proteasome led Eytan and coworkers to search for the deubiquitinating enzyme responsible for the activity (1993). They found that the 26S proteasome itself harbored the activity and that it required ATP hydrolysis, was insensitive to ubiquitin-aldehyde, and was inhibited by the metal chelator *o*-phenanthroline. The search for the deubiquitnating subunit of the 26S proteasome ceased for nearly a decade until the relationship between the CSN and 19S lid was uncovered.

Given that deneddylation activity is tightly associated with CSN5 and that CSN5 and Rpn11 are highly conserved, it seemed likely that the deneddylating and deubiquitinating activities of these complexes is directly catalyzed by these subunits. In a multifaceted effort, our laboratory pursued this possibility through bioinformatic, genetic, biochemical, and structural studies (Cope *et al.* 2002, Verma *et al.* 2002, Ambroggio *et al.* 2004). The bioinformatic studies of CSN5 homologs revealed a conserved motif composed of histidines and an aspartic acid residue in a manner reminiscent of zinc hydrolases (Cope *et al.* 2002). We gave this motif the acronym of JAMM for Jab1/MPN metalloenzyme. In the following chapters we present structural and biochemical studies of JAMM domain proteins and the complexes to which they belong.

References

Ambroggio, X. I., Rees, D. C., and Deshaies, R. J. (2004). JAMM: A Metalloprotease-Like Zinc Site in the Proteasome and Signalosome. PLoS Biology *2*, e2.

Aravind, L., and Koonin, E. V. (2000). The U box is a modified RING finger - a common domain in ubiquitination. Curr Biol *10*, R132-134.

Cope, G. A., Suh, G. S., Aravind, L., Schwarz, S. E., Zipursky, S. L., Koonin, E. V., and Deshaies, R. J. (2002). Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cul1. Science *298*, 608-611.

Deshaies, R. J. (1999). SCF and Cullin/Ring H2-based ubiquitin ligases. Annu Rev Cell Dev Biol *15*, 435-467.

Eytan, E., Armon, T., Heller, H., Beck, S., and Hershko, A. (1993). Ubiquitin C-terminal hydrolase activity associated with the 26 S protease complex. J Biol Chem *268*, 4668-4674.

Feldman, R. M., Correll, C. C., Kaplan, K. B., and Deshaies, R. J. (1997). A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. Cell *91*, 221-230.

Glickman, M. H., Rubin, D. M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, V. A., and Finley, D. (1998). A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. Cell *94*, 615-623.

Hershko, A., Heller, H., Elias, S., and Ciechanover, A. (1983). Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. J Biol Chem *258*, 8206-8214.

Hofmann, K., and Bucher, P. (1998). The PCI domain: a common theme in three multiprotein complexes. Trends Biochem Sci *23*, 204-205.

Lyapina, S., Cope, G., Shevchenko, A., Serino, G., Tsuge, T., Zhou, C., Wolf, D. A., Wei, N., and Deshaies, R. J. (2001). Promotion of NEDD-CUL1 conjugate cleavage by COP9 signalosome. Science *292*, 1382-1385.

Morimoto, M., Nishida, T., Honda, R., and Yasuda, H. (2000). Modification of cullin-1 by ubiquitin-like protein Nedd8 enhances the activity of SCF(skp2) toward p27(kip1). Biochem Biophys Res Commun *270*, 1093-1096.

Podust, V. N., Brownell, J. E., Gladysheva, T. B., Luo, R. S., Wang, C., Coggins, M. B., Pierce, J. W., Lightcap, E. S., and Chau, V. (2000). A Nedd8 conjugation pathway is essential for proteolytic targeting of p27Kip1 by ubiquitination. Proc Natl Acad Sci U S A *97*, 4579-4584.

Read, M. A., Brownell, J. E., Gladysheva, T. B., Hottelet, M., Parent, L. A., Coggins, M. B., Pierce, J. W., Podust, V. N., Luo, R. S., Chau, V., and Palombella, V. J. (2000).

Nedd8 modification of cul-1 activates SCF(beta(TrCP))-dependent ubiquitination of IkappaBalpha. Mol Cell Biol *20*, 2326-2333.

Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J., and Harper, J. W. (1997). F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. Cell *91*, 209-219.

Verma, R., Aravind, L., Oania, R., McDonald, W. H., Yates, J. R., 3rd, Koonin, E. V., and Deshaies, R. J. (2002). Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. Science *298*, 611-615.

Wu, K., Chen, A., and Pan, Z. Q. (2000). Conjugation of Nedd8 to CUL1 enhances the ability of the ROC1-CUL1 complex to promote ubiquitin polymerization. J Biol Chem *275*, 32317-32324.

Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M.*, et al.* (2002). Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. Nature *416*, 703-709.

Figures

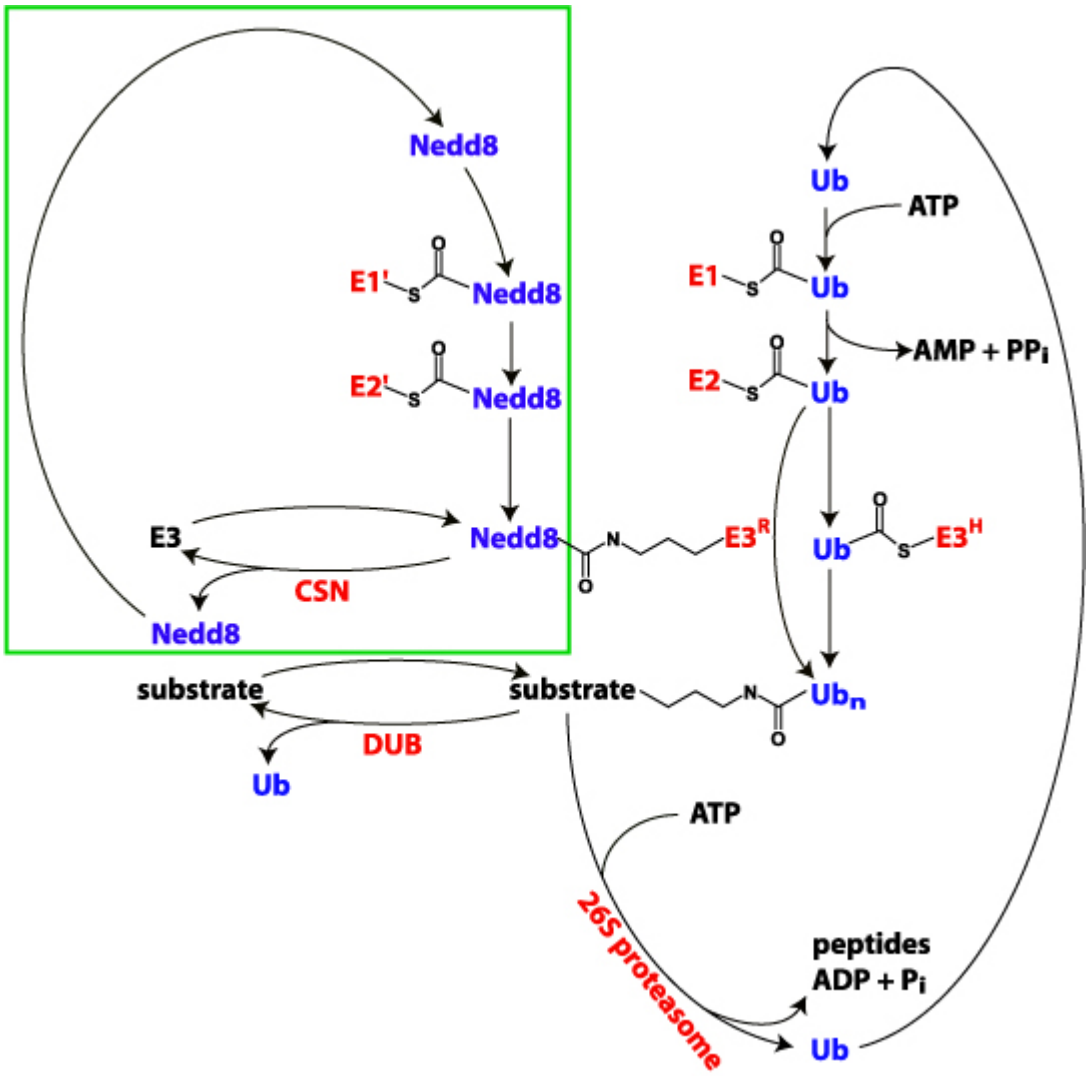

Figure 1.1 **Schematic of the ubiquitin system**

Beginning at the upper right with free ubiquitin: A ubiquitin activating enzyme (E1) uses ATP to convert ubiquitin to an adenylate and subsequently form a thioester linkage to a specific cysteine residue; ubiquitin is transferred to a cysteine of a ubiqutin-carrier protein (E2); ubiquitin is either transferred to a cysteine residue of a HECT-type ubiquitin ligase ($E3^H$) or directly to the substrate with aid from a RING-type ubiquitin ligase ($E3^R$); an isopeptide bond is formed between the ε-amino group of a lysine residue in the

substrate and the carboxy-terminus of ubiquitin; the substrate is degraded in an ATP dependent reaction by the 26S proteasome resulting in the release of peptides and free ubiquitin. Substrates may evade degradation by having their ubiquitin chains trimmed or removed by cellular deubiquitinating enzymes (DUB). (Inset) RING-type ubiquitin ligases ($E3^R$) are activated by modification with the ubiquitin-like protein, Nedd8. Nedd8 is conjugated to RING-type ubiquitin ligases ($E3^R$) in a manner similar to ubiquitin conjugation of proteasome substrates. The COP9 signalosome (CSN) can remove the Nedd8 modification. See text for details.

Chapter 2: JAMM, A Metalloprotease-Like Zinc Site in the Proteasome and Signalosome

Xavier I. Ambroggio[1], Douglas C. Rees[2,3], Raymond J. Deshaies[1,3*]

1 Division of Biology, California Institute of Technology, Pasadena, California, United States of America, **2** Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California, United States of America, **3** Howard Hughes Medical Institute, Chevy Chase, Maryland, United States of America

Received August 29, 2003; Accepted October 9, 2003; Published November 24, 2003

DOI: 10.1371/journal/pbio.0020002

Abbreviations**:** AfJAMM, *A. fulgidus* JAMM protein; AMSH, associated molecule with SH3 domain of STAM; CDA, cytidine deaminase; CSN, COP9 signalosome; Cul1, Cullin 1; DUB, deubiquitinating enzyme; JAMM, JAB1/MPN/Mov34 metalloenzyme; MAD, multiwavelength anomalous dispersion; MPN, Mpr1p Pad1p N-terminal domain;

Nedd8, neural precursor cell expressed, developmentally downregulated 8; RMS, root-mean squared; Rpn11, regulatory particle number 11; SCF, Skp1/Cdc53/Cullin/F-box receptor; ScNP, *S. caespitosus* zinc endoprotease; Ub, ubiquitin; Ubls, ubiquitin-like proteins; UBP, ubiquitin-specific protease; UCH, ubiquitin C-terminal hydrolase

Academic Editor: Hidde L. Ploegh, Harvard Medical School

*To whom correspondence should be addressed. E-mail: deshaies@caltech.edu

Abstract

The JAMM (JAB1/MPN/Mov34 metalloenzyme) motif in Rpn11 and Csn5 underlies isopeptidase activities intrinsic to the proteasome and signalosome, respectively. We show here that the archaebacterial protein AfJAMM possesses the key features of a zinc metalloprotease, yet with a distinct fold. The histidine and aspartic acid of the conserved EX_nHS/THX_7SXXD motif coordinate a zinc, whereas the glutamic acid hydrogen-bonds an aqua ligand. By analogy to the active site of thermolysin, we predict that the glutamic acid serves as an acid-base catalyst and the second serine stabilizes a tetrahedral intermediate. Mutagenesis of Csn5 confirms these residues are required for Nedd8 isopeptidase activity. The active site-like architecture specified by the JAMM motif motivates structure-based approaches to the study of JAMM domain proteins and the development of therapeutic proteasome and signalosome inhibitors.

Introduction

Many cellular proteins are degraded by the proteasome after they become covalently modified with a multiubiquitin chain. The 26S proteasome is a massive protein composed of a 20S core and two 19S regulatory particles (Voges et al. 1999). The 20S core can be subdivided into a dimer of heptameric rings of β subunits—which contain the proteolytic active sites responsible for the protein degradation activity of the proteasome—flanked by heptameric rings of α subunits. The 19S regulatory particle can be divided into a base thought to comprise a hexameric ring of AAA ATPases and a lid composed of eight or more distinct subunits. Whereas 20S core particles and AAA ATPase rings have been found in compartmentalized proteases in prokaryotes, the lid domain of the 19S regulatory particle is unique to eukaryotes and provides the specificity of 26S proteasomes for ubiquitinated substrates (Glickman et al. 1998). Ubiquitin (Ub), an 8 kD protein, is conjugated by Ub ligases to proteasome substrates via an isopeptide bond that links its carboxyl terminus to the amino sidechain of a lysine residue in the substrate. Ub-like proteins (Ubls), of which there are several, are conjugated to their target proteins in a similar manner. Ubls typically do not promote degradation of their targets by the proteasome, but rather regulate target activity in a more subtle manner reminiscent of protein phosphorylation (Hershko and Ciechanover 1998; Peters et al. 1998).

As is the case for protein phosphorylation, the attachment of Ub and Ubls to target proteins is opposed by isopeptidase enzymes that undo the handiwork of Ub ligases. For example, removal of the Ubl Nedd8 (neural precursor cell expressed, developmentally downregulated 8) regulatory modification from the Cullin 1 (Cul1)

subunit of the SCF (Skp1/Cdc53/Cullin/F-box receptor) Ub ligase is catalyzed by the COP9 signalosome (CSN) (Lyapina et al. 2001). The CSN was identified in *Arabidopsis thaliana* from genetic studies of constitutively photomorphogenic mutant plants (Osterlund et al. 1999). It later became evident that CSN and the proteasome lid are paralogous complexes (Glickman et al. 1998; Seeger et al. 1998; Wei et al. 1998). Csn5 of CSN and Rpn11 (regulatory particle number 11) of the proteasome lid are the subunits that are most closely related between the two complexes. CSN-dependent isopeptidase activity is sensitive to metal ion chelators, and Csn5 contains a conserved, putative metal-binding motif (EX_nHS/THX_7SXXD), referred to as the JAMM motif, that is embedded within the larger JAB1/MPN/Mov34 domain (hereafter referred to as the JAMM domain) and is critical for Csn5 function in vivo (Cope et al. 2002). Removal of Ub from proteasome substrates is also promoted by a metal ion-dependent isopeptidase activity associated with the proteasome (Verma et al. 2002; Yao and Cohen 2002). The JAMM/MPN^+ motif of Rpn11 is critical for its function in vivo (Maytal-Kivity et al. 2002; Verma et al. 2002; Yao and Cohen 2002), and proteasomes that contain Rpn11 bearing a mutated JAMM motif are unable to promote deubiquitination and degradation of the proteasome substrate Sic1 (Verma et al. 2002). Taken together, these observations suggested that the JAMM motif specifies a catalytic center that in turn defines a novel family of metalloisopeptidases. Interestingly, the JAMM motif is found in proteins from all three domains of life (Cope et al. 2002; Maytal-Kivity et al. 2002), indicating that it has functions beyond the Ub system. In this study, we present the crystal structure of the *Archaeoglobus fulgidus AF2198* gene product AfJAMM and explore the implications of its novel metalloprotease architecture.

Results and Discussion

We proposed that the subset of JAMM domain proteins that contain a JAMM motif comprise a novel family of metallopeptidases (Cope et al. 2002). To gain a clearer understanding of these putative enzymes—in particular the pertinent subunits of the proteasome lid and signalosome (Figure 2.1)—we cloned and expressed in *Escherichia coli* a variety of JAMM motif-containing proteins to find a suitable candidate for crystallographic analysis. The expression of all candidates except for AfJAMM led to insoluble aggregates. Unlike many JAMM proteins that contain an additional domain, the AfJAMM protein consists entirely of the JAMM domain. We were able to purify and crystallize native and selenomethionine-substituted AfJAMM; the latter was used for phasing by employing the multiwavelength anomalous dispresion (MAD) technique (see Table 1 for statistics).

AfJAMM consists of an eight-stranded β sheet (β1–β8), flanked by a long α helix (α1) between the first and second strand, and a short α helix (α2) between the fourth and fifth strand. This β sheet resembles a β barrel halved longitudinally and curled around α1 (Figure 2.2A). The α2 helix is oriented lengthwise on the convex surface of the β sheet. The zinc-binding site is adjacent to a loop that spans the end of β4 to the beginning of α2 and is stabilized by a disulfide bond between C74 from this loop to C95 on β5. Although disulfide bonds are scarce in intracellular proteins, they are often present in homologous proteins found in hyperthermophiles (Mallick et al. 2002). The overall fold resembles that of the zinc metalloenzyme cytidine deaminase (CDA). CDA from *Bacillus subtilis* (Johansson et al. 2002) can be superimposed onto AfJAMM with a root-mean squared

(RMS) deviation of 3.0 Å over 79 α carbons, despite only 9% sequence identity over structurally aligned residues. The catalytic zinc ions of AfJAMM and CDA, 4.9 Å apart in the superposition, occupy the same general vicinity in the tertiary structures but are coordinated by entirely different protein ligands, two histidines and an aspartic acid in AfJAMM compared to three cysteines in CDA, located at different positions in the sequence (Figure 2.2A). Consequently, the JAMM fold represents a departure from the papain-like cysteine protease architecture that underlies the deubiquitinating activity of the most thoroughly characterized deubiquitinating enzymes (DUBs), the Ub carboxy-terminal hydrolases (UCHs) (Johnston et al. 1997) and Ub-specific proteases (UBPs) (Hu et al. 2002).

The two AfJAMM subunits in the asymmetric unit are connected through a parallel β sheet formed at the dimer interface (Figure 2.2B). The subunits are related by a 2-fold screw axis along the crystallographic c-axis with a translation of 3.38 Å, corresponding to a displacement of one residue along the β3 strand. AfJAMM behaves as a monomer during size exclusion chromatography, suggesting that the dimer observed in the asymmetric unit is an artifact of crystallization. Yet the residues of β3 are highly conserved among JAMM proteins (see Figure 2.1) and predominantly hydrophobic, which makes it difficult to regard the observed interaction as completely insignificant. Flanking β3 to the carboxy-terminal side, there is a striking covariation of residues, MPQSGTG in Rpn11 orthologues and LPVEGTE in Csn5 orthologues. The potential of β3 and the flanking region to mediate specific protein–protein interactions, such as the assembly of Rpn11 and Csn5 into their respective complexes or their specificity towards Ub or Nedd8, warrants further investigation.

The zinc-binding site of AfJAMM is located in a furrow formed by the convex surface of the β2–β4 sheet and α2. The catalytic zinc has a tetrahedral coordination sphere (Figure 2.3A), with ligands provided by $N^{\varepsilon 2}$ of H67 and H69 on β4, the carboxylate of D80 on α2, and a water molecule. The latter hydrogen-bonds to the sidechain of E22 on β2. Thus, the crystal structure confirms previous predictions that the histidine and aspartic acid residues in the JAMM motif are ligands for a metal (Cope et al. 2002; Verma et al. 2002; Yao and Cohen 2002). It must be noted that the identity of the physiological metal in AfJAMM and eukaryotic JAMM homologues is still unknown. The majority of metalloproteases naturally employ zinc but show altered activities with other substituted metals (Auld 1995).

The arrangement of zinc ligands in AfJAMM resembles that found in thermolysin, the *Streptomyces caespitosus* zinc endoprotease (ScNP), and neurolysin, a mammalian metalloprotease (Kurisu et al. 2000; Brown et al. 2001; English et al. 2001). Thermolysin, neurolysin, and ScNP are homologues that have the classical HEXXH metalloprotease motif and adopt the same core fold. In contrast, the sequence, zinc-binding motif, and fold adopted by AfJAMM are entirely distinct. Nonetheless, the active site metal and ligand atoms of thermolysin and ScNP can be superimposed on those of AfJAMM with an RMS deviation of approximately 0.4–0.5 Å (Figure 2.3B).

While this manuscript was under revision, an independent report of a crystal structure of the *AF2198* gene product appeared (Tran et al. 2003). These authors used the fold similarity to CDA as a framework to evaluate the function of the JAMM motif. Given the biochemical data supporting the JAMM motif's role in proteolysis, the common active site architecture seen in AfJAMM and thermolysin, and the similarity of

zinc ligands between thermolysin and AfJAMM, we believe that the extensive body of mechanistic studies on thermolysin and related metalloproteases provide a better framework for the analysis of JAMM function than CDA. In addition to the correspondence between zinc ligands, the glutamic acid residue (E166) downstream of the HEXXH motif of thermolysin is functionally equivalent to the aspartic acid ligand of AfJAMM (D80). E22 in AfJAMM is functionally equivalent to the glutamic acid in thermolysin's HEXXH motif, which serves as the general acid-base catalyst. The conserved serine between the histidine ligands interacts with E22 through a sidechain–main chain hydrogen bond. In more distant JAMM relatives, the serine is replaced by a threonine or asparagine (Aravind and Ponting 1998), both of which are capable of the same bracing function. Meanwhile, the γ-hydroxyl group of the highly conserved S77 in AfJAMM occupies a position similar to $N^{\varepsilon 2}$ of H231 in thermolysin. This atom flanks the 'oxyanion hole' and is implicated in stabilizing the tetrahedral intermediate formed during hydrolysis of the scissile bond (Matthews 1988; Lipscomb and Strater 1996).

AfJAMM was tested for the ability to hydrolyze a number of substrates, including Ub derivatives, resofurin-labeled casein, and D-alanine compounds. Unfortunately, none of the in vitro assays yielded positive results. As nothing is known about AfJAMM in the context of *A. fulgidus* biology, these negative results do not rule out the possibility that AfJAMM functions as a peptide hydrolase in vivo. To validate the suitability of the AfJAMM structure as a basic model for eukaryotic JAMM proteins, we performed site-directed mutagenesis of *Schizosaccharomyces pombe csn5*$^{+}$. The zinc ligands of Csn5 were previously established to be essential for its role in sustaining cleavage of the isopeptide bond that links Nedd8 to Cul1 (Cope et al. 2002). Alanine substitutions for the

putative general acid-base catalyst (E56A) and the catalytic serine (S128) in the JAMM motif of Csn5 likewise abolished its ability to remove the Nedd8 moiety from Cul1 in a *csn5*$^{+}$ background (Figure 2.4A). The E56A mutation had no effect on the assembly of Csn5 with Csn1^{myc13}, while assembly with S128A was slightly hindered (Figure 2.4A). Mutation of the equivalent serine codon in *RPN11* destroyed complementing activity without altering assembly of Rpn11 into the lid. However, the effect of this mutation on Rpn11 isopeptidase activity was not evaluated (Maytal-Kivity et al. 2002). Alanine substitutions for a catalytic residue (E56) or zinc ligands (H118A, D131N) exerted a modest dominant-negative phenotype in *csn5*$^{+}$ cells (Figure 2.4B).

We have been able to assign biochemical functions to Csn5 and Rpn11 (Cope et al. 2002; Verma et al. 2002; Yao and Cohen 2002), but the functions of other eukaryotic JAMM proteins (Figure 2.4C) such as AMSH and C6.1A, as well as the prokaryotic protein RadC and the viral phage λ tail assembly protein K, remain unknown. The structure of AfJAMM provides a useful tool for dissecting the functions of JAMM motifs in these varied contexts and inspires the search for specific JAMM active site inhibitors. The mechanistic implications of the AfJAMM structure explain why the deubiquitinating activity of the lid was unaffected by inhibitors of classical DUBs, the UCHs and UBPs. In classical DUBs, the nucleophile that attacks the carbon of the scissile bond is provided by a cysteine residue in the active site. This property is exploited by using the irreversible inhibitor Ub–aldehyde, which forms a nonhydrolyzable bond to the nucleophilic cysteine (Johnston et al. 1999). In contrast, JAMM proteins likely hydrolyze Ub conjugates in a manner similar to thermolysin, in which the zinc-polarized aqua ligand serves as the nucleophile (Lipscomb and Strater 1996). In the case of thermolysin, metal chelators and

phosphonamidate peptides are effective inhibitors (Bartlett and Marlowe 1987), whereas other zinc metalloproteases are sensitive to peptidomimetic substrates bearing a hydroxamate group (Skiles et al. 2001). Metal chelators have been shown to be effective inhibitors of JAMM proteins (Cope et al. 2002; Verma et al. 2002); it would be interesting to see whether phosphonamidate and hydroxamate peptide mimics of Ub conjugate isopeptides would be equally effective.

The proteasome inhibitor PS-341 has gained attention for its novelty and effectiveness in treating various forms of cancer (Adams 2002). PS-341 was recently approved by the United States' Food and Drug Administration for treatment of relapsed multiple myeloma, thereby validating the proteasome as a target for anticancer therapies. The active site of JAMM proteins is an intriguing target for second-generation therapeutics targeted at the Ub–proteasome pathway for two reasons: the JAMM motif in the proteasome lid is essential for the proteasome to function and the JAMM motif in the CSN specifically regulates the activity of a critical family of E3 Ub ligases (Nalepa and Harper 2003). Inhibition of SCF and other Cullin-based ligases by way of the JAMM motif may be a more specific means of modulating levels of key proteasome substrates in cancer cells.

Materials and Methods

The gene for *A. fulgidus* JAMM (Ponting et al. 1999), open reading frame *AF2198*, was cloned from genomic DNA (ATCC #49558D; American Type Culture Collection, Manassas, Virginia, United States) into the pCRT7 vectors (Invitrogen, Carlsbad, California, United States). During cloning, the alternate start codon, GTG, was

replaced with the canonical start codon, ATG. The construct was expressed in BL21(DE3)pLysS cells (Novagen, Madison, Wisconsin, United States). The cells were grown to midlog phase in terrific broth media and induced with 0.5 mM IPTG. The cells were lysed by sonication and the protein was isolated by immobilized metal ion chromatography using a Ni-NTA resin (Qiagen, Valencia, California, United States). The protein was further purified by gel filtration on a Sephacryl S100 column (Amersham Pharmacia Biotech, Chalfont St Giles, United Kingdom) and concentrated. The amino-terminal tag was removed by limited digestion with trypsin. Mass spectrometry analysis revealed that trypsin only cut AfJAMM in the amino-terminal tag region, and only a single band was evident on a Coomassie-stained polyacrylamide gel. The tag and uncut protein were removed with Ni-NTA resin followed by anion-exchange chromatography with SOURCE 30Q resin (Amersham Pharmacia Biotech). The processed protein was then concentrated to approximately 30 mg/ml by ultrafiltration. The selenomethioine protein was produced as described elsewhere (Van Duyne et al. 1993) and purified using the same protocol as for the native protein.

Protein crystals were obtained in 100 mM $NH_4H_2PO_4$, 200 mM sodium citrate (pH 5) using vapor diffusion with sitting drops and hanging drops. Crystals were incubated for approximately 1 min in a cryo-solution of equal volumes of reservoir solution and 35% meso-erythritol for the selenomethioine crystals and supplemented with 5 mM $ZnCl_2$ for the native crystals. The crystals belonged to the space group $P6_5$, with cell dimensions of a = b = 76 Å, c = 94 Å and two subunits per asymmetric unit. Data for the selenomethioine crystals were collected on Beamline 9.2 at the Stanford Synchrotron Radiation Laboratory (SSRL) (Stanford, California, United States) and data for the native

crystals were collected on Beamline 8.2.1 at the Advanced Light Source (ALS) (Lawrence Berkeley National Laboratory, Berkeley, California, United States) (see Table 1).

Phases were obtained by the MAD technique using data collected from selenomethionine-substituted crystals (see Table 1). Three Se atoms were located by SOLVE (Terwilliger and Berendzen 1999) and used to calculate the initial phases. Phasing was subsequently improved by noncrystallographic symmetry averaging, using operators derived from the Se positions, and solvent flattening in RESOLVE (Terwilliger 2000). The polypeptide model was built in O (Jones et al. 1991) and refined with CNS (Brünger et al. 1998).

Since two monomers in the unit cell are related by a fractional translation along c of approximately 0.54, the intensities of the diffraction pattern are modulated by a factor of $\cos^2(0.54\pi l)$. As a result, reflections with l-indices such that $\cos^2(0.54\pi l) < ½$ are systematically weak, leading to an R-factor higher than would be expected for a nonpseudocentered crystal structure. However, when only the reflections with $\cos^2(0.54\pi l) > ½$ (which will have a more normal intensity distribution) are used for the R-factor calculation, reasonable values for R are obtained.

The geometry of the final model was analyzed with PROCHECK (Morris et al. 1992). The Ramachandran plot shows 98.9% of the residues in the allowed regions and 1.1% in the disallowed regions. The main chain of K66, which constitutes the residue in the disallowed region, was modeled on segments taken from well-refined, high-resolution structures. The Protein Data Bank was searched for structural neighbors of AfJAMM using the DALI server (Holm and Sander 1993). The superpositions with cytidine

deaminase (1JTK), thermolysin (1FJQ), and ScNP (1C7K) were done using the LSQKAB program of the CCP4 distribution (CCP4 1994). All structural figures were made with PyMOL (DeLano 2000).The experiments with *S. pombe* were performed as previously described by Cope et al. (2002).

Supporting Information

Accession Numbers

The accession numbers for the proteins discussed in this paper are 20S proteasomes (PDB ID 1RYP), AfJAMM (Entrez Protein ID NP_071023), AMSH (Entrez Protein ID NP_006454), AtCSN5/AJH1 (Entrez Protein ID NP_173705), AtRpn11 (Entrez Protein ID NP_197745), C6.1A (Entrez Protein ID NP_077308), CeCSN5 (Entrez Protein ID NP_500841), CeRpn11 (Entrez Protein ID NP_494712), Csn5 (Entrez Protein ID NP_593131), Cul1 (Entrez Protein ID NP_594259), cytidine deaminase (PDB ID 1JTK), DmCsn5/CH5 (Entrez Protein ID NP_477442), DmRpn11/p37b (Entrez Protein ID AAF08394), EcRadC (Entrez Protein ID NP_418095), HsAMSH (Entrez Protein ID NP_006454), HsC6.1A (Entrez Protein ID NP_077308), HsCsn5 (Entrez Protein ID NP_006828), HsRpn11/POH1 (Entrez Protein ID NP_005796), JAB1 (Entrez Protein ID AAC17179), lambdaK (Entrez Protein ID AAA96551), Mov34 (Entrez Protein ID NP_034947), Mpr1p (Entrez Protein ID AAN77865), Nedd8 (Swiss-Prot ID Q15843), neurolysin (PDB ID 1I1I), Pad1p (Entrez Protein ID NP_594014), phage λ tail assembly protein K (Entrez Protein ID AAA96551), RadC (Entrez Protein ID NP_418095), Rpn11 (Entrez Protein ID AAN77865), SCF (PDB ID 1LDK), ScNP (PDB ID 1C7K), ScRpn11 (Entrez Protein ID AAN77865), Sic1 (Entrez Protein ID 1360441), SpCsn5 (Entrez

Protein ID NP_593131), SpRpn11/Pad1 (Entrez Protein ID NP_594014), thermolysin (PDB ID 1FJQ), ubiquitin (Swiss-Prot ID P04838), UBP (PDB ID 1NB8), and UCH (PDB ID 1UCH).

These databases may be found at http://www.ncbi.nlm.nih.gov/entrez (Entrez Protein), http://www.rcsb.org/pdb (Protein Data Bank [PDB]), and http://us.expasy.org/sprot/ (Swiss-Prot).

Acknowledgements

This work was supported by the National Science Foundation Graduate Research Fellowship and the Gordon Moore Foundation (to XIA), as well as the Howard Hughes Medical Institute (to DCR and RJD). We would like to thank the staff at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the United States Department of Energy, Office of Basic Energy Sciences, and the Advanced Light Source, which is supported by the Director of the Office of Science, Office of Basic Energy Sciences, Materials Sciences Division of the United States Department of Energy under contract number DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory. Special thanks go to R. Verma and G. Cope for assistance with the associated biochemistry, T. Yeates and O. Einsle for insights concerning the treatment of the pseudocentered crystals, K. Locher and P. Strop for helpful discussions, and J. Ambroggio for back massages and constant support.

References

Adams J (2002) Development of the proteasome inhibitor PS-341. Oncologist 7: 9–16.

Aravind L, Ponting CP (1998) Homologues of 26S proteasome subunits are regulators of transcription and translation. Protein Sci 7: 1250–1254.

Auld DS (1995) Removal and replacement of metal-ions in metallopeptidases. Methods Enzymol 248: 228–242.

Bartlett PA, Marlowe CK (1987) Possible role for water dissociation in the slow binding of phosphorus-containing transition-state-analogue inhibitors of thermolysin. Biochemistry 26: 8553–8561.

Brown CK, Madauss K, Lian W, Beck MR, Tolbert, WD, et al. (2001) Structure of neurolysin reveals a deep channel that limits substrate access. Proc Natl Acad Sci U S A 98: 3127–3132.

Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, et al. (1998) Crystallography and NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 54: 905–921.

Collaborative Computational Project Number 4 (CCP4) (1994) The CCP4 suite: Programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50: 760–763.

Cope GA, Suh, GS, Aravind L, Schwarz SE, Zipursky SL, et al. (2002) Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cul1. Science 298: 608-611.

DeLano WL (2000) The PyMOL molecular graphics system. Available at http://pymol.sourceforge.net/overview/tsld001.htm via the Internet. Accessed 3 November 2003.

English AC, Groom CR, Hubbard RE (2001) Experimental and computational mapping of the binding surface of a crystalline protein. Protein Eng 14: 47–59.

Glickman MH, Rubin DM, Coux O, Wefes I, Pfeifer G, et al. (1998) A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. Cell 94: 615–623.

Hershko A, Ciechanover A (1998) The ubiquitin system. Annu Rev Biochem 67: 425–479.

Holm L, Sander C (1993) Protein structure comparison by alignment of distance matrices. J Mol Biol 233: 123–138.

Hu M, Li P, Li M, Li W, Yao T, et al. (2002) Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. Cell 111: 1041–1054.

Johansson E, Mejlhede N, Neuhard J, Larsen S (2002) Crystal structure of the tetrameric cytidine deaminase from *Bacillus subtilis* at 2.0 Å resolution. Biochemistry 41: 2563–2570.

Johnston SC, Larsen CN, Cook W, Wilkinson KD, Hill CP (1997) Crystal structure of a deubiquitinating enzyme (human UCH-L3) at 1.8 Å resolution. EMBO J 16: 3787–3796.

Johnston SC, Riddle SM, Cohen RE, Hill CP (1999) Structural basis for the specificity of ubiquitin C-terminal hydrolases. EMBO J 18: 3877–3887.

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 A 47: 110–119.

Kurisu G, Kai Y, Harada S (2000) Structure of the zinc-binding site in the crystal structure of a zinc endoprotease from *Streptomyces caespitosus* at 1 Å resolution. J Inorg Biochem 82: 225–228.

Lipscomb WN, Strater N (1996) Recent advances in zinc enzymology. Chem Rev 96:

2375–2434.

Lyapina S, Cope G, Shevchenko A, Serino G, Tsuge T, et al. (2001) Promotion of NEDD–CUL1 conjugate cleavage by COP9 signalosome. Science 292: 1382–1385.

Mallick P, Boutz DR, Eisenberg D, Yeates TO (2002) Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds. Proc Natl Acad Sci U S A 99: 9679–9684.

Matthews BW (1988) Structural basis of the action of thermolysin and related zinc peptidases. Acc Chem Res 21: 333–340.

Maytal-Kivity V, Reis N, Hofmann K, Glickman MH (2002) MPN^{+}, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. BMC Biochem 3: 28–39.

Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) Stereochemical quality of protein structure coordinates. Proteins 12: 345–364.

Nalepa G, Harper JW (2003) Therapeutic anti-cancer targets upstream of the proteasome. Cancer Treat Rev 29: 49–57.

Osterlund MT, Ang LH, Deng XW (1999) The role of COP1 in repression of Arabidopsis photomorphogenic development. Trends Cell Biol 9: 113-118.

Peters JM, Harris JR, Finley D (1998) Ubiquitin and the biology of the cell. New York: Plenum Press. 472 p.

Ponting CP, Aravind L, Schultz J, Bork P, Koonin EV (1999) Eukaryotic signaling domain homologues in archaea and bacteria: Ancient ancestry and horizontal gene transfer. J Mol Biol 289: 729–745.

Seeger M, Kraft R, Ferrell K, Bech-Otschir D, Dumdey R, et al. (1998) A novel protein complex involved in signal transduction possessing similarities to 26S proteasome subunits. FASEB J 12: 469–478.

Skiles JW, Gonnella NC, Jeng AY (2001) The design, structure, and therapeutic application of matrix metalloproteinase inhibitors. Curr Med Chem 8: 425–474.

Terwilliger TC (2000) Maximum-likelihood density modification. Acta Crystallogr D Biol Crystallogr 56: 965–972.

Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. Acta Crystallogr D Biol Crystallogr 55: 849–861.

Tran HJ, Allen MD, Lowe J, Bycroft M (2003) Structure of the Jab1/MPN domain and its implications for proteasome function. Biochemistry 42: 11460–11465.

Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J Mol Biol 229: 105–124.

Verma R, Aravind L, Oania R, McDonald WH, Yates JR III, et al. (2002) Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. Science 298: 611–615.

Voges D, Zwickl P, Baumeister W (1999) The 26S proteasome: A molecular machine designed for controlled proteolysis. Ann Rev Biochem 68: 1015–1068.

Wei N, Tsuge T, Serino G, Dohmae N, Takio K, et al. (1998) The COP9 complex is conserved between plants and mammals and is related to the 26S proteasome regulatory complex. Curr Biol 8: 919–922.

Yao T, Cohen RE (2002) A cryptic protease couples deubiquitination and degradation by the proteasome. Nature 419: 403–407.

Figures and Tables

		Selenomethionine MAD			
		Crystal 1			Crystal 2
	Native (Zn^{2+})	Peak	Inflection	Remote	Peak (2)
Beamline	ALS 8.2.1	SSRL 9.2	SSRL 9.2	SSRL 9.2	SSRL 9.2
Wavelength (Å)	1.1271	0.9790	0.9792	0.9184	0.9790
Resolution (Å)	38-2.3	30-2.5	30-2.5	30-2.5	30-2.3
R_{sym} (%)	9.4 (37.9)	7.5 (43.9)	6.8 (35.7)	6.9 (40.0)	7.3 (42.8)
Completeness (%)	99.9 (99.9)	100 (100)	99.9 (99.8)	100 (100)	100 (100)
I/σI	4.7 (1.8)	20.0 (3.4)	13.9 (2.5)	14.0 (2.6)	34.0 (6.1)
All reflections	278,220	111,979	55,743	56,206	285,937
Redundancy	20.6	10.5	5.2	5.3	21.1
Refinement					
R_{cryst}/R_{free} (%)	26.1/30.4				
$R_{N\text{-}cryst}/R_{N\text{-}free}$[a] (%)	22.6/27.0				
Number of protein atoms	1634				
Number of Zn^{2+} atoms	2				
Number of waters	89				
RMS deviation bonds (Å)	0.006				
RMS deviation angles (°)	1.21				

[a]Owing to pseudocentering, reflections with l values such that $\cos^2(0.54\pi l) < ½$ are systematically weak, leading to an R-factor higher than would be expected for a nonpseudocentered crystal structure. R_N are the R-factors calculated with only the

reflections with $\cos^2(0.54\pi l) > ½$ (see Materials and Methods). Barring rearrangements of sidechains in the vicinity of the zinc atom, no significant changes were seen between the native and selenomethionine forms.

DOI: 10.1371/journal/pbio.0020002.t001

Table 2.1 **Data Collection Statistics**

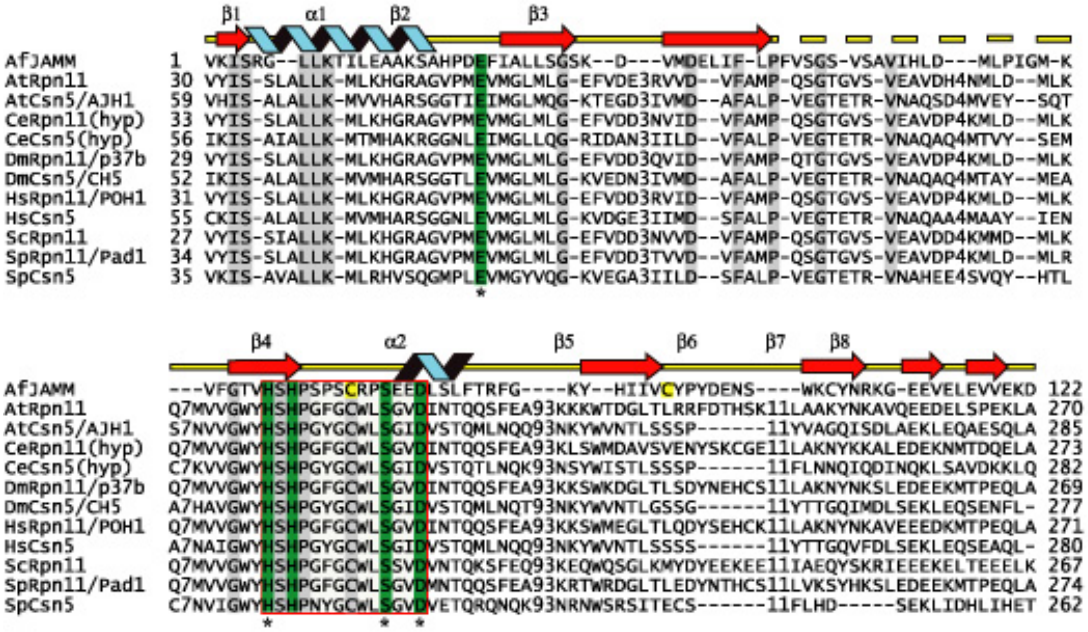

Figure 2.1 **Alignment of Eukaryotic JAMM Domains with AfJAMM**

Eukaryotic JAMM domain proteins were aligned with AfJAMM using ClustalX and manually refined. Sequences are named with a two-letter code corresponding to the genus and species of the respective organism followed by the name of the protein (see Supporting Information for accession numbers), and 'hyp' is an abbreviation for hypothetical. The JAMM motif comprises the residues highlighted in green (E22, H67, H69, S77, and D80), and the active site core is surrounded by a red box. Conserved residues are highlighted in gray. The disulfide cysteine residues are highlighted in yellow

(C74, C95). Active site residues that were mutated in *S. pombe* Csn5 are marked with an asterisk beneath the alignment. The secondary structure of AfJAMM is indicated above the sequence; helices are blue, sheets are red arrows, and loops are yellow lines. The dashed yellow line indicates a loop (F42–G58) that is disordered in the crystal.

DOI: 10.1371/journal/pbio.0020002.g001

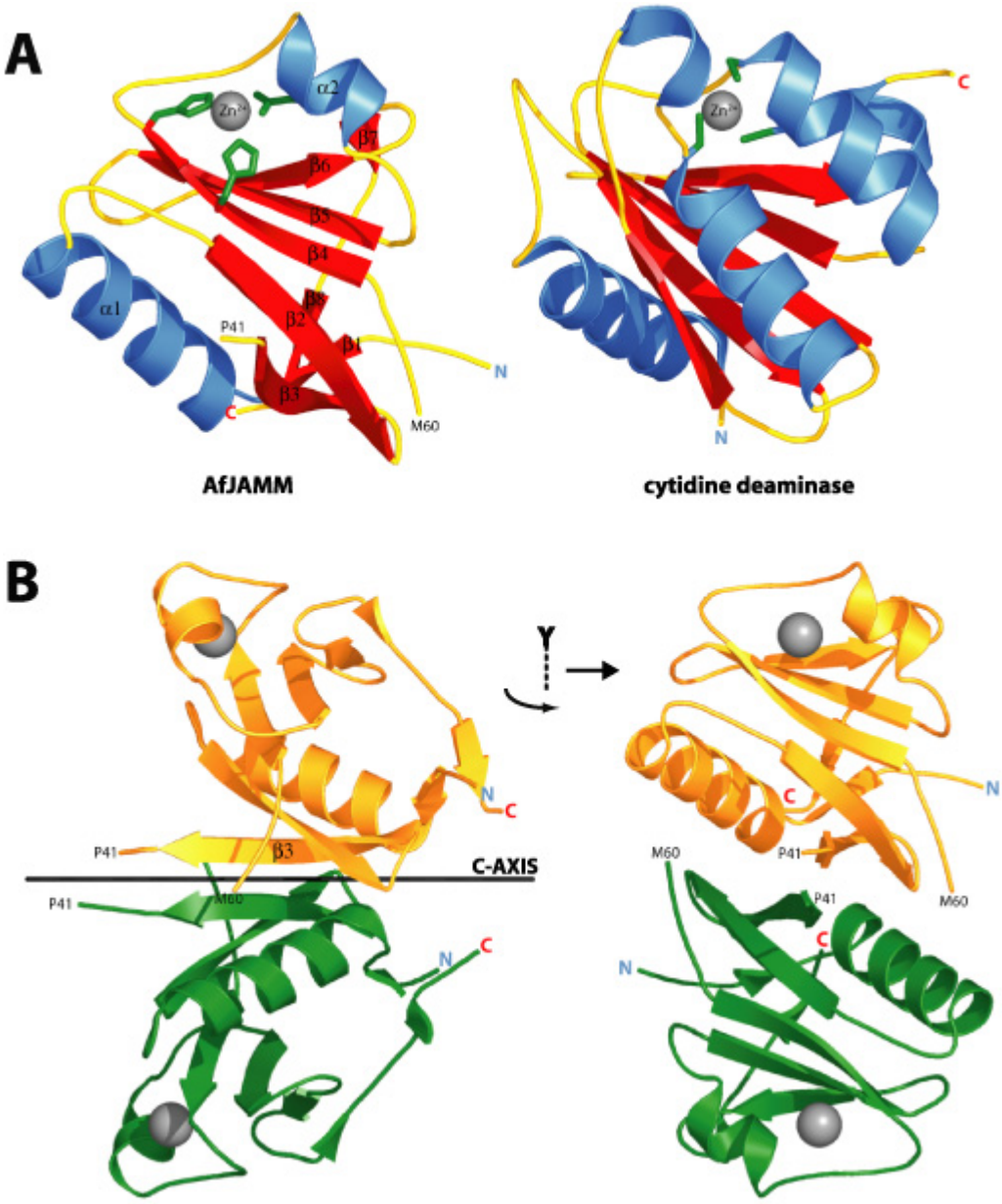

Figure 2.2 **Crystal Structure of AfJAMM**

(A) On the left, the AfJAMM protomer is presented. The amino and carboxyl termini are marked by N and C. The catalytic zinc atom is depicted as a gray sphere. The zinc ligands (H67, H69, and D80) are colored in green. Secondary structure elements are numbered α1–α2 and β1–β8. The amino acids that mark the beginning and end of the disordered loop (P41–M60) are labeled. On the right, the crystal structure of the cytidine deaminase protomer is shown in the same orientation as AfJAMM to highlight the fold likeness as well as the similarly situated zinc-binding sites. The zinc ligands (C53, C86, and C89) are colored in green.

(B) The dimer in the asymmetric unit of AfJAMM crystals. The side view is obtained by rotating the monomer in (A) by 90° as indicated by the quarter-arrow around the y-axis. The gold protomer is related to the green protomer by a 180° rotation around the crystallographic c-axis (shown as a black bar in the side view) and a translation of 3.38 Å.

DOI: 10.1371/journal/pbio.0020002.g002

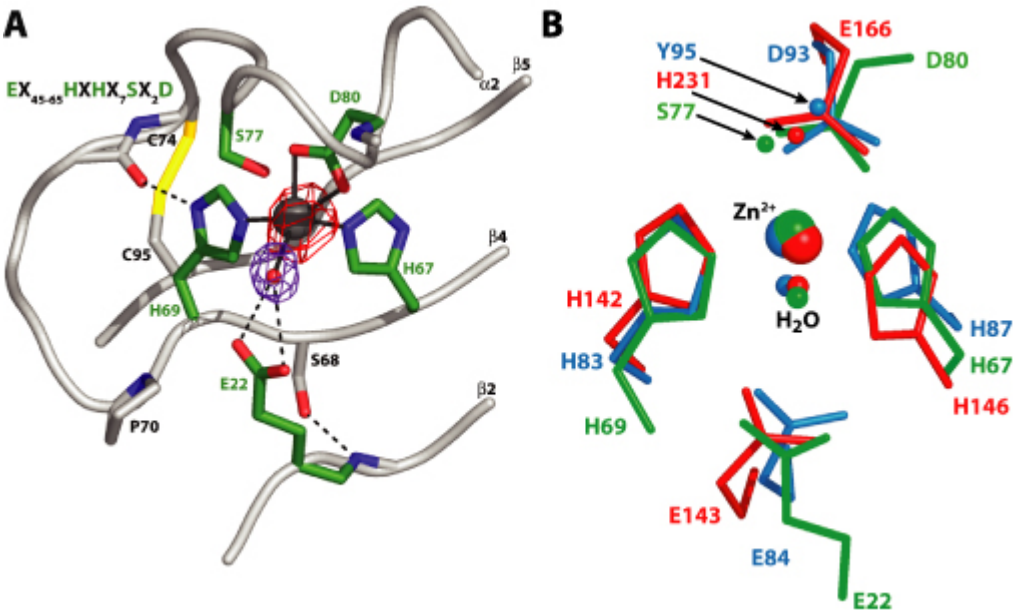

Figure 2.3 **Metalloprotease-Like Active Site of AfJAMM**

(A) The active site of AfJAMM is shown centered around the catalytic zinc ion, which is represented as a dark gray sphere surrounded by anomalous cross Fourier difference density (contoured at 9.5 σ) colored in red. The aqua ligand, which lies at 2.9 Å from the zinc, is shown as a red sphere surrounded by purple density (contoured at 3 σ) of an $F_{obs} - F_{calc}$ map, in which the aqua ligand was omitted from the calculation. Residues that underlie isopeptide bond cleavage are shown in green. The carboxylate oxygen atoms of D80 lie 2.2 Å from the zinc. The $N^{\varepsilon 2}$ atoms of H67 and H69 lie 2.1 Å from the zinc. The carboxylate oxygen atoms of E22 lie 3.2–3.5 Å from the aqua ligand and 4.5–5.0 Å from the zinc. Ancillary active site residues and the backbone (ribbon diagram) are shown in grey. The disulfide bond that links C74 to C95 is shown in yellow. The JAMM motif is shown in the upper lefthand corner for reference.

(B) Superimposition of active site residues in ScNP, thermolysin, and AfJAMM. AfJAMM is in green, ScNP in blue, and thermolysin in red. For clarity only, the sidechains from the residues that bind the zinc or aqua ligands are shown in their entirety. In addition, atoms that stabilize the putative tetrahedral intermediate are shown. These include O^{γ} of S77 in AfJAMM, O^{η} of Y95 in ScNP, and the $N^{\varepsilon 2}$ of H231 in thermolysin.

DOI: 10.1371/journal/pbio.0020002.g003

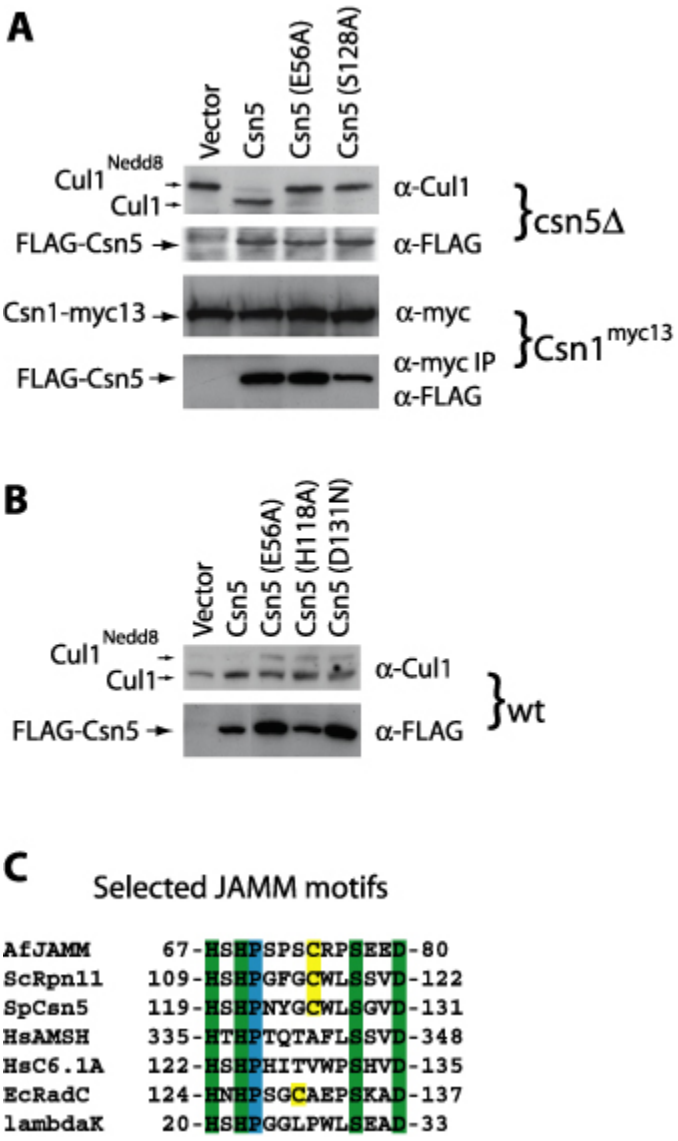

Figure 2.4 **Mutations in the JAMM Motif of Csn5 Abrogate the Deneddylating Activity of the CSN**

(A) Mutations in the glutamic acid (E56A) that positions the aqua ligand and in the proposed catalytic serine (S128A) of Csn5 disrupt deneddylation of Cul1 by CSN but have no effect on assembly with Csn1. A *csn5Δ* strain of *S. pombe* was transformed with an empty pREP-41 plasmid (lane 1) or with the plasmid encoding FLAG tagged: Csn5 (lane 2), $Csn5^{E56A}$ (lane 3), or $Csn5^{S128A}$ (lane 4). Whole-cell lysates were used for Western blot analysis with anti-Cul1 antibodies (top gel) and anti-FLAG antibodies

(second from top). A strain with a *myc13*-tagged Csn1 was transformed with the above plasmids, and whole-cell lysates were used for Western blot analysis. Antibodies to the Myc tag were used to detect Csn1^{myc13} (third from top), and were used to pull down Csn1^{myc13} and subsequently blot with anti-FLAG antibodies to detect coprecipitated Csn5 mutant proteins (bottom gel).

(B) Mutations in the JAMM motif display a modest dominant-negative phenotype. Western blot analysis of crude cell lysates was performed as described in (A).

(C) Selected JAMM motifs from proteins of diverse functions. The canonical JAMM motif residues are highlighted in green. The conserved proline is highlighted in blue, and semiconserved cysteine is highlighted in yellow.

DOI: 10.1371/journal/pbio.0020002.g004

Chapter 3: Purification, Characterization, and Crystallization of the Lid Subcomplex of the 19S Proteasome Regulatory Particle

Abstract

The lid particle is a subcomplex of the 19S proteasome regulatory particle composed of eight subunits (Rpn3, Rpn5, Rpn6, Rpn7, Rpn8, Rpn9, Rpn11, Rpn12) that together with the base complex confers specificity for ubiquitinated substrates to the 26S proteasome. A strain of *S. cerevisiae*, RJD2909, has been constructed to incorporate a C-terminal, twelve histidine tag at the chromosomal *RPN11* locus. Using RJD2909, 19S proteasome regulatory particles and the lid subcomplex can be rapidly purified to homogeneity in milligram quantities. The lid subcomplex preparation has been crystallized and the crystals have been determined to contain zinc, an integral metal of the Rpn11 subunit.

Introduction

Cellular proteins that are modified by the addition of ubiquitin chains of length four or greater are ultimately degraded into short peptides by the 26S proteasome. The 26S proteasome is a large protein complex made up of a cylindrical 20S core particle flanked

by 19S proteasome regulatory particles at both openings (Voges *et al.*, 1999). The 19S proteasome regulatory particle is a multi-subunit complex that enables the 20S proteasome core particle to degrade ubiquitinated substrates in an ATP dependent manner. This ability is the resultant of a variety of functions including: recognition of ubiquitinated substrates, cleavage of ubiquitin chains, unfolding and translocation of substrates, gating the 20S core particle, and other functions that have yet to be discovered.

The 19S regulatory particle is composed of two stable subcomplexes, known as the base and the lid. The base particle has eight subunits, six of which, Rpt1 to Rpt6 (regulatory particle triple-A), contain AAA ATPase domains, and as such are thought to form a hexameric ring characteristic of the AAA ATPase family. The two remaining base subunits, Rpn1 and Rpn2 (regulatory particle non-ATPase), are the largest subunits of the 19S regulatory particle with an approximate molecular mass of 100 kDa. It was previously thought that the fold of these subunits would resemble the horseshoe motif of the ribonuclease inhibitor due to their leucine-rich repeats (Lupas et al. 1997). However recent structures of proteins containing leucine-rich repeats and analysis of their sequences and those of Rpn1 and Rpn2 suggests that the repeats of Rpn1 and Rpn2 may form an all α-helical toroidal structure that resides within the ATPase ring (Kajava 2002). The base particle and 20S core particle are homologous to well studied protease complexes in prokaryotic cells, such as the ClpXP proteases and HslUV (Pickart and Cohen, 2004).

The lid particle is also composed of eight subunits (Rpn3, Rpn5, Rpn6, Rpn7, Rpn8, Rpn9, Rpn11, Rpn12), though unlike the base and core particles there are no prokaryotic

counterparts to the lid particle. In eukaryotic cells however, the lid particle is homologous to two distinct complexes, the COP9 signalosome and the translation initiation factor eIF3 (Glickman *et al.*, 1998; Seeger *et al.*, 1998; Wei *et al.*, 1998). Of all the subunits of the lid particle, Rpn11 is the most thoroughly characterized. Rpn11 was demonstrated to be responsible for removal of the ubiquitin chain from proteasome substrates (Verma *et al.*, 2002; Yao and Cohen, 2002). This is accomplished by the JAMM motif that was shown to fold into a metalloprotease-like active site on a cytidine deaminase fold in AfJAMM, a homologous archaeal protein (Ambroggio *et al.*, 2004). Rpn8 and Rpn11 share homology in their JAMM domain, though Rpn8 is missing the conserved residues of the JAMM motif. All of the remaining subunits except Rpn12 contain C-terminal PCI (proteasome, COP9, initiation factor 3) domains, which are predicted to be α-helical in character and function in protein-protein interactions (Hofmann and Bucher, 1998). In addition to the PCI domain, Rpn3 contains an N-terminally adjacent PAM (PCI associated module) domain predicted to form a TPR-like α-helical structure that has a homologous counterpart in the COP9 signalosome subunit, Csn2 (Ciccarelli *et al.*, 2003). Interestingly, the homolog of Csn2 in the lid particle is not Rpn3 but Rpn6, which was found to interact with Csn2 in a yeast two-hybrid screen (Lier and Paululat, 2002). Yeast mutants of *RPN12* were shown to be necessary for the G1/S and G2/M transitions of the cell cycle (Bailly and Reed, 1999) and *Arabidopsis RPN12* mutants were later found to be responsible for the stability of factors involved in cytokinin regulation, which governs the same cell cycle transition in plants (Smalle *et al.*, 2002). In order to understand how the lid particle is assembled and how it functions in the context of the 19S regulatory

particle and in the greater context of the cell, we have crystallized this particle with the aim of obtaining an atomic model through X-ray diffraction studies.

Expression and Purification

Plasmid pJS-TM53H (Seol *et al.* 2001), which contains a tagging cassette encoding two sequential tobacco etch virus protease cleavage sites followed by nine consecutive Myc epitopes, the CDC53 transcription terminator, and *HIS3* of *Schizosaccharomyces pombe*, was modified by overlap PCR to replace the Myc epitopes with twelve histidine codons resulting in plasmid pT2H12. The *S. cerevisiae* RJD415 strain [*MAT***a** *can1-100 leu2-3,112 his3 trp1-1 ura3-1 ade2-1 pep4::TRP1 bar1::LEU2*] was transformed with a PCR product from a reaction with pT2H12 as the template using a forward primer homologous to the 42 nucleotides preceding the stop codon of *RPN11* and a reverse primer with 41 nucleotides that are homologous to the region following the stop codon. The last 21 nucleotides of the forward and reverse primers are respectively homologous to the first and last regions of the tagging cassette of pT2H12. Transformants were selected on SD –HIS agar plates. Colonies were used to inoculate 50 mL growths in YPD and screened for integration by nickel precipitation using Ni-NTA Agarose (Qiagen) followed by Western blot analysis of the eluates with anti-His (Amersham Biosciences) and anti-Pad1 (*S. pombe* Rpn11) antibodies (Figure 3.1A). A positive transformant was selected and designated as RJD2909 and an aliquot of the corresponding eluate from nickel precipitation was saved for later analysis.

The RJD2909 strain was grown in YPD and levels of $Rpn11^{H12}$ were monitored through late stationary phase by nickel precipitation and detection by anti-His antibodies (Figure 3.1B). Expression levels of $Rpn11^{H12}$ remained constant through all phases of growth. A large (170 L) growth of RJD2909 in YPD was performed at the UCLA Fermentor Facility. The cells were harvested at an OD_{600} of 18.4 yielding a 1.7 kg cell pellet and frozen at –80°C.

A cell pellet of 70 g was thawed in 140 mL of chilled lysis buffer (25 mM Tris pH 7.5, 200 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol, 0.3% Triton X-100, 50 mM NaF, 1 mM PMSF, 3 complete EDTA-free protease inhibitor cocktail tablets (Roche)) and added to 200 mL of chilled 500 μm glass beads inside a stainless steel chamber of a BeadBeater (BioSpec Products). The chamber jacket was filled with ice water and the BeadBeater was run for one minute and let cool for one minute for six cycles at 4°C. The contents of the chamber were applied with vacuum to a filter unit (Nalgene) equipped with a 180 μm nylon net filter (Millipore) to separate the glass beads from the supernatant. The filtrate was subsequently spun at 5000 RPM in an Eppendorf 5804 centrifuge for 20 minutes to remove intact cells and precipitated material.

The cell lysis solution was applied to a column of Ni-NTA Superflow resin (Qiagen) at a flow rate of 2 mL/min. The resin was subsequently washed in wash buffer (25 mM Tris pH 7.5, 500 mM NaCl, 0.3% Triton X-100, 50 mM imidazole, 10 mM β-mercaptoethanol) and starting buffer (25 mM Tris pH 7.5, 200 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol) and the bound protein was eluted with elution buffer (25 mM Tris pH 7.5, 100 mM NaCl, 300 mM imidazole, 10 mM β-mercaptoethanol). Peak fractions were collected and concentrated to a final volume of ≥ 2

mL by ultra-centrifugation in a Centricon concentrator (Millipore), with a molecular weight cut-off of 100 kDa.

The sample was loaded and run on a 26/60 Sephacryl S-400 gel filtration column (Amersham Biosciences) at a flow rate of 0.5 mL/min with gel filtration buffer (50 mM Tris pH 7.5, 200 mM NaCl, 5% glycerol, 1 mM DTT). The lid particle elutes in three peaks at ~156 mL (peak 1), ~164 mL (peak 2), and ~196 mL (peak 3) (Figure 3.2A). Based on calculations derived from calibration standards, the theoretical molecular masses of the three peaks are ~1,200 kDa for peak 1, ~800 kDa for peak 2, and ~500 kDa for peak 3. In peaks 1 and 2, the lid particle elutes as part of the 19S proteasome regulatory particle. In peak 3, the lid particle elutes in a relatively pure form. Fractions corresponding to peaks 2 and 3 were concentrated to ~200 μL by ultra-centrifugation in an Amicon Ultra-15 concentrator (Millipore) with a molecular weight cut-off of 100 kDa. The protein concentrations of peak 2 and 3 were 4 mg/mL and 10 mg/mL respectively, as determined by a DC protein assay (Biorad).

A fourth peak elutes at ~220 mL (peak 4) that contains a mixture of species of SNF1 complexes (see below). SNF1 is a metabolic protein with a molecular weight of ~70 kDa that naturally occurs with thirteen consecutive histidine residues in the N-terminus and therefore co-elutes from immobilized metal affinity resins, such as Ni-NTA Superflow. SNF1 together with SNF4 forms a hetero-trimer with either SIP1, SIP2, or GAL83 (Jiang and Carlson, 1997). SNF1 contamination could be responsible for the common observation that eluates of yeast cell lysates from immobilized metal affinity resins are ‘dirty’. Aside from the SNF1 complexes, there appear to be few species that cannot be removed by routine washes.

Characterization

The eluate from nickel precipitation was analyzed by multidimensional, protein identification technology (MudPIT) (Link *et al.*, 1999). All of the known stoichiometric components of the 19S regulatory particle were identified as well as all of the known components of SNF1 complexes (Table 1).

The 19S sample (peak 2) and the lid sample (peak 3) were analyzed by SDS-PAGE followed by silver staining (Figure 3.2B). All bands were excised from the gel and subjected to in gel tryptic digestion followed by tandem mass spectrometry at the Caltech Protein/Peptide Micro-Analytical Analysis Laboratory. For the 19S sample, all of the subunits of the lid particle were identified as well as four of the eight base particles. The ATPase subunits of the base are of similar molecular mass and have redundant peptides, which may have resulted in the apparent absence of Rpt3 to Rpt6. As they were present in the MudPIT analysis, it seems likely that they are also present in the 19S sample. In the lid sample, all of the lid subunits are present.

Crystallization

Hampton Crystal Screens were used to setup sitting-drop crystallization screens for the lid sample. For all screens, 800 μL of reservoir solution was used and mixed 1:1 with 1 μL of lid sample. After four days, crystals appeared in condition 41 of Hampton Screen

I (0.1 M HEPES pH 7.5, 10% v/v iso-Propanol, 20% w/v Polyethylene Glycol 4000) grown at 293 K. The crystals were approximately 10 μm^3 and grew as multiple attached crystals in an octaheadral-like fashion (Figure 3.3 inset). The crystals were readily reproducible in a narrow range surrounding the initial condition.

Fluorescence Data Collection

Fluorescence data was collected at 100 K from a crystal cluster at Beamline 9.2 of the Stanford Synchrotron Radiation Laboratory (Stanford, California). The looped crystals were cryoprotected by supplementing the mother liquor to a final concentration of 25 % v/v isopropanol. Three data sets were collected at energies adjacent to the K edges of iron, cobalt, and zinc. Only a signal for zinc was detected (Figure 3.3). As there was no zinc added to the expression media or during the purification, the positive signal for zinc is interpreted as evidence that the crystals are composed lid particles since the Rpn11 subunit is a putative zinc metalloprotein.

Acknowledgements

We would like to thank Thomas Sutherland of the UCLA Fermentor Facility, Johannes Graumann for assistance with the MUDPit analysis, Gary Hathaway of the Protein/Peptide Micro-Analytical Analysis Laboratory, and Akif Tezcan for assistance with the fluorescence scans.

References

Ambroggio, X. I., Rees, D. C., and Deshaies, R. J. (2004). JAMM: A Metalloprotease-Like Zinc Site in the Proteasome and Signalosome. PLoS Biol *2*, E2.

Bailly, E., and Reed, S. I. (1999). Functional characterization of rpn3 uncovers a distinct 19S proteasomal subunit requirement for ubiquitin-dependent proteolysis of cell cycle regulatory proteins in budding yeast. Mol Cell Biol *19*, 6872-6890.

Ciccarelli, F. D., Izaurralde, E., and Bork, P. (2003). The PAM domain, a multi-protein complex-associated module with an all-alpha-helix fold. BMC Bioinformatics *4*, 64.

Glickman, M. H., Rubin, D. M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, V. A., and Finley, D. (1998). A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. Cell *94*, 615-623.

Hofmann, K., and Bucher, P. (1998). The PCI domain: a common theme in three multiprotein complexes. Trends Biochem Sci *23*, 204-205.

Jiang, R., and Carlson, M. (1997). The Snf1 protein kinase and its activating subunit, Snf4, interact with distinct domains of the Sip1/Sip2/Gal83 component in the kinase complex. Mol Cell Biol *17*, 2099-2106.

Kajava, A. V. (2002). What curves alpha-solenoids? Evidence for an alpha-helical toroid structure of Rpn1 and Rpn2 proteins of the 26 S proteasome. J Biol Chem *277*, 49791-49798.

Lier, S., and Paululat, A. (2002). The proteasome regulatory particle subunit Rpn6 is required for Drosophila development and interacts physically with signalosome subunit Alien/CSN2. Gene *298*, 109-119.

Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., 3rd (1999). Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol *17*, 676-682.

Lupas, A., Baumeister, W., and Hofmann, K. (1997). A repetitive sequence in subunits of the 26S proteasome and 20S cyclosome (anaphase-promoting complex). Trends Biochem Sci *22*, 195-196.

Pickart, C. M., and Cohen, R. E. (2004). Proteasomes and their kin: proteases in the machine age. Nat Rev Mol Cell Biol *5*, 177-187.

Seeger, M., Kraft, R., Ferrell, K., Bech-Otschir, D., Dumdey, R., Schade, R., Gordon, C., Naumann, M., and Dubiel, W. (1998). A novel protein complex involved in signal transduction possessing similarities to 26S proteasome subunits. Faseb J *12*, 469-478.

Seol, J. H., Shevchenko, A., and Deshaies, R. J. (2001). Skp1 forms multiple protein complexes, including RAVE, a regulator of V-ATPase assembly. Nat Cell Biol *3*, 384-391.

Smalle, J., Kurepa, J., Yang, P., Babiychuk, E., Kushnir, S., Durski, A., and Vierstra, R. D. (2002). Cytokinin growth responses in Arabidopsis involve the 26S proteasome subunit RPN12. Plant Cell *14*, 17-32.

Verma, R., Aravind, L., Oania, R., McDonald, W. H., Yates, J. R., 3rd, Koonin, E. V., and Deshaies, R. J. (2002). Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. Science *298*, 611-615.

Voges, D., Zwickl, P., and Baumeister, W. (1999). The 26S proteasome: a molecular machine designed for controlled proteolysis. Annu Rev Biochem *68*, 1015-1068.

Wei, N., Tsuge, T., Serino, G., Dohmae, N., Takio, K., Matsui, M., and Deng, X. W. (1998). The COP9 complex is conserved between plants and mammals and is related to the 26S proteasome regulatory complex. Curr Biol *8*, 919-922.

Yao, T., and Cohen, R. E. (2002). A cryptic protease couples deubiquitination and degradation by the proteasome. Nature *419*, 403-407.(Yao and Cohen, 2002)

Figures and Tables

Protein	*m.w.(Da)*	*Uniq.(all)*	*S.C.(%)*
Base Particle Subunits			
Rpt1	51,983	21(26)	44.8
Rpt2	48,828	18(27)	45.8
Rpt3	47,894	14(28)	40.2
Rpt4	49,408	10(25)	20.4
Rpt5	48,256	20(47)	44.5
Rpt6	45,272	12(28)	43.2
Rpn1	109,492	26(60)	33.8
Rpn2	104,232	23(28)	28.6
Total	505,365		
Lid Particle Subunits			
Rpn3	60,393	8(10)	17.6
Rpn5	51,769	13(18)	29.0
Rpn6	49,774	19(97)	34.3
Rpn7	48,959	12(17)	26.3
Rpn8	38,313	20(45)	51.5
Rpn9	45,783	22(46)	44.3
Rpn11*	34,398	16(58)	49.0
Rpn12	31,919	12(27)	63.1

Total	361,308		
SNF1 Complex Subunits			
Snf1	72,045	18(42)	33.2
Snf4	36,401	12(25)	33.9
Sip1	96,259	5(7)	10.5
Sip2	46,405	7(14)	29.2
Gal83	46,648	11(22)	41.5
1,4:Sip1	204,705		
1,4:Sip2	154,851		
1,4:Gal83	155,094		

Table 3.1 **MUDPit analysis of eluates from nickel precipitations of RJD2909 extracts.**

The first column gives the standard name of the protein; the second column gives the molecular weight of the protein calculated from the amino acid sequence; the third column gives the number of unique peptides followed by the total number of peptides in parenthesis; the fourth column gives the percent of the sequence covered by recovered peptides. Calculated totals are given for each complex. SNF1 forms a dimmer with SNF4 and either SIP1, SIP2, or GAL83. *Rpn11 was present in tagged form ($Rpn11^{H12}$) with a predicted molecular weight of 37,818 Da. Peptides resulting from digestion of the tag were not included in the analysis.

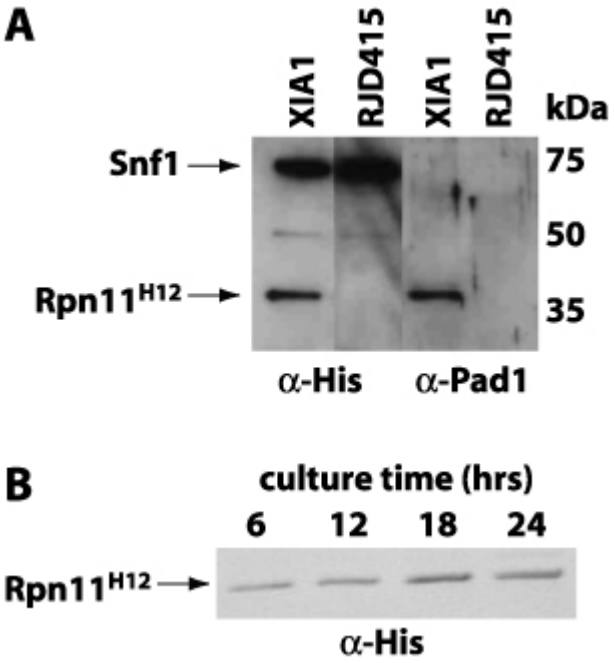

Figure 3.1 **Western blot analysis of RJD2909 strain cell extracts.**

(A) RJD2909 strain and RJD415 parent strain were grown in YPD to the logarithmic phase and cell extracts were prepared for western blot analysis using anti-His (lanes 1 and 2) and anti-Pad1 antibodies (lanes 3 and 4) to verify proper integration of the carboxy-terminal twelve histidine tag. Pad1 is the *S. pombe* Rpn11 homolog and the anti-Pad1 anitbodies cross react with *S. cerevisiae* Rpn11. Snf1 cross reacts with the anti-His antibodies due to the presence of thirteen consecutive histidines in the amino-terminus (see Table 3.1). Molecular weight markers are shown to the right. (B) Time course of RJD2909 growth monitored by western blot analysis of cell lysates using the anti-His antibody.

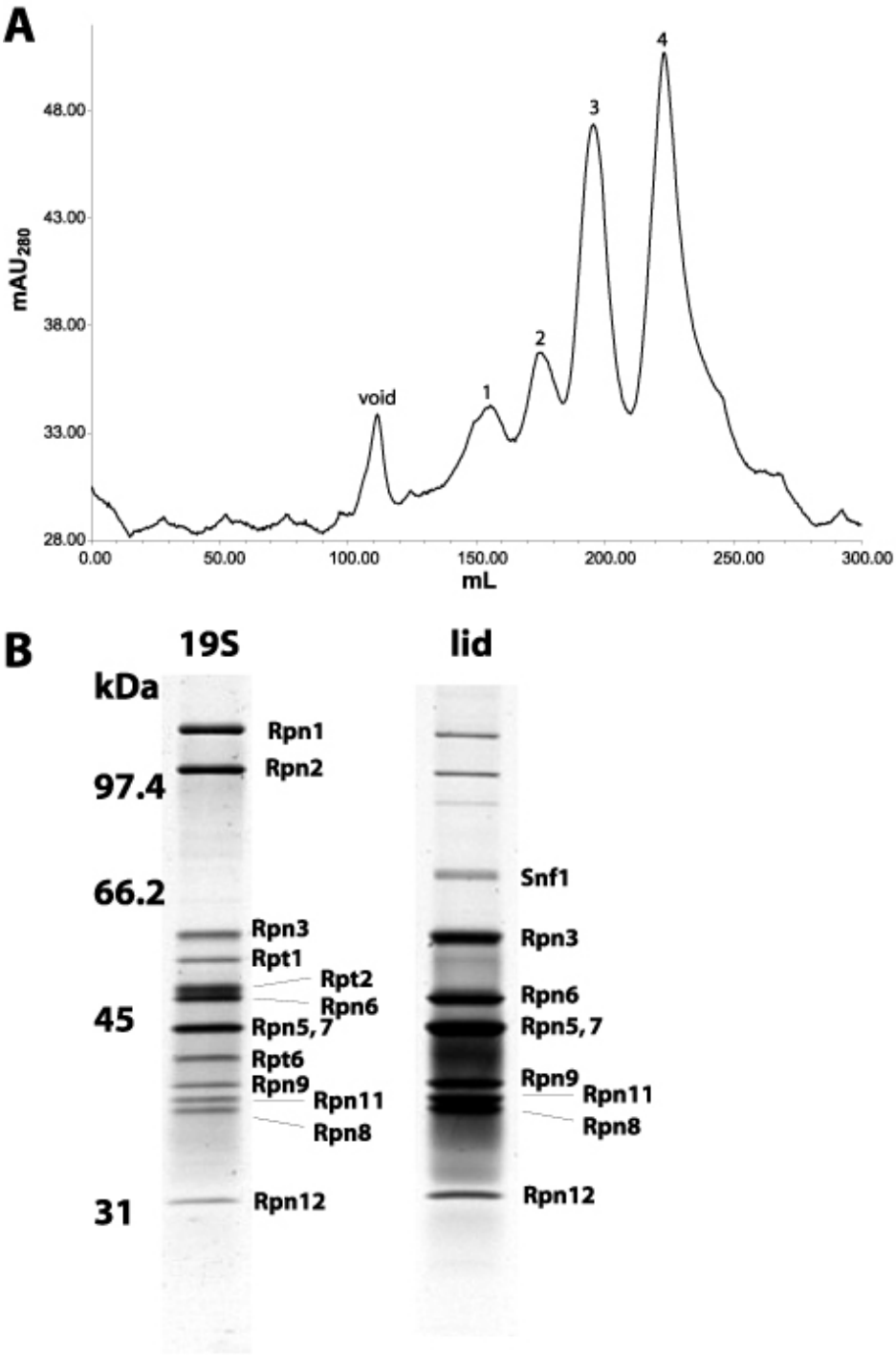

Figure 3.2 **Purification and characterization of the 19S proteasome regulatory particle and the lid subcomplex**

(A) Separation of 19S proteasome regulatory particle species, the lid subcomplex, and SNF1 complexes on Sephacryl S-400 resin (Amersham Biosciences). Peaks 1 and 2

correspond to distinct subpopulations of cellular 19S proteasome regulatory particles. Peak 3 corresponds to the lid subcomplex and peak 4 corresponds to a collection of SNF1 complexes. (B) Analysis of peak 2 (19S) and 3 (lid) fractions by SDS-PAGE with silver staining and tandem mass spectroscopy of excised gel fragments digested with trypsin. Minor contamination of the lid subcomplex fraction by Snf1 and base components can be seen. These contaminants are not present in detectable amounts in the final concentrated sample by analysis with SDS-PAGE with Coomassie staining (data not shown). Molecular weight markers are shown to the left.

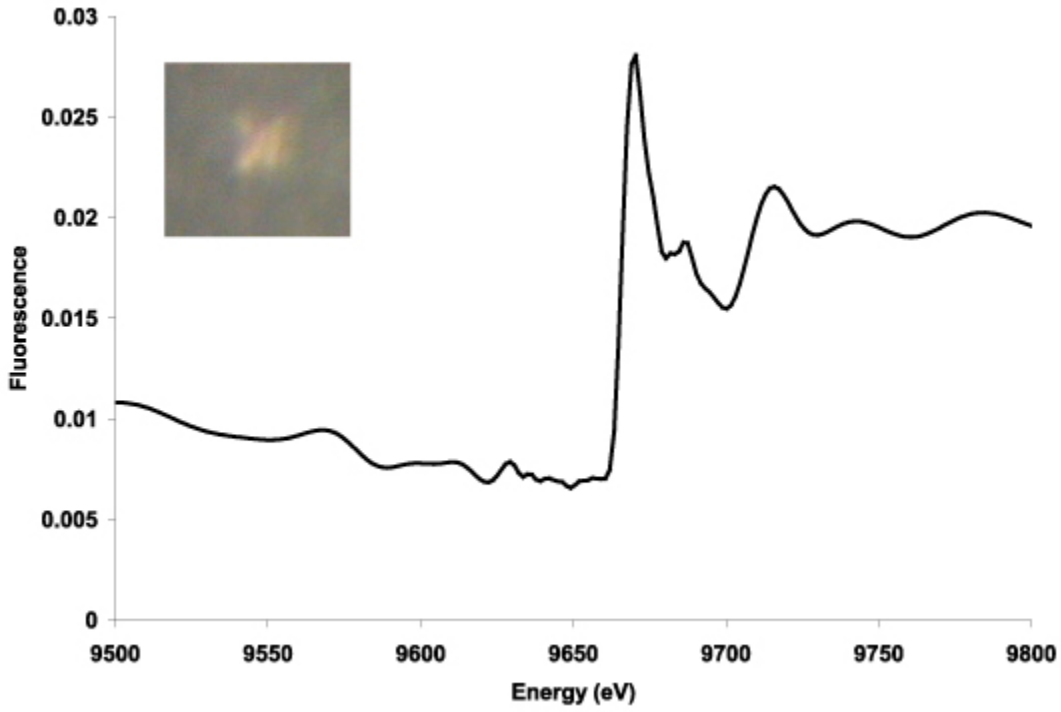

Figure 3.3 **Zinc fluorescence scan of putative lid subcomplex crystals.** Fluorescence scans for cobalt and iron yielded no signals. (Inset) A cluster of crystals (10 μm^3) prepared from a 10 mg/mL sample of lid subcomplex.

Appendix 1: High Resolution Crystal Structures of the Wild Type and Cys-55 → Ser and Cys-59 → Ser Variants of the Thioredoxin-like [2Fe-2S] Ferredoxin from *Aquifex aeolicus*[*]

Andrew P. Yeh[‡], Xavier I. Ambroggio[§], Susana L. A. Andrade[‡¶], Oliver Einsle[‡¶], Claire Chatelet[‖], Jacques Meyer[‖], and Douglas C. Rees[‡¶**]**

From the[‡] Division of Chemistry and Chemical Engineering, the [§] Division of Biology, and the [¶] Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125 and the[‖] Département de Biologie Moléculaire et Structurale, Commissariat à l'Energie Atomique, Grenoble F-38054, France

Received for publication, May 23, 2002, and in revised form, June 25, 2002

Abstract

The [2Fe-2S] ferredoxin (Fd4) from *Aquifex aeolicus* adopts a thioredoxin-like polypeptide fold that is distinct from other [2Fe-2S] ferredoxins. Crystal structures of the Cys-55 → Ser (C55S) and Cys-59 → Ser (C59S) variants of this protein have been determined to 1.25 Å and 1.05 Å resolution, respectively, whereas the resolution of the wild type (WT) has been extended to 1.5 Å. The improved WT structure provides a detailed description of the [2Fe-2S] cluster, including two features that have not been

noted previously in any [2Fe-2S] cluster-containing protein, namely, pronounced distortions in the cysteine coordination to the cluster and a Cα-H-Sγ hydrogen bond between cluster ligands Cys-55 and Cys-9. These features may contribute to the unusual electronic and magnetic properties of the [2Fe-2S] clusters in WT and variants of this ferredoxin. The structures of the two variants of Fd4, in which single cysteine ligands to the [2Fe-2S] cluster are replaced by serine, establish the metric details of serine-ligated Fe-S active sites with unprecedented accuracy. Both the cluster and its surrounding protein matrix change in subtle ways to accommodate this ligand substitution, particularly in terms of distortions of the Fe_2S_2 inorganic core from planarity and displacements of the polypeptide chain. These high resolution structures illustrate how the interactions between polypeptide chains and Fe-S active sites reflect combinations of flexibility and rigidity on the part of both partners; these themes are also evident in more complex systems, as exemplified by changes associated with serine ligation of the nitrogenase P cluster.

Introduction

The low potential iron-sulfur (Fe-S) electron carriers known as ferredoxins (Fds)[1] are found in three distinct classes, the [3Fe-4S]/[4Fe-4S] bacterial type Fds, the plant and mammalian type [2Fe-2S] Fds, and the thioredoxin-like [2Fe-2S] Fds (1). The first two classes were discovered nearly 40 years ago (2, 3) and have since been characterized in considerable detail, including by high resolution x-ray crystallography (4, 5). The third class of Fd is more sparsely distributed and, therefore, has not been investigated as

thoroughly (1). The best characterized members of that group are [2Fe-2S] Fds from the bacteria *Clostridium pasteurianum* (6), *Azotobacter vinelandii* (7), and *Aquifex aeolicus* (8). The high level of similarity of these proteins allows for easy transfer of structural information among them. For instance, many properties of molecular variants of the *C. pasteurianum* [2Fe-2S] Fd (6, 9, 10) could be rationalized from the crystal structure of *A. aeolicus* Fd4 (11). This structure established the unexpected thioredoxin-like fold of these Fds and confirmed that they are distinct from the other two ferredoxin classes (1). The 2.3 Å resolution of the Fd4 structure, however, was not among the highest currently reported (about 1 Å) for metalloenzymes. Indeed, such high resolution structures are of utmost interest because they bring forth precise geometries of metal sites (4) and may allow description of redox transitions (5). Additional efforts have therefore been made, both on wild type (WT) thioredoxin-like Fds and on several of the molecular variants that were produced over the years (6, 10, 12), with the aim of improving the crystallographic data.

In some modified forms of *C. pasteurianum* Fd, cysteine ligands of the Fe-S cluster were replaced by serine (9, 13). In their reduced $[2Fe\text{-}2S]^{+}$ level, these serine-ligated active sites were found to assume a delocalized mixed valence state having a ground spin state of 9/2 (14, 15). This unprecedented occurrence in binuclear iron-sulfur clusters has stimulated experimental and theoretical work (16, 17). The development of these investigations has been hampered, however, by the absence of structural data on serine-ligated [2Fe-2S] active sites and their environment. Because only the Fd4 from *A. aeolicus* has been crystallized, we have repeated amino acid substitutions in that protein previously performed on *C. pasteurianum* Fd and produced the C55S and C59S

variants that contain serine-ligated [2Fe-2S] clusters. We report here the high resolution structures for both of these variants as well as WT protein which provide accurate metric details for serine-ligated Fe-S clusters in proteins.

Experimental Procedures

Protein Samples-- Fd4 from *A. aeolicus* was purified as described by Chatelet *et al*. (8). The C55S and C59S variants were prepared by site-directed mutagenesis as described for the C56S and C60S counterparts from *C. pasteurianum* (13). The mutagenic oligonucleotides were 5'-cacgcgttcatgGaaccggtgggag-3' (hybridizing to the coding strand, mutated base in underlined uppercase) and 5'-ggttgcatgaacgcgtCtatgatgggaccg-3' (hybridizing to the noncoding strand), for C55S and C59S, respectively. The mutated genes were overexpressed in *Escherichia coli*, and the C55S and C59S proteins were purified as described for the WT (8).

Crystallization-- Crystals of oxidized WT, C55S, and C59S Fd4 were prepared by the sitting drop vapor diffusion method. Although anaerobic conditions were employed to minimize degradation of the cluster by exposure to atmospheric oxygen, no reductants were present during the crystallizations, so that the proteins should remain in the oxidized state. In the case of C55S, crystals were obtained by equilibrating 2 µl of reservoir solution and 2 µl of ~83 mg/ml Fd4 C55S (in 20 mM Tris-HCl buffer at pH 8.0 and 0.2 M NaCl) against a reservoir solution containing 30% (w/v) polyethylene glycol 4000, 0.2 M ammonium acetate, and 0.1 M sodium acetate at pH 4.6. Crystals of Fd4 C59S were obtained by equilibrating 2 µl of reservoir solution and 2 µl of ~67 mg/ml Fd4 C59S (in 20 mM Tris-HCl buffer at pH 8.0 and 0.2 M NaCl) against a reservoir solution containing

1.0 M 1,6-hexanediol, 0.01 M cobalt chloride and 0.1 M sodium acetate at pH 4.6. WT Fd4 was crystallized by equilibrating 2 µl of ~10 mg/ml Fd4 (in 10 mM Tris-HCl buffer at pH 8.0 and 0.2 M NaCl) and 2 µl of reservoir solution against a reservoir solution containing 0.01 M zinc sulfate heptahydrate, 0.1 M MES buffer at pH 6.5, and 25% polyethylene glycol monomethyl ether 550. Despite the different crystallization conditions, in all three cases nearly isomorphous crystals were obtained in space group *C2* (C55S: a = 67.3 Å, b = 59.8 Å, c = 46.9 Å,β = 109.8°; C59S: a = 67.3 Å, b = 59.8 Å, c = 46.8 Å,β = 109.3°; WT: a = 67.2 Å, b = 59.8 Å, c = 47.2 Å,β = 110.3°), with one dimeric Fd4 molecule/asymmetric unit.

Data Collection, Structure Determination, and Refinement-- Diffraction data to 1.25 Å resolution for C55S and 1.05 Å resolution for C59S were collected under cryogenic conditions on beamline 9-2 at the Stanford Synchrotron Radiation Laboratory on an Area Detector Systems Corp. Quantum-4 CCD detector controlled by the distributed control system software BLU-ICE. C55S and C59S data sets were processed and scaled using DENZO and SCALEPACK (18). Diffraction data to 1.5 Å resolution for WT Fd4 were collected under cryogenic conditions at the Stanford Synchrotron Radiation Laboratory beamline 9-2 on an Area Detector Systems Corp. Quantum-315 CCD detector and were processed and scaled using MOSFLM and SCALA (19). A summary of the data collection statistics is listed in Table A.1.

The Fd4 C55S structure was solved by molecular replacement using EPMR (20) with the original WT Fd4 structure determined at 2.3 Å resolution (11) as the search model. Multiple rounds of positional refinement and individual isotropic B-factor refinement with crystallography NMR software (21) were alternated with model rebuilding in the

molecular graphics program O (22) against $2F_o - F_c$ σ_a-weighted and $F_o - F_c$ σ_a-weighted maps (23). The [2Fe-2S] cluster geometry was not restrained during refinement. Upon solvent addition and completion of refinement with CNS, positional and anisotropic B-factor refinements of the model were performed using the programs SHELX97 (24) and REFMAC5 (25), which resulted in a final *R*-factor and *R*-free of 14.4 and 19.6%, respectively. A final round of refinement in SHELX97 yielded the standard uncertainties in atomic coordinates, bond lengths, and angles. The final Fd4 C55S model comprises two subunits (2 × 101 residues, 1,575 atoms), two [2Fe-2S] clusters (8 atoms), and 215 water molecules. Because of the absence of electron density for the first 2 residues at the N terminus and the last 7 residues at the C terminus, these residues were not modeled. The Fd4 C55S model without water molecules was used as the starting model for the Fd4 C59S model. The Fd4 C59S model was refined using a protocol similar to that outlined above for Fd4 C55S to an *R*-factor and *R*-free of 13.8 and 16.2%, respectively. As with Fd4 C55S, no electron density was present for the first 2 residues at the N terminus, and the last 7 residues at the C terminus were absent. The final Fd4 C59S model consists of two subunits (2 × 101 residues, 1,576 atoms), two [2Fe-2S] clusters (8 atoms), and 198 water molecules.

The Fd4 C55S model without water molecules was also used as the starting model for the high resolution WT structure. The model was refined using CNS (21) with a protocol similar to that described above to an *R*-factor and *R*-free of 18.4 and 21.6%, respectively. The cluster geometry was not restrained during refinement. During the solvent addition process, 4 zinc ions and 1 sulfate anion were modeled into difference electron density peaks that were significantly higher than those corresponding to the water molecules. In

contrast to the variant structures, the WT temperature factors were not refined anisotropically because of the lower resolution of the data. The final model comprises two subunits (103 residues in subunit A, 101 residues in subunit B, 1,590 atoms), two [2Fe-2S] clusters (8 atoms), 187 water molecules, 4 zinc ions, and 1 sulfate anion. Final refinement statistics for all three structures are listed in Table A.1.

Results and Discussion

Wild Type Fd4

Overall Structure-- The WT and serine-substituted forms of Fd4 analyzed in this study were crystallized in a monoclinic space group, *C2*, which is distinct from the original tetragonal form solved at 2.3 Å resolution (11). As observed originally, the current structure of WT Fd4 determined at 1.5 Å resolution exists as a homodimer, with each monomer adopting a thioredoxin-like fold (Figure A.1*A*). The two noncrystallographic symmetry-related subunits are nearly identical, with a root mean square deviation of 0.27 Å between 101 Cα atoms. As a consequence of the differences in crystal packing between the original and present WT structures, the two monomers in the Fd4 dimer undergo a slight rigid body shift relative to each other (Figure A.1*A*). With the A subunits of both WT forms superimposed, the shift in the B subunit of the new WT form relative to that of the old form can be characterized quantitatively as a 3.8° rotation about an axis oriented ~74° from the 2-fold rotation axis relating subunits in the dimer. The axis about which this 3.8° rotation occurs passes near residue Thr-B53, which along with Pro-B52, Gly-B54, and the corresponding residues in subunit A, form a short stretch of antiparallel β-sheet which stabilizes the dimer interface. As a result of this change in dimer packing,

the hydrogen bonding geometries of residues in this antiparallel arrangement of β-strands are slightly modified. Because cluster ligand Cys-55 is adjacent to this region, it is possible that these alterations in subunit-subunit packing could be coupled to changes in cluster environment. The differences in crystal interactions are also reflected in changes in conformations in two flexible loop regions spanning residues 13-20 and residues 39-46 (Figure A.1*A*), the former of which is near the cluster ligands Cys-9 and Cys-22.

Iron-Sulfur Cluster-- The [2Fe-2S] cluster, located near the surface of each monomer, is coordinated by 4 cysteines, with Cys-9 and Cys-22 ligating Fe1 and Cys-55 and Cys-59 ligating Fe2 (Figure A.1*B*). At the resolution of the current study, it was possible to conduct the refinement without restraining the cluster geometry, resulting in stereochemical parameters that are both more accurate and also less biased than in the previous WT model.

A [2Fe-2S] cluster coordinated by four sulfhydryl groups may be idealized as a framework consisting of two edge-sharing tetrahedral iron sites with a planar Fe_2S_2 inorganic core. The geometries of real [2Fe-2S] clusters, as observed in both model compounds and in proteins, generally reflect this expectation, although deviations from this idealization are evident (26). Table A.2 summarizes the average stereochemical parameters for [2Fe-2S] clusters in the Fd4 structures described here as well as in proteins refined at high resolution and in model compounds. In general, the bond distances and angles observed in protein-bound [2Fe-2S] clusters agree well with each other and with model compounds. A notable deviation from the idealized symmetry in the protein-bound clusters, however, is the nonplanarity of the Fe_2S_2 core, which can be characterized by the average absolute value of ~175° for the Fe-S-S-Fe torsion angle, where 180° would

correspond to exact planarity. For further comparison, the Fe_2S_2 unit present in [4Fe-4S] clusters is significantly more nonplanar than in [2Fe-2S] clusters, with an average value for this torsion angle of only 162.0° (27). The pronounced puckering of this unit in [4Fe-4S] clusters reflects primarily a more compressed Fe-S-Fe angle (71.7°) relative to [2Fe-2S] clusters (75.5°).

A detailed examination of the geometrical parameters of the [2Fe-2S] cluster in WT Fd4 (Table A.3) reveals two pronounced outliers: (i) the relatively small Cys-55 Sγ-Fe2 Cys-59 Sγ bond angle and (ii) a relatively long Cys-55 Sγ-Fe2 bond. The Cys-55 Sγ-Fe2 Cys-59 Sγ angle averages 90.5° in the two crystallographically independent subunits of WT Fd4, a value that is significantly smaller than the average Cys-9 Sγ-Fe2-Cys-22 Sγ angle of 104.6° observed in the same structure and the consensus value of 105.1° observed in other well refined [2Fe-2S] protein structures (Table A.2). The more compressed Cys-55 Sγ-Fe2-Cys-59 Sγ angle has not been observed previously in other well refined ferredoxins with Cys ligands but is similar to the His Nδ-Fe-His Nδ bond angles of ~94° observed in Rieske type [2Fe-2S] clusters (28, 29). The other unusual feature of the Fd4 cluster geometry involves the Cys-55 Sγ-Fe2 bond, which is longer on average by ~0.06 Å than the other three Fe-Sγ bonds (Table A.3). Constraints on the position of Cys-55 may contribute to this elongated bond because the residues around Cys-55 are well defined and appear to be relatively rigid, as reflected by the lower average temperature factors in this region (14.9 $Å^2$ for residues 53-57 *versus* 21.4 $Å^2$ for all protein atoms). An associated phenomenon may be the increased distortion of the peptide bond torsion angles (ω) of the residues surrounding Cys-55, particularly Thr-53, Met-56, and Ala-58, whose ω torsion angles are on average 166.0°, 171.8°, and 189.1°, respectively

(2.4, 1.4, and 1.6 standard deviations from the ideal value of 180°). With an average Fe1-S1-S2-Fe2 torsion angle of -173°, the extent of nonplanarity of the [2Fe-2S] cluster in WT Fd4 is comparable with those observed in other well refined [2Fe-2S] protein structures.

Cys-55 also participates in an unusual interaction involving the probable formation of a Cα-H-Sγ hydrogen bond between the Cys-55 Cα-H and Cys-9 Sγ (Figure A.2*A*). This interaction is identified on the basis of the Cys-55 Cα- Cys-9 Sγ and the Cys-55 Cα-H-Cys-9 Sγ distances of 3.6 and 2.7 Å, respectively, with a Cα-H-Sγ angle of 149°. For the purposes of this calculation, hydrogen positions were generated with the CCP4 program HGEN (19). Identification of this interaction as a hydrogen bond is consistent with the criteria used to identify potential Cα-H-O hydrogen bonds (a Cα-H-O distance of 2.7 Å (30), which does not take into account the increased van der Waals radius of S relative to O). To our knowledge, this type of hydrogen bond has not been described previously in [2Fe-2S] cluster-containing proteins, although they have been described elsewhere (31, 32). For this interaction to occur, the Cys-55 Cα must be positioned over the [2Fe-2S] cluster, with the hydrogen directed toward the Sγ ligand (Cys-9) of the adjacent iron. Examination of the [2Fe-2S] cluster-containing proteins used in the analysis for Table A.2, as well as of [4Fe-4S] cluster-containing proteins described in Ref. 27, suggests that there are two related side chain conformations that can potentially achieve this interaction. These conformations may be defined in terms of three torsion angles: χ_1, the N-Cα-Cβ-Sγ angle; χ_2, which describes the Cα-Cβ-Sγ-Fe angle; and χ_3, which describes the Cβ-Sγ-Fe-S angle. Because there are either two or three cluster sulfides that can be used to define this latter angle, for [2Fe-2S] and [4Fe-4S] clusters, respectively,

the convention that will be used is to adopt the angle whose absolute value is closest to 0°. With these definitions, the side chain conformations that place the ligand Cα over the ring are (χ_1, χ_2, χ_3) ~(180°, 60°, 30°) and (-60°, −60°, −30°); the positions of the Cα, Cβ, and Sγ atoms for these two conformations are related by a mirror plane that passes through the two iron sites, perpendicular to the Fe_2S_2 plane (Figure A.2). The first solution corresponds to that observed for Cys-55 of Fd4, whereas the latter is observed in [4Fe-4S] cluster-containing proteins such as high potential iron proteins. In the latter case, however, the Cα-H-Sγ bond cannot be formed because the significant pucker in the Fe_2S_2 moiety of [4Fe-4S] clusters leads to an increase of the Cα-Sγ distance to ~5.3 Å, a separation which is too great for this interaction to occur (Figure A.2*B*).

With the exception of the Cα-H-Sγ interaction just described, the hydrogen bonding network in the [2Fe-2S] cluster environment of Fd4 (Table A.4) is altogether not unlike that occurring in plant type Fds (5, 33), which assume distinct protein folds. It is interesting to note that in both cases, the Sγ atoms of the cysteine ligands to the more reducible iron are more solvent exposed and are collectively involved in a larger number of hydrogen bonds than the cysteine ligands of the nonreducible iron (9, 14, 34). The hydrogen bonding network around the cluster in Fd4 does appear to exhibit some variability, as evidenced by the ability of Arg-13 in subunit B to adopt two different conformations: one in which the guanidino moiety can donate a hydrogen bond to Cys-59 Sγ and another in which this group is far enough away such that no such interaction can occur. In the former conformation, Arg-13 is in a position that could shield the [2Fe-2S] cluster from the solvent.

Structural Effects of the Cysteine to Serine Substitutions

Previous spectroscopic characterizations of the C56S and C60S variants of the homologous [2Fe-2S] ferredoxin from *C. pasteurianum* revealed that substitution of the cysteine with a serine at either of these positions resulted in stable protein with serine-coordinated [2Fe-2S] clusters (9, 13). The high resolution crystal structures of the corresponding C55S (1.25 Å) and C59S (1.05 Å) mutated forms of the *A. aeolicus* Fd4 confirm this serine coordination and furthermore allow the consequences of this serine substitution on the [2Fe-2S] cluster geometry to be accurately assessed.

The overall structures of both C55S and C59S Fd4 are nearly identical to that of the WT structure of Fd4 at 1.5 Å resolution (Figure A.1*A*). The root mean square deviations between the corresponding subunits A and B of WT and C55S are 0.17 and 0.16 Å and of WT and C59S are 0.18 and 0.23 Å. Although there are no gross perturbations between the mutant and WT structures, there are subtle differences in the cluster geometry and its local environment resulting from the cysteine to serine substitutions. For all three structures, the distances and angles pertaining to the [2Fe-2S] cluster and hydrogen bonds involving the cluster are provided in Tables A.2 and A.3.

As expected, the overall geometry of the cluster in both mutated forms is very similar to that of WT Fd4, with the most pronounced differences being the shorter length of the Fe2-Oγ bond in both variants compared with that of the Fe-Sγ bond in WT (Table A.3 and Figure A.3). The average Fe2-Oγ distance is 1.99 Å in C55S and 1.94 Å in C59S, whereas the average length of the Fe-Sγ bonds in WT Fd4 is 2.31 Å. The Fe-O bond lengths in these structures are comparable with those seen in the structures of an

Anabaena [2Fe-2S] ferredoxin in which one of the cysteine ligands, Cys-49, was replaced by a serine (34) (PDB 1QOA; average Fe-O bond length of 1.91 Å) as well as of *C. pasteurianum* rubredoxin in which one of the iron ligands, Cys-42, was replaced by a serine (35) (PDB 1BE7; Fe-O bond length of 1.94 Å (36)). These bonds, however, are shorter than the value of 2.16 Å reported for the Fe-O bond observed in the structure of a C77S ligand-exchanged form of the *Chromatium vinosum* [4Fe-4S] high potential iron protein (37). The relatively longer Fe-O bond length observed in that structure was speculated to result from the structural rigidity of the polypeptide backbone. Aside from the shorter Fe-O bond length, also associated with the substitution of the Ser Oγ for Cys Sγ is an increase of ~6-9° in the Oγ-Fe-Sγ angle relative to the compressed values observed for the corresponding Cys-55 Sγ-Fe-Cys-59 Sγ angle in WT (Table A.2). In the C49S variant of the *Anabaena* ferredoxin (34), an increase in the Oγ-Fe-Sγ angle of ~7° on average was observed relative to the native Sγ-Fe-Sγ value (5), whereas changes from approximately -4° to +5° in the three Oγ-Fe-Sγ angles relative to the corresponding Sγ-Fe-Sγ angles were observed (36) in the C42S variant of *C. pasteurianum* rubredoxin (35). Accommodation of the cysteine to serine substitution in both C55S and C59S is facilitated by structural changes in the protein backbone/side chain as well as the inorganic iron-sulfur core itself. The main displacements are those of the Oγ and Fe2 atoms toward each other as well as a somewhat lesser movement of the Sγ atom of the other Fe2 ligand, which follows the movement of the iron. The details of these structural changes in each case (C55S and C59S) differ, however, most likely as a result of the different flexibility of the polypeptide chain near each of these cysteine residues. In the case of C55S, adaptation to the substitution occurs primarily through movement of the

Ser-55 Oγ atom toward the iron-sulfur core (Figure A.3). Relative to the position of the corresponding Cys-55 Sγ atom in the WT structure, the Ser-55 Oγ atom shifts by ~0.64 Å in the C55S structure, with little accompanying movement in the backbone atoms of this residue. There is also a ~0.2 Å shift of Fe2 toward the Ser-55 Oγ atom to which it is coordinated, which results in an even more pronounced distortion from planarity of the [2Fe-2S] core than in the WT structure, as reflected by the average Fe1-S1-S2-Fe2 torsion angle of −171° compared with the WT value of −173°. Other changes in the cluster geometry include a decrease of ~0.04 Å in the Fe1-Fe2 distance which is associated with changes of approximately +1° and -2° in the average S-Fe-S and Fe-S-Fe angles, respectively (Table A.2). The third structural change that arises from the C55S mutation involves Cys-59, the other Fe2 ligand. As a result of the slight movement of Fe2 toward Ser-55, both the side chain and main chain atoms of Cys-59 are pulled toward Fe2 to maintain an Fe2-Sγ bond length of ~2.3 Å. The greater degree to which Cys-59 is structurally perturbed compared with Ser-55, particularly in terms of movement of the backbone atoms, highlights the apparently greater structural rigidity in the region surrounding residue 55.

The main structural perturbations resulting from the C59S mutation, as with C55S, also involve residues 55, 59, and Fe2 of the inorganic core. The shorter Fe-O bond in the C59S structure is accommodated by the movement of Ser-59 and Fe2 toward each other (Figure A.3). Relative to the corresponding atoms in the WT structure (*i.e.* Cys-59 Cα and Sγ), the Cα and Oγ atoms of Ser-59 have shifted 0.42 and 0.70 Å toward Fe2, whereas Fe2 has shifted by ~0.17 Å toward Ser-59. As a result, the Fe-S core becomes more planar (torsion angle of −177°), as opposed to the increase in distortion observed in

C55S (Table A.2). Aside from a shift in the Fe2 position, the only other discernible difference in the inorganic core is a decrease in the Fe1-Fe2 distance by ~0.04 Å, similar to what was observed in the C55S structure. And as in C55S, this compression is associated with changes of approximately +1° and -1° in the values of the S-Fe-S and Fe-S-Fe angles, respectively. Reflecting the same trends as seen for the Cys-55 Sγ-Fe and Cys-59 Sγ-Fe bonds in the WT protein, the Fe-O bonds in C59S are shorter than those in C55S by an average of ~0.05 Å (Table A.3). Again, a likely explanation for this observation is the greater flexibility in the region surrounding position 59 compared with that surrounding position 55, as discussed previously.

Hydrogen bonding interactions between the protein and the [2Fe-2S] cluster in the serine variants are generally similar to those observed in the WT structure (Table A.4). One distinction, however, is that in both the C55S and C59S structures, the side chain of Arg-13 adopts the minor conformation of the WT structure, with a water molecule forming a hydrogen bond to the Cys-59 Sγ/Oγ atom in place of the Arg-13 guanidino group.

Conclusions

The crystal structure of the WT form of Fd4 at the higher resolution of 1.5 Å reveals metric details of the [2Fe-2S] cluster which could not be assigned confidently in the initial study at 2.3 Å resolution. Some of these features are shared with other structurally characterized [2Fe-2S] proteins (5, 33, 38), even though the polypeptide folds are distinct. For instance, in the available high resolution structures, the Fe_2S_2 inorganic cores of most protein-bound clusters are distorted from planarity by ~5-10°. The active site moiety containing the reducible iron (55Sγ-Fe2-59Sγ in the case of *A. aeolicus* Fd4) is more

solvent exposed and is involved in a larger number of hydrogen bonds than the other moiety (9, 14, 34). Furthermore, as noted in the *Anabaena* [2Fe-2S] ferredoxin (5), the Fe-S bond lengths involving the more solvent-exposed Fe tend to be slightly longer than those to the buried site (Table A.3). The [2Fe-2S] active site of *A. aeolicus* Fd4 is also notable for features that have not been observed previously in any [2Fe-2S] protein: the long Cys-55 Sγ-Fe2 bond, the compressed Cys-55 Sγ-Fe2-Sγ-Cys-59 angle, and the Cys-55 Cα-H-Cys-9 Sγ hydrogen bond. These unique features are at least in part consequences of the rigid protein environment around residue 55 and are most probably relevant to the spectroscopic idiosyncrasies of thioredoxin-like [2Fe-2S] Fds, particularly in the ways in which they differ from the plant type [2Fe-2S] Fds (1, 9).

The structures of the C55S and C59S variants have been obtained at even higher resolution (1.25 and 1.05 Å, respectively), and the resolution of the C59S variant is the highest to date for any [2Fe-2S] protein. These structures confirm in each case the presence of a Ser Oγ-Fe2 bond that had been inferred previously from spectroscopic data (9, 13). More importantly, they reveal in considerable detail the conformational changes, in both the inorganic core and the polypeptide chain, that take place to accommodate the shortening of the Fe-Sγ bond upon replacement of sulfur by oxygen. The main structural perturbations observed in each case involve positional shifts of both Fe2 ligands, as well as adjustments to the nonplanarity of the Fe-S core. Interestingly, the structural rearrangements in the C55S and C59S variants differ in ways that are in keeping with the greater rigidity of the polypeptide chain around Cys-55.

The types of structural accommodations associated with changes observed for cluster ligands in Fd4 are also evident in more complex systems, such as nitrogenase. The P

cluster of nitrogenase is a [8Fe-7S] metallocenter that exhibits structurally distinct oxidation states (39, 40). In the dithionite-reduced form assigned to the P^N oxidation state, the P cluster may be considered as two [4Fe-4S] clusters that share a common, hexacoordinate, sulfur. This assembly is coordinated to the nitrogenase MoFe-protein through six cysteine ligands, four of which coordinate a single iron, whereas the remaining two cysteines bridge two irons. In an oxidized form identified as the P^{OX} state, two of the irons move away from the central hexacoordinate sulfur, and these interactions are replaced with protein ligands, an amide nitrogen of one of the cluster cysteines and the side chain hydroxyl of Ser-β188. The structural rearrangements associated with the switch between these two forms of the P-cluster are primarily restricted to an increase in planarity of several Fe_2S_2 faces (Figure A.4) as the relevant Fe change positions; these correspond to an increase in the magnitude of the Fe-S-S-Fe torsion angles from ~145° in P^N (near that of [4Fe-4S] clusters) to ~175° in P^{OX} (near that of [2Fe-2S] clusters). Because these transitions are associated with little change in positions of the coordinating residues, the P^N to P^{OX} conversion more closely resembles the consequences for the [2Fe-2S] cluster geometry of serine ligation at residue Cys-55.

The high resolution structures of the C55S and C59S variants of *A. aeolicus* Fd4 may help illuminate a puzzling property of the counterpart C56S and C60S variants of the homologous protein from *C. pasteurianum*. In the one-electron reduced $[2Fe\text{-}2S]^+$ level, these mutated proteins, but not the WT, assume a delocalized mixed valence state resulting in a ground spin state $S = 9/2$, whereas in all other known cases, $[2Fe\text{-}2S]^+$ clusters display localized mixed valence states with an $S = 1/2$ ground spin state (14, 15). Although the structures reported here are those of the $[2Fe\text{-}2S]^{2+}$ redox level, they may

nevertheless be used, with due caution, in the present discussion. Indeed, high resolution structures of both redox levels of a plant type Fd have shown that no major structural reorganization of the [2Fe-2S] cluster occurs upon reduction (5). The structural features favoring the appearance of the delocalized mixed valence pair may therefore be present, at least incipiently, in the $[2Fe\text{-}2S]^{2+}$ structures reported here. The distortion of the Fe_2S_2 inorganic core from planarity is unlikely to play a role because it is larger in C55S, and smaller in C59S, compared with WT (Table A.2). In contrast, the shortening of the Fe-Fe distance, albeit small (about 0.04 Å), is a unique feature of these serine-ligated [2Fe-2S] clusters. It should be noted that this shortening of the Fe-Fe distance is not a universal consequence of the substitution of Ser for Cys, however, as indicated by the slight increase in this distance in the Ser-49 variant of the *Anabaena* ferredoxin (34) and the absence of any significant change in the series of model compounds prepared by Coucouvanis and co-workers (41). It is feasible that this slight decrease in the Fe-Fe distance may favor the occurrence of the delocalized mixed valence state, especially because it is consistent with the prediction that transition from the localized to the delocalized valence state is determined by subtle structural modifications (15). Another potentially relevant feature occurs in both the WT and serine-ligated structures, namely, the unique distortion of the Cys-55 Sγ-Fe2-Sγ-Cys-59 moiety. Although this strain is likely to enhance the differences in coordination environment between the Fe1 and Fe2 sites in the WT protein, the Sγ/Oγ substitution may perhaps rebalance the electronic properties of Fe1 and Fe2 and thus favor the setup of double exchange and valence delocalization. These questions clearly beg for structural data on the reduced levels of both the WT and serine-ligated forms of *A. aeolicus* Fd4.

Sequence similarities indicate that several large redox enzymes, in particular hydrogenases and complex I of respiratory chains (1), contain subunits or domains that are predicted to assume structures similar to that of *A. aeolicus* Fd4. These subunits or domains presumably function as electron transfer agents and differ in at least two ways from *A. aeolicus* Fd4: (i) a single Fd-like module appears to be present, unlike the dimeric structure of Fd4, and (ii) the protruding loop in the vicinity of the [2Fe-2S] cluster is absent. The latter observation suggests that this loop in *A. aeolicus* Fd4 (and homologs in other bacteria, *e.g. C. pasteurianum* and *A. vinelandii*) may serve a possibly more sophisticated function than just electron transfer. In that respect, it should be emphasized that very rigid (near Cys-55) and more flexible (near Cys-59 or Cys-22) regions of the polypeptide chain, as well as a structurally constrained [2Fe-2S] metal site, are combined at the base of the protruding loop. This enhances the likelihood of tight interactions between the conformation of the polypeptide chain and the electronic structure (*e.g.* redox level) of the metal site. The role of these structural idiosyncrasies in the yet mysterious function of the thioredoxin-like [2Fe-2S] Fds will be the aim of future research.

Acknowledgements

Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the United States Department of Energy, Office of Basic Energy Sciences. The Stanford Synchrotron Radiation Laboratory Structural Molecular Biology Program is supported by the Department of Energy and the National Institutes of Health.

Footnotes

* This work was supported by National Institutes of Health Grant GM45062 (to D. C. R.) and by a predoctoral fellowship from the National Science Foundation (to X. I. A.).The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and the structure factors (code *1M2A*, *1M2B* and *1M2D*) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (*http://www.rcsb.org/*).

** To whom correspondence may be addressed: Division of Chemistry and Chemical Engineering, 147-75CH, California Institute of Technology, Pasadena, CA 91125 (for D. C. R.). E-mail: dcrees@caltech.edu or DBMS-BECP, CEA-Grenoble, 38054 Grenoble, France (for J. M.). E-mail: jacques.meyer@cea.fr.

Published, JBC Papers in Press, June 27, 2002, DOI 10.1074/jbc.M205096200

Abbreviations

The abbreviations used are: Fd(s), ferredoxin(s); Fd4, [2Fe-2S] ferredoxin 4 from *A. aeolicus*; MES, 4-morpholineethanesulfonic acid; WT, wild type.

References

1. Meyer, J. (2001) *FEBS Lett.* **509**, 1-5

2. Mortenson, L. E., Valentine, R. C., and Carnahan, J. E. (1962) *Biochem. Biophys. Res. Commun.* **7**, 448-452

3. Tagawa, K., and Arnon, D. I. (1962) *Nature* **195**, 537-543

4. Dauter, Z., Wilson, K. S., Sieker, L. C., Meyer, J., and Moulis, J. M. (1997) *Biochemistry* **36**, 16065-16073

5. Morales, R., Charon, M. H., Hudry-Clergeon, G., Pétillot, Y., Nørager, S., Medina, M., and Frey, M. (1999) *Biochemistry* **38**, 15764-15773

6. Golinelli, M. P., Chatelet, C., Duin, E. C., Johnson, M. K., and Meyer, J. (1998) *Biochemistry* **37**, 10429-10437

7. Chatelet, C., and Meyer, J. (1999) *J. Biol. Inorg. Chem.* **4**, 311-317

8. Chatelet, C., Gaillard, J., Pétillot, Y., Louwagie, M., and Meyer, J. (1999) *Biochem. Biophys. Res. Commun.* **261**, 885-889

9. Meyer, J., Fujinaga, J., Gaillard, J., and Lutz, M. (1994) *Biochemistry* **33**, 13642-13650

10. Golinelli, M. P., Akin, L. A., Crouse, B. R., Johnson, M. K., and Meyer, J. (1996) *Biochemistry* **35**, 8995-9002

11. Yeh, A. P., Chatelet, C., Soltis, S. M., Kuhn, P., Meyer, J., and Rees, D. C. (2000) *J. Mol. Biol.* **300**, 587-595

12. Chatelet, C., and Meyer, J. (2001) *Biochim. Biophys. Acta* **1549**, 32-36

13. Fujinaga, J., Gaillard, J., and Meyer, J. (1993) *Biochem. Biophys. Res. Commun.* **194**, 104-111

14. Crouse, B. R., Meyer, J., and Johnson, M. K. (1995) *J. Am. Chem. Soc.* **117**, 9612-9613

15. Achim, C., Golinelli, M. P., Bominaar, E. L., Meyer, J., and Münck, E. (1996) *J. Am. Chem. Soc.* **118**, 8168-8169

16. Bominaar, E. L., Achim, C., and Borshch, S. A. (1999) *J. Chem. Phys.* **110**, 11411-11422

17. Achim, C., Bominaar, E. L., Meyer, J., Peterson, J., and Münck, E. (1999) *J. Am. Chem. Soc.* **121**, 3704-3714

18. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307-326

19. Bailey, S. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760-763

20. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 484-491

21. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905-921

22. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47**, 110-119

23. Read, R. J. (1986) *Acta Crystallogr. Sect. A* **42**, 140-149

24. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol.* **277**, 319-343

25. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr.* **53**, 240-255

26. Berg, J. M., and Holm, R. H. (1982) in *Iron Sulfur Proteins* (Spiro, T. G., ed) , pp. 1-66, John Wiley & Sons, New York

27. Strop, P., Takahara, P. M., Chiu, H.-J., Hayley, C., Angove, C., Burgess, B.K., and Rees, D. C. (2001) *Biochemistry* **40**, 651-656

28. Colbert, C. L., Couture, M. M. J., Eltis, L. D., and Bolin, J. T. (2000) *Structure* **8**, 1267-1278

29. Iwata, S., Saynovits, M., Link, T. A., and Michel, H. (1996) *Structure* **4**, 567-579

30. Derewenda, Z. S., Lee, L., and Derewenda, U. (1995) *J. Mol. Biol.* **252**, 248-262

31. van Wart, H. E., and Scheraga, H. A. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 13-17

32. Langen, R., Oh, K. J., Cascio, D., and Hubbell, W. L. (2000) *Biochemistry* **39**, 8396-8405

33. Bes, M. T., Parisini, E., Inda, L. A., Saraiva, L. M., Peleato, M. L., and Sheldrick, G. M. (1999) *Structure* **7**, 1201-1211

34. Hurley, J. K., Weber-Main, A. M., Hodges, A. E., Stankovich, M. T., Benning, M. M., Holden, H. M., Cheng, H., Xia, B., Markley, J. L., Genzor, C., Gomez-Moreno, C., Hafezi, R., and Tollin, G. (1997) *Biochemistry* **36**, 15109-15117

35. Xiao, Z. G., Lavery, M. J., Ayhan, M., Scrofani, S. D. B., Wilce, M. C. J., Guss, J. M., Tregloan, P. A., George, G. N., and Wedd, A. G. (1998) *J. Am. Chem. Soc.* **120**, 4135-4150

36. Meyer, J., and Moulis, J.-M. (2001) in *Handbook of Metalloproteins* (Messerschmidt, A. , Huber, R. , Poulos, T. , and Wieghardt, K., eds), Vol. 1 , pp. 505-517, John Wiley & Sons, Chichester, UK

37. Mansy, S. S., Xiong, Y., Hemann, C., Hille, R., Sundaralingam, M., and Cowan, J. A. (2002) *Biochemistry* **41**, 1195-1201

38. Rebelo, J. M., Dias, J. M., Huber, R., Moura, J. J. G., and Romao, M. J. (2001) *J. Biol. Inorg. Chem.* **6**, 791-800

39. Mayer, S. M., Lawson, D. M., Gormal, C. A., Roe, S. M., and Smith, B. E. (1999) *J. Mol. Biol.* **292**, 871-891

40. Peters, J. W., Stowell, M. H. B., Soltis, S. M., Finnegan, M. G., Johnson, M. K., and Rees, D. C. (1997) *Biochemistry* **36**, 1181-1187

41. Salifoglou, A., Simopoulos, A., Kostikas, A., Dunham, R. W., Kanatzidis, M. G., and Coucouvanis, D. (1988) *Inorg. Chem.* **27**, 3394-3406

42. Cruickshank, D. W. J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 583-601

43. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283-291

44. Mayerle, J. J., Denmark, S. E., DePamphilis, B. V., Ibers, J. A., and Holm, R. H. (1975) *J. Am. Chem. Soc.* **97**, 1032-1045

45. Parisini, E., Capozzi, F., Lubini, P., Lamzin, V., Luchinat, C., and Sheldrick, G. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 1773-1784

Figure and Tables

	WT	C55S	C59S
Data collection statistics			
Maximum resolution (Å)	1.50	1.25	1.05
Wavelength (Å)	0.9918	0.8860	0.9580
Total reflections	98,554	183,676	358,550
Unique reflections	27,758	48,256	77,950
Completeness (%)[a]	98.9 (97.1)	99.7 (99.5)	95.6 (92.0)
I/σ (I)	7.8 (1.5)	28.5 (3.8)	38.2 (6.4)
R_{sym} (%)[b]	5.1 (40.4)	4.7 (26.1)	3.5 (21.6)
Refinement statistics			
Resolution limits (Å)	31.5-1.50	44.3-1.25	31.2-1.05
R-factor[c]	0.184	0.144	0.138
R-free	0.216	0.196	0.162
Estimated coordinate error (Å)[d]	0.09	0.05	0.03
RMS deviations from ideal values			
Bond lengths (Å)	0.024	0.015	0.017
Bond angles (°)	2.192	3.056	2.666
Dihedral angles (°)	25.11	26.48	26.38

Improper torsion angles (°)	1.60	1.64	1.83
Average temperature factor (Å^2)			
Protein	22.1, 20.6	21.0, 19.9	18.1, 16.1
Iron-sulfur	13.0, 12.8	12.9, 12.2	9.5, 9.7
Water	36.8	39.3	31.9
Zinc	37.1		
Sulfate	44.3		
Ramachandran plot,[e] residues in			
Most favored regions (%)	90.2	93.2	93.2
Additional allowed regions (%)	9.1	6.8	6.8
Generously allowed regions (%)	0.6	0.0	0.0
Disallowed regions (%)	0.0	0.0	0.0

[a] Numbers in parentheses correspond to values in the highest resolution shell.

[b] $R_{sym} = (\Sigma_{hkl}\Sigma_i|I_i(hkl) - < I(hkl) > |)/(\Sigma_{hkl}\Sigma_i I(hkl))$.

[c] *R*-factor = Σ ($|F_{obs}| - |F_{calc}|)/\Sigma |F_{obs}|$, *R*-free is the *R*-factor calculated for a 3% test set of reflections excluded from the refinement calculation.

[d] Coordinate errors were obtained from the diffraction component precision index, calculated from the values of *R*-free by the method of Cruickshank (42).

[e] As determined by PROCHECK (43).

Table A.1 **Summary of data collection and refinement statistics**

Parameter	WT	C55S	C59S	Protein standards	Model compound
Bonds (Å)					
Fe-S	2.23	2.24	2.23	2.23 ± 0.03	2.201
Fe-Fe	2.73	2.69	2.69	2.73 ± 0.02	2.691
S-S	3.52	3.59	3.55	3.51 ± 0.04	3.483
Angles (°)					
S-Fe-S	104.5	106.2	105.6	104.3 ± 1.8	104.6
Fe-S-Fe	75.4	73.5	74.4	75.5 ± 1.0	75.4
9Sγ-Fe-22Sγ	104.6	105.1	106.1	105.1 ± 1.7	111.2
55S/Oγ-Fe-59S/Oγ	90.5	96.4	99.1		
|Fe-S-S-Fe| torsion angle	173	171	177	175	180

Table A.2 **Average stereochemical parameters for Fd4 structures (from Table A.1)**

Protein standards are the average values for four independent [2Fe-2S] clusters coordinated by four cysteine ligands in protein structures refined at resolutions ≤1.4 Å. These structures include PDB entries 1QT9, 1AWD, and the two distinct clusters in 1HLR (5, 33, 38). The model compound parameters are derived from the structure of $(Fe_2S_2\ (SC_6H_4CH_3)_4)^{2-}$ (44).

		WT		C55S		C59S	
Parameter	Residue	A	B	A	B	A	B
Bonds (Å)							
Fe1-Fe2		2.74	2.72	2.692	2.681	2.697	2.690
S1-S2		3.52	3.52	3.590	3.592	3.548	3.546
Fe1-S1		2.24	2.22	2.262(7)	2.244(7)	2.216(4)	2.214(3)
Fe1-S2		2.21	2.19	2.215(9)	2.222(8)	2.215(4)	2.217(4)
Fe2-S1		2.23	2.24	2.20(1)	2.200(9)	2.232(4)	2.220(4)
Fe2-S2		2.25	2.25	2.305(8)	2.308(8)	2.217(4)	2.253(4)
Sγ-Fe1	C9	2.27	2.26	2.330(7)	2.328(6)	2.303(4)	2.302(4)
Sγ-Fe1	C22	2.31	2.28	2.302(8)	2.306(8)	2.302(4)	2.304(4)
Sγ-Fe2	C55	2.34	2.37	1.97(1)	2.01(2)	2.318(4)	2.315(5)
Sγ-Fe2	C59	2.29	2.33	2.300(9)	2.296(9)	1.940(9)	1.942(8)
Angles (°)							
S1-Fe1-S2		104.8	106.0	106.7(3)	107.1(3)	106.4(2)	106.3(2)
S1-Fe2-S2		103.7	103.3	105.6(3)	105.6(3)	104.6(2)	104.9(2)
Fe1-S1-Fe2		75.6	75.1	74.2	74.2	74.7	74.7
Fe1-S2-Fe2		75.7	75.3	73.1	72.5	74.3	74.0
Sγ-Fe1-Sγ	C9/C22	105.0	104.3	105.7(3)	104.6(2)	106.1(1)	106.1(1)
O/Sγ-Fe2-O/Sγ	C/S55-C/S59	90.5	90.4	96.6(5)	96.3(5)	100.0(3)	98.2(3)

Sγ-Fe1-S1	C9	105.4	106.4	103.9(3)	104.6(3)	104.2(1)	104.4(1)
Sγ-Fe1-S2	C9	116.7	116.2	115.7(3)	115.2(2)	116.5(1)	116.1(1)
Sγ-Fe1-S1	C22	115.5	113.7	115.7(3)	114.5(3)	115.4(1)	115.4(1)
Sγ-Fe1-S2	C22	109.9	110.6	109.5(3)	110.9(3)	108.5(1)	108.9(1)
O/Sγ-Fe2-S1	C/S55	115.4	115.7	109.2(6)	109.6(6)	114.7(2)	114.9(2)
O/Sγ-Fe2-S2	C/S55	112.7	113.5	112.7(5)	112.0(5)	112.8(1)	112.7(2)
O/Sγ-Fe2-S1	C/S59	113.2	112.2	110.8(4)	111.8(4)	107.6(3)	107.9(3)
O/Sγ-Fe2-S2	C/S59	121.8	122.3	121.6(3)	121.2(3)	117.5(3)	118.6(3)
Torsion Angle (°)							
Fe1-S1-S2-Fe2		−174.4	−172.0	−170.8	−170.6	−176.7	−176.7

Table A.3 **Cluster geometries in molecules A and B for WT, C55S, and C59S Fd4**

Numbers in parentheses correspond to standard uncertainties in the last digit. Because of the lower resolution of the WT structure, coordinate uncertainties of the individual atoms were not calculated.

		WT		C55S		C59S	
		A	B	A	B	A	B
D	A	D-A	D-A	D-A	D-A	D-A	D-A
C22N	S1	3.39	3.44	3.43	3.43	3.40	3.40
M56N	S2	3.26	3.29	3.35	3.34	3.34	3.35

A58N	S2	3.55	3.55	3.57	3.45	3.63	3.58
R13NH1	S1	3.54	3.42				
Wat A	S1			3.56	3.62	3.76	3.73
Q11N	Cys9Sγ	3.58	3.52	3.57	3.48	3.55	3.54
V64N	Cys22Sγ	3.51	3.6	3.48	3.53	3.54	3.57
Wat B	Cys22Sγ	3.11	3.19	3.19	3.25	3.30	3.28
R13NH2	55S/Oγ	3.17	3.06	3.68	3.74	3.42	3.48
N57N	55S/Oγ	3.25	3.29	3.51	3.54	3.30	3.35
Wat C	55S/Oγ					3.84	3.73
Wat D	55S/Oγ			2.76			
Wat E	55S/Oγ					3.63	4.15
R13NH1	59S/Oγ	3.46	3.39				
Wat A	59S/Oγ			3.36	3.35	2.93	2.95
Wat C	59S/Oγ					2.63	2.58
Wat D	59S/Oγ			3.77			
Wat F	59S/Oγ	3.45	3.42				
Wat G	59S/Oγ		3.67		3.79		

Table A.4 **Hydrogen bonding geometry in the [2Fe-2S] cluster environment**

D (donor), A (acceptor); separation distance is measured in Å. Although not specifically indicated, the D-H-A angles for these interactions are all >125 °.

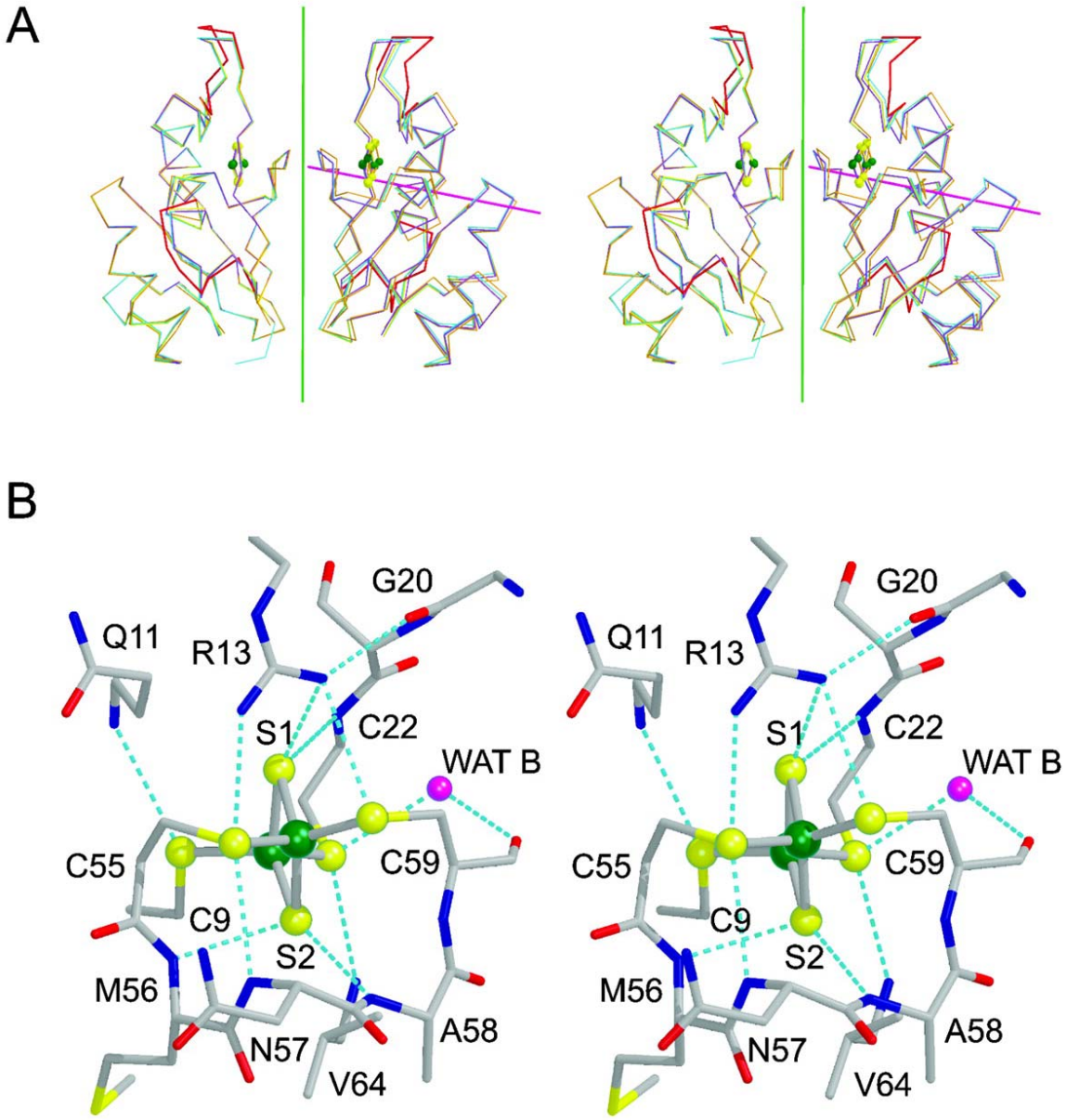

Figure A.1

A, stereoview of superimposed Fd4 dimers from the original WT (*orange*), current WT (*cyan*), C55S (*yellow*), and C59S (*purple*) structures. The A subunits of these molecules were used in the superposition. The *green line* denotes the 2-fold rotation axis relating subunits A (*left*) and B (*right*) in a dimer; the *magenta line* corresponds to the axis about

which the B subunit of the previous WT structure (11) is rotationally shifted relative to the B subunit of the current WT form. This rigid body rotational shift as well as regions of the old form which differ significantly from the new form (residues 13-20 and 39-46, *highlighted* in *red*) are likely the results of the different crystal packing between the two WT forms. *B*, stereoview of the [2Fe-2S] cluster and its immediate environment in WT Fd4, illustrating the hydrogen bonding network involving the cluster, ligands, and surrounding residues, with the exception of the Cys-55 CαH-Cys-9 Sγ hydrogen bond detailed in Figure A.2*A*. Iron and sulfur atoms are colored *green* and *yellow*, respectively. Fe1 is coordinated by Cys-9 and Cys-22, whereas Fe2 is ligated by Cys-55 and Cys-59.

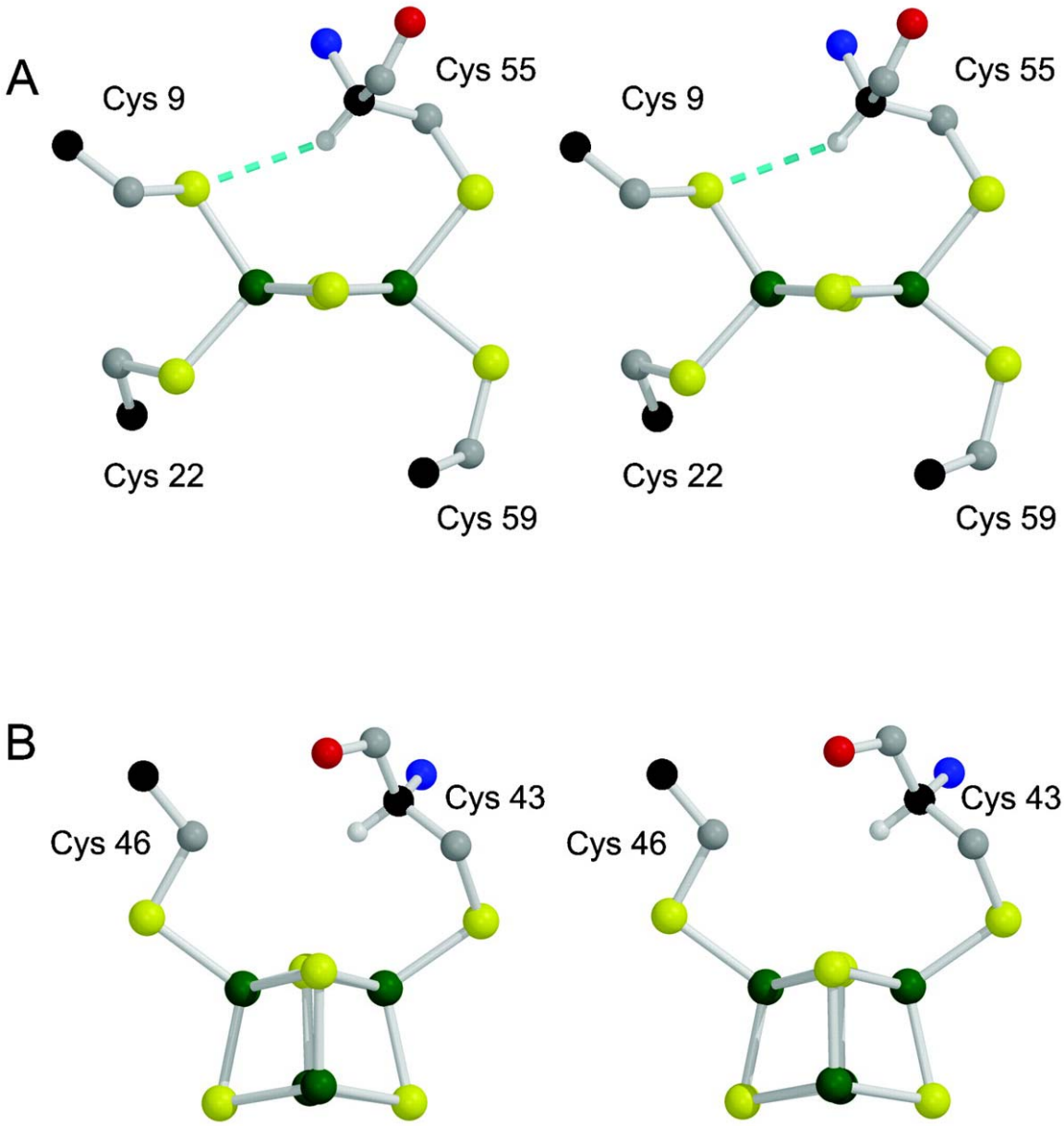

Figure A.2 **Stereoviews of the potential interactions between Cys ligands across the Fe2S2 face of iron-sulfur clusters.**

A, the [2Fe-2S] cluster of Fd4, illustrating the conformation of Cys-55 in which the Cα-H group is positioned over the cluster to hydrogen bond (*dashed line*) with the Cys-9 Sγ. *B*, a similar configuration is observed in the high potential iron protein from *C. vinosum* ((45) PDB 1CKU). Whereas Cys-43 and Cys-46 have the same general relationship to the

common face of the [4Fe-4S] cluster in this protein, the increased pucker of the Fe_2S_2 core precludes the formation of a Cα-H-Sγ hydrogen bond. For clarity, two of the cysteine ligands in high potential iron protein have been omitted from the figure. Iron, sulfur, nitrogen, oxygen, and Cα atoms are colored *green*, *yellow*, *blue*, *red*, and *black*, respectively.

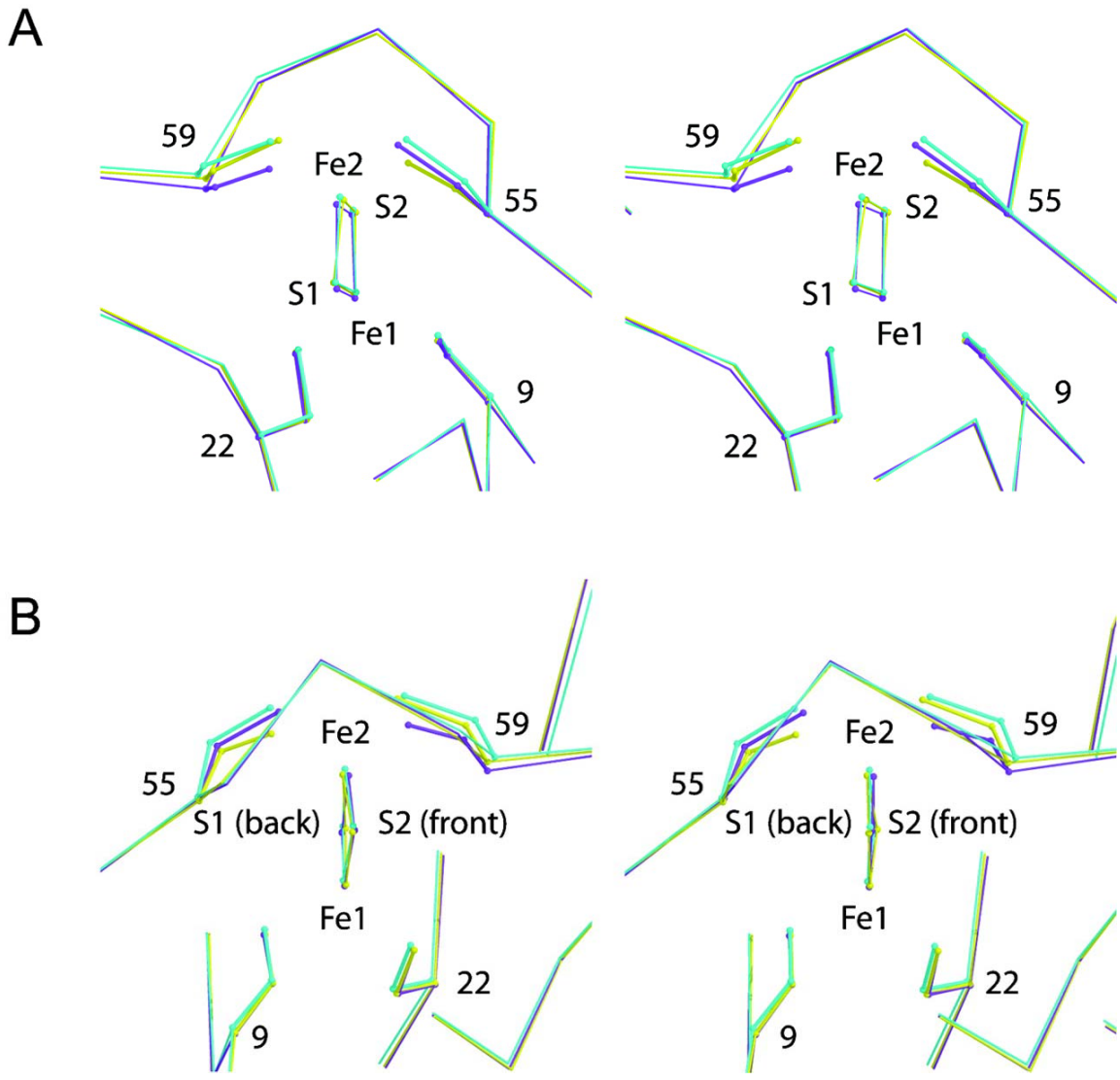

Figure A.3

A, stereoview of the [2Fe-2S] cluster and its ligands from the WT (*cyan*), C55S (*yellow*), and C59S (*purple*) structures upon superposition of the corresponding 101 Cα atoms of the three structures, showing the varying degree of positional shifts that occur in the inorganic core as well as residues 55 and 59 caused by the cysteine to serine substitutions. *B*, stereoview of the same region, viewed from a direction perpendicular to that in *A*, illustrates the varying degrees to which the inorganic core is distorted in each structure. The color scheme is the same as in *A*.

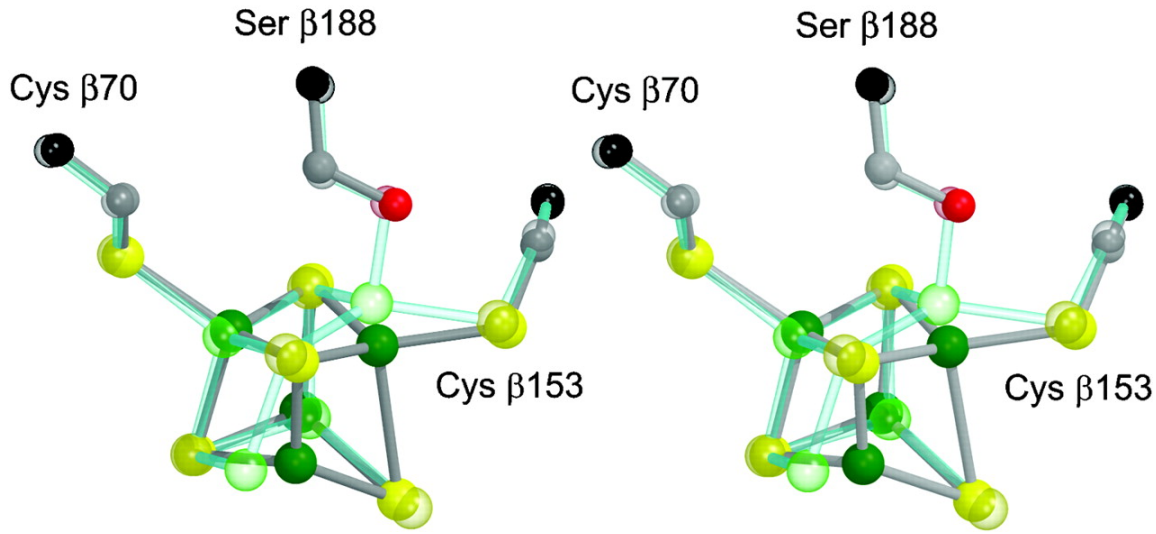

Figure A.4 **Stereoview comparing the P cluster of nitrogenase in the oxidized (*transparent ball-and-stick* model in *cyan*) and reduced (*solid ball-and-stick* model in *gray*) states (**39**,** 40**)**

In the oxidized state, one of the irons is coordinated by the side chain of Ser-β188, whereas in the reduced state this iron is shifted and coordinates an inorganic sulfur in the cluster instead. The coloring scheme is as in Figure A.3. PDB entries 2MIN (oxidized) and 3MIN (reduced) were used for this figure.