This is to certify that the
thesis entitled

NEW INSIGHTS IN THE UREASE ACTIVATION PROCESS OBTAINED BY CHARACTERIZATION OF APOUREASE COMPLEXES AND THE UreG ACCESSORY PROTEIN OF *KLEBSIELLA AEROGENES*

presented by

Soledad De Los Ángeles Quiroz Valenzuela

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Biochemistry and Molecular Biology

Robert P. Hausinger
Major Professor's Signature

April 24, 2008
Date

MSU is an affirmative-action, equal-opportunity employer

NEW INSIGHTS IN THE UREASE ACTIVATION PROCESS OBTAINED BY CHARACTERIZATION OF APOUREASE COMPLEXES AND THE UreG ACCESSORY PROTEIN OF *KLEBSIELLA AEROGENES*

By

Soledad De Los Ángeles Quiroz Valenzuela

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHiLOSOPHY

Department of Biochemistry and Molecular Biology

2008

ABSTRACT

NEW INSIGHTS IN THE UREASE ACTIVATION PROCESS OBTAINED BY CHARACTERIZATION OF APOUREASE COMPLEXES AND THE UreG ACCESSORY PROTEIN OF *KLEBSIELLA AEROGENES*

By

Soledad De Los Ángeles Quiroz Valenzuela

The metallocenter assembly pathway of nickel-containing urease from *Klebsiella aerogenes* has served as a paradigm for understanding the biosynthesis of other metalloproteins. Our current understanding of the process is that urease apoprotein $(UreABC)_3$ interacts with a urease-specific molecular chaperone (UreDFG) and a metallochaperone (UreE) that delivers nickel ions. At each of the three active sites, a specific lysine residue is carbamylated and dinuclear nickel centers are formed in a GTP-dependent process. Urease activation is accompanied by dissociation of the accessory proteins. This thesis expands upon current knowledge by focusing on structural comparisons of the $(UreABC)_3$, $(UreABC\text{-}UreD)_3$, and $(UreABC\text{-}UreDF)_3$ apoprotein complexes, and by characterizing the UreG accessory protein.

The structure of urease apoprotein is known while those of the two activation complexes are unknown. Computational analysis of the urease crystal structure suggested that a hinge in the amino terminal region of UreB could allow a repositioning of this subunit in one of these complexes to open up the nascent active site. I tested the effects of hindering the flexibility of this peptide by substituting two glycine residues with prolines, revealing an important role for Gly11. In addition, the results of small angle X-ray scattering analyses allowed

me to conclude that the UreD and UreF accessory proteins are positioned at the vertices of the urease structure, very close to the UreB subunit. Significantly, the data were better fit to a model in which UreB was repositioned according to the proposed conformational change, thus opening the active site for activation.

The UreG accessory protein, responsible for hydrolyzing GTP during the activation process exists as an independent protein in addition to being part of the activation complex. I created a biotin-tagged version of UreG, allowing for an improved purification procedure. I showed that the biotin tag had no adverse effect on UreG's ability to activate urease and created a series of mutants of UreG in which conserved residues were replaced by alanine. Most of the mutations resulted in the complete loss of urease activity *in vivo*, although substitution of Cys72, His74, and Ser111 (which correspond to metal ligands in the related protein HypB) exhibited negligible changes in metal binding parameters. In additional studies, I show the ability of the protein to form a novel complex with urease apoprotein, UreD, UreF, UreG, and UreE. Furthermore, I demonstrated that the D80A variant of UreG interacted only with UreE, suggesting that Asp80 stabilizes the larger complex. The two new complexes uncovered here set the stage for future studies involving their characterization.

A mis padres, por su apoyo incondicional.

A mi hermana Marcela y mi hermano Carlos, por su cariño.

A mi hija Mila, por sus besos, abrazos y sonrisas.

ACKNOWLEDGEMENTS

This has been quite a journey. Life gave me the chance to become a better scientist and a better person at the same time, which made it exhausting at times. That's why, nothing presented here could have been possible without the incredible people who surrounded me, starting with my family, to whom I dedicated this dissertation. I was also very lucky to find a great mentor: Dr. Robert Hausinger, Bob. He gave me the freedom, guidance and support I needed to develop a project I believe in. Working with him was an immense privilege that I will treasure forever. Bob was also very good putting together a great group of people: Scott, our obligated first stop to learn about anything; Piotr and Tina, two very knowledgeable people that were always willing to help with suggestions and talk about life and food; Mukta, Kim, Jana, Meng, Thalia, Jodi, Erick and Nicholas, my lab mates whom I shared so much time, ideas, jokes, complaints, everything!; the best undergrads: Kimberly and Rachel, who helped me with the experiments and not to get bored in the lab; Andrea, Bruce and Aaron, whom I enjoyed getting to know.

I was also very lucky to belong to the biochemistry department in MSU, were the graduate secretaries Jessica and Julie became my friends; my committee members and other faculty who supported me: Dr. Christoph Benning, Dr. Charlie Hoogstraten, Dr. Leslie Kuhn, Dr. Dennis Arvidson, Dr. Michael Garavito, Dr. Michael Feig and Dr. Bill Wedemeyer; the program director, Dr. Jon Kaguni and the former chair, Dr. Shelagh Ferguson-Miller. Also my classmates James, Colleen, Cora, Chetan, Alec and Dean.

Finally my friends: Clarisa, Mariah, Andres, Karen and Soren, they made my life easier and funnier, which was sometimes a big challenge.

To each of you, a piece of my heart.

Thank you.

TABLE OF CONTENTS

Images in this dissertation are presented in color

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ACS	acetyl-coenzyme A synthase
ARD	aci-reductone dioxygenase
CD	circular dichroism
CH_3-S-CoM	methyl-*S*-coenzyme M
CoA-SH	coenzyme A
CoB-SH	coenzyme B, *N*-7-mercaptoheptanoylthreonine phosphate
CODH	carbon monoxide dehydrogenase
Co-FeSP	corrinoid-iron-sulfur protein
CoM-SH	2-thioethanesulfonate
CP	carbamoyl phosphate
Glx	glyoxylase
GSH	glutathione
GTP	guanosine triphosphate
MBP	maltose binding protein
NMR	nuclear magnetic resonance
ORF	open reading frame
PAR	4-(2-pyridylazo)-resorcinol
PDB	Protein Data Bank
SAXS	small-angle x-ray scattering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis.
SOD	superoxide dismutase.
$(UreABC)_3$	urease apoprotein

$(UreABC\text{-}UreD)_3$	complex of UreD bound to urease apoprotein
$(UreABC\text{-}UreDF)_3$	complex of UreD and UreF bound to urease apoprotein
$(UreABC\text{-}UreDFG)_3$	complex formed by UreD, UreF, and UreG bound to urease apoprotein

CHAPTER 1

Introduction

Portions of this Introduction were derived from "Chaperones of Nickel Metabolism", Chapter 14 of Nickel and Its Surprising Impact in Nature (2007) John Wiley & Sons, Ltd., written by
Soledad Quiroz, Jong K. Kim, Scott B. Mulrooney, and Robert P. Hausinger

INTRODUCTION TO NICKEL METABOLISM

Nickel is an essential micronutrient of many organisms where it serves as a cofactor for enzymes involved in several critical metabolic processes (*1, 2*). Like other transition metal ions, excess Ni is toxic to cells; thus, synthesis of these Ni-enzymes requires the presence of carefully controlled Ni-processing mechanisms that range from selective transport of Ni into the cells to productive insertion of Ni into the apoproteins. Various accessory proteins participate in these processes and are required for the biosynthesis of several Ni-dependent enzymes. Here, I provide an overview of the catalytic activity, biological role, and active site architecture of urease and briefly describe seven other structurally characterized Ni-dependent enzymes. In addition, I summarize what is known about activation of these Ni-enzymes and emphasize two particular accessory protein roles: metallochaperones that bind and deliver Ni to the apoprotein forms of the enzymes, and molecular chaperones that ensure productive conformations of the apoproteins for Ni incorporation.

1. General features of nickel incorporation into proteins

1.1. Transport of Ni into the Cell

The first required step for synthesis of any Ni-enzyme is for the cell to take up the metal ion from the environment in a regulated manner. Bacteria have developed two major types of high-affinity Ni transport systems for efficient Ni uptake (*3, 4*): ABC-type transporters and Ni-specific permeases.

The best-characterized ABC-type transporter system is that encoded by the *E. coli nikABCDE* operon (*5*), which is regulated by the product of a downstream gene, *nikR*. NikA is a periplasmic Ni-binding receptor protein with a Ni dissociation constant (K_d) of less than 0.1 μM (ten-fold lower than for Co, Cu, or Fe) (*6*). Two crystal structures of metal-bound NikA reveal pronounced differences at the metal-binding site: in one case, a penta-hydrated Ni ion is suggested to be bound (with a single polar interaction) within a large cavity of the protein (*7*), whereas the second study reported the binding of a monohydrated Fe-EDTA complex at the same site of the protein using many specific interactions between the chelator and the protein side chains (*8*). The long (2.7 Å) average Ni-O bond distance of the first crystal structure is inconsistent with spectroscopic data suggesting a much shorter distance (2.06 Å) (*9*), whereas the latter structural data more easily accommodate this distance. It is likely that a yet unidentified chelated form of Ni, rather than the ion itself, is the physiological species recognized by NikA. NikB and NikC are hydrophobic transmembrane proteins forming a pore for passage of the metal. NikD and NikE bind and hydrolyze ATP, and couple this energy release to the transport process. Ni homeostasis is achieved by use of NikR, a Ni-specific transcriptional repressor that binds to the NikR box in the promoter of the *nik* operon in the presence of Ni, resulting in suppression of Ni uptake (*10*).

The first Ni-specific permease (encoded by *hoxN*) was identified in *Ralstonia eutropha* (*11*). HoxN is an integral membrane protein containing eight membrane-spanning segments according to membrane topology analyses, and

is the prototype of a novel family of transition metal permeases. Transport assays showed that HoxN has a high affinity for Ni with a transport constant (K_t) of ~ 20 nM, but with very low capacity. HoxN homologues have been reported in many bacteria (e.g., HupN in *Bradyrhizobium japonicum*, NixA in *H. pylori*, NicT in *Mycobacterium tuberculosis*, and NhlF in *Rhodococcus rhodochrous*) and Nic1p in the fission yeast, *Schizosaccharomyces pombe* (*12*). The absence of HupN resulted in low levels of hydrogenase activity in *B. japonicum* under Ni-limiting conditions. Nic1p and NixA are essential for urease activity, and NixA exhibits a Ni K_t of 11 nM. NhlF was originally identified as a Co transporter in *R. rhodochrous* J1, but subsequent reinvestigation revealed that the permease transports both Ni and Co with a slight preference for Co. Regulation of the genes encoding these permeases is not well studied.

1.2. Additional Processing of Ni in the Cell

Once Ni enters the cell, it must be delivered and incorporated into the correct binding sites of Ni enzymes. This process may require metallochaperones, molecular chaperones, and a variety of other assembly steps.

The term metallochaperone refers to a protein that reversibly binds a metal ion, transports it within the cell, and provides it for metallocenter assembly to the target apoprotein. Molecular chaperones are proteins that prevent misfolding or assist in re-folding of other proteins, often by using energy derived from nucleotide hydrolysis. For example, the best studied molecular chaperones

are the GroES:GroEL chaperonin, Hsp70, and Hsp40 that act on many cellular proteins (*13, 14*). Recent studies have shown that SlyD, a Ni-binding protein, similarly exhibits a diverse molecular chaperone role (*15-18*). Such non-specific molecular chaperones are likely to stimulate the proper folding of many Ni-containing enzymes, as evidenced by the diminished hydrogenase activity in *groEL* or *groES* mutants and by the specific binding of GroEL to the HycE precursor protein (*19*). While these housekeeping proteins are rather non-specific in their action, this section focuses on molecular chaperones that are specific to individual Ni-enzyme activation systems. These proteins may drive the reaction by coupling metal insertion to nucleotide hydrolysis and/or they may use a metallochaperone rather than the free metal ion. In general, such proteins appear to be more essential to the activation processes than the metallochaperones, whose absence often can be overcome by excess Ni. A general scheme for metallochaperone and molecular chaperones function is presented in Figure 1.1.

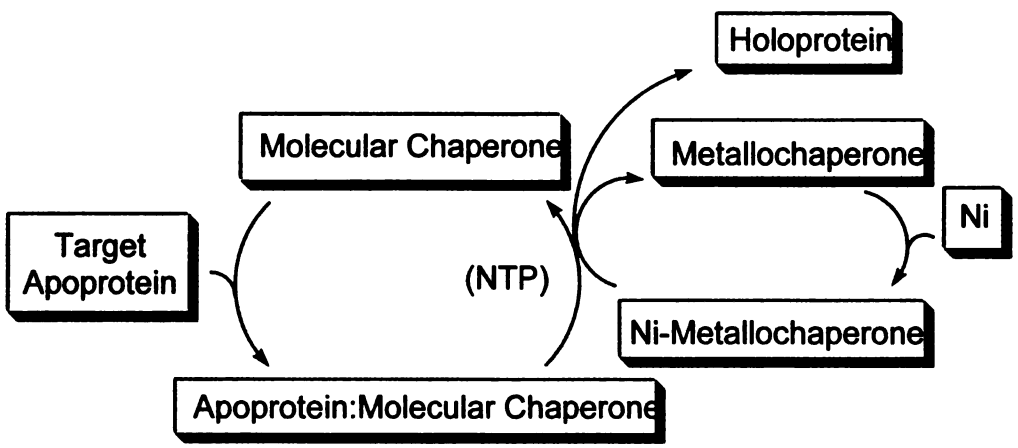


Figure 1.1. Generalized mechanism of metallochaperones and molecular chaperones. NTP, nucleotide triphosphate; required in some, but not all, molecular chaperones.

Accessory proteins can be involved in several other processes. For example, apoprotein proteolysis is associated with synthesis of hydrogenases and Ni-SOD. Cofactor synthesis is required prior to incorporation of the F_{430} tetrapyrrole into methyl coenzyme M reductase. Finally, many enzymes require the incorporation of another constituent prior to addition of Ni, such as the lysine carbamate of urease, the $Fe(CN)_2(CO)$ site of hydrogenase, and the iron-sulfur cluster components of CODH and ACS. In the following sections, I discuss the specifics of these processes for urease and then for selected other examples.

2. Nickel-containing enzymes and their activation process

2.1. Hydrogenase

Hydrogenases catalyze the reversible oxidation of molecular hydrogen into protons and electrons (Eq. 1.1). These enzymes provide a mechanism for many microorganisms to use H_2 as an energy source by generating a proton gradient or to remove excess reducing power in the form of molecular hydrogen (*20*).

$$H_2 \rightleftharpoons 2\,H^+ + 2\,e^- \qquad (1.1)$$

Three distinct classes of hydrogenases are defined by the metal content of their active sites: [NiFe]-hydrogenases, [Fe]-hydrogenases, and [iron-sulfur-cluster-free]-hydrogenases (*20, 21*). The crystal structures of several [NiFe]-hydrogenases have been resolved, including those of *Desulfovibrio gigas* and *Desulfomicrobium baculatum* (*22-24*). Each heterodimeric protein has three iron-sulfur clusters in its small subunit and a [NiFe] active site in its large subunit. The active center contains Ni coordinated by four Cys residues (or three Cys and a

selenocysteine in the *D. baculatum* enzyme), two of which bridge to the Fe that is also liganded by one carbon monoxide and two cyanide groups (Figure 1.2).

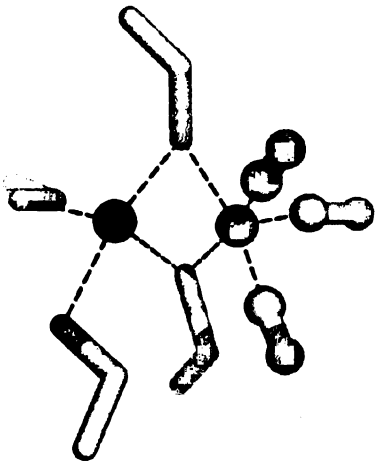


Figure 1.2. Dinuclear Ni-Fe active site of the [NiFe] hydrogenase from *Desulfovibrio baculatus* (PDB code 1CC1). The Ni is bound to a seleno-Cys and three Cys (or to four Cys in related enzymes), two of which also coordinate the Fe. The Fe-bound diatomic ligands are two cyanide and one carbon monoxide molecules.

2.2. Hydrogenase Activation

Seven accessory proteins are required to synthesize the *Escherichia coli* HycGE NiFe-hydrogenase (*25*), the paradigm system for defining the activation process of these enzymes (*26, 27*). These accessory proteins are the products of the six *hyp* genes (*hypABCDEF*) and another gene encoding a specific endopeptidase (*hycI*). The current model of HycE (large subunit) maturation includes a complicated series of steps involving (1) HypDEF-mediated formation of an $Fe(CN)_2(CO)$ site in a process facilitated by HypC (*28*); (2) insertion of Fe and its ligands into the precursor of the large subunit (retaining its C-terminal extension) when in complex with HypC (*29*); (3) GTP-dependent addition of Ni to

the active center mediated by HypAB; and (4) proteolytic processing of the C-terminus of HycE by HybD, leading to internalization of the catalytic center.

2.2.1. Metallochaperones: HypA, HypB, and SlyD

Of the many proteins involved in maturation of [NiFe] hydrogenases, three are known to directly bind Ni and may function as metallochaperones: HypA, HypB, and SlyD.

HypA designates the ~13-kDa protein required for activation of hydrogenase 3 of *E. coli* and the corresponding protein in many other hydrogenase systems. Homologues are termed HupA or HybF when referring to the protein used for *E. coli* hydrogenase systems 1 and 2. Purified HypA from *H. pylori* binds two Ni ions in a cooperative manner (*30*). Site-directed mutagenesis studies revealed a single His residue is required for binding Ni, and introduction of the corresponding His to Ala mutation resulted in substantial loss of hydrogenase activity (*30*). Subsequent investigations of *E. coli* HypA showed that it binds stoichiometric amounts of Ni and Zn, with μM and nM affinities, respectively (*31*). Based on UV/visible spectroscopic results indicating thiolate ligation, the bound Zn is proposed to have a structural role. Similar mutagenesis and binding studies of *E. coli* HybF found a single histidine is necessary for Ni binding, but mutagenesis of this residue resulted in a protein that retained the ability to bind Zn (*32*).

HypB, alternatively termed HupB in certain microorganisms, is a ~30-kDa protein that contains a nucleotide-binding motif and possesses low levels of

GTPase activity (*30, 33-36*). Site-directed mutations in the GTP-binding motifs result in elimination of hydrogenase activity (*34*). In addition to its GTPase activity, HypB proteins of selected organisms contain His-rich regions that are capable of binding several Ni ions. When this motif is deleted from *B. japonicum* HypB, the protein still binds one equivalent of Ni and retains competence in activating hydrogenase (*35, 37, 38*). The best-studied HypB is that from *E. coli*. This protein lacks a His-rich region, yet it tightly binds one Ni per monomer (K_d of 0.12 pM) using a CXXCGC motif at the N-terminus (*39*). Furthermore, the *E. coli* protein has a second metal-binding site located in the GTP-binding domain that has weaker affinity for either Ni or Zn ions (*39*).

In addition to the myriad studies of the individual HypA and HypB proteins and their homologues, there is significant evidence that the two proteins interact. Chemical cross-linking studies showed that a stable HypA-HypB complex is formed for these proteins (*30*) (*31*). In contrast to these results, chemical cross-linking studies carried out with a Strep-tagged variant of HybF failed to detect an association with HypB (*32*). Although the above studies of HypA and HypB proteins have greatly added to our knowledge, the question of how they function in Ni delivery and/or insertion into hydrogenase remains to be discovered.

SlyD is a ~21-kDa protein possessing an N-terminal region (146 amino acids with similarity to FK506-binding proteins) that contains peptidyl-prolyl cis/trans-isomerase activity and a short C-terminal metal-binding region rich in His, Asp, Glu, and Cys (*40, 41*). Evidence suggests that SlyD may act as a chaperone for several proteins (*15-18*). More pertinent to this discussion, SlyD

reversibly binds Ni in its C-terminal region such that the peptidyl/prolyl isomerase activity is inhibited (*41*). The ability of SlyD to bind Ni was suggested by its co-purification with several recombinant His-tagged proteins when using Ni-nitrilotriacetic acid affinity chromatography (*17, 42-44*). Of particular interest with regard to maturation of the [NiFe] hydrogenases, *E. coli* SlyD was shown to interact with HypB from the same source (*45*). Furthermore, cells deleted in *slyD* have greatly reduced activity of all three hydrogenases and reduced intracellular concentrations of Ni. These findings suggest that SlyD has a role in the Ni insertion step of hydrogenase activation (*45*).

2.2.2. Hydrogenase Molecular Chaperones: HypC and HypB

The complex biosynthetic pathway of *E. coli* hydrogenase 3 includes two molecular chaperone-like proteins: HypC and HypB. *E. coli* contains a second HypC-like protein, termed HybG, which specifically binds to hydrogenase 2 (both HypC and HybG bind to hydrogenase 1, but only the former facilitates activation) (*46*). Homologues to HypC and HypB are found in many other organisms where they are sometimes given alternative designations (e.g. HupB and HupC). HypC (or its homolog) plays a central and multifaceted role in [NiFe]-hydrogenase biosynthesis. HypC forms a complex with HypD, especially when the synthesis of CP is reduced so that the formation of the $Fe(CN)_2(CO)$ center is hindered (*29*). The HypC-HypD species additionally binds HypE in such a manner as to allow its carbamoylation by HypF (*28*). HypD dissociates from HypC as a new complex, HypC-HycE, is formed (*47, 48*) containing the $Fe(CN)_2(CO)$ center, formed in the

earlier complex, but still lacking Ni (*47*). The interaction between the chaperone and the large subunit precursor requires the N-terminal Cys residue of HypC and a particular large subunit Cys residue, which eventually coordinates Ni in the active site (*49*). Ni is provided by the action of HypA/HypB metallochaperone (*50*), with HypB additionally having a molecular chaperone role. After the complete set of metallocenter components is in place, HypC dissociates, HycI binds to the HycE-bound Ni and becomes proteolytically active (*51*). The large subunit extension is removed, resulting in the [NiFe] site becoming buried in the protein.

Nickel insertion into the $Fe(CN)_2(CO)$-containing and HypC-bound hydrogenase precursor requires the HypA/HypB metallochaperone, with HypB also exhibiting a GTP-dependent molecular chaperone role (*34*). HypB is homologous to the urease accessory protein UreG and likewise contains a nucleotide-binding motif. Native HypB exhibits weak GTPase activity, but a mutation affecting the P-loop eliminates the GTPase activity and greatly decreases the hydrogenase activity (*30, 52*). The HypB proteins of some organisms contain His-rich termini that are able to bind Ni; removal of this sequence has only a small effect on hydrogenase activity while having a larger affect on cellular Ni content. The HypB molecular chaperone appears to drive Ni insertion into the hydrogenase subunit by coupling this reaction to GTP hydrolysis (*30, 33-36*).

2.3. Carbon Monoxide Dehydrogenase (CODH)

CODHs catalyze the reversible oxidation of carbon monoxide to carbon dioxide (Eq. 1.2). Organisms possessing these enzymes play critical roles in the global carbon cycle and the degradation of environmental pollutants (*53*).

$$CO + H_2O \rightleftharpoons CO_2 + 2H^+ + 2e^- \qquad (1.2)$$

Crystal structures are known for CODHs from *Carboxydothermus hydrogenoformans* and *Rhodospirillum rubrum* (*54, 55*). Both proteins are ~ 130-kDa homodimers containing five metal-sulfur clusters of three types (B, C, and D) in a C-B'-D-B-C' arrangement where the D cluster bridges the two subunits. While the B, B' and D sites are the same cubane type [4Fe-4S] clusters in both proteins, the structures of the active site clusters (C and C') slightly differ in the two proteins. The C cluster of *R. rubrum* is essentially a [1Ni-3Fe-4S] cubane bridged to a mononuclear Fe site, whereas the structure of the *C. hydrogenoformans* enzyme can be viewed as a [3Fe-4S] cluster fused with a [Ni-S-Fe] fragment containing a bridging sulfide (Figure 1.3A).

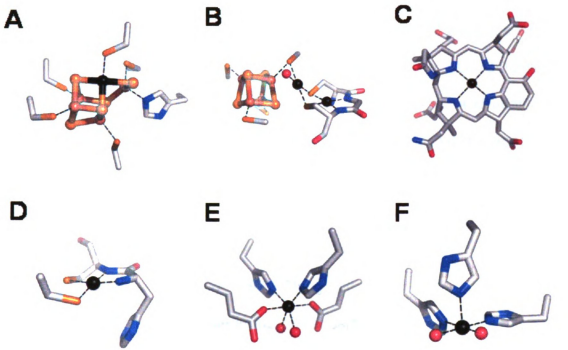


Figure 1.3. Active sites of structurally-characterized Ni-containing enzymes. In each case, Ni is a solid black sphere, nitrogen atoms are blue, sulfur orange, oxygen red, and carbon white. A. [Ni-Fe_4-S_5] cluster of *Carboxydothermus hydrogenoformans* CODH (PDB code 1SU8). The structure of this cluster slightly varies in other CODH sites. B. [4Fe-4S]-Ni-Ni site of *Carboxydothermus hydrogenoformans* ACS (PDB code 1RU3). The fourth ligand on the central Ni is water. C. F_{430} Ni-tetrapyrrole of *Methanobacterium thermoautotrophicum* methyl coenzyme M reductase (PDB code 1MRO). D. The active site of *Streptomyces coelicolor* Ni-SOD (PDB code 1T6U). The imidazole nitrogen of His1 is a ligand in the active enzyme, when the Ni is oxidized. E. *E. coli* Ni-glyoxylase showing two bound water molecules (PDB code 1F9Z). His5 and Glu56 are derived from one subunit and His74 and Glu122 from the second subunit in the symmetric dimer. The two water molecules are displaced by substrate. F. Ni-containing form of ARD from *Klebsiella oxytoca* as derived by a combination of solution structure analysis and homology modeling (PDB code 1M4O). The non-side chain ligands of the metal are water molecules.

2.4. Acetyl-CoA Synthase/CODH (ACS/CODH)

The CODH activity described above is found in another set of enzymes isolated from acetogenic bacteria and methanogenic archaea. The ACS/CODHs are bifunctional catalysts that exhibit the activity shown in Eq. 1.2 and additionally synthesize (or decompose) acetyl-coenzyme A (CoA-SH) using the remarkable chemistry shown in Eq. 1.3. The CODH site of ACS/CODH reduces CO_2 to CO and then this gaseous molecule traverses a molecular tunnel within the protein to reach the ACS site where it is joined to CoA-SH and the methyl group from the corrinoid-iron-sulfur protein (Co-FeSP). Along with the monofunctional CODHs, these enzymes play a major role in the global carbon cycle and in the formation and removal of greenhouse gases (*56*).

$$CO + CoA\text{-}SH + CH_3\text{-}Co(III)\text{-}FeSP \rightleftharpoons CH_3C(O)\text{-}S\text{-}CoA + Co(I)\text{-}FeSP \quad (1.3)$$

Crystallographic studies of *Moorella thermoacetica* ACS/CODH revealed that the tetrameric protein contains the dimeric CODH subunits at its core and one ACS subunit on each end (*57, 58*). The ACS metallocenter is a [4Fe-4S]-Ni-Ni site called the A-cluster. The [4Fe-4S] cluster is bridged to one Ni via a Cys side chain, and this metal is in turn bridged by two Cys residues to a second Ni, that is additionally bound by two backbone amides. The central Ni in the A-cluster is subject to metal substitution, resulting in inactive Cu-Ni and Zn-Ni species that were critical to identifying closed and open conformations of the protein. The [4Fe-4S]-Ni-Ni cluster (Figure 1.3B) also was observed in the structure of the monomeric *C. hydrogenoformans* ACS (*59*).

2.5. CODH and ACS Activation

Information regarding the mechanism of Ni insertion into CODH is available for *R. rubrum* where the *cooCTJ* gene cluster (*60*), located downstream of the *cooS* structural gene, is known to be involved. The CooC protein, which contains a nucleotide-binding motif, acts as an ATP/GTP-dependent molecular chaperone, while CooJ delivers Ni by using its histidine-rich C-terminal motif.

Little is known about the mechanism concerning metallocenter assembly in ACS/CODHs. Since the enzyme has two different sets of Ni-containing active sites, it is anticipated that several accessory proteins are required for biosynthesis. Consistent with this notion, ACS/CODH gene clusters contain several non-subunit open reading frames (ORFs). In particular, AcsF encodes a CooC-like protein that is further described below.

2.5.1. CODH Metallochaperone: CooJ

The *R. rubrum* protein CooJ contains 115 residues of which 16 of the C-terminal 34 amino acids are His (*61*), an arrangement similar to the His-rich regions in sequences of some HypB and UreE proteins. Cells containing a chromosomal deletion that eliminates the His-rich region display Ni-dependence for growth on CO that is identical to wild-type strain, while cells with an insertional mutation of *cooJ* required 1000-fold higher Ni than wild type for optimal growth (*60*). These findings suggest that the His-rich C-terminal region is not required for CooJ function, and the functional Ni-binding site lies elsewhere in the protein.

2.5.2. CODH and ACS Molecular Chaperones: CooC/AcsF

The CODH operon of *R. rubrum* encodes the suspected molecular chaperone CooC. This membrane-bound homodimer of 62 kDa is related in sequence to UreG of urease activation and HypB of hydrogenase biosynthesis (*60*). CooC contains a P-loop motif in its N-terminus, and the purified protein hydrolyzes both GTP and ATP with similar K_m values, but with a 10-fold greater V_{max} for ATP. Mutation of residues in the P-loop motif prevents Ni insertion into CODH and abolishes the ATPase activity, both *in vivo* and *in vitro*. Ni is not present in the purified protein (*62*).

The gene cluster encoding the ACS/CODH bifunctional enzyme of *M. thermoacetica* includes a gene encoding AscF that resembles CooC and the other nucleotide-dependent molecular chaperones described above. This protein contains a P-loop and has five conserved Cys in a motif characteristic of iron coordination. Despite the suspected importance of this gene for ACS/CODH activation, its deletion had no effect on enzyme activity. It remains possible that a second copy of the gene is present in the genome (*63*).

2.6. Methyl Coenzyme M Reductase

Methyl coenzyme M reductase catalyzes the reaction of methyl-*S*-coenzyme M (CH_3-S-CoM, where CoM-SH is 2-thioethanesulfonate) with coenzyme B (CoB-SH, *N*-7-mercaptoheptanoylthreonine phosphate) to form methane and the heterodisulfide, CoM-S-S-CoB (Eq. 1.4). This is the final step of

methane formation in methanogenic archaea growing on simple molecules such as acetate, methanol, formate, and carbon dioxide plus hydrogen gas (*64*).

$$CH_3\text{-S-CoM} + \text{CoB-SH} \rightarrow CH_4 + \text{CoB-S-S-CoM} \qquad (1.4)$$

The X-ray crystal structure of methyl coenzyme M reductase, first obtained from *Methanothermobacter marburgensis* (*65*), reveal that the protein is a 300-kDa heterohexamer of three different subunits ($\alpha_2\beta_2\gamma_2$) containing two molecules of the Ni-containing tetrapyrrole, F_{430} (Figure 1.3C). This cofactor, named on the basis of its characteristic absorbance maximum at 430 nm when in the Ni(II) state, must be in the Ni(I) state for the enzyme to be active. Each active site F_{430} is buried deep in the protein and accessible from the surface by a 50 Å long channel composed of mainly hydrophobic amino acids through which CH_3-S-CoM can enter, and which is blocked by the binding of CoB-SH. An interesting aspect of this enzyme is the presence of five post-translationally modified amino acids near the active site: thio-Gly, *N*-methyl-His, *S*-methyl-Cys, 5-methyl-Arg, and 2-methyl-Gln. Labeling studies have shown that the methyl groups are derived by methyl group transfer from *S*-adenosylmethionine, and not from the methyl group of CH_3-*S*-CoM.

An enzyme of related interest is found in methanotrophic archaea (*66*), such as those located in microbial mats that catalyze the anaerobic oxidation of methane (*67*). These prokaryotes, closely related to methanogens in the order *Methanosarcinales*, contain homologues of genes encoding methyl coenzyme M reductase (*68*) and possess an F_{430}-like molecule with a 46 Da mass increase

(*67*). The mechanism by which these microbes essentially reverse the last step of methanogenesis remains unclear.

2.7. Methyl Coenzyme M Reductase Activation

The biosynthetic pathway of F_{430} is an offshoot of those for other biological tetrapyrroles (*69*). Early labeling studies demonstrated that F_{430} is derived from dihydrosirohydrochlorin, which is also the precursor of siroheme and corrinoids. The dihydrosirohydrochlorin is formed from 5-aminolevulinic acid via uroporphyrinogen III. The conversion of dihydrosirohydrochlorin to F_{430} requires several steps including amidation of acetate groups on two rings, reduction of two double bonds, cyclization of an acetamide to form the five-membered ring, cyclization of a propionic acid to form the six-membered ring, and insertion of Ni. However, the order of these steps and the mechanism underlying the Ni insertion and F_{430} incorporation into the protein remain unknown.

2.8. Superoxide Dismutase (SOD)

SODs are ubiquitous metalloenzymes that function to protect biological molecules from oxidative damage by catalyzing the dismutation of superoxide anion radicals to peroxide and molecular oxygen (Eq. 1.5). In addition to the well-known Cu,Zn-, Fe-, and Mn-containing SODs, recent studies have revealed the existence of Ni-SODs in *Streptomyces* species and some cyanobacteria.

$$2\ O_2^- + 2\ H^+ \rightarrow H_2O_2 + O_2 \qquad (1.5)$$

Crystal structures of Ni-SODs have been solved for *S. seoulensis* and *S. coelicolor* enzymes (*70, 71*). The proteins are homohexamers consisting of four-helix bundle subunits. The N-terminal loop coordinates the active site Ni(III) in square pyramidal geometry using two thiolate side chains (Cys-2 and Cys-6), two backbone amides (His-1 and Cys-2), and the His-1 side chain ligand at the axial position. The axial ligand is lost in the reduced state, with Ni(II) becoming square planar (Figure 1.3D). Apoprotein structures show that the residues involved in binding Ni are disordered.

Ni-SODs in *Streptomyces* species are products of *sodN*, which encodes a preprotein with an N-terminal extension of 14 amino acids. During SOD maturation, proteolytic cleavage precedes Ni binding and results in the creation of the six-residue Ni-binding site. Recently, ORFs with significant similarity to Ni-SODs were identified in the genomes of several cyanobacteria including *Prochlorococcus marinus* MIT9313 (*72*). In this microbe, an ORF located downstream of *sodN* and named *sodX* was suggested to be the peptidase for maturation of the Ni-SOD. Coexpression of *sodX* and *sodN* in an oxygen-sensitive *E. coli* strain restored oxygen tolerance in a Ni-dependent manner, indicating the production of a catalytically active enzyme and providing confirmatory evidence for the importance of SodX in Ni-SOD maturation. Ni-SOD activity in *S. seoulensis* is stimulated by the overproduction of CbiXhp, a Ni-binding protein, suggesting that it too many function in metallocenter assembly (*73*). Contrary to this notion, the gene encoding CbiXhp is located between two genes suggested to function in cobalamin biosynthesis. Further studies are

needed to elucidate the detailed maturation steps of Ni-SOD activation, including the mechanism of Ni incorporation to the enzyme.

2.9. Glyoxylase

Glyoxylase I is the first of two enzymes in the pathway to convert cytotoxic methylglyoxal into non-toxic α-hydroxycarboxylic acids. It converts the hemimercaptal substrate, formed nonenzymatically from methylglyoxal and glutathione (GSH, Eq. 1.6), to non-toxic *S*-D-lactoylglutathione (Eq. 1.7), which is the substrate for Glyoxylase II (Eq. 1.8). These enzymes are important for cellular protection because methylglyoxal can exert toxic effects by reacting with DNA, RNA and proteins.

$$CH_3\text{-}CO\text{-}CHO + GSH \rightleftharpoons CH_3\text{-}CO\text{-}C(OH)\text{-}SG \quad (1.6)$$

$$CH_3\text{-}CO\text{-}C(OH)\text{-}SG \rightarrow CH_3\text{-}CH(OH)\text{-}CO\text{-}SG \quad (1.7)$$

$$CH_3\text{-}CH(OH)\text{-}CO\text{-}SG + H_2O \rightarrow CH_3\text{-}CH(OH)\text{-}COOH + GSH \quad (1.8)$$

Unlike the case for glyoxylase I of humans, *Saccharomyces cerevisiae*, and *Pseudomonas putida*, where the active site metal is zinc, glyoxylase I from *E. coli* is completely inactive in the presence of Zn and is maximally active with Ni (*74*). Reduced activity is found in the enzyme substituted with Co, Cd, and Mn. Crystallographic analyses revealed that catalytically active forms of *E. coli* glyoxylase I with Ni, Co, and Cd each have octahedral metal coordination (Figure 1.3E), which is also observed in the Zn-containing human enzyme, whereas the inactive Zn-containing *E. coli* protein displays a five-coordinate metal site (*75*). Several other pathogenic microorganisms are hypothesized to possess a Ni-

containing glyoxylase on the basis of sequence comparisons (*76*). The cellular mechanism of Ni incorporation into glyoxylase I is unknown.

2.10. Aci-Reductone Dioxygenase (ARD)

Many microorganisms utilize the methionine salvage pathway to regenerate methionine from methylthioadenosine, produced during polyamine biosynthesis from *S*-adenosylmethionine. Aci-reductone is a key intermediate of this pathway, and is oxidized to two different sets of products in *Klebsiella pneumoniae*. One oxidation pathway leads to production of formate and the ketoacid precursor of methionine. The other route of oxidation, a non-productive pathway, converts the aci-reductone to formate, carbon monoxide, and methylthiobutyric acid. Remarkably, the two reactions are carried out by the same enzyme, ARD, depending on which metal is bound at the active site (Fe or Ni, respectively) (*77*).

The solution structure of *K. pneumoniae* Ni-ARD was determined by NMR methods (*78*). The enzyme is a monomer containing two β-sheets that hinge together to form a jellyroll. Unfortunately, paramagnetism of the bound Ni causes broadening of the ^{1}H resonance lines from residues near the metal center, thus hindering the structural characterization of the active site. Biophysical studies suggest the presence of three His ligands to the Ni, along with three other nitrogen or oxygen atoms (*79*). Homology modeling of the active center, based on the structure of jack bean canavalin (another member of the cupin family), provides a reasonable model of the active site (Figure 1.3F). The mechanism of Ni insertion into the enzyme is unknown.

2.11. Other Potential Metallochaperones

Several other proteins possess His-rich regions and/or tightly bind Ni; however, none of these has been convincingly shown to facilitate Ni metallocenter assembly. For example, CbiXhp of *S. seoulensis* contains a carboxyl terminus in which 11 of 19 residues are His, and the cellular overproduction of this protein stimulated Ni-SOD activity (*73*). On the other hand, this protein is more likely to be involved in cobalamin biosynthesis on the basis of flanking genes and its presence in cells that lack a Ni-SOD. The Hpn protein of *H. pylori* is worth a few comments because of the high levels of Ni-containing urease and the important role of hydrogenase in this microorganism (*80, 81*). Hpn is comprised of only 60 amino acids, 28 of which are His (*82*). Deletion of the corresponding gene has no effect on urease activity for cells grown on blood agar, but does make the cells more susceptible to growth inhibition at high Ni concentrations (*83*). Recent studies have shown Hpn binds 5 Ni per monomer (K_d 7.1 μM) and provided evidence that the concentrations of this protein correlate to the intracellular Ni concentration and to the cell's ability to tolerate high Ni concentrations (*84*). The authors suggested that Hpn may function in Ni storage, Ni donation, and Ni detoxification. It will be of interest to learn whether follow-up studies confirm the putative metallochaperone role for this protein and to monitor whether Ni metallochaperones for other enzyme systems are identified.

3. Urease

Urease catalyzes the hydrolysis of urea to produce ammonia and carbamate. The latter molecule spontaneously decomposes to yield another molecule of ammonia and carbonic acid (Eqs. 1.9 and 1.10). This enzyme, found in plants, fungi and bacteria, has several biological roles including its participation in recycling of environmental nitrogen and its use as a virulence factor in pathogenic microorganisms that are associated with gastric ulceration and urinary stone formation (*85*).

$$H_2N\text{-}CO\text{-}NH_2 + H_2O \rightarrow H_2N\text{-}COOH + NH_3 \qquad (1.9)$$

$$H_2N\text{-}COOH + H_2O \rightarrow H_2CO_3 + NH_3 \qquad (1.10)$$

Crystallographic analyses have revealed that most bacterial ureases possess three structural subunits (encoded by *ureA*, *ureB*, and *ureC*) associated into a trimer of trimers [$(UreABC)_3$] (Figure 1.4, left), with each UreC subunit containing a dinuclear Ni active site bridged by a carbamylated Lys residue (*86-88*) (Figure 1.4, right). Some species, such as *Helicobacter pylori*, have only two subunits (UreA, corresponding to a fusion of the small subunits in other bacteria, and the large subunit, labeled UreB) in a $(UreA_3UreB_3)_4$ supramolecular structure (*89*). Plants and fungi have a single subunit corresponding to a fusion of all of the bacterial subunits, and form an Ure_6 urease structure (*90*) that resembles a dimer of the bacterial enzyme.

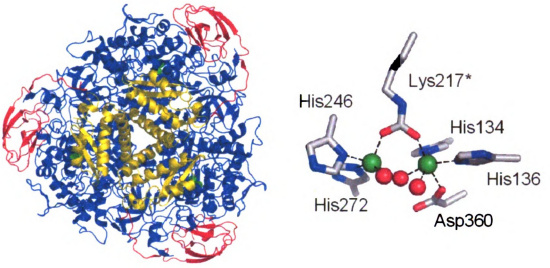


Figure 1.4. Left: Crystal structure of *K. aerogenes* urease. UreA is colored in yellow, UreB in red and UreC in blue. The green spheres are the nickel ions at the active site. Right: Dinuclear Ni-Ni active site of urease (PDB code 1FWJ). The metal-bridging side chain is a carbamylated Lys and the three red spheres coordinated to the metals are water molecules.

3.1. Urease mechanism

The breakdown of urea requires access of the molecule into the buried active site. It has been suggested that a conformational change in the flap formed by residues 308-336 of UreC could allow easy access of the substrate to the active site (*86, 91*). Two mechanisms have been proposed for urease, with both initiating by urea displacement of a water molecule coordinated to Ni1 (the one bound by His246 and His272). Then in one case, a water molecule bound to Ni2 attacks urea to form a tetrahedral intermediate and His320 protonates the amido nitrogen, promoting the formation of ammonia and carbamate, the final products of the reaction. The alternative mechanism suggests that the oxygen from urea coordinates to Ni1 and one amido group to Ni2; then, the bridging water attacks

the carbonyl group to form a tetrahedral intermediate and donates the proton to the distal amido group, with Asp360 assisting the protonation (*92*). Until now, there is no evidence to clarify the origin of the water molecule attacking urea and distinguishing whether the amido group binds to Ni2.

3.2. Urease activation

The urease gene cluster of most bacteria is composed of both structural genes (*ureABC*) and accessory genes (typically including *ureDEFG*, with additional urease-related genes present in some species). The structural gene products assemble into an apoprotein that requires activation by the accessory proteins. The best-studied urease activation system is that found in *Klebsiella aerogenes*, which contains the *ureDABCEFG* gene cluster (*93, 94*). Using this system, UreD, UreF, and UreG were identified as forming a GTP-dependent molecular chaperone that binds urease apoprotein (*95*), while UreE was shown to function as a metallochaperone that delivers Ni (*96, 97*). A scheme that shows this process is depicted in figure 1.5.

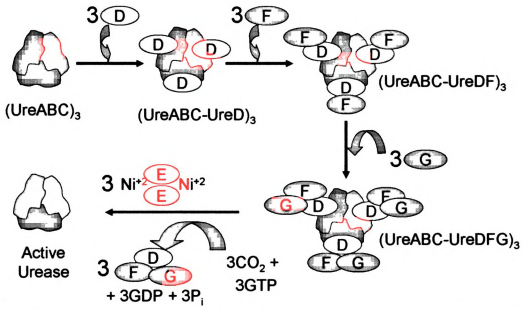


Figure 1.5. Proposed urease activation process. The *K. aerogenes* UreA, UreB and UreC urease subunits assemble into the $(UreABC)_3$ apoprotein (depicted simply as a trimeric species since UreA plus UreB or all three subunits are fused together in ureases from some sources). UreD, UreF and UreG sequentially bind to form the $(UreABC\text{-}UreD)_3$, $(UreABC\text{-}UreDF)_3$, and $(UreABC\text{-}UreDFG)_3$ activation complexes. CO_2 adds to the active site Lys as Ni^{+2} ions are delivered to $(UreABC\text{-}UreDFG)_3$ by the dimeric UreE metallochaperone in a process that requires GTP hydrolysis, with UreE and $(UreDFG)_3$ being released from the activated urease.

3.2.1. Urease Metallochaperone: UreE

Among the multiple accessory genes required for urease activation in most urealytic organisms, UreE appears to function as a metallochaperone that delivers Ni to the urease apoprotein. The first suggestion that UreE functions as a Ni-binding protein was provided by the sequences of the *K. aerogenes* and *Proteus mirabilis* urease operons (*93, 98*). The carboxyl termini of these proteins contain His-rich regions consisting of 10 His in the last 15 residues for *K. aerogenes* and 9 of the last 10 residues in the case of *P. mirabilis*, indicative of a potential metal binding role. Subsequent sequences of *ureE* genes from other

sources reveal that the His-rich C-terminal region is common, but it is completely absent in some organisms (*99*). Equilibrium dialysis studies of *K. aerogenes* UreE showed that about 6 Ni bind per dimeric protein (*100*), while metal-binding studies of *Bacillus pasteurii* UreE, which contains only two conserved His residues in this region, found a single Ni bound per dimer (*101*). Purified UreE proteins also bind other metal ions, such as Cu and Zn, demonstrating that the specificity of urease for Ni does not reside solely with this delivery protein (*102*). Using site-directed mutation methods to create a truncated form of *K. aerogenes* UreE with the last 15 amino acids removed (His144*UreE), the His-rich region was demonstrated to be non-essential; i.e., the truncated protein still binds 2-3 Ni ions per dimer and is still competent in facilitating Ni-dependent activation of urease *in vivo* (*103, 104*). In a complementary study, the native *H. pylori ureE* gene, which does not encode a protein with a His-rich C-terminal tail, was fused to several extensions to produce different His-rich regions (*105*). The resulting His-rich variants show increased Ni-binding and cells containing these variants have increased urease levels; thus, the C-terminal His-rich region has a Ni-sequestering function that aids in urease activation.

Several lines of evidence indicate that UreE interacts with other accessory proteins during Ni-dependent activation of urease. *In vitro* studies showed that a complex of urease apoprotein with UreD, UreF, and UreG is fully activated only by including UreE in a mixture containing GTP, bicarbonate, and Ni (*97*), thus providing strong evidence that UreE functions as a metallochaperone to deliver Ni to the UreDFG-urease apoprotein complex. These studies also showed that

UreE does not simply function as a reversible Ni-binding protein because activation occurred even when metal ion chelators were included in the reaction (*97*). Additional work involving yeast two-hybrid analysis demonstrated an interaction between *H. pylori* UreE and UreG proteins (*106, 107*).

Site-directed mutagenesis and structural studies provided detailed insights into the metal-binding properties of UreE. Variants of *K. aerogenes* His144*UreE affecting His-110 or His-112 exhibit reduced Ni binding while not greatly affecting urease activation, whereas a variant affecting His-96 binds less Ni and abolishes UreE's capacity to activate urease (*104*). These results are easily rationalized by the crystal structures of Cu-bound *K. aerogenes* H144*UreE (*108*) and Zn-bound *B. pasteurii* UreE (*109*). (Figure 1.6) The overall structures are nearly identical, but contain three and one metal-binding sites, respectively. Both proteins bind a metal at the dimer interface using the symmetric pair of critical His-96 residues in the *K. aerogenes* protein or the pair of His-100 residues for *B. pasteurii* UreE. This metal site is essential for UreE's function in urease activation. In addition, each subunit of *K. aerogenes* UreE binds a metal at sites involving His-110 and His-112, residues that are substituted with other side chains in the *B. pasteurii* protein.

The UreE crystal structures also reveal the presence of two distinct domains in the proteins. The metal-binding domains, located in the C-terminal half of each molecule, resemble the structure of the yeast copper metallochaperone Atx1 (*110*). The N-terminal domains have structural similarities to a domain of the yeast Hsp40 molecular chaperone Sis1 (*111*), suggesting that

this domain may be involved in molecular recognition and binding to other urease accessory proteins and/or urease apoprotein. Arguing against this conclusion are results from studies involving a construct that produced only the metal-binding domain of *K. aerogenes* UreE (residues 70-143); the single domain of UreE is capable of delivering Ni to the urease apoprotein and the N-terminal domain is not required (*112*).

Complicating the *B. pasteurii* UreE structure described above, the crystallization conditions promote oligomerization of the protein to form a dimer of dimers $(\alpha_2)_2$ in which all four His100 side chains serve as ligands to a single Zn (*109*). The tendency of *B. pasteurii* UreE to aggregate was further examined by protein NMR and equilibrium dialysis approaches (*113*). Those studies showed that the tetramer form is favored only at high protein concentrations and provided evidence for a second Ni-binding site in the C-terminal region of the dimeric form of UreE, which probably binds a total of 3 Ni ions.

A final comment about urease metallochaperones focuses on the situation in *H. pylori* where HypA and HypB, normally associated with hydrogenase activation, are also required for urease activity. Deletions of either gene encoding these proteins results in cells with very low urease activity; however, the urease activity can be restored by addition of excess Ni (*30, 52*). Thus, it is possible that HypA, HypB, or a complex of these proteins function as a metallochaperone and assists in urease activation. These proteins are discussed further below.

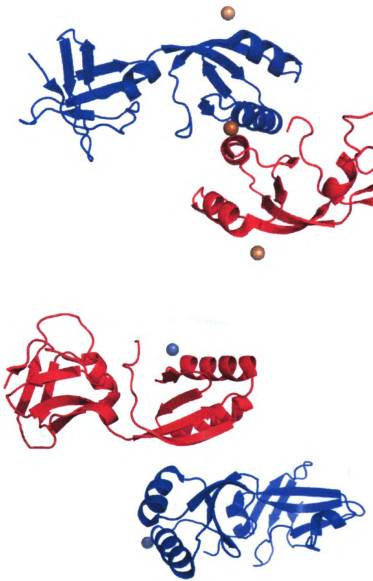


Figure 1.6. *K. aerogenes* UreE (upper structure, PDB Code 1GMU) and *B. pasteurii* UreE (PDB Code 1EAR). The three copper ions coordinated by *K. aerogenes* UreE are shown as brown spheres and a zinc ion is shown in gray at the center of the *B. subtilis* dimer.

3.2.2. Urease Molecular Chaperones: UreD, UreF and UreG

Structural studies have revealed that the Ni active site of urease is buried within the enzyme (*86*), and this site is also relatively inaccessible in the apoprotein (*114*). These results are consistent with the need for one or more urease-specific molecular chaperone(s) to alter the urease protein conformation and allow Ni to gain access to the active site. From studies with the *K. aerogenes* urease system, three proteins, each of which is required for *in vivo* enzyme activation, are proposed (*1, 94*) to act together to fulfill this role: UreD, UreF and UreG.

As expected of molecular chaperone proteins, each of the UreD, UreF, and UreG accessory proteins are found in complexes that include the urease apoprotein. Thus, UreD-, UreDF-, and UreDFG-urease apoprotein complexes have been described (*115-117*). These complexes possess distinct properties when compared to those of the urease apoprotein alone, especially with regard to their activation properties. Approximately 15% of the apoprotein is activated *in vitro* by addition of 100 μM Ni and 100 mM bicarbonate (needed to carbamylate the active site Lys) (*118, 119*). In contrast, about 30% of the UreD-urease apoprotein is activated by these conditions, demonstrating that UreD directly enhances this process (*115*). Furthermore, the UreDF-urease apoprotein is activated to the same extent by using nearly 1000-fold lower concentrations of bicarbonate, and activation of this complex is resistant to the detrimental effects of high concentrations of Ni compared to the apoprotein alone or to UreD-urease apoprotein (*116*). A UreDFG-urease apoprotein complex forms upon addition of

UreG to UreDF-urease apoprotein (*95*) and is normally present in cells expressing the intact *K. aerogenes* urease gene cluster (*117*). Significantly, this species exhibits GTP-dependent urease activation—shown by mutagenesis studies to be associated with the nucleotide-binding (P-loop) motif located within UreG (*95*). Urease activation is not achieved with a non-hydrolyzable analog of GTP. When the UreE metallochaperone is provided to this complex along with GTP and near-physiological levels of Ni plus bicarbonate, fully active urease is generated (*97*). Three of the urease apoprotein species were probed by a chemical cross-linking/proteolysis/mass spectrometric approach to examine the sites of binding of UreD and UreF to urease (*120*). Additional evidence from these studies suggests that UreF interacts with UreD-urease apoprotein to give rise to a conformational change within urease that may enhance access of Ni to its ligand-binding residues. Finally, we note that a UreDFG complex is generated in the absence of the structural proteins, and this species was enriched by binding to an ATP-linked agarose resin (*121*). Further structural and mechanistic studies are required to better define the individual roles of UreD, UreF, and UreG within the heterotrimeric molecular chaperone that couples GTP hydrolysis to urease activation.

In addition to participating as a component of the urease molecular chaperone, UreG has been studied in its purified form. The *K. aerogenes* protein is a monomer of 21.9-kDa that, despite having a P-loop motif, fails to bind or hydrolyze GTP (*121*). In contrast, UreG from *B. pasteurii*, which is over 50% identical to the *K. aerogenes* protein, is dimeric and has weak GTPase activity

(*122*). NMR studies of the *B. pasteurii* protein suggest that it is intrinsically disordered; however, this UreG was purified from a heterologous overproduction system and required urea dissolution of inclusion bodies, so the disorder may be artifactual. Significantly, the purified *B. pasteurii* UreG binds 2 Zn per dimer (K_d 42 μM) or binds 4 Ni per dimer with weak affinity (K_d 360 μM).

Two other potential molecular chaperones for urease exist in *H. pylori*. First, is the GroES-homolog called HspA (*123*). This heat shock protein possesses an N-terminal domain resembling the broad specificity molecular chaperones, but the protein additionally contains a 27 amino acid extension that is rich in His. Not unexpectedly, HspA binds 2 Ni with high specificity. Significantly, when *hspA* is co-expressed with the *H. pylori* urease genes in *E. coli*, urease activity is enhanced 4-fold in accordance with a possible urease-specific function (*123*). The second example of a possible *H. pylori*-specific molecular chaperone involves the typical hydrogenase accessory protein HypB which is required for both urease and hydrogenase activity in this organism. Since HypB is suggested to function as a molecular chaperone of hydrogenase, it may also serve this role in urease activation in *H. pylori*. Alternatively, HypB along with HypA may possess a metallochaperone role in urease activation in this microorganism.

CONCLUSIONS AND REMAINING QUESTIONS

Accessory proteins are clearly shown to play critical roles in the biosynthesis of several Ni-containing enzymes. Metallochaperones often are used to deliver Ni to the target apoprotein and nucleotide-dependent molecular chaperones can have a role in altering the apoprotein conformation to allow access by the metal ion. Indeed, the two functions are usually closely linked. Even when such proteins have been identified, their precise mechanisms of action remain unclear. For example, it is unknown how Ni becomes bound to metallochaperones and how the Ni is delivered to the target protein. Similarly, the mechanism by which nucleotide hydrolysis is coupled to Ni insertion remains a mystery. Furthermore, these two themes are not universal among all Ni enzymes, as exemplified by Ni-dependent glyoxylase and aci-reductone dioxygenase, for which no accessory proteins have been observed; on the other hand, it remains possible that future investigations will reveal the existence of such accessory proteins for these cases. Finally, it is important to note that exceptions to the metallocenter assembly mechanism exist within the particular enzyme systems. For example, *Bacillus subtilis* synthesizes sufficient levels of Ni-containing urease to allow for growth on urea as sole nitrogen source, even though its genome lacks homologues to *ureDEFG* (124). The mechanism by which this organism generates active urease in the apparent absence of any accessory protein remains unknown.

This thesis aims to shed light on the activation of *Klebsiella aerogenes* urease by focusing on selected activation complexes and the accessory protein UreG. In

chapter 2, I present structural studies of urease activation complexes (UreABC-UreD)$_3$ and (UreABC-UreDF)$_3$ performed by small angle X-ray scattering, flexibility analysis, and site-directed mutagenesis in a combined effort of the laboratories of Dr. William Heller (Oak Ridge National Laboratory), Dr. Leslie Kuhn (Michigan State University) and us. Combining the three methods we were able to propose that a conformational change in UreB would provide access to the active site of urease, facilitating metallocenter assembly. In chapter 3 I describe the characterization of a biotinylated form of the accessory protein UreG and the effect on urease activity caused by single mutations in conserved residues of UreG. Finally, in chapter 4, I present general conclusions and outline possible directions for future research.

REFERENCES

1. Mulrooney, S. B., and Hausinger, R. P. (2003) "Nickel uptake and utilization by microorganisms" *FEMS Microbiol Rev 27*, 239-261.

2. Hausinger, R. P. (1993) *Biochemistry of Nickel*, Plenum Publishing, New York.

3. Eitinger, T., and Mandrand-Berthelot, M.-A. (2000) "Nickel transport systems in microorganisms" *Arch Microbiol 173*, 1-9.

4. Rodionov, D. A., Hebbeln, P., Gelfand, M. S., and Eitinger, T. (2006) "Comparative and functional genomic analysis of prokaryotic nickel and cobalt uptake transporters: evidence for a novel group of ATP-binding cassette transporters" *J Bacteriol 188*, 317-327.

5. Navarro, C., Wu, L.-F., and Mandrand-Berthelot, M.-A. (1993) "The *nik* operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel" *Molec Microbiol 9*, 1181-1191.

6. de Pina, K., Navarro, C., McWalter, L., Boxer, D. H., Price, N. C., Kelly, S. M., Mandrand-Berthelot, M.-A., and Wu, L.-F. (1995) "Purification and characterization of the periplasmic nickel-binding protein NikA of *Escherichia coli* K12" *Eur J Biochem 227*, 857-865.

7. Heddle, J., Scott, D. J., Unzai, S., Park, S.-Y., and Tame, J. R. H. (2003) "Crystal structures of the liganded and unliganded nickel-binding protein NikA from *Escherichia coli*" *J Biol Chem 278*, 50322-50329.

8. Cherrier, M. V., Martin, L., Cavazza, C., Jacquamet, L., Lemaire, D., Gaillard, J., and Fontecilla-Camps, J. C. (2005) "Crystallographic and spectroscopic evidence for high affinity binding of $FeEDTA(H_2O)^-$ to the periplasmic nickel transporter NikA" *J Am Chem Soc 127*, 10075-10082.

9. Allan, C. B., Wu, L.-F., Gu, Z., Choudhury, S. B., Al-Mjeni, F., Sharma, M. L., Mandrand-Berthelot, M.-A., and Maroney, M. J. (1998) "An X-ray absorption spectroscopic structural investigation of the nickel site in *Escherichia coli* NikA protein" *Inorg Chem 37*, 5952-5955.

10. de Pina, K., Desjardin, V., Mandrand-Berthelot, M.-A., Giordano, G., and Wu, L.-F. (1999) "Isolation and characterization of the *nikR* gene encoding a nickel-responsive regulator in *Escherichia coli*" *J Bacteriol 181*, 670-674.

11. Eitinger, T., and Friedrich, B. (1991) "Cloning, nucleotide sequence, and heterologous expression of the high-affinity nickel transport gene from *Alcaligenes eutrophus*" *J Biol Chem 266*, 3222-3227.

12. Eitinger, T., Degen, O., Böhnke, U., and Müller, M. (2000) "Nic1p, a relative of bacterial transition metal permeases in *Schizosaccharomyces pombe*, provides nickel ion for urease biosynthesis" *J Biol Chem 275*, 18029-18033.

13. Hartl, F. U., and Hayer-Hartl, M. (2002) "Protein folding - Molecular chaperones in the cytosol: from nascent chain to folded protein" *Science 295*, 1852-1858.

14. Wittung-Stafshede, P. (2002) "Role of cofactors in protein folding" *Acc Chem Res 35*, 201-208.

15. Yan, S. Z., Beeler, J. A., Chen, Y., Shelton, R. K., and Tang, W. J. (2001) "The regulation of type 7 adenylyl cyclase by its C1b region and *Escherichia coli* peptidylprolyl isomerase, SlyD" *J Biol Chem 276*, 8500-8506.

16. Satumba, W. J., and Mossing, M. C. (2002) "Folding and assembly of lambda Cro repressor dimers are kinetically limited by proline isomerization" *Biochemistry 41*, 14216-14224.

17. Scholz, C., Schaarschmidt, P., Engel, A. M., Andres, H., Schmitt, U., Faatz, E., Balbach, J., and Schmid, F. X. (2005) "Functional solubilization of aggregation-prone HIV envelope proteins by covalent fusion with chaperone modules" *J Mol Biol 345*, 1229-1241.

18. Butland, G., Peregrin-Alvarez, J. M., Li, J.-G., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., Davey, M., Parkinson, J., Greenblatt, J., and Emili, A. (2005) "Interaction network containing conserved and essential protein complexes in *Escherichia coli*" *Nature 433*, 531-537.

19. Rodrigue, A., Batia, N., Müller, M., Fayet, O., Böhm, R., Mandrand-Berthelot, M.-A., and Wu, L.-F. (1996) "Involvement of the GroE chaperonins in the nickel-dependent anaerobic biosynthesis of NiFe-hydrogenases of *Escherichia coli*" *J Bacteriol 178*, 4453-4460.

20. Vignais, P. M., and Colbeau, A. (2004) "Molecular biology of microbial hydrogenases" *Curr Issues Molec Biol 6*, 159-188.

21. Vignais, P. M., Billoud, B., and Meyer, J. (2001) "Classification and phylogeny of hydrogenases" *FEMS Microbiol Rev 25*, 455-501.

22. Volbeda, A., Charon, M.-H., Piras, C., Hatchikian, E. C., Frey, M., and Fontecilla-Camps, J. C. (1995) "Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*" *Nature (London) 373*, 580-587.

23. Garcin, E., Vernede, X., Hatchikian, E. C., Volbeda, A., Frey, M., and Fontecilla-Camps, J. C. (1999) "The crystal structure of a reduced [NiFeSe] hydrogenase provides an image of the activated catalytic center" *Structure 7*, 557-566.

24. Volbeda, A., Garcin, E., Piras, C., de Lacey, A. L., Fernandez, V. M., Hatchikian, E. C., Frey, M., and Fontecilla-Camps, J. C. (1996) "Structure of the [NiFe] hydrogenase active site: evidence for biologically uncommon Fe ligands" *J Am Chem Soc 118*, 12989-12996.

25. Lutz, A., Jacobi, A., Schlensog, V., Böhm, R., Sawers, G., and Böck, A. (1991) "Molecular characterization of an operon (*hyp*) necessary for the activity of the three hydrogenase isoenzymes in *Escherichia coli*" *Molec Microbiol 5*, 123-135.

26. Blokesch, M., Paschos, A., Theodoratou, E., Bauer, A., Hube, M., Huth, S., and Böck, A. (2002) "Metal insertion into NiFe-hydrogenases" *Biochem Soc Trans 30*, 674-680.

27. Casalot, L., and Rousset, M. (2001) "Maturation of [NiFe] hydrogenases" *Trends Microbiol 9*, 228-237.

28. Blokesch, M., Albracht, S. P. J., Matzanke, B. F., Drapal, N., Jacobi, A., and Böck, A. (2004) "The complex between hydrogenase-maturation

proteins HypC and HypD is an intermediate in the supply of cyanide to the active site iron of [NiFe]-hydrogenases" *J Mol Biol 344*, 155-167.

29. Blokesch, M., and Böck, A. (2002) "Maturation of [NiFe]-hydrogenases in *Escherichia coli*: the HypC cycle" *J Mol Biol 324*, 287-296.

30. Mehta, N. S., Olson, J. W., and Maier, R. J. (2003) "Characterization of *Helicobacter pylori* nickel metabolism accessory proteins needed for maturation of both urease and hydrogenase" *J Bacteriol 185*, 726-734.

31. Atanassova, A., and Zamble, D. B. (2005) "*Escherichia coli* HypA is a zinc metalloprotein with a weak affinity for nickel" *J Bacteriol 187*, 4689-4697.

32. Blokesch, M., Rohrmoser, M., Rode, S., and Böck, A. (2004) "HybF, a zinc-containing protein involved in NiFe hydrogenase biosynthesis" *J Bacteriol 186*, 2603-2611.

33. Maier, T., Jacobi, A., Sauter, M., and Böck, A. (1993) "The product of the *hypB* gene, which is required for nickel incorporation into hydrogenases, is a novel guanine nucleotide-binding protein" *J Bacteriol 175*, 630-635.

34. Maier, T., Lottspeich, F., and Böck, A. (1995) "GTP hydrolysis by HypB is essential for nickel insertion into hydrogenases of *Escherichia coli*" *Eur J Biochem 230*, 133-138.

35. Fu, C., Olson, J. W., and Maier, R. J. (1995) "HypB protein of *Bradyrhizobium japonicum* is a metal-binding GTPase capable of binding 18 divalent nickel ions per dimer" *Proc Natl Acad Sci USA 92*, 2333-2337.

36. Rey, L., Imperial, J., Palacios, J.-M., and Ruiz-Argüeso, T. (1994) "Purification of *Rhizobium leguminosarum* HypB, a nickel-binding protein required for hydrogenase synthesis" *J Bacteriol 176*, 6066-6073.

37. Olson, J. W., and Maier, R. J. (2000) "Dual roles of *Bradyrhizobium japonicum* nickelin protein in nickel storage and GTP-dependent Ni mobilization" *J Bacteriol 182*, 1702-1705.

38. Olson, J. W., Fu, C., and Maier, R. J. (1997) "The HypB protein from *Bradyrhizobium japonicum* can store nickel and is required for the nickel-

dependent transcriptional regulation of hydrogenase" *Molec Microbiol 24*, 119-128.

39. Leach, M. R., Sandal, S., Sun, H., and Zamble, D. B. (2005) "Metal binding activity of the *Escherichia coli* hydrogenase maturation factor HypB" *Biochemistry 44*, 12229-12238.

40. Roof, W. D., Horne, S. M., Young, K. D., and Young, R. (1994) "*slyD*, a host gene required for PhiX174 lysis, is related to the FK506-binding protein family of peptidyl-prolyl *cis-trans*-isomerases" *J Biol Chem 269*, 2902-2910.

41. Hottenrott, S., Schumann, T., Pluckthun, A., Fischer, G., and Rahfeld, J. U. (1997) "The *Escherichia coli* SlyD is a metal ion-regulated peptidyl-prolyl cis/trans-isomerase" *J Biol Chem 272*, 15697-15701.

42. Roche, E. D., and Sauer, R. T. (2001) "Identification of endogenous SsrA-tagged proteins reveals tagging at positions corresponding to stop codons" *J Biol Chem 276*, 28509-28515.

43. Finzi, A., Cloutier, J., and Cohen, E. A. (2003) "Two-step purification of His-tagged Nef protein in native condition using heparin and immobilized metal ion affinity chromatographies" *J Virol Methods 111*, 69-73.

44. Mukherjee, S., Shukla, A., and Guptasarma, P. (2003) "Single-step purification of a protein-folding catalyst, the SlyD peptidyl prolyl isomerase (PPI), from cytoplasmic extracts of *Escherichia coli*" *Biotechnol Appl Biochem 37*, 183-186.

45. Zhang, J. W., Butland, G., Greenblatt, J. F., Emili, A., and Zamble, D. B. (2005) "A role for SlyD in the *Escherichia coli* hydrogenase biosynthetic pathway" *J Biol Chem 280*.

46. Blokesch, M., Magalon, A., and Böck, A. (2001) "Interplay between the specific chaperone-like proteins HybG and HypC in maturation of hydrogenases 1, 2, and 3 from *Escherichia coli*" *J Bacteriol 183*, 2817-2822.

47. Drapal, N., and Böck, A. (1998) "Interaction of the hydrogenase accessory protein HypC with HycE, the large subunit of *Escherichia coli* hydrogenase 3 during enzyme maturation" *Biochemistry 37*, 2941-2948.

48. Maier, T., Drapal, N., Thanbichler, M., and Böck, A. (1998) "Stept-tag II affinity purification: an approach to study intermediates of metalloenzyme biosynthesis" *Anal Biochem 259*, 68-73.

49. Magalan, A., and Böck, A. (2000) "Analysis of the HypC-HycE complex, a key intermediate in the assembly of the metal center of *Escherichia coli* hydrogenase 3" *J Biol Chem 275*, 21114-21120.

50. Hube, M., Blokesch, M., and Böck, A. (2002) "Network of hydrogenase maturation in Escherichia coli: role of accessory proteins HypA and HybF" *J Bacteriol 184*, 3879-3885.

51. Theodoratou, E., Paschos, A., Magalon, A., Fritsche, E., Huber, R., and Böck, A. (2000) "Nickel serves as a substrate recognition motif for the endopeptidase involved in hydrogenase maturation" *Eur J Biochem 267*, 1995-1999.

52. Olson, J. W., Mehta, N. S., and Maier, R. J. (2001) "Requirement of nickel metabolism proteins HypA and HypB for full activity of both hydrogenase and urease in *Helicobacter pylori*" *Molec Microbiol 39*, 176.

53. Lindahl, P. A. (2002) "The Ni-containing carbon monoxide dehydrogenase family: light at the end of the tunnel?" *Biochemistry 41*, 2097-2105.

54. Dobbek, H., Svetlitchnyi, V., Gremer, L., Huber, R., and Meyer, O. (2001) "Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni-4Fe-5S] cluster" *Science 293*, 1281-1285.

55. Drennan, C. L., Heo, J., Sintchak, M. D., Schreiter, E., and Ludden, P. W. (2001) "Life on carbon monoxide: X-ray structure of *Rhodospirillum rubrum* Ni-Fe-S carbon monoxide dehydrogenase" *Proc Natl Acad Sci USA 98*, 11973-11978.

56. Drennan, C. L., Doukov, T. I., and Ragsdale, S. W. (2004) "The metalloclusters of carbon monoxide dehydrogenase/acetyl-CoA synthase: a story in pictures" *J Biol Inorg Chem 9*, 511-515.

57. Doukov, T. I., Iverson, T. M., Seravalli, J., Ragsdale, S. W., and Drennan, C. L. (2002) "A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase" *Science 298*, 567-272.

58. Darnault, C., Volbeda, A., Kim, E. J., Legrand, P., Vernède, X., Lindahl, P. A., and Fontecilla-Camps, J. C. (2003) "NiZn[Fe_4S_4] and NiNi[Fe_4S_4] clusters in closed and open α subunits of acetyl-CoA synthase/carbon monoxide dehydrogenase" *Nature Struct Biol 10*, 271-279.

59. Svetlitchnyi, V., Dobbek, H., Meyer-Klaucke, W., Meins, T., Thiele, B., Römer, P., Huber, R., and Meyer, O. (2004) "A functional Ni-Ni-[4Fe-4S] cluster in the monomeric acetyl-CoA synthase from *Carboxydothermus hydrogenoformans*" *Proc Natl Acad Sci USA 101*, 446-451.

60. Kerby, R. L., Ludden, P. W., and Roberts, G. P. (1997) "In vivo nickel insertion into carbon monoxide dehydrogenase of *Rhodosprillum rubrum*: molecular and physiological characterization of *cooCTJ*" *J Bacteriol 179*, 2259-2266.

61. Watt, R. K., and Ludden, P. W. (1998) "The identification, purification and characterization of CooJ. A nickel-binding protein that is co-regulated with the Ni-containing CO dehydrogenase from *Rhodospirillum rubrum*" *J Biol Chem 273*, 10019-10025.

62. Jeon, W. B., Cheng, J., and Ludden, P. W. (2001) "Purification and characterization of membrane-associated CooC protein and its functional role in the insertion of nickel into carbon monoxide dehydrogenase from *Rhodospirillum rubrum*" *J Biol Chem 276*, 38602-38609.

63. Loke, H.-K., and Lindahl, P. A. (2003) "Identification and preliminary characterization of AcsF, a putative Ni-insertase used in the biosynthesis of acetyl-CoA synthase from *Clostridium thermoaceticum*" *J Inorg Biochem 93*, 33-40.

64. Thauer, R. K. (1998) "Biochemistry of methanogenesis: a tribute to Marjory Stephenson" *Microbiol 144*, 2377-2406.

65. Ermler, U., Grabarse, W., Shima, S., Goubeaud, M., and Thauer, R. K. (1997) "Crystal structure of methyl-coenzyme M reductase: the key enzyme of biological methane formation" *Science 278*, 1457-1462.

66. Shima, S., and Thauer, R. K. (2005) "Methyl-coenzyme M reductase and anaerobic oxidation of methane in methanotrophic archaea" *Curr Opinion Microbiol 8*, 643-648.

67. Krüger, M., Meyerdierks, A., Glöckner, F. O., Amann, R., Widdel, F., Kube, M., Reinhardt, R., Kahnt, J., Böcher, R., Thauer, R. K., and Shima, S. (2003) "A conspicuous nickel protein in microbial mats that oxidize methane anaerobically" *Nature 426*, 878-881.

68. Hallam, S. J., Putnam, N., Preston, C. M., Detter, J. C., Rokhsar, D., Richardson, P. M., and DeLong, E. F. (2004) "Reverse methanogenesis: testing the hypothesis with environmental genomics" *Science 305*, 1457-1462.

69. Thauer, R. K., and Bonacker, L. G. (1994) "Biosynthesis of coenzyme F430, a nickel porphinoid involved in methanogenesis" *Ciba Foundation Symposium 180*, 210-227.

70. Wuerges, J., Lee, J.-W., Yim, Y.-I., Kang, S. O., and Carugo, K. D. (2004) "Crystal structure of nickel-containing superoxide dismutase reveals another type of active site" *Proc Natl Acad Sci USA 101*, 8569-8574.

71. Barondeau, D. P., Kassman, C. J., Bruns, C. K., Tainer, J. A., and Getzoff, E. D. (2004) "Nickel superoxide dismutase structure and mechanism" *Biochemistry 43*, 8038-8047.

72. Eitinger, T. (2004) "In vivo production of active nickel superoxide dismutase from *Prochlorococcus marinus* MIT9313 is dependent on its cognate peptidase" *J Bacteriol 186*, 7812-7825.

73. Kim, I.-K., Yim, Y.-I., Kim, Y.-M., Lee, J.-W., Yim, H.-S., and Kang, S.-O. (2003) "CbiX-homologous protein (CbiXhp), a metal-binding protein, from Streptomyces seoulensis is involved in expression of nickel-containing superoxide dismutase" *FEMS Microbiol Lett 228*, 21-26.

74. Clugston, S. L., Barnard, J. F. J., Kinach, R., Miedema, D., Ruman, R., Daub, E., and Honek, J. F. (1998) "Overproduction and characterization of a dimeric non-zinc glyoxylase I from *Escherichia coli*: evidence for optimal activation by nickel ions" *Biochemistry 37*, 8754-8763.

75. He, M. M., Clugston, S. L., Honek, J. F., and Matthews, B. W. (2000) "Determination of the structure of *Escherichia coli* glyoxylase I suggests a structural basis for differential metal activation" *Biochemistry 39*, 8719-8727.

76. Clugston, S. L., and Honek, J. F. (2000) "Identification of sequences encoding the detoxification metalloisomerase glyoxylase I in microbial genomes from several pathogenic organisms" *J Molec Evol 50*, 491-495.

77. Dai, Y., Wensink, P. C., and Abeles, R. H. (1999) "One protein, two enzymes" *J Biol Chem 274*, 1193-1195.

78. Pochapsky, T. C., Pochapsky, S. S., Ju, T., Mo, H., Al-Mjeni, F., and Maroney, M. J. (2002) "Modeling and experiment yields the structure of acireductone dioxygenase from *Klebsiella pneumoniae*" *Nature Struct Biol 9*, 966-972.

79. Al-Mjeni, F., Ju, T., Pochapsky, T. C., and Maroney, M. J. (2002) "XAS investigation of the structure and function of Ni in acireductone dioxygenase" *Biochemistry 41*, 6761-6769.

80. Weeks, D. L., Eskandar, S., Scott, D. R., and Sachs, G. (2000) "A H^+-gated urea channel: the link between *Helicobacter pylori* urease and gastric colonization" *Science 287*, 482-485.

81. Olson, J. W., and Maier, R. J. (2002) "Molecular hydrogen as an energy source for *Helicobacter pylori*" *Science 298*, 1788-1790.

82. Gilbert, J. V., Ramakrishna, J., Sunderman, F. W., Jr., Wright, A., and Plaut, A. G. (1995) "Protein Hpn: cloning and characterization of a histidine-rich metal-binding polypeptide in *Helicobacter pylori* and *Helicobacter mustelae*" *Infect Immun 63*, 2682-2688.

83. Mobley, H. L. T., Garner, R. M., Chippendale, G. R., Gilbert, J. V., Kane, A. V., and Plaut, A. G. (1999) "Role of Hpn and NixA of *Helicobacter pylori* in susceptibility and resistance to bismuth and other metal ions" *Helicobacter 4*, 162-169.

84. Ge, R., Watt, R. M., Sun, X., Tanner, J. A., He, Q.-Y., Huang, J.-D., and Sun, H. (2006) "Expression and characterization of the histidine-rich

protein, Hpn: potential for nickel storage in *Helicobacter pylori*" *Biochem J 393*, 285-293.

85. Hausinger, R. P., and Karplus, P. A. (2001) in *Handbook of Metalloproteins* (Wieghardt, K., Huber, R., Poulos, T. L., and Messerschmidt, A., Eds.) pp 867-879, John Wiley & Sons, Ltd., West Sussex, U.K.

86. Jabri, E., Carr, M. B., Hausinger, R. P., and Karplus, P. A. (1995) "The crystal structure of urease from *Klebsiella aerogenes*" *Science 268*, 998-1004.

87. Pearson, M. A., Michel, L. O., Hausinger, R. P., and Karplus, P. A. (1997) "Structure of Cys319 variants and acetohydroxamate-inhibited *Klebsiella aerogenes* urease" *Biochemistry 36*, 8164-8172.

88. Benini, S., Rypniewski, W. R., Wilson, K. S., Miletti, S., Ciurli, S., and Mangani, S. (1999) "A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from *Bacillus pasteurii*: why urea hydrolysis costs two nickels" *Structure 7*, 205-216.

89. Ha, N.-C., Oh, S.-T., Sung, J. Y., Cha, K.-A., Lee, M. H., and Oh, B.-H. (2001) "Supramolecular assembly and acid resistance of *Helicobacter pylori* urease" *Nature Structure Biology 8*, 505-509.

90. Sheridan, L., Wilmont, C. M., Cromie, K. D., van der Logt, P., and Phillips, S. E. V. (2002) "Crystallization and preliminary X-ray structure determination of jack bean urease with a bound antibody fragment" *Acta Crystallogr D58*, 374-376.

91. Moncrief, M. B. C., Hom, L. G., Jabri, E., Karplus, P. A., and Hausinger, R. P. (1995) "Urease activity in the crystalline state" *Prot Science 4*, 2234-2236.

92. Pearson, M. A., Park, I.-S., Schaller, R. A., Michel, L. O., Karplus, P. A., and Hausinger, R. P. (2000) "Kinetic and structural characterization of urease active site variants" *Biochemistry 39*, 8575-8584.

93. Mulrooney, S. B., and Hausinger, R. P. (1990) "Sequence of the *Klebsiella aerogenes* urease genes and evidence for accessory proteins facilitating nickel incorporation" *J Bacteriol 172*, 5837-5843.

94. Lee, M. H., Mulrooney, S. B., Renner, M. J., Markowicz, Y., and Hausinger, R. P. (1992) "*Klebsiella aerogenes* urease gene cluster: sequence of *ureD* and demonstration that four accessory genes (*ureD*, *ureE*, *ureF*, and *ureG*) are involved in nickel metallocenter biosynthesis" *J Bacteriol 174*, 4324-4330.

95. Soriano, A., and Hausinger, R. P. (1999) "GTP-dependent activation of urease apoprotein in complex with the UreD, UreF, and UreG accessory proteins" *Proc Natl Acad Sci USA 96*, 11140-11144.

96. Colpas, G. J., and Hausinger, R. P. (2000) "In vivo and in vitro kinetics of metal transfer by the *Klebsiella aerogenes* urease nickel metallochaperone, UreE" *J Biol Chem 275*, 10731-10737.

97. Soriano, A., Colpas, G. J., and Hausinger, R. P. (2000) "UreE stimulation of GTP-dependent urease activation in the UreD-UreF-UreG-urease apoprotein complex" *Biochemistry 39*, 12435-12440.

98. Jones, B. D., and Mobley, H. L. T. (1989) "*Proteus mirabilis* urease: nucleotide sequence determination and comparison with jack bean urease" *J Bacteriol 171*, 6414-6422.

99. Musiani, F., Zambelli, B., Stola, M., and Ciurli, S. (2004) "Nickel trafficking: insights into the fold and function of UreE, a urease metallochaperone" *J Inorg Biochem 98*, 803-813.

100. Lee, M. H., Pankratz, H. S., Wang, S., Scott, R. A., Finnegan, M. G., Johnson, M. K., Ippolito, J. A., Christianson, D. W., and Hausinger, R. P. (1993) "Purification and characterization of *Klebsiella aerogenes* UreE protein: a nickel-binding protein that functions in urease metallocenter assembly" *Prot Science 2*, 1042-1052.

101. Ciurli, S., Safarof, N., Miletti, S., Dikiy, A., Christensen, S. K., Kornetzky, K., Bryant, D. A., Vandenberghe, I., Devreese, B., Samyn, B., Remaut, H., and Van Beeumen, J. (2002) "Molecular characterization of *Bacillus pasteurii* UreE, a metal-binding chaperone for the assembly of the urease active site" *J Biol Inorg Chem 7*, 623-631.

102. Colpas, G. J., Brayman, T. G., McCracken, J., Pressler, M. A., Babcock, G. T., Ming, L.-J., Colangelo, C. M., Scott, R. A., and Hausinger, R. P. (1998) "Spectroscopic characterization of metal binding by *Klebsiella aerogenes* UreE urease accessory protein" *J Biol Inorg Chem 3*, 150-160.

103. Brayman, T. G., and Hausinger, R. P. (1996) "Purification, characterization, and functional analysis of a truncated *Klebsiella aerogenes* UreE urease accessory protein lacking the histidine-rich carboxyl terminus" *J Bacteriol 178*, 5410-5416.

104. Colpas, G. J., Brayman, T. G., Ming, L.-J., and Hausinger, R. P. (1999) "Identification of metal-binding residues in the *Klebsiella aerogenes* urease nickel metallochaperone, UreE" *Biochemistry 38*, 4078-4088.

105. Benoit, S., and Maier, R. J. (2003) "Dependence of *Helicobacter pylori* urease activity on the nickel-sequestering ability of the UreE accessory protein" *J Bacteriol 185*, 4787-4795.

106. Rain, J.-C., Selig, L., de Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schächter, V., Chemama, Y., Labigne, A., and Legrain, P. (2001) "The protein-protein interaction map of *Helicobacter pylori*" *Nature 409*, 211--215.

107. Voland, P., Weeks, D. L., Marcus, E. A., Prinz, C., Sachs, G., and Scott, D. (2003) "Interactions among the seven *Helicobacter pylori* proteins encoded by the urease gene cluster" *Liver Physiol 284*, G96-G106.

108. Song, H. K., Mulrooney, S. B., Huber, R., and Hausinger, R. P. (2001) "Crystal structure of *Klebsiella aerogenes* UreE, a nickel-binding metallochaperone for urease activation" *J Biol Chem 276*, 49359-49364.

109. Remaut, H., Safarof, N., Ciurli, S., and Van Beeumen, J. (2001) "Structural basis for Ni^{2+} transport and assembly of the urease active site by the metallo-chaperone UreE from *Bacillus pasteurii*" *J Biol Chem 276*, 49365-49370.

110. Rosenzweig, A. C., Huffman, D. L., Hou, M. Y., Wernimont, A. K., Pufahl, R. A., and O'Halloran, T. V. (1999) "Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution" *Structure 7*, 605-617.

111. Sha, B., Lee, S.-K., and Cyr, D. M. (2000) "The crystal structure of the peptide-binding fragment from yeast Hsp40 protein Sis1" *Structure 8*, 799-807.

112. Mulrooney, S. B., Ward, S. K., and Hausinger, R. P. (2005) "Purification and properties of the *Klebsiella aerogenes* UreE metal-binding domain, a functional metallochaperone of urease" *J Bacteriol 187*, 3581-3585.

113. Won, H.-S., Lee, Y.-H., Kim, J.-H., Shin, I. S., Lee, M. H., and Lee, B.-J. (2004) "Structural characterization of the nickel-binding properties of *Bacillus pasteurii* UreE in solution" *J Biol Chem 279*, 17466-17472.

114. Jabri, E., and Karplus, P. A. (1996) "Structures of the *Klebsiella aerogenes* urease apoprotein and two active-site mutants" *Biochemistry 35*, 10616-10626.

115. Park, I.-S., Carr, M. B., and Hausinger, R. P. (1994) "*In vitro* activation of urease apoprotein and role of UreD as a chaperone required for nickel metallocenter assembly" *Proc Natl Acad Sci USA 91*, 3233-3237.

116. Moncrief, M. B. C., and Hausinger, R. P. (1996) "Purification and activation properties of UreD-UreF-urease apoprotein complexes" *J Bacteriol 178*, 5417-5421.

117. Park, I.-S., and Hausinger, R. P. (1995) "Evidence for the presence of urease apoprotein complexes containing UreD, UreF, and UreG in cells that are competent for in vivo enzyme activation" *J Bacteriol 177*, 1947-1951.

118. Park, I.-S., and Hausinger, R. P. (1995) "Requirement of carbon dioxide for in vitro assembly of the urease nickel metallocenter" *Science 267*, 1156-1158.

119. Park, I.-S., and Hausinger, R. P. (1996) "Metal ion interactions with urease and UreD-urease apoproteins" *Biochemistry 35*, 5345-5352.

120. Chang, Z., Kuchar, J., and Hausinger, R. P. (2004) "Chemical crosslinking and mass spectrometric identification of sites of interaction for UreD, UreF, and urease" *J Biol Chem 279*, 15305-15313.

121. Moncrief, M. B. C., and Hausinger, R. P. (1997) "Characterization of UreG, identification of a UreD-UreF-UreG complex, and evidence suggesting that a nucleotide-binding site in UreG is required for in vivo metallocenter assembly of *Klebsiella aerogenes* urease" *J Bacteriol 179*, 4081-4086.

122. Zambelli, B., Stola, M., Musiani, F., De Vriendt, K., Samyn, B., Devreese, B., Van Beeumen, J., Dikiy, A., Bryant, D. A., and Ciurli, S. (2005) "UreG, a chaperone in the urease assembly process, is an intrinsically unstructured GTPase that specifically binds Zn^{2+}" *J Biol Chem 280*, 4684-4695.

123. Kansau, I., Guillain, F., Thiberge, J.-M., and Labigne, A. (1996) "Nickel binding and immunological properties of the C-terminal domain of the *Helicobacter pylor*i GroES homologue (HspA)" *Molec Microbiol 22*, 1013-1023.

124. Kim, J. K., Mulrooney, S. B., and Hausinger, R. P. (2005) "Biosynthesis of active *Bacillus subtilis* urease in the absence of known urease accessory proteins" *J Bacteriol 187*, 7150-7154.

Chapter 2

The Structure of Urease Activation Complexes Examined by Flexibility Analysis, Mutagenesis, and Small-Angle X-Ray Scattering Approaches

The computational analysis included in this chapter were performed by Sai Chetan K. Sukuru, under the supervision of Dr. Leslie A. Kuhn. The small angle X-ray scattering data collection and analysis was done by Dr. William T. Heller.

ABSTRACT

Conformational changes of *Klebiella aerogenes* urease apoprotein $(UreABC)_3$ induced upon binding of the UreD and UreF accessory proteins were examined by a combination of flexibility analysis, mutagenesis, and small-angle x-ray scattering (SAXS). These studies build on prior work reporting a chemical cross-link between UreB Lys76 and UreC Lys382 in the $(UreABC\text{-}UreDF)_3$ complex that was interpreted in terms of a conformational change involving UreB. ProFlex analysis of urease provided evidence that the major domain of UreB can move in a hinge-like motion to account for the cross-linking result. Rigidification of the UreB hinge region in the G11P variant was found to reduce the extent of urease activation, in part by decreasing the nickel content of the mutant enzyme, and to sequester a portion of the urease apoprotein in an activation complex that includes all of the accessory proteins. SAXS analyses of urease, $(UreABC\text{-}UreD)_3$, and $(UreABC\text{-}UreDF)_3$ are most consistent with UreD and UreF binding near UreB. Notably, improved fits were observed for $(UreABC\text{-}UreDF)_3$ models where UreB is repositioned in line with the predicted conformational change. Significantly, the predicted structures of $(UreABC\text{-}UreDF)_3$ containing the domain-shifted UreB conformations allow CO_2 and nickel ions to gain access to the nascent active site, compatible with a mechanism for urease activation.

INTRODUCTION

Urease is a nickel-containing enzyme that hydrolyzes urea (*1, 2*). Crystallographic analyses of ureases from bacterial and plant sources (*3-7*) reveal a basic trimeric structure with three active sites, each composed of two nickel ions coordinated by a carbamylated Lys, four His and an Asp. Genetic and biochemical studies carried out with plants, fungi, and bacteria [reviewed in (*8-10*)] have shown that additional genes encoding accessory proteins are required for proper assembly of the urease metallocenter, with the possible exception of *Bacillus subtilis* (*11*). The current model for urease metallocenter assembly (Figure 2.1) derives primarily from studies involving expression of the *Klebsiella aerogenes ureDABCEFG* gene cluster in *Escherichia coli* [reviewed in (*8, 12*)]. The active enzyme possesses three copies of each of three subunits (UreA, UreB, and UreC of M_r 11,086, 11,695, and 60,304, respectively)(*13*). Deletions within *ureD*, *ureE*, *ureF*, or *ureG* eliminate urease activity due to production of the inactive $(UreABC)_3$ urease apoprotein (*14*). Expression of *ureDABC* produces $(UreABC\text{-}UreD)_3$, with UreD (M_r 29,300) in complex with urease apoprotein (*15*). Co-expression of *ureF* (encoding a protein of M_r 25,221) with *ureDABC* produces the $(UreABC\text{-}UreDF)_3$ complex (*16*). The soluble protein UreG (M_r 21,943) reversibly binds to $(UreABC\text{-}UreDF)_3$ forming $(UreABC\text{-}UreDFG)_3$ (*17, 18*). Urease activity is generated by incubating these complexes with high concentrations of bicarbonate (to supply the CO_2 needed for Lys carbamylation) and nickel ions, but the required levels of these additives (100 mM and 100 μM, respectively) are not physiologically relevant and only a portion of the proteins

are activated (*19, 20*). In contrast, fully active urease is generated with only 100 μM bicarbonate and 20 μM nickel ions using $(UreABC\text{-}UreDFG)_3$ plus UreE (M_r 17,558) and GTP (*21*). UreE is a nickel-binding protein that delivers the metal ion to the targeted protein (*22, 23*), and GTP is hydrolyzed by UreG when present in the $(UreABC\text{-}UreDFG)_3$ complex (*24*). Although UreE is often referred to as a metallochaperone (*25, 26*) and UreDFG has been termed a urease-specific molecular chaperone (*9*), the mechanism of urease metallocenter assembly has remained obscure.

The near identity in structure of the $(UreABC)_3$ apoprotein (*27*) and the holoenzyme (*3*) indicate that conformational changes are required to introduce the metal ions and CO_2 into the deeply buried nascent active site. Chemical cross-linking of $(UreABC\text{-}UreDF)_3$ (*28*) identified a cross-link between UreB Lys76 and UreC Lys382 that provided evidence for a conformational change of the protein, since UreB Lys76 is positioned far from UreC Lys382 in the $(UreABC)_3$ crystal structure. Here, we use computational flexibility analysis to identify a hinge region that allows the main UreB domain to shift to a position that allows formation of the critical intra-urease cross-link. Furthermore, we show that one of two amino acid changes affecting this hinge region leads to a large reduction in urease activation, partly due to decreasing the extent of nickel incorporation, while also sequestering a large percentage of the urease protein in a complex with the accessory proteins. Finally, using SAXS methods we obtain best-fit models of $(UreABC\text{-}UreD)_3$ and $(UreABC\text{-}UreDF)_3$ that depict UreD and UreF binding together with UreB at the perimeter of the disk formed by $(UreAC)_3$.

Notably, improved fits were observed for $(UreABC\text{-}UreDF)_3$ models where UreB is repositioned in line with the predicted conformational change. These results are compatible with earlier urease activation studies and suggest that the combined action of UreD and UreF serves to expose the nascent active site of urease.

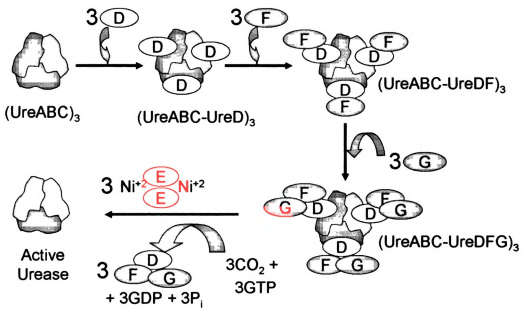


FIGURE 2.1. Proposed pathway of urease activation. The *K. aerogenes* UreA, UreB and UreC urease subunits assemble into the $(UreABC)_3$ apoprotein (depicted simply as a trimeric species since UreA plus UreB or all three subunits are fused together in ureases from some sources). UreD, UreF and UreG sequentially bind to form the $(UreABC\text{-}UreD)_3$, $(UreABC\text{-}UreDF)_3$, and $(UreABC\text{-}UreDFG)_3$ activation complexes. CO_2 adds to the active site Lys as Ni^{+2} ions are delivered to $(UreABC\text{-}UreDFG)_3$ by the dimeric UreE metallochaperone in a process that requires GTP hydrolysis, with UreE and $(UreDFG)_3$ being released from the activated urease.

EXPERIMENTAL PROCEDURES

Protein Purification. $(UreABC\text{-}UreD)_3$, $(UreABC\text{-}UreDF)_3$, and urease holoenzyme were produced in *E. coli* DH5α carrying pKAUD2 (*15*), *E. coli* DH5α pKAUD2F+Δ*ureG* (*16*), or *E. coli* HMS174(DE3) carrying pKK17 (*25*) and purified as previously described (*29*). HEDG buffer (25 mM HEPES, pH 7.4, 1 mM EDTA, 1 mM DTT, 1 % glycerol) was used as a final storage buffer unless noted. The homogeneity of samples was assessed by densitometric analysis (AlphaImager) of Coomassie-stained gels after sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (*30*). The expression level of urease subunits in cell extracts was assessed by SDS-PAGE followed by electroblotting the sample onto Immobilon-P polyvinylidene difluoride membrane, probing with anti-*K. aerogenes* urease antibodies (*31*), and visualizing with anti-rabbit immunoglobulin G-alkaline phosphatase conjugates. In a similar manner, the identity of a contaminating band in one sample was examined by doing a Western blot with anti-*K. aerogenes* UreE antibodies (*32*).

Site-Directed Mutagenesis and Activity Assay. Plasmid pKK17 (*25*) containing the entire urease gene cluster was cut with *Bam*HI and the smaller of two fragments (3.3 kbp) containing *ureB* was ligated into *Bam*HI-restricted pUC19 (New England BioLabs), producing pUCB. Mutations of *ureB* were introduced by PCR using primers 5'- GAA TAT CAC GTT AAG CCC CCA CAG ATA GCC CTG AAT ACC -3' and its complement to introduce the UreB G11P mutation and 5'- CAG ATA GCC CTG AAT ACC CCA CGG GCA ACC TGT CGC GTG -3' and its complement for the UreB G18P mutation (in each case the

mutated codon is underlined). The PCR reaction (18 cycles of 50 s at 95 °C, 50 s at 50 °C and 8 min at 72 °C) was performed with 12.5 μL of *PfuTurbo*® Hotstart PCR master mix (Stratagene), 10 μM of each primer, and the pUCB plasmid as template, followed by incubation for 1 h at 37 °C with 0.5 μL of *DpnI*. DH5α cells were transformed with 5 μL of the digested PCR reaction. Plasmids from putative clones were purified, sequenced to confirm the mutations, and digested with *Bam*HI to recover the 3.3-kbp fragments. These fragments were cloned back into pKK17 to create pKKBG11P and pKKBG18P.

E. coli cells containing pKK17, pKKBG11P, or pKKBG18P were grown in Luria-Bertani medium containing 1 mM $NiCl_2$ for three h and induced overnight with 0.1 mM isopropyl-β-D-thiogalactopyranoside. The stationary phase cells were harvested by centrifugation, sonicated, and clarified by ultracentrifugation. Cell extracts were tested for expression of the urease genes by denaturing gel electrophoresis (*30*) and subjected to protein analyses (*33*) and urease activity assays (*34*) using standard procedures.

Metal Quantification. The nickel content of selected samples was assessed by using inductively coupled plasma-mass spectrometry at the University of Georgia Chemical Analysis Laboratory.

Flexibility Analysis. We used the graph theoretic algorithm ProFlex to analyze the flexibility of urease (Protein Data Bank (PDB) entry 1FWJ). The program identifies the flexible and rigid regions in a given structure (which bonds are constrained and which bonds remain free to rotate) based on analysis of constraints posed by the protein's network of covalent bonds, hydrogen bonds,

salt bridges, and hydrophobic interactions (*35*). ProFlex calculations have been shown to predict the conformational flexibility of proteins reliably from a single 3D structure (*35-37*). The ProFlex code was modified and extended to allow processing of the very large urease structure (~22000 atoms in the trimer of trimers).

SAXS Measurements and Analysis. Small-angle X-ray scattering (SAXS) data were obtained using the ORNL Center for Structural Molecular Biology 4m SAXS instrument, described previously (*38*). Sample intensity patterns were collected for native urease, $(UreABC\text{-}UreD)_3$, and $(UreABC\text{-}UreDF)_3$ plus backgrounds consisting of the buffer solution. Protein concentrations were 3.8 mg/mL for native urease, 5.4 mg/mL for $(UreABC\text{-}UreD)_3$, and 2.0 mg/mL for $(UreABC\text{-}UreDF)_3$. These low concentrations made it impractical to measure a concentration series, but also make it unlikely that interparticle interference effects are significantly influencing the data. Multiple sample runs were averaged together, which enabled testing for time-dependent aggregation indicative of radiation damage; none was found. For $(UreABC\text{-}UreD)_3$ and $(UreABC\text{-}UreDF)_3$, four 4-hour runs were summed together, while five 4-hour runs were summed together for the native urease complex. These measurements included runs with fresh material and runs in which the sample was exposed for an additional 4 hours to check for radiation damage. No artifacts due to radiation damage were observed. Data were reduced, azimuthally averaged and scaled into absolute units (1/cm) according to previously published procedures (*38*) to provide the 1D

intensity profile $I(q)$ vs. q, where $q = 4\pi\sin(\theta)/\lambda$, 2θ is the scattering angle from the direct beam, and λ is the wavelength of the X-ray radiation (1.542 Å).

Small-angle X-ray scattering analysis and modeling. Data were subjected to Guinier analysis (*39*) for the radius of gyration, R_g, and for the pair-distance distribution function $P(r)$. $I(q)$ and $P(r)$ are related through the Fourier transform shown in Equation 2.1.

$$P(r) = \frac{1}{2\pi^2} \int_0^\infty qr \cdot I(q) \cdot \sin(qr) \cdot dr \qquad (2.1)$$

The program GNOM (*40*) uses an indirect transform method to find $P(r)$ from an input maximum linear dimension, d_{max}. The optimum d_{max} is found by trial and error, based on the quality of fit to the input data. The $P(r)$ fitting also provides a secondary measure of the R_g, which is the second moment of $P(r)$.

The program ORNL_SAS (*41*) was employed to compare the scattering profiles calculated from the urease structure and various models of complexes against the measured SAXS profiles of the enzyme, $(UreABC\text{-}UreD)_3$, and $(UreABC\text{-}UreDF)_3$. To model the $(UreABC\text{-}UreD)_3$ and $(UreABC\text{-}UreDF)_3$ complexes, ellipsoids were used in place of the unknown structures of UreD and UreF. The structures of the higher-order complexes were built by placing three identical ellipsoids with the $(UreABC)_3$ structure and using the same three-fold symmetry axis around which the trimer of trimers is formed. The translation coordinates were chosen randomly from a range of values that made it possible to produce complexes that extended beyond the experimentally determined d_{max}.

To ensure the proper volume for the added proteins, two of the ellipsoidal semiaxes were randomly chosen from a range of 10 Å to 35 Å, and the third was initially picked to produce the correct expected volume based on the amino acid sequence of the subunit. In the event that the third semiaxis was found to be less than 10 Å, a new set of semiaxes was generated. The ellipsoids were placed around the $(UreABC)_3$ structure and the volumes of the ellipsoids that did not overlap with either the $(UreABC)_3$, or the set of UreD ellipsoids in the case of $(UreABC\text{-}UreDF)_3$, were determined. If the amount of overlapping volume exceeded 1% of the correct ellipsoidal volume, the ellipsoidal semiaxes were scaled to provide the correct volume. As the specific overlap region with the other structures changes as the semiaxes are scaled, an iterative process was employed until the volume of the overlap regions was less than 1% of the correct volume. Only the portions of the ellipsoid that did not overlap were retained for the intensity calculations. Models found to have R_g values consistent with the experimental data were input into ORNL_SAS for comparison against the experimental data. ORNL_SAS was configured to treat the density of the scattering particle as uniform because no atomic-resolution structures are available for UreD and UreF. A 3 Å thick hydration layer, assumed to be 10 % more dense than the surrounding solution, was used for the ORNL_SAS intensity calculation. The thickness and density of the hydration layer were not parameters in the data fitting.

The quality of the fit of the model intensity profiles to the experimental data was evaluated using the reduced χ^2 parameter, defined in Equation 2.2.

$$\chi^2 = \frac{1}{\left(\sum_j N_{j,pts}\right) - N_f} \sum_j \sum_{N_{j,pts}} \frac{\left(I_j(q) - I_{m,j}(q)\right)^2}{\sigma_j(q)^2} \qquad (2.2)$$

$N_{j,pts}$ is the number of data points modelled against in the measured intensity $I_j(q)$. $\sigma_j(q)$ is the experimental uncertainty in the measured intensity $I_j(q)$. N_f is the number of degrees of freedom, and was 2, which accounts for the scaling of the model intensity profile to the data input into ORNL_SAS. ORNL_SAS, being a general intensity calculator (*41*), does not have a mechanism to account for the ellipsoidal structural parameters in N_f. The number of data points is a great deal larger than the number of degrees of freedom in any of the models tested, so the impact on χ^2 is relatively small. Additionally, each model is tested relative to models generated with the same number of free parameters, so the relative comparisons are not affected. In order to judge the range of structures that fit the experimental data collected for $(UreABC\text{-}UreD)_3$ and $(UreABC\text{-}UreDF)_3$, the best 25 models found were maintained in an ordered list that was updated as better models were found, in a manner similar to previous work (*42*), making it possible to judge the reproducibility of the modeling.

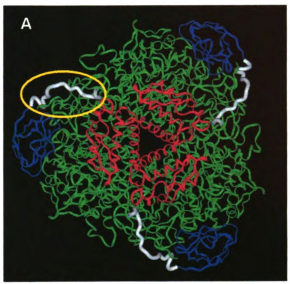
RESULTS

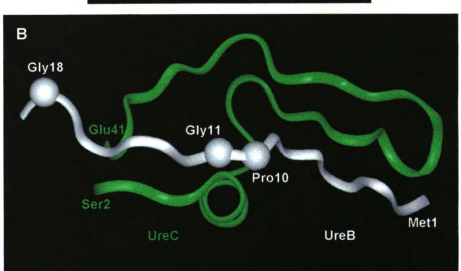
Flexibility Analysis of Urease. ProFlex, the software designed to analyze flexibility of proteins (*35*), was used to examine the flexibility within the native enzyme trimer of trimers (PDB entry 1FWJ; Figure 2.2 and Figure 2.3 top panels), identifying a total of ~3100 hydrogen bonds and ~1500 hydrophobic interactions. The regions of the protein defined as rigid or flexible were found to vary little with the choice of hydrogen-bond energy cutoff in ProFlex (between -1 and -2 kcal/mol), defining the set of hydrogen bonds and salt bridges incorporated in the network. In the crystal structure of urease, UreB is anchored by six N-terminal residues that add to the edge of a beta sheet in UreC (Figure 2.2, region 1). A salt bridge and at least six hydrophobic interactions between UreB residues 2-8 and UreC residues 6-29 reinforce the attachment (Tables 2.1 and 2.2). ProFlex predicted UreB residues 11-19 to form a flexible hinge (Figure 2.2, region 2; Tables 2.1 and 2.4) between the N-terminal anchor and the relatively rigid domain formed by UreB residues 20-101. The latter domain includes polar and hydrophobic interactions with UreC (Tables 2.5 and 2.6), but these are few in number compared to the interactions with regions 1 and 2 and consistent with the possibility of domain movement. The anchored and hinge residues of the N-terminal region of UreB (residues 1-19) fit into a groove of the N-terminal region of UreC formed by residues C2-C41 (Figure 2.2).

Chemical modification results (*28*) indicate that UreB Lys76 and UreC Lys382 can be cross-linked when in the $(UreABC\text{-}UreDF)_3$ species. This requires bringing their side chains to within 10 Å, although they are 50 Å apart in the

urease crystal structure. Thus, we probed whether the flexibility of UreB residues 11-19 would allow these two Lys residues to move to within cross-linking distance while maintaining favorable packing between UreB and UreC. In the first approach, UreB Gly11 and Gly18 were of special interest due to the prevalence of Gly in flexible regions of proteins; i.e., Gly residues have no constraints on main-chain bond rotations (Φ and Ψ angle torsions) due to the absence of side-chain induced steric hindrance. The torsion angles of UreB Gly11 and Gly18 were manually changed to reduce the distance between UreB Lys76 and UreC Lys382 and attain reasonable packing between UreB and UreAC. The resulting distance between the Cα atoms of UreB Lys76 and UreC Lys382 was 19.8 Å, close enough to allow cross-linking of their side chains. This motion involved a rotation of +131 degrees in Φ and +110 degrees in Ψ for Gly11, with 7 degree changes in both Φ and Ψ for Gly18, creating UreB conformation 1 (Figure 2.3, middle panels). In a second approach, we cut the tether at UreB Gly11, docked UreB Lys76 within cross-linking distance of UreC Lys282 while maintaining good packing between the subunits, and reconnected the tether. This approach created UreB conformation 2 (Figure 2.3, bottom panels). A close-up view highlighting the repositioning of UreB to achieve conformation 2 and allow cross-linking is depicted in Figure 2.4.

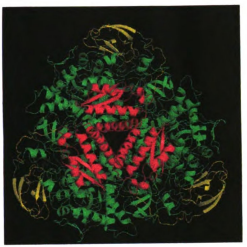
FIGURE 2.2. Tether and hinge regions between UreB and UreC from the crystallographic structure of urease. (A) The native urease structure, with ribbons colored red for UreA, blue for UreB (except for its hinge and tether to UreC shown in white), and green for UreC. (B) An expanded view of the region encircled in yellow in panel A. The N-terminus of UreB (residues 2-8) forms the terminal strand of a beta sheet with UreC. UreB residues 11-19 together with UreC residues 2-6 and 13-41 form a flexible linkage between the main domain of UreB (blue ribbons in panel A) and the disk formed by $(UreAC)_3$ (red and green ribbons in panel A). Sites relevant to flexibility probing mutations, UreB Pro10, Gly11 and Gly18, are rendered as beads. (C) The same view as panel B, colored in terms of ProFlex flexibility analysis of the crystal structure (PDB entry 1FWJ). The N-terminus of UreB partitions from a rigid region (colored blue; region 1) to a flexible hinge (colored gold; region 2) which connects to the globular domain of UreB (shown in blue ribbons in panel A). The terminus of UreC is highly flexible (red), whereas residues in UreC that intervene between regions 1 and 2 are isostatic, or barely rigid, as shown in grey.

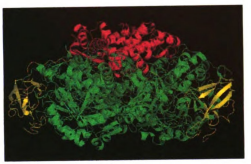

A

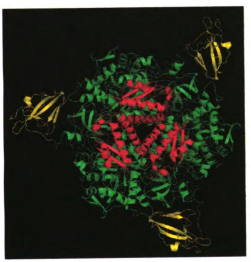

B
Gly18
Glu41
Gly11
Pro10
Ser2
Met1
UreC
UreB

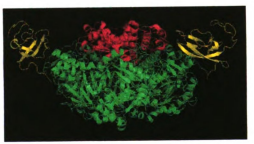

C
Region 2
Region 1
Mutually Rigid
Mutually Flexible

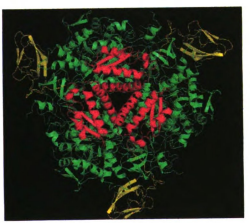
Figure 2.3: Two views of (A) the native conformation of urease, (B) UreB conformation 1 (torsionally adjusted UreB Gly11 and Gly18 residues), and (C) UreB conformation 2 (severed linker, docked domain, and reconnected linker). UreA is depicted in red, UreB in yellow, and UreC in green.

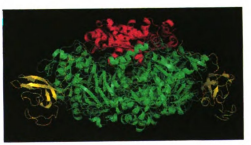


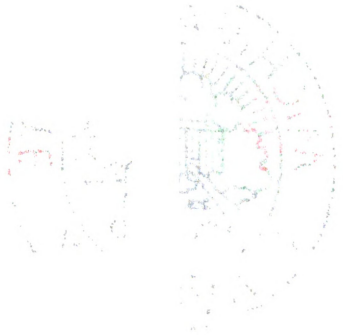












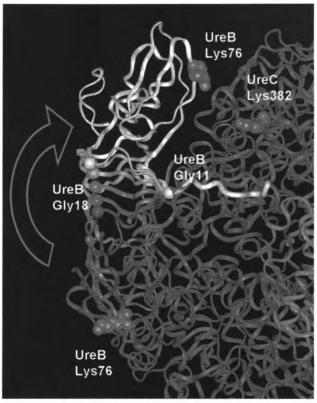


FIGURE 2.4. Close-up of the repositioning of UreB from its crystallographic position (dark blue; PDB 1FWJ) to a position (white) in which UreB Lys76 can cross link with UreC Lys382 (pink CPK spheres), opening access to the active site. The range of motion of UreB hinge residues resulting in this rotation of UreB is shown by the series of blue to lighter blue conformations of residues 11-19 between the UreB crystallographic and cross-linked open positions.

Table 2.1: Polar Interactions of Region 1 (UreB residues 2 – 8)

Residue 1	Donor Atom	Residue 2	Acceptor Atom	Energy (Kcal/mol)
UreB Gly4	N	UreC Ala24	O	-2.73
UreC Lys20	N	UreB His7	O	-3.62
UreB His7	N	UreC Lys20	O	-4.58
UreC Arg22	N	UreB Glu5	O	-6.52
UreC Arg22	NH1	UreB Glu5	OE2	-8.11

Table 2.2: Hydrophobic Interactions of Region 1 (UreB residues 2 – 8)

Residue 1	Atom 1	Residue 2	Atom 2
UreB Glu5	CG	UreC Trp29	CH2
UreB Tyr6	CD1	UreC Val21	CG2
UreB His7	CB	UreC Trp29	CZ3
UreB Val8	CG1	UreC Arg6	CB
UreB Val8	CG2	UreC Ala10	CB

Table 2.3: Polar Interactions of Region 2 (UreB residues 11 – 19)

Residue 1	Donor Atom	Residue 2	Acceptor Atom	Energy (Kcal/mol)
UreC Arg6	N	UreB Gly11	O	-3.45
UreC Ile4	N	UreB Ile13	O	-4.53
UreB Leu15	N	UreC Ser2	O	-4.73
UreB Asn16	N	UreC Tyr39	O	-4.87
UreB Ile13	N	UreC Ile4	O	-7.18
UreB Arg19	NH2	UreC Glu41	OE2	-7.34
UreB Arg19	NH1	UreC Glu41	OE2	-8.88

Table 2.4: Hydrophobic Interactions of Region 2 (UreB residues 11 – 19)

Residue 1	Atom 1	Residue 2	Atom 2
UreB Ile13	CB	UreC Ile4	CG2
UreB Ile13	CG2	UreC Tyr39	CG

Table 2.5: Polar Interactions of Region 3 (UreB residues 20 – 101)

Residue 1	Donor Atom	Residue 2	Acceptor Atom	Energy (Kcal/mol)
UreB His39	NE2	UreC Glu41	OE2	-1.03
UreB Arg60	NH2	UreC Glu41	OE1	-1.43
UreB Arg60	NE	UreC Glu41	OE2	-1.84
UreC Lys49	NZ	UreB Gly66	O	-2.11
UreB His87	N	UreC Pro102	O	-2.20
UreB Ala85	N	UreC Ile104	O	-2.84
UreB Ala89	N	UreC Asp103	O	-3.09
UreB His87	ND1	UreC Asp103	OD1	-4.27
UreB His39	NE2	UreC Glu41	OE1	-8.55
UreB Arg60	NH2	UreC Glu41	OE2	-8.98

Table 2.6: Hydrophobic Interactions of Region 3 (UreB residues 20 – 101)

Residue 1	Atom 1	Residue 2	Atom 2
UreB Tyr40	CD2	UreC Met55	CE
UreB Phe84	CG	UreC Ile104	CG2
UreB Phe84	CD1	UreC Ile104	CB
UreB Phe84	CE1	UreC Ile104	CD1
UreB Phe91	CB	UreC Gln59	CG
UreB Phe93	CE1	UreC Met55	CE
UreB Phe93	CE1	UreC Met55	SD
UreB Phe93	CE1	UreC Met55	CG
UreB Phe93	CZ	UreC Met55	CG

Both approaches yielded substantially similar placement of UreB at the periphery of $(UreAC)_3$ due to the strong constraints placed by maintaining the anchoring interactions of UreB residues 2-10 while meeting the cross-linking distance between UreB Lys76 and UreC Lys382.

Mutagenesis of Hinge Residues. To directly test the importance of putative UreB hinge region residues Gly11 and Gly18 in urease activation, their codons were independently modified to encode Pro residues that would restrict hinge flexibility. Constructs encoding the G11P and G18P variants of UreB were created and used to substitute for the wild-type sequence in a plasmid containing the complete urease gene cluster. The mutated plasmids were transformed into host *E. coli* cells, and urease overexpression was shown to be comparable in the control and mutant strains by using Western blots (data not shown). Urease activity in cell extracts containing the G18P variant of UreB was similar to that for extracts containing wild-type enzyme; in contrast, extracts containing the G11P mutant displayed 15-50 % (depending on the preparation) of the activity of the control strain.

Urease containing UreB G11P was purified from the mutant strain and subjected to metal analysis. Whereas control enzyme exhibits a specific activity of 2,200 ± 200 μmol min^{-1} (mg $protein^{-1}$) and contains 2.1 ± 0.3 nickel ions per active site (*43*), the purified UreB G11P variant protein possessed a specific activity of approximately 440 μmol min^{-1} (mg $protein^{-1}$) and only contained 1.67 nickel ions per active site (single determination with an estimated error of <10 %). For comparison, 2.13 to 1.74 nickel ions per active site were present after treating $(UreABC)_3$ with the metal ion or nickel plus bicarbonate yielding specific activities of 0 and 442 μmol min^{-1} (mg $protein)^{-1}$ (*20*); thus, high nickel content can be associated with inactive protein. These results suggest both a deficiency in nickel incorporation and formation of a less effective dinuclear site in the

mutant protein. Significantly, the mutant urease protein was resolved into two fractions during phenyl-Sepharose chromatography (Figure 2.5). The highly purified urease analyzed above was obtained by elution with buffer lacking salt, as in the case of wild-type enzyme. In addition, a nearly inactive urease-containing fraction was obtained by subsequent washing of the resin with water; such a second pool of enzyme is not apparent when purifying wild-type urease. The second pool of urease contained four major contaminating proteins that co-migrated with UreD (M_r 29,807), UreG (M_r 21,943), UreF (M_r 25,221), and UreE (M_r 17,558) (note that the peptides do not migrate precisely according to their known size). A Western blot analysis with anti-UreE antibodies confirmed the identity of UreE in this sample. The finding of this apparent complex is compatible with the need for flexibility in the hinge region of UreB to achieve accessory protein dissociation. The deleterious effects on urease activity, nickel content, and accessory protein dissociation that come from restricting the motion of UreB Gly11 by Pro substitution is consistent with the observation that large changes in main-chain Φ and Ψ values of UreB Gly11 are needed to place Lys76 of this subunit within cross-linking distance of Lys382 in UreC. The neighboring residue, UreB Pro10, already limits the accessible Φ angles so the G11P mutant would severely restrict the conformations available to the hinge. We hypothesize that the hinge-like motion of UreB relative to UreC upon binding of UreD and UreF is associated with the opening of the active site for activation.

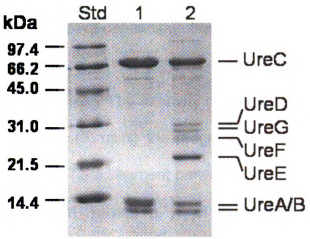


FIGURE 2.5. Two pools of the UreB G11P mutant urease resolved by phenyl-Sepharose chromatography. Molecular weight standards (Std), the purified active mutant urease (lane 1), and the very low activity complex containing mutant urease (lane 2) were examined by SDS-PAGE using a 13.5% acrylamide gel and stained with Coomassie brilliant blue.

Small-Angle X-Ray Scattering Measurements and Analyses. SAXS data collected for the three complexes studied are shown in Figure 2.6. Instrument stability issues, primarily due to temperature fluctuations in the facility, caused the differences in usable minimum q shown in the graph. The inset curves in Figure 2.6 are the Guinier regions for the three data sets, and correspond to R_g of 32.7 ± 2.4 Å , 40.3 ± 2.3 Å, and 50.6 ± 2.5 Å for native urease, (UreABC-UreD)$_3$, and (UreABC-UreDF)$_3$, respectively. In all cases, the Guinier regions are linear, indicative of monodisperse scattering particles. The $P(r)$ curves derived from the SAXS data are shown in Figure 2.7. The R_g for urease determined from the $P(r)$ fitting was 35.7 ± 0.8 Å, with a d_{max} of 95 ± 5 Å. The values of R_g for the (UreABC-UreD)$_3$ and (UreABC-UreDF)$_3$ complexes were 44.9 ± 0.7 Å and 53.7 ± 1.4 Å, respectively. The d_{max} of the (UreABC-UreD)$_3$ complex was 130 ± 8 Å, while that of the (UreABC-UreDF)$_3$ complex was 155 ± 10 Å. The agreement

between the Guinier- and GNOM-derived R_g values is reasonable considering the very different methods of obtaining the values and estimating the uncertainties.

Models of the complexes. The intensity profile calculated from the wild-type urease crystal structure (*3*) using the program ORNL_SAS (*41*) is shown with the data in Figure 2.6. The agreement between the measured data and the simulated profile is excellent, having a χ^2 of 0.493. The uncertainties in measured SAXS intensities derive from specific assumptions about the counting statistics. In cases of relatively low count rates, the error propagation can result in uncertainties that overestimate the true uncertainty in the measurement, making it possible to have χ^2 significantly less than one. An inspection of the fidelity of the model profile to the data is required to ensure that the quality of the fit is truly excellent, as is the case here.

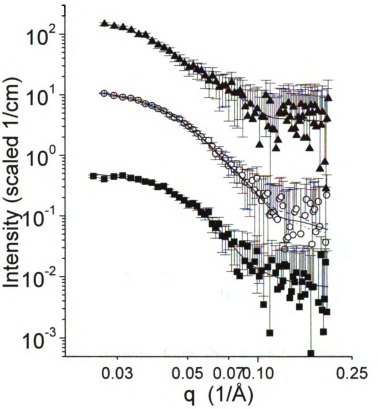


FIGURE 2.6. $I(q)$ curves derived from the scattering data for urease (■), $(UreABC\text{-}UreD)_3$ (○) and $(UreABC\text{-}UreDF)_3$ (▲). The lines are the model fits to the data using the crystal structure of urease (PDB 1FWJ) (solid line), with UreB Lys11/Lys18 torsionally adjusted to allow cross-linking of UreB Lys76 to UreC Lys382 (dashed line), and UreB docked to UreAC from the crystal structure, allowing cross-linking of UreB Lys76 to UreC Lys382 (dotted line). The curves have been offset by a multiplicative factor for clarity. The curves have been offset for clarity, and the region of data covered by the line indicates the range of data used for the fitting.

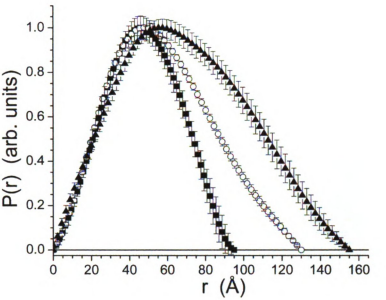
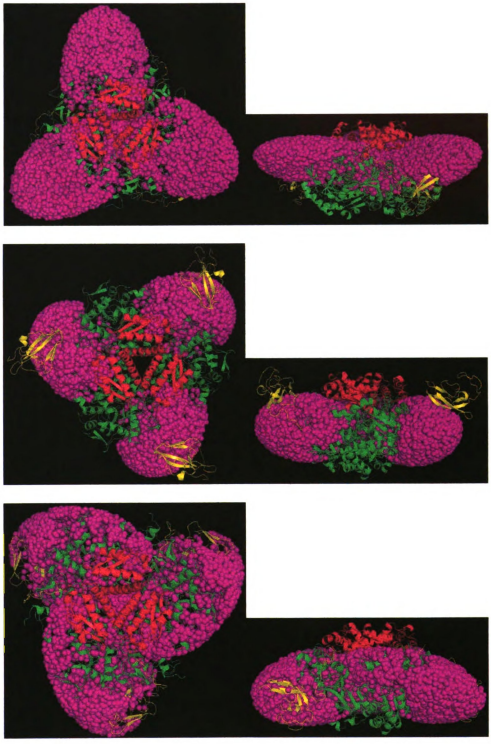


FIGURE 2.7. $P(r)$ curves derived from the scattering data for urease (■), $(UreABC\text{-}UreD)_3$ (○), and $(UreABC\text{-}UreDF)_3$ (▲). To simplify comparison, the curves have been scaled to have a value of 1.0 at the peak.

Models of $(UreABC\text{-}UreD)_3$ were generated by adding UreD ellipsoids to the wild-type urease structure and to $(UreABC)_3$ with the two alternative UreB conformations (one from changes in torsional angles of Gly11 and Gly18 in the hinge and the other from cleaving the tether, docking of the major UreB domain, and reconnecting the linker). Ellipsoids were used because no structure or homology model is available for any UreD. In all cases, the overall structures of the final complexes were very similar. The best models had UreD ellipsoids added to the vertices of $(UreABC)_3$ near the UreB subunit such that the total structure has a planar, triangular character, as can be seen in three pairs of panels in Figure 2.8. The best three model intensity profiles for the three different starting structures have χ^2 of 0.218, 0.252 and 0.224 when starting with the

Figure 2.8: Four views (two with UreABC in ribbon and two with UreABC in spacefilling representation) of the best models of $(UreABC\text{-}UreD)_3$ generated by adding ellipsoids for UreD to the (A) native urease conformation, (B) UreB conformation 1, and (C) UreB conformation 2. UreA is depicted in red, UreB in yellow, UreC in green, and UreD in purple.



native structure, UreB conformation 1 (torsionally-adjusted), and UreB conformation 2 (docked), respectively. In all cases, the fits of the profiles to the data are excellent and suggest that all of the structures are reasonable. It is important to note that the three models all have the same general shape, which is the most reliable result of the modeling considering the method of building the models and the quality of the data. The addition of UreD results in a planar, triangular structure. The specific details of the interaction of UreD with UreB cannot be effectively differentiated with the SAXS data and modeling, in spite of the differences in χ^2, because the way in which the ellipsoids were allowed to conform to the surface of the starting structure enabled them to fill space in such a way that the final structures have the same general shape. Higher quality data would not have completely eliminated this ambiguity from the modeling results. Only a reliable high-resolution structure of UreD, which does not exist, would have made it possible to differentiate between the different UreB models based on the SAXS data alone. The $(UreABC\text{-}UreD)_3$ results are in agreement with UreD interacting with UreB as suggested by chemical cross-linking (*28*), but would require additional flexibility to accommodate the observed cross-linking of UreD to UreC Lys401.

The best models for $(UreABC\text{-}UreDF)_3$ were created by adding ellipsoids to represent appropriate molecular volumes of UreD and UreF to the $(UreABC)_3$ crystal structure and the two alternative UreB conformations (to allow for cross-linking of UreB Lys76 with UreC Lys382). As above, no structure or model is available for UreD; however, a homology model was reported for UreF from

Bacillus pasteurii (*44*). Given the high E-value (4.21) from the 3D-PSSM server and the fact the model does not represent *K. aerogenes* UreF, we felt justified in using ellipsoids to represent this protein. The two alternate UreB conformations, one produced by docking and the other by torsional adjustments in UreB residues Gly11 and Gly18, are similar, and in fact resulted in similar placements of UreD and UreF in the best-fitting SAXS models (shown for the docked conformation in Figure 2.9). The best UreB conformation 1 (torsionally-adjusted) and UreB conformation 2 (docked) structures fit the scattering data very well and have χ^2 of 0.093 and 0.094, respectively, which is slightly better than the χ^2 of 0.096 observed for the native structure. The fit of the model profiles to the data are all excellent, so it is not possible to discriminate between the SAXS models produced from the different UreB models for the reasons provided above. The overall shape of the complex, which can be reliably extracted from the data, is very consistent between the three models, having a planar, triangular character where the additional mass corresponding to UreD and UreF are located near the vertices, slightly above the plane defined by the rest of the structure. Reliable high-resolution structures of the two subunits could provide an effective means to discriminate between the different UreB structures based on the SAXS data alone, but no such structures exist for either UreD or UreF. The model depicted in Figure 2.9 appears to build on the models of $(UreABC\text{-}UreD)_3$, with the UreD and UreF ellipsoids positioned pair wise at the vertices of the $(UreABC)_3$ structure. In this case UreB, UreD, and UreF essentially add onto the edge of the disk formed primarily by the UreC trimer, in which UreA forms the hub (Figure

2.2). These structures are consistent with immunological results that show anti-UreD antibodies recognize UreD within $(UreABC\text{-}UreD)_3$, but not within $(UreABC\text{-}UreDF)_3$, suggesting that UreF partially masks UreD (*16*).

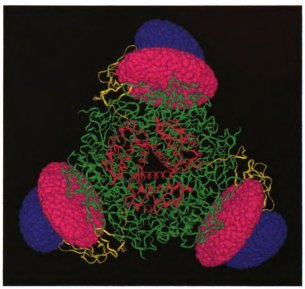


FIGURE 2.9. Predicted positioning of UreD and UreF relative to the crystallographic structure of $(UreABC)_3$, based on best-fit models to SAXS data. The best-fit models resulted in packing of UreD and UreF against UreB near a vertex of the $(UreAC)_3$ disk. A representative example is illustrated. UreA, UreB, and UreC are rendered in red, yellow, and green ribbons, respectively. UreD and UreF from SAXS results are rendered as solid ellipsoids colored purple and magenta, respectively. The non-interpenetrating volumes of the UreD and UreF ellipsoids accounts for the appropriate molecular weight of each subunit.

DISCUSSION

In this work we combined multi-scale modeling and sparse experimental constraints to obtain insight into a flexible molecular assembly, the urease activation complex. In particular, we used flexibility analysis to provide evidence that the major domain of UreB can move in a hinge-like motion to allow sufficiently close juxtaposition of UreB Lys76 with UreC Lys382 to form the

reported chemical cross-link between these residues, as previously hypothesized (*28*). The UreB G11P variant, which is likely to rigidify the hinge region, was shown to lead to reduced levels of urease activation and lower nickel content while also sequestering a significant portion of the urease apoprotein in an ineffective activation complex that includes all four of the known *K. aerogenes* accessory proteins. These results support the importance of the hinge region in urease activation, although we cannot exclude alternative explanations to account for the properties of the mutant protein. SAXS analysis of urease, $(UreABC\text{-}UreD)_3$, and $(UreABC\text{-}UreDF)_3$ were used to provide evidence consistent with UreD and UreF binding near UreB. Notably, the $(UreABC\text{-}UreDF)_3$ data are best fit by models in which UreB occupies an altered conformation that would account for the previously described cross-linking results (*28*). Significantly, the predicted structures of $(UreABC\text{-}UreDF)_3$ containing the alternative UreB conformations provide access to the nascent active site and may be a critical step in urease activation.

Comparison of the *H. pylori* urease structure (PDB entry 1E9Z) with that of *K. aerogenes* urease discussed here provides additional support for the proposed sites of UreD and UreF interaction with UreB, at the periphery of the $(UreAC)_3$ disk. The *H. pylori* UreA subunit (corresponding to a fusion of UreA and UreB in the *K. aerogenes* enzyme) contains a fold that matches the *K. aerogenes* UreB fold, but also contains residues that add to one side of this shared fold in a similar position to where we predict UreD and UreF bind. A viral protein (PDB entry 1C5E) also contains this fold with an additional domain in the

same region as the added domain in *H. pylori* UreA. The residues buried on the part of the *H. pylori* subunit in common with *K. aerogenes* UreB are the same in both the *H. pylori* and viral proteins, suggesting that this region of UreB has evolved to interact with other domains or proteins.

ACKNOWLEDGMENTS

I would like to thank Scott Mulrooney and Kimberly Anderson for their assistance.

REFERENCES

1. Hausinger, R. P., and Karplus, P. A. (2001) Urease, in *Handbook of Metalloproteins* (Wieghardt, K., Huber, R., Poulos, T. L., and Messerschmidt, A., Eds.), pp 867-879, John Wiley & Sons, Ltd., West Sussex, U.K.

2. Ciurli, S., and Mangani, S. (2001) Nickel-containing enzymes, in *Handbook on Metalloproteins* (Bertini, I., Sigel, A., and Sigel, H., Eds.), pp 669-708, Marcel Dekker, New York, NY.

3. Jabri, E., Carr, M. B., Hausinger, R. P., and Karplus, P. A. (1995) The crystal structure of urease from *Klebsiella aerogenes*, *Science 268*, 998-1004.

4. Pearson, M. A., Michel, L. O., Hausinger, R. P., and Karplus, P. A. (1997) Structure of Cys319 variants and acetohydroxamate-inhibited *Klebsiella aerogenes* urease, *Biochemistry 36*, 8164-8172.

5. Benini, S., Rypniewski, W. R., Wilson, K. S., Miletti, S., Ciurli, S., and Mangani, S. (1999) A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from *Bacillus pasteurii*: why urea hydrolysis costs two nickels, *Structure 7*, 205-216.

6. Ha, N.-C., Oh, S.-T., Sung, J. Y., Cha, K.-A., Lee, M. H., and Oh, B.-H. (2001) Supramolecular assembly and acid resistance of *Helicobacter pylori* urease, *Nature Structure Biology 8*, 505-509.

7. Sheridan, L., Wilmont, C. M., Cromie, K. D., van der Logt, P., and Phillips, S. E. V. (2002) Crystallization and preliminary X-ray structure determination of jack bean urease with a bound antibody fragment, *Acta Crystallogr. D58*, 374-376.

8. Mulrooney, S. B., and Hausinger, R. P. (2003) Nickel uptake and utilization by microorganisms, *FEMS Microbiol. Rev. 27*, 239-261.

9. Quiroz, S., Kim, J. K., Mulrooney, S. B., and Hausinger, R. P. (2007) Chaperones of nickel metabolism, in *Metal Ions in Life Sciences* (Sigel, A., Sigel, H., and Sigel, R. K. O., Eds.), pp 519-544, John Wiley & Sons, New York.

10. Hausinger, R. P., and Zamble, D. B. (2007) Microbial physiology of nickel and cobalt, in *Molecular Microbiology of Heavy Metals* (Nies, D. H., and Silver, S., Eds.), Springer.

11. Kim, J. K., Mulrooney, S. B., and Hausinger, R. P. (2005) Biosynthesis of active *Bacillus subtilis* urease in the absence of known urease accessory proteins, *J. Bacteriol. 187*, 7150-7154.

12. Hausinger, R. P., Colpas, G. J., and Soriano, A. (2001) Urease: a paradigm for protein-assisted metallocenter assembly, *ASM News 67*, 78-84.

13. Mulrooney, S. B., and Hausinger, R. P. (1990) Sequence of the *Klebsiella aerogenes* urease genes and evidence for accessory proteins facilitating nickel incorporation, *J. Bacteriol. 172*, 5837-5843.

14. Lee, M. H., Mulrooney, S. B., Renner, M. J., Markowicz, Y., and Hausinger, R. P. (1992) *Klebsiella aerogenes* urease gene cluster: sequence of *ureD* and demonstration that four accessory genes (*ureD*, *ureE*, *ureF*, and *ureG*) are involved in nickel metallocenter biosynthesis, *J. Bacteriol. 174*, 4324-4330.

15. Park, I.-S., Carr, M. B., and Hausinger, R. P. (1994) *In vitro* activation of urease apoprotein and role of UreD as a chaperone required for nickel metallocenter assembly, *Proc. Natl. Acad. Sci. USA 91*, 3233-3237.

16. Moncrief, M. B. C., and Hausinger, R. P. (1996) Purification and activation properties of UreD-UreF-urease apoprotein complexes, *J. Bacteriol. 178*, 5417-5421.

17. Park, I.-S., and Hausinger, R. P. (1995) Evidence for the presence of urease apoprotein complexes containing UreD, UreF, and UreG in cells that are competent for in vivo enzyme activation, *J. Bacteriol. 177*, 1947-1951.

18. Soriano, A., and Hausinger, R. P. (1999) GTP-dependent activation of urease apoprotein in complex with the UreD, UreF, and UreG accessory proteins, *Proc. Natl. Acad. Sci. USA 96*, 11140-11144.

19. Park, I.-S., and Hausinger, R. P. (1995) Requirement of carbon dioxide for in vitro assembly of the urease nickel metallocenter, *Science 267*, 1156-1158.

20. Park, I.-S., and Hausinger, R. P. (1996) Metal ion interactions with urease and UreD-urease apoproteins, *Biochemistry 35*, 5345-5352.

21. Soriano, A., Colpas, G. J., and Hausinger, R. P. (2000) UreE stimulation of GTP-dependent urease activation in the UreD-UreF-UreG-urease apoprotein complex, *Biochemistry 39*, 12435-12440.

22. Song, H. K., Mulrooney, S. B., Huber, R., and Hausinger, R. P. (2001) Crystal structure of *Klebsiella aerogenes* UreE, a nickel-binding metallochaperone for urease activation, *J. Biol. Chem. 276*, 49359-49364.

23. Mulrooney, S. B., Ward, S. K., and Hausinger, R. P. (2005) Purification and properties of the *Klebsiella aerogenes* UreE metal-binding domain, a functional metallochaperone of urease, *J. Bacteriol. 187*, 3581-3585.

24. Moncrief, M. B. C., and Hausinger, R. P. (1997) Characterization of UreG, identification of a UreD-UreF-UreG complex, and evidence suggesting that a nucleotide-binding site in UreG is required for in vivo metallocenter assembly of *Klebsiella aerogenes* urease, *J. Bacteriol. 179*, 4081-4086.

25. Colpas, G. J., Brayman, T. G., Ming, L.-J., and Hausinger, R. P. (1999) Identification of metal-binding residues in the *Klebsiella aerogenes* urease nickel metallochaperone, UreE, *Biochemistry 38*, 4078-4088.

26. Musiani, F., Zambelli, B., Stola, M., and Ciurli, S. (2004) Nickel trafficking: insights into the fold and function of UreE, a urease metallochaperone, *J. Inorg. Biochem. 98*, 803-813.

27. Jabri, E., and Karplus, P. A. (1996) Structures of the *Klebsiella aerogenes* urease apoprotein and two active-site mutants, *Biochemistry 35*, 10616-10626.

28. Chang, Z., Kuchar, J., and Hausinger, R. P. (2004) Chemical crosslinking and mass spectrometric identification of sites of interaction for UreD, UreF, and urease, *J. Biol. Chem. 279*, 15305-15313.

29. Todd, M. J., and Hausinger, R. P. (1989) Competitive inhibitors of *Klebsiella aerogenes* urease. Mechanisms of interaction with the nickel active site, *J. Biol. Chem. 264*, 15835-15842.

30. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature (London) 227*, 680-685.

31. Mulrooney, S. B., Pankratz, H. S., and Hausinger, R. P. (1989) Regulation of gene expression and cellular localization of cloned *Klebsiella aerogenes* (*K. pneumoniae*) urease, *J. Gen. Microbiol. 135*, 1769-1776.

32. Lee, M. H., Pankratz, H. S., Wang, S., Scott, R. A., Finnegan, M. G., Johnson, M. K., Ippolito, J. A., Christianson, D. W., and Hausinger, R. P. (1993) Purification and characterization of *Klebsiella aerogenes* UreE protein: a nickel-binding protein that functions in urease metallocenter assembly, *Prot. Science 2*, 1042-1052.

33. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248-254.

34. Weatherburn, M. W. (1967) Phenol-hypochlorite reaction for determination of ammonia, *Anal. Chem. 39*, 971-974.

35. Jacobs, D. J., Rader, A. J., Kuhn, L. A., and Thorpe, M. F. (2001) Protein flexibility predictions using graph theory, *Proteins 44*, 150-165.

36. Hespenheide, B. M., Rader, A. J., Thorpe, M. F., and Kuhn, L. A. (2002) Identifying protein folding cores from the evolution of flexible regions during unfolding, *J. Molec. Graphics Modeling 21*, 195-207.

37. Rader, A. J., Hespenheide, B. M., Kuhn, L. A., and Thorpe, M. F. (2002) Protein unfolding: rigidity lost, *Proc. Natl. Acad. Sci. USA 99*, 3540-3545.

38. Woodward, J. D., Pickel, J. M., Anovitz, L. M., Heller, W. T., and Rondinone, A. J. (2006) Self-assembled colloidal crystals from ZrO_2 nanoparticles, *J. Phys. Chem. B 100*, 19456-19460.

39. Guinier, A., and Fournet, G. (1955) *Small-Angle Scattering of X-rays*, John Wiley & Sons, New York, NY.

40. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria, *J. Appl. Crystallogr. 25*, 495-503.

41. Tjioe, E., and Heller, W. T. (2007) ORNL_SAS: software for calculation of small-angle scattering intensities from bio-macromolecular structures, *J. Appl. Crystallogr. 40*, 782-785.

42. Heller, W. T. (2006) Ellstat: Shape modeling for solution small-angle scattering of proteins and protein complexes with automated statistical characterization, *J. Appl. Crystallogr. 39*, 671-675.

43. Todd, M. J., and Hausinger, R. P. (1987) Purification and characterization of the nickel-containing multicomponent urease from *Klebsiella aerogenes*, *J. Biol. Chem. 262*, 5963-5967.

44. Salomone-Stagni, M., Zambelli, B., Musiani, F., and Ciurli, S. (2007) A model-based proposal for the role of UreF as a GTPase-activating protein in the urease active site biosynthesis, *Proteins 68*, 749-761.

Chapter 3

Mutagenesis of the *Klebsiella aerogenes* UreG Urease Accessory Protein: Effects on UreG Properties and Urease Activation

ABSTRACT

UreG is a GTPase required for assembling the nickel-containing active site of urease. This urease accessory protein was previously purified from *Klebsiella aerogenes*, *Bacillus subtilis,* and *Mycobacterium tuberculosis*, with pronounced differences observed in their respective properties. This work describes an improved purification method for *K. aerogenes* UreG that utilizes a biotin tag, where the fusion peptide was shown to not interfere with urease activation. Although UreG can form a disulfide-linked dimer, we show that the dimer is not required for the function of UreG *in vivo*. The monomeric protein binds 2 nickel ions per molecule (K_d = 12.5 μM), whereas the oxidized protein exhibits greatly reduced metal binding capacity. Several residues were targeted for mutagenesis, including four (Cys72, His74, Ser111, and Ser115) with possible sequence similarity to the dinuclear zinc ligands of the structurally characterized HypB GTPase of *Methanocaldococcus jannaschii* and others (Lys20, Asp49, Glu68, Asp80) thought to function in GTPase activity or metal binding. Single and double substitutions of these UreG amino acids exhibited little effect on nickel binding, but most of these alterations abolished UreG's ability to activate urease. The biotin tag on UreG was used to isolate a novel complex containing urease apoprotein along with UreD, UreF, UreG, and UreE. In contrast, the D80A variant form of UreG interacted only with UreE revealing a new heterodimeric species. These results suggest a critical role for Asp80 in stabilizing the larger activation complex.

INTRODUCTION

Urease, a nickel-containing metalloenzyme found in plants and microorganisms, catalyzes the hydrolysis of urea to form ammonia and carbamate, which spontaneously decomposes to carbon dioxide and ammonia (*1*). The structure of the enzyme has been characterized for several species (*2-5*), and in all cases the dinuclear nickel metallocenters are deeply buried in structural subunits that exhibit three-fold symmetry. With the possible exception of the protein of *Bacillus subtilis* (*6*), ureases require a series of accessory proteins to assemble their active sites (*7, 8*). The proposed urease activation process (depicted in Fig. 3.1) begins with the structural subunits (UreA, UreB and UreC in the case of the enterobacterium *Klebsiella aerogenes*, our model system) assembling into the urease apoprotein $(UreABC)_3$ (*9, 10*). The UreD, UreF, and UreG accessory proteins sequentially associate with the apoprotein to form the $(UreABC\text{-}UreD)_3$ (*11*), $(UreABC\text{-}UreDF)_3$ (*12*), and $(UreABC\text{-}UreDFG)_3$ (*13*) activation complexes. Finally, in a process that requires GTP, CO_2, and the metallochaperone UreE (which specifically delivers the nickel ions for urease activation), the active site is assembled and the accessory proteins are released from the active enzyme (*14, 15*).

The precise roles of the UreD, UreF and UreG accessory proteins are not well understood. *K. aerogenes* UreD is insoluble when expressed by itself and no sequence-related proteins have been structurally characterized, thus preventing a better understanding of its function. UreF from *K. aerogenes* also is insoluble when synthesized separately from the other urease gene products; however, a

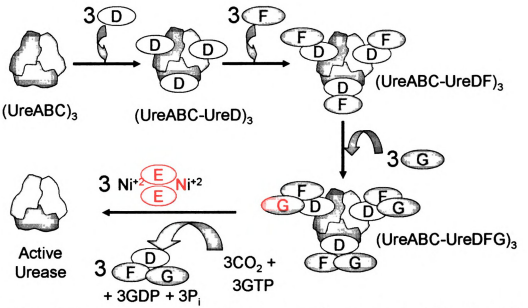


Figure 3.1. Proposed urease activation process. Urease apoprotein $(UreABC)_3$ is synthesized with the nascent active site lacking nickel and carbamylation of Lys217. Urease accessory proteins UreD, UreF, and UreG bind the apoprotein in a sequential manner to form the $(UreABC\text{-}UreDFG)_3$ activation complex. Urease activation requires carbamylation of Lys217 by CO_2, provision of nickel ions by the UreE metallochaperone, and GTP hydrolysis accompanied by the release the accessory proteins.

UreE-UreF fusion protein is soluble and has been partially characterized (*16*). A fold recognition method was used to create a homology model for UreF of *Bacillus pasteurii*, and it was proposed to function as a GTPase-activating protein (*17*). Purified recombinant UreG proteins (subunit M_r 22,000 - 23,000) of *K. aerogenes*, *B. pasteurii*, and *Mycobacterium tuberculosis* are soluble and contain motifs found in GTPases, although the GTPase activity is very low or not detectable (*13, 18, 19*). Mutation of Lys20 or Thr21 in the GXGKT P-loop motif (one of the GTPase motifs) of the *K. aerogenes* protein abolishes urease activation (*13*) and this region is critical to *in vitro* activation of the $(UreABC\text{-}UreDFG)_3$ complex (*15*). Whereas the *K. aerogenes* protein is monomeric (*13*),

the other two UreG proteins are dimeric with the subunits joined by a disulfide bridge involving Cys68 (*B. pasteurii*) and probably Cys90 (*M. tuberculosis)* (*19, 20*). UreG of *B. pasteurii* binds two zinc ions per dimer (K_d 42 μM) or four nickel ions per dimer (K_d 360 μM), perhaps utilizing Glu64, Cys68 (i.e., the residue proposed to participate in a disulfide), and His70 as metal ligands (*18*) although no experiments were performed to support these assignments. No crystal structure is available for any UreG; however, the crystal structure of the related protein HypB from *Methanocaldococcus jannaschii* was reported in 2006 (*21*). HypB is an accessory protein that participates in the metallocenter assembly of Ni-Fe hydrogenases (reviewed in (*7, 22*) and chapter 1). The crystal structure reveals two types of zinc binding sites: a mononuclear site in each subunit involving His100 and His104 (corresponding residues are not present in UreG sequences) and a non-symmetrical dinuclear binding site at the subunit interface (Fig. 3.2). The metal-binding residues of the dinuclear site in *M. jannaschii* HypB (Cys95, His96, and Cys127) most likely correspond to Cys72, His74, and either Ser111 or Ser115 in *K. aerogenes* UreG (or Cys68, His70, and Ser107 or Ser111 in the *B. pasteurii* protein). Of interest, the *Escherichia coli* HypB sequence retains the ligands of the dinuclear center (Cys166, His167, and Cys198), but lacks the His residues associated with the mononuclear site. In addition, the *E. coli* protein contains an amino-terminal extension with a CXXCGC motif, not found in *M. jannaschii* HypB or in UreG proteins, responsible for high affinity (sub-picomolar K_d) binding of a nickel ion (23) .

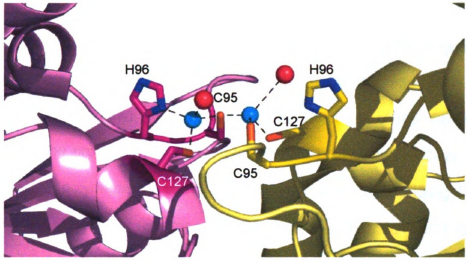


Figure 3.2. HypB dinuclear zinc site. HypB from *M. jannaschii* (PDB code 2HF8) is a dimeric protein (with the individual subunits depicted in yellow and pink) that forms an asymmetrical dinuclear site coordinated by Cys95, His96, and Cys127. The zinc atoms are shown as cyan spheres and the water molecules are depicted as red spheres. Nitrogen atoms are colored in blue and sulfur atoms in orange.

In this chapter, I describe an improved purification procedure for *K. aerogenes* UreG that utilizes a biotin tag, I reexamine the quaternary structure of the protein, and I explore the effects of mutating several UreG residues on the properties of the protein, the ability to form activation complexes, and urease activation.

EXPERIMENTAL PROCEDURES

Vector Construction, Cell Growth, and Purification of Biotin-Tagged UreG – The *ureG* sequence was subcloned into pASK-IBA3plus and pASK-IBA5plus plasmids (IBA, Germany) to create vectors pIBA3+Gb and pIBA5+Gb (Table 3.1) that encode the protein with a biotin tag (denoted UreGb) at the C- or N-termini, respectively. A polymerase chain reaction (PCR) was performed using *Pfu*Turbo® Hotstart PCR Master Mix (Stratagene, USA) and the primers 5'-TACT GTC CCG CGG GATG AAC TCT TAT AAA CAC-3' and 5'-TACT GTC CTG CAG TTT GCC AAG CAT GCC TTT-3'. The first primer contains a *Sac*II restriction site and the second a *Pst*I restriction site that were used to subclone the fragment into pASK-IBA3plus. In a similar manner, the primers 5'-TACT GTC CCG CGG GG AAC TCT TAT AAA CAC CCG-3' and 5'- TACT GTC GGA TCC CTA TTT GCC AAG CAT GCC-3', containing restriction sites for *Sac*II and *Bam*HI respectively, were used to subclone the fragment into pASK-IBA5plus. The plasmids and PCR products were digested with the corresponding restriction enzymes and ligated to produce plasmids pIBA3+G and pIBA5+G. Isolated colonies of *E. coli* DH5α were transformed with the plasmids and grown at 37°C overnight in 10 mL of Luria broth (LB) media supplemented with 300 µg/mL of ampicillin. These cultures were used to inoculate 1 L of LB media supplemented with 300 µg/mL of ampicillin. The cultures were grown at 37°C for 4 h and induced overnight with 100 µl of 2 mg/ml anhydrotetracycline. The cells were harvested by centrifugation and resuspended in 1 mL of buffer W (100 mM Tris/HCl, pH 8.0, containing 150 mM NaCl and 1 mM EDTA) per g of cells and supplemented with 1 mM

TABLE 3.1. Plasmids used in this study

Plasmid	Description	*E. coli* strain	Reference
pKK17	*K. aerogenes ureDABCEFG* gene cluster inserted into pKK223-3	JM109	(*24*)
pKKGb	Modified pKK17 encoding UreGb	DH5α	This work
pKKGK20Ab, pKKGD49Ab, pKKGE68Ab, pKKGC72Ab, pKKGH74Ab, pKKGH74Cb, pKKGH74Nb, pKKGD80Ab, pKKGS111Ab, and pKKGS115Ab	Modified pKKGb encoding the K20A, D49A, E68A, C72A, H74A, H74C, H74N, D80A, S111A, and S115A variant forms of UreGb	DH5α	This work
pASK-IBA3plus	Plasmid for creating fusion proteins with a biotin tag at the C-terminus	DH5α	IBA
pASK-IBA5plus	Plasmid for creating fusion proteins with a biotin tag at the N-terminus	DH5α	IBA
pIBA3+Gb	Modified pASK-IBA3plus to encode biotin-tagged UreGb	DH5α	This work
pIBA5+Gb	Modified pASK-IBA5plus to encode biotin-tagged UreGb	DH5α	This work
pIBA3+GK20Ab, pIBA3+GD49Ab, pIBA3+GE68Ab, pIBA3+GC72Ab, pIBA3+GH74Ab, pIBA3+GH74Cb, pIBA3+GH74Nb, pIBA3+GS111Ab, pIBA3+GC72A/H74Ab, pIBA3+GC72A/S111A, and pIBA3+GH74A/S111Ab	Modified pIBA3+Gb to encode the K20A, D49A, E68A, C72A, H74A, H74C, H74N, S111A, C72A/H74A, C72A/S111A, and H74A/S111A variants of UreGb	BL21	This work

phenylmethylsulphonyl fluoride (PMSF) before sonification (Branson 450 sonifier, 5 repetitions, each of 2 min, at 3 W output power and 50% duty cycle). The lysate was centrifuged at 100,000 *g* at 4°C for 45 min and the supernatant was loaded onto a 1 ml Strep-Tactin column (IBA, Germany) previously equilibrated in buffer W. This column has an engineered streptavidin ligand that binds biotin with high affinity. The protein was eluted with desthiobiotin according to the manufacturer's instructions. For comparative studies, native UreG was purified as previously described (*13*).

Fractions containing UreGb, UreG, or mutants were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (*25*) using gels prepared with 13.5% acrylamide and stained with Coomassie brilliant blue. Molecular weight markers were obtained from Bio-Rad (Hercules, CA). Protein concentrations were determined by using a commercial assay (Bio-Rad, Hercules, CA) with bovine serum albumin as the standard. In some cases the protein was loaded onto a preparative Superdex-75 column (65 cm x 1.5 cm diam., Amersham, USA) equilibrated in 50 mM HEPES buffer, pH 7.4, containing 200 mM NaCl and chromatographed at 1 mL/min in this buffer for further purification.

Site-Directed Mutagenesis - Mutations were generated in pIBA3+G by using overlapping oligonucleotides containing the desired mutation (see Table 3.2). The PCR was performed with *Pfu*Turbo® Hotstart PCR Master Mix and the corresponding oligonucleotides. The products were digested with *Dpn*I for one h at 37°C and used to transform chemically competent *E. coli* DH5α cells. The mutations were confirmed by sequencing (Davis sequencing, Davis, CA, USA). The

mutated plasmids were purified and used to transform *E. coli* BL21 Gold competent cells (Stratagene, USA). Double mutants (C72A/H74A, H74/S111A and C72A/S111A) were prepared by using as a template a previous mutant. All mutants were expressed and purified as described for UreGb.

TABLE 3.2. Oligonucleotides used to generate *ureG* mutations

Purpose	Sequence
UreG mutation D49A	5' - GAC ATC TAT ACC AAA GAA *GCG* CAG CGC ATC CTC ACC GAA - 3'
UreG mutation E68A	5' - GAA CGC ATC GTC GGT GTG *GCG* ACC GGC GGC TGC CCG CAT - 3'
UreG mutation C72A	5' - GTC GGT GTG GAA ACC GGC GGC *GCG* CCG CAT ACG GCG ATC CGC GAA - 3'
UreG mutation H74A	5' - GAA ACC GGC GGC TGC CCG *GCA* ACG GCG ATC CGC GAA GAT - 3'
UreG mutation H74C	5' - GAA ACC GGC GGC TGC CCG *TGC* ACG GCG ATC CGC GAA GAT - 3'
UreG mutation H74N	5' - GAA ACC GGC GGC TGC CCG *AAT* ACG GCG ATC CGC GAA GAT - 3'
UreG mutation D80A	5' - CAT ACG GCG ATC CGC GAA *GCG* GCC TCA ATG AAC CTC GCC - 3'
UreG mutation S111A	5' - GAA AGC GGC GGC GAT AAC CTG *GCC* GCC ACC TTC AGC CCG GAG CTG - 3'
UreG mutation S115A	5' - AAC CTG AGC GCC ACC TTC *GCC* CCG GAG CTG GCG GAT CTG - 3'
Double UreG mutation C72A/H74A	5' - GTC GGT GTG GAA ACC GGC GGC *GCG* CCG GCA ACG GCG ATC CGC GAA - 3'

Circular Dichroism (CD) - Proteins were purified and concentrated to 0.2 mg/mL in 15 mM phosphate buffer, pH 7.6, containing 1 mM dithiothreitol (DTT). A 100 µL sample was placed into a Jasco J-710 spectropolarimeter and data collected between 180 and 300 nm with a 1 cm path length. The data were analyzed with the DICHROWEB server (*26*). The best fit was obtained using CDSSTR and set 4.

Analytical Gel Filtration Chromatography - Superdex-75 (45 cm x 1.0 cm diam., Amersham, USA) was used for analytical hydrodynamic radius assays. The buffer contained 50 mM HEPES, pH 7.4, containing 200 mM NaCl and other additives as indicated, with the flow at 1 ml/min.

Metal Quantification - The nickel content of freshly purified UreG and UreGb was assessed by using inductively coupled plasma-mass spectrometry at the University of Georgia Chemical Analysis Laboratory.

Nickel Binding - Purified proteins were dialyzed overnight against 50 mM HEPES buffer, pH 7.4, containing 200 mM NaCl, 10 mM EDTA, and 1 mM DTT, followed by dialysis 4 times against 50 mM HEPES buffer, pH 7.4, containing 200 mM NaCl. The samples were incubated for 30 min at 4°C with different concentrations of nickel, and centrifuged at 14,000 *g* in a tabletop centrifuge for 20 min using a Microcon® (Millipore, USA) centrifuge unit with a nominal molecular weight cut off of 10 kDa. A 100, 50 or 20 µL aliquot of the flow-through fraction was adjusted to 100 µL (as needed), mixed with 900 µL of 100 µM 4-(2-pyridylazo)-resorcinol (PAR), and analyzed spectrophotometrically to determine the amount of metal. The data were plotted and analyzed in Sigma Plot using the following equation where [Ni] is the concentration of free nickel ion, B_{max} is the maximum number of nickel ions bound per UreG peptide, Ni_b is the number of nickel ions bound per UreG and K_d is the dissociation constant.

$$Ni_b = (B_{max} \times [Ni]) / (K_d + [Ni])$$

Analysis of Cells Expressing the Urease Operon Encoding the UreGb Variants - Plasmid pKK17(*24*), which contains the entire *ureDABCEFG* urease

gene cluster under the control of the *lac* promoter, was modified to encode UreGb and its mutant forms by replacing a *PstI/KpnI* fragment. For analysis of urease activity in cell extracts, a single colony containing the desired plasmid was inoculated into 1 mL of LB media supplemented with 300 µg/mL of ampicillin and 1 mM $NiCl_2$ (unless noted) and grown overnight at 37°C with agitation. A 0.5 mL aliquot of the culture was used to inoculate 50 mL of LB containing 300 µg/mL of ampicillin and 1 mM $NiCl_2$ (unless noted) and grown for 2.5 h at 37°C with agitation. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to 0.1 mM to induce the expression of the operon overnight at 37°C with agitation. Cells were harvested by centrifugation for 10 min at 5,000 *g* and 4°C and resuspended in 1 mL of 25 mM HEPES buffer, pH 7.4, for urease activity assays. If the samples were to be used for pull-down assays, cells were resuspended in 750 µl of buffer W. PMSF was added to 0.1 mM, the cells were sonicated (Branson 450 sonifier, 5 repetitions, each of 45 sec, at 1 W output power and 50% duty cycle), and the disrupted cells were centrifuged 10 min at 4°C and 16,000 *g* in a microcentrifuge. The cell extracts were used to test urease activity and perform pull down assays.

Urease Activity Assays - Urease activities were measured by quantifying the rate of ammonia release from urea by formation of indophenol, which was monitored at 625 nm (27). One unit of urease activity was defined as the amount of enzyme required to hydrolyze 1 µmole of urea per min at 37°C. The standard assay buffer consisted of 50 mM HEPES, pH 7.8, and 50 mM urea.

Pull-Down Assays - Samples (0.5 mL) of cell extracts (from *E. coli* DH5α containing pKK17 with the modified versions of *ureG*) were loaded onto a 0.3 mL

Strep-Tactin column (IBA, Germany) equilibrated in buffer W. Proteins were eluted according to the manufacturer's instructions and analyzed using 13.5% SDS-PAGE.

Western Blot - Proteins resolved by SDS-PAGE were transferred to an Immobilon-P polyvinylidene difluoride membrane (Millipore, USA). ExtrAvidin®–alkaline phosphatase conjugate (1:2500 dilution, Sigma, USA) was used as a probe to bind to biotin-tagged forms of UreG. BCIP®/NBT-Blue Liquid Substrate (Sigma, USA) was added to develop the color. To detect UreE or urease, the membranes were incubated for 45 min with anti-UreE IgG (1:10,000 dilution) or anti-urease antibody (1:5,000 dilution) in TBS buffer (150 mM NaCl, 100 mM Tris, pH 7.4) containing 1% Tween 20. After washing the membranes four times with TBS, they were incubated for 45 min with anti-rabbit IgG conjugated to alkaline phosphatase (Sigma, USA) diluted 30,000 times. The membranes were washed again and the BCIP®/NBT-Blue Liquid Substrate was added to develop the color. Prestained molecular weight markers were obtained from Bio-Rad (Hercules, CA).

RESULTS

Characterization of Biotin-Tagged UreG (UreGb) – The native form of *K. aerogenes* UreG was previously purified from recombinant *E. coli* cells by using a series of three different columns (*13*); however, the tendency of the protein to elute from ion exchange resins over a large number of fractions led to low overall yields. To overcome the low yield problem and to facilitate a single-step purification of UreG variants, we examined a new purification system involving a

fusion peptide sequence that becomes biotinylated. This tag was specifically designed to allow affinity purification without introduction of metal-binding residues as in the commonly used His_6 tag (*28*). The *ureG* sequence was subcloned into plasmid pASK-IBA3plus and pASK-IBA5plus so as to encode UreG fused with a biotin-tagged peptide at the C- or N-termini, respectively. The protein derived from the pASK-IBA3plus vector was overproduced in higher amounts by cells, so this plasmid was selected for further experiments. For comparative analyses, native UreG also was obtained by using the previously described protocol (*13*).

UreGb was highly purified by single-step chromatography on a Strep-Tactin column, and essentially homogeneous protein was obtained by subsequent gel filtration chromatography (Fig. 3.3). The biotin-tagged UreGb possessed nearly the same secondary structure as the native UreG according to CD measurements (60% α helix, 18% β strands, 4% turns, and 18% random coil for UreGb, versus 65% α helix, 15% β strands, 5% turns and 15% random coil for native UreG, each with a normalized root mean square deviation of 0.001; Fig. 3.4). Results obtained by size exclusion chromatography were consistent with UreGb being strictly a monomeric protein (Fig. 3.5), even when 0.5 mM nickel or zinc ions were added to the buffer (data not shown). This finding contrasts with the situation for native *K. aerogenes* UreG which gives rise to features consistent with both monomeric and dimeric species, although the latter is present in small amounts (Fig. 3.5). The dimer disappears when the protein is dialyzed overnight with 1 mM DTT, suggesting the presence of some disulfide-linked subunits. The

ratio of dimer to monomer does not change in the presence of nickel or zinc (data not shown), but it increases after several days of protein storage, consistent with formation of an oxidation product.

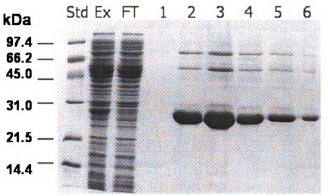


Figure 3.3. UreGb purification. UreGb was isolated by use of a Strep-Tactin affinity column followed by gel filtration chromatography. The purified sample was subjected to SDS-PAGE analysis followed by staining with Coomassie brilliant blue. Lanes: Std, standard proteins used as size markers; Ex, cell extracts; FT, flow through; 1-6, fractions recovered by elution with desthiobiotin.

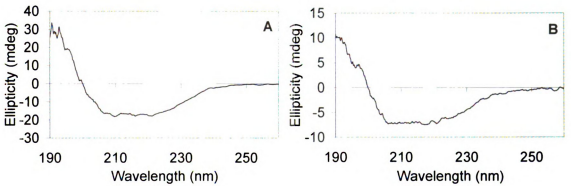


Figure 3.4. CD spectra for UreG and UreGb. The proteins were analyzed in the 190-260 nm range at 0.2 mg/ml and 0.1 mg/ml respectively. Panel A shows UreG spectra and panel B shows UreGb.

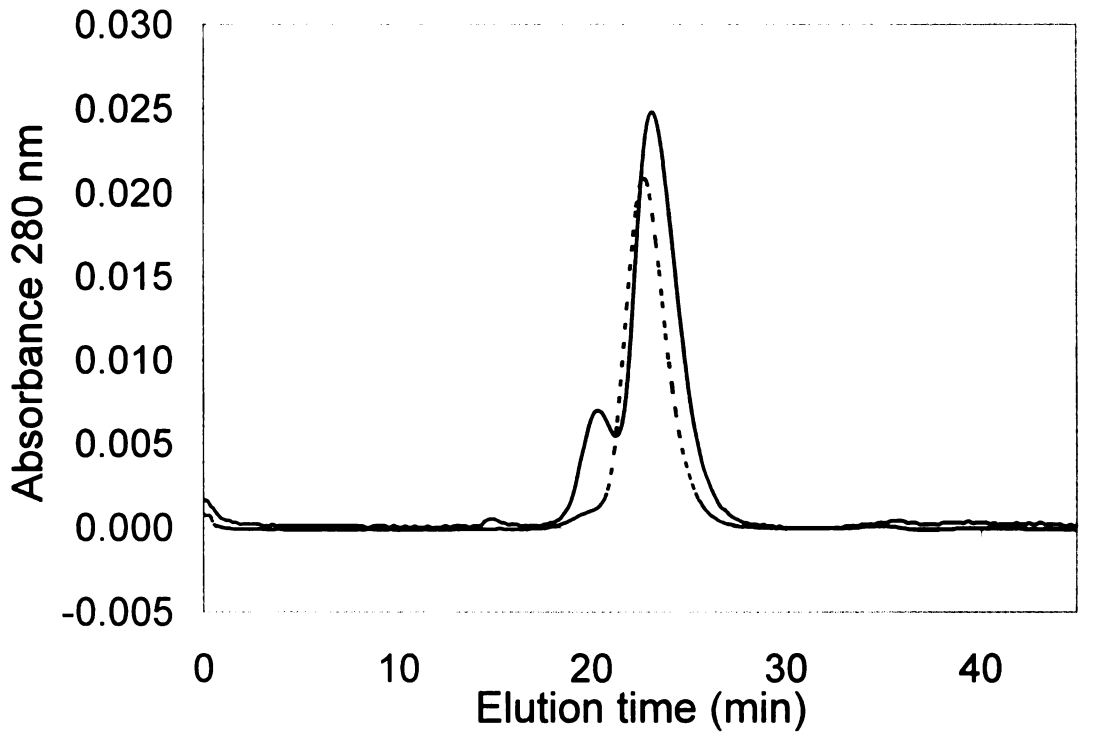


Figure 3.5. Size exclusion profile of native UreG and UreGb. 300 µl of a 2 mg/ml (app. 90 µM) solution of UreG or UreGb were loaded onto a 35 ml Superdex-75 column (1.0 x 45 cm). The buffer contained 50 mM HEPES, pH 7.4, 200 mM NaCl and 0.5 µM $NiNO_2$ .The solid line corresponds to UreG and the dashed line corresponds to UreGb.

Targeting Residues for Mutagenesis – Two criteria were used to select amino acids for mutagenesis. First, a sequence alignment that included sequences of UreG and HypB from 30 different organisms published in Pfam-B, release 4.0 (http://pfam.sanger.ac.uk/) was used to find highly conserved residues. The high level of identity between UreG proteins of different species (over 50%; see complete sequence comparisons in (*18, 19*)) precludes the identification of critical amino acid residues by simple sequence alignment; however, UreG sequence comparison with the related protein HypB highlights fewer amino acids. For example, the P-loop motif (GSGKT at positions 17-21 in

K. aerogenes UreG or residues 43-47 of *M. jannaschii* HypB), the signature motif for the SIMBI G3E family of GTPases (*29*) (ESGG at positions 104-107 of UreG or ENVG at 120-123 of HypB), and the guanine specificity loop (NKTD at positions 151-154 of UreG or NKID at residues 167-170 of HypB) were conserved as expected. Second, the crystal structure of *M. jannaschii* HypB (*21*) showed a dinuclear zinc binding site. We aimed to identify the corresponding amino acids to verify if the same metal binding site exists in UreG. Therefore, the residues targeted for mutagenesis included: Lys20, previously shown to be a critical P-loop residue (*13*) and earlier changed to K20A; Asp49, equivalent to the Mg^{+2}-coordinating Asp75 in *M. jannaschii* HypB (*21*) was changed to D49A; Glu68, the corresponding residue of which was suggested to be a metal ligand for the *B. pasteurii* protein (*18*), was changed to E68A; Cys72, likely to correspond to the Cys95 metal ligand at the dinuclear site of *M. jannaschii* HypB (*21*) was changed to C72A; His74, likely to correspond to the His96 dinuclear center ligand of HypB (*21*) was changed to H74A, H74C, and H74N; Asp80, corresponding to Asp98 of HypB and highly conserved in both proteins, was changed to D80A, and Ser111 and Ser115, likely to correspond to the Cys127 ligand of the dinuclear site in HypB were changed to S111A and S115A. Three double mutants were also constructed: C72A/H74A, C72A/S111A, and H74A/S111A.

Nickel Binding to UreG, UreGb and Variants –UreG and UreGb contained no significant levels of bound nickel or zinc ions as freshly purified by the methods described above (data not shown). The nickel ion-binding properties of

UreG, UreGb, and selected mutant proteins were examined by a centrifugation/PAR reactivity approach (see Experimental Procedures). This method revealed that the fully reduced and monomeric native form of UreG bound 2.0 ± 0.1 nickel ions per molecule with a K_d of 12.5 ± 3.3 μM (Table 3.3). In contrast, a partially oxidized sample of UreG bound 1 nickel ion per molecule with a K_d of 46.6± 13.6, compatible with thiol group participation in metal binding. The strictly monomeric UreGb binds 2.0 ± 0.1 nickel ions per molecule with a K_d of 29.9 ± 19.1 μM (Table 3.3 and Fig. 3.6). This result is consistent with the biotin tag leading to decreased metal ion binding affinity as well as prevention of dimerization, potentially due to interference with a thiol group. The C72A and S111A single mutants of UreGb bound 1.5 ± 0.1 nickel ions per molecule and the H74A variant binds 1.7 ± 0.2 nickel ions per molecule, all with lower K_d than observed for UreGb and comparable to that noted for UreG (Table 3.3 and Fig. 3.6). Similarly, the C72A/S111A and H74A/S111A double mutants of UreGb exhibited the same properties (Table 3.3). The C72A/H74A UreGb double mutant was isolated in very low amounts, thus preventing metal-binding analysis.

Table 3.3. Thermodynamics of nickel ion binding to UreG proteins.

Protein*	K_d	B_{max}
UreG (monomer) (1)	12.5 ± 3.3	2.0 ± 0.1
UreG (partial dimer) (3)	46.6 ± 13.6	1.0 ± 0.1
UreGb (3)	29.9 ± 19.1	2.0 ± 0.1
UreGC72Ab (3)	15.8 ± 3.6	1.5 ± 0.1
UreGH74Ab (3)	9.2 ± 2.3	1.7 ± 0.2
UreGS111Ab (2)	12.5 ± 3.6	1.5 ± 0.1
UreGC72A/S111Ab (1)	10.9 ± 2.8	1.5 ± 0.1
UreGH74A/S111AAb (3)	12.1 ± 4.9	1.6 ± 0.1

*The numbers in parentheses indicate the number of experiments.

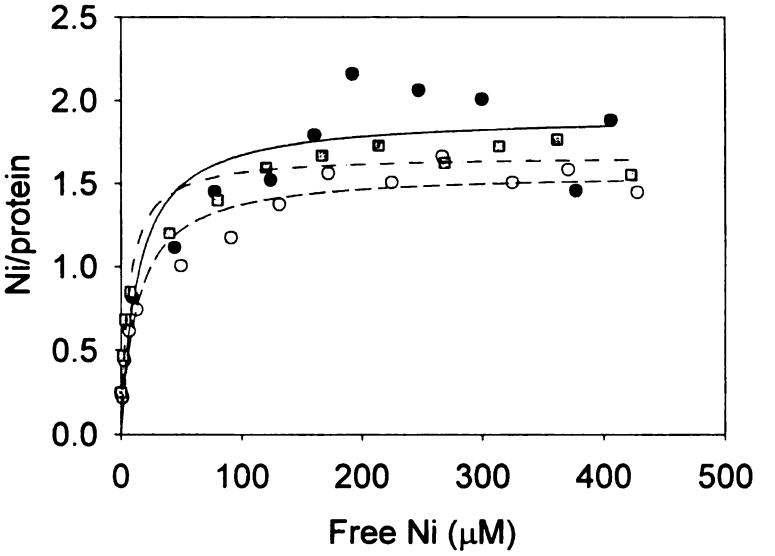


Figure 3.6. Metal binding to UreGb and selected variants. Samples were incubated with varied concentrations of nickel ions, and portions of the buffers containing the free metal ion were separated from proteins using a Microcon centrifuge unit. The nickel ion concentrations of the protein-free solutions were determined spectrophotometrically using the colorimetric reagent PAR. UreG samples included: UreGb (closed circles, data fitting in solid line), C72A (open circles, data fitting in dash line) and H74A (gray squares, data fitting in dots and dash).

Effect of UreGb Variants on Urease Activity in Cell Extracts –Selected versions of *ureG* were expressed as part of the urease operon, and the levels of the encoded UreGb variants were shown to be indistinguishable when cell extracts were examined by Western blot (Fig. 3.7). The urease activity measured in extracts of cells producing UreGb was indistinguishable from that of extracts from cells containing native UreG (Fig. 3.8). In contrast, the cells producing K20A, D49A, C72A, H74A, H74C, H74N, D80A, and S111A variants of UreGb led to nearly undetectable levels of urease activity. Extracts from cells containing the E68A UreGb variant possessed about 18% of the wild-type level of urease activity. Finally, the S115A mutant had no effect on urease activation.

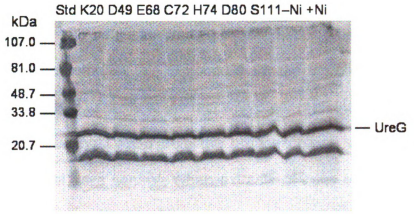


Figure 3.7. Analysis of mutant UreGb levels in cell extracts. Extracts of *E. coli* BL21 cells containing derivatives of pKKGb were subjected to SDS-PAGE, the proteins were transferred to an Immobilon-P membrane, and phosphatase-conjugated avidin was used to assess the content of biotin-tagged proteins in the samples. The lower band likely corresponds to the endogenous carboxyl carrier protein in the cells, whereas the upper band corresponds to the UreGb variants. Lanes: Std, prestained molecular weight standards; UreGb variants cell extracts; -Ni and +Ni correspond to cells containing pKKGb plasmid grown without or with nickel respectively.

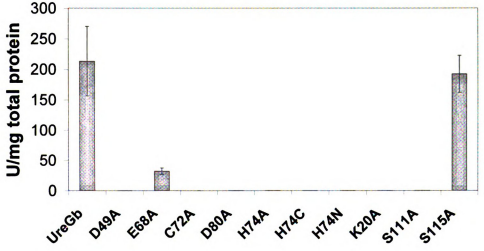


Figure 3.8. Urease activity in cell extracts expressing UreGb and its mutants. The urease activity was examined in extracts of *E. coli* DH5α cells with the variant pKKGb plasmids encoding the various forms of biotin-tagged UreG. For comparison, activity of cells containing pKAU17 (also containing the complete urease operon) was 198 U/mg (*30, 31*). The error bars indicate the standard deviation of at least three independent measurements.

Pull-down Assays – A lack of urease activity in extracts of cells containing altered forms of UreGb could be caused by an inability to form the (UreABC-UreDFG)$_3$ activation complex. We exploited the biotin tag on UreGb to examine the ability of the UreGb variants to interact with other proteins to form complexes *in vivo*. Cells expressing UreGb from the urease operon were grown with or without added nickel as a control. When extracts of cells containing the entire urease operon and encoding UreGb were loaded onto the Strep-Tactin column, washed, and eluted, several bands were observed by SDS-PAGE in addition to UreGb (Fig. 3.9). Urease structural subunits were identified by their characteristic size and by Western blot using anti-urease antibodies (data not shown). Additional bands migrated at positions expected for the UreD and UreF accessory proteins. Finally, an extra band was identified as UreE by using anti-UreE antibodies in a Western blot (Fig. 3.10). This band was shown to be present in all samples. Of potential interest, UreE was present in smaller amounts for cultures grown in the absence of added nickel ions than for the sample grown in its presence, suggesting that the metal-bound form of UreE preferentially binds to this complex. The sample derived using extracts containing the D80A UreGb variant possessed very low quantities of the urease structural subunits and appeared to lack UreD and UreF, yet the band corresponding to UreE was present. This finding is consistent with a direct UreG-UreE interaction. The samples obtained for extracts containing the K20A, H74A and S111A UreGb variants bound weakly to the column, as if the biotin tag was inaccessible in these UreGb-containing complexes, perhaps indicating that the proteins in the

complexes mask or obscure the UreGb carboxyl terminus. In contrast, the samples containing each of the three His74 UreGb variants exhibited complexes that bound to the resin, with the H74A UreGb sample appearing identical to the non-mutated UreGb. It is interesting to note that variant H74C of UreGb led to a more intense band corresponding to UreE in that gel, as if this mutation led to binding stabilization of the UreE apoprotein. Finally, the E68A, C72A and H74N variants of UreGb resulted in complexes that behaved like those for the non-mutated UreGb.

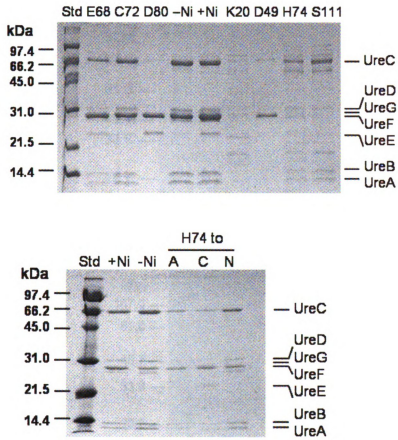


Figure 3.9. Pull-down assays. Extracts of *E. coli* BL21 cells containing the various pKKGb plasmids that encode variant forms of UreGb along with all other urease structural and accessory proteins were chromatographed on Strep-Tactin columns. The samples that were eluted with desthiobiotin were subjected to SDS-PAGE and stained with Coomassie brilliant blue. Lanes: Std, molecular weight standards; +Ni, sample from pKKGb cultures grown with nickel; -Ni, sample from pKKGb cultures grown without nickel; the corresponding mutants are indicated in each lane.

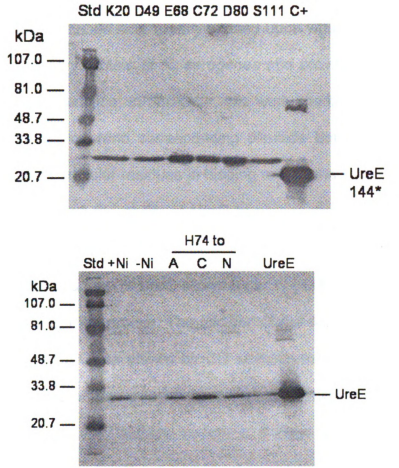


Figure 3.10. Analysis of the UreE content in pull down samples. Western blotting with anti-UreE polyclonal antibodies. Std, prestained molecular weight standards; C+, purified sample of the UreE144* (15-residue truncated variant, (*30*)); +Ni, sample from pKKGb cultures grown with nickel; -Ni, sample from pKKGb cultures grown without nickel; UreE, purified sample of full-length UreE; the corresponding mutants are indicated in each lane.

DISCUSSION

The studies described here greatly expand upon what is known about the UreG urease accessory protein of *K. aerogenes* and provide new insights into critical aspects of its function. In particular, this work sheds light on the lack of necessity of dimerization (and corresponding disulfide bond formation) in the protein, the roles of particular residues in binding metal ions or facilitating *in vivo* activation of urease, and the ability of UreG to participate in a newly identified activation complex.

The biotin-tagged form of UreG allows for easy purification of the wild-type and mutant versions of the protein. The addition of the tag has no effect on the tertiary structure of UreG, as shown by CD spectroscopy, but it does prevent formation of a dimeric protein even in the presence of metal ions. Significantly, the ability of UreGb to fully activate urease is a clear demonstration that the dimer is not relevant for UreG function *in vivo*. The wild-type protein exhibits identical gel filtration chromatography behavior regardless of the presence of metal ions, but it slowly undergoes dimer formation concomitant with a decrease in nickel ion binding capacity. This finding suggests that one of the two Cys residues in UreG (located at positions 28 and 72) is involved in both dimerization and metal binding. These results call into question the metal-binding results obtained with dimeric, disulfide-containing UreG proteins previously characterized from *B. pasteurii* (*20*) and *M. tuberculosis* (*19*).

Several additional differences exist between the *K. aerogenes* UreG and the corresponding proteins from other organisms. UreG from *K. aerogenes* has

only 15% disordered structure as determined by CD spectroscopy, as opposed to 30% or 45% for UreG proteins from *B. pasteurii* (*20*) and *M. tuberculosis* (*19*), suggesting that the *K. aerogenes* protein is more structured and perhaps better suited to crystallographic studies. We found that Glu68, previously hypothesized to be an important metal-binding residue at the comparable position in *B. pasteurii* UreG (*18*), was nonessential for urease activation. In contrast, we used mutagenesis approaches to demonstrate that several residues (Lys20, Asp49, Cys72, His74, Asp80, and Ser111) in *K. aerogenes* UreG are critical for urease activation, even though the variant proteins displayed little effect on their metal ion binding properties. In the case of Lys20 and Asp49, roles in MgGTP binding are likely on the basis of the HypB crystal structure. The finding that C72A, H74A, and S111A UreGb variants bind nickel ions with parameters much like the native protein suggests that the metal is simultaneously coordinated by many residues, so that any single or certain double mutants exhibit no effects, or that the metal binding sites are located elsewhere on the protein.

The newly created biotin-tagged form of UreG also was used to provide evidence of novel protein-protein interactions in the cell. In particular, the pull-down assays revealed the formation of a new urease activation complex with all of the structural and accessory gene products including UreE. The use of this tag will allow the purification and characterization of this key complex. Also of great interest, the D80A UreGb variant possessed lower affinity for most proteins in the complex, but retained the ability to bind UreE thus demonstrating a unique UreG-UreE complex that can also be further studied. An interaction between UreG and

UreE was previously proposed on the basis of results from yeast two-hybrid studies carried out with the *H. pylori* proteins (*32, 33*).

UreG and HypB share similarities in function (i.e., activation of a nickel-containing enzyme), as revealed by the multiple sequence alignments. The *K. aerogenes* UreG residues Cys72, His74, and Ser111 are likely to be equivalent to *M. jannaschii* HypB residues Cys166, His167, and Cys198, and mutations to the *K. aerogenes* residues abolished urease activity. Whereas these amino acids coordinate a dinuclear zinc site in the *M. jannaschii* HypB crystal structure and were identified as metal binding amino acids by mutagenesis studies of the *E. coli* protein (23), our mutants show only a slightly smaller amount of nickel bound per molecule (1.5 Ni/molecule compared to 2.0 Ni/molecule for UreGb). In addition, the GTPase domain of HypB binds only one nickel atom per protein, suggesting that at least one nickel binding site in UreG is found only in this protein.

ACKNOWLEDGMENTS

I would like to thank Scott Mulrooney, Kimberly Anderson, and Rachel Morr for their assistance.

REFERENCES

1. Hausinger, R. P., and Karplus, P. A. (2001) Urease, in *Handbook of Metalloproteins* (Wieghardt, K., Huber, R., Poulos, T. L., and Messerschmidt, A., Eds.), pp 867-879, John Wiley & Sons, Ltd., West Sussex, U.K.

2. Jabri, E., Carr, M. B., Hausinger, R. P., and Karplus, P. A. (1995) The crystal structure of urease from *Klebsiella aerogenes*, *Science 268*, 998-1004.

3. Benini, S., Rypniewski, W. R., Wilson, K. S., Miletti, S., Ciurli, S., and Mangani, S. (1999) A new proposal for urease mechanism based on the crystal structures of the native and inhibited enzyme from *Bacillus pasteurii*: why urea hydrolysis costs two nickels, *Structure 7*, 205-216.

4. Ha, N.-C., Oh, S.-T., Sung, J. Y., Cha, K.-A., Lee, M. H., and Oh, B.-H. (2001) Supramolecular assembly and acid resistance of *Helicobacter pylori* urease, *Nature Structure Biology 8*, 505-509.

5. Sheridan, L., Wilmont, C. M., Cromie, K. D., van der Logt, P., and Phillips, S. E. V. (2002) Crystallization and preliminary X-ray structure determination of jack bean urease with a bound antibody fragment, *Acta Crystallogr. D58*, 374-376.

6. Kim, J. K., Mulrooney, S. B., and Hausinger, R. P. (2005) Biosynthesis of active *Bacillus subtilis* urease in the absence of known urease accessory proteins, *J. Bacteriol. 187*, 7150-7154.

7. Quiroz, S., Kim, J. K., Mulrooney, S. B., and Hausinger, R. P. (2007) Chaperones of nickel metabolism, in *Metal Ions in Life Sciences* (Sigel, A., Sigel, H., and Sigel, R. K. O., Eds.), pp 519-544, John Wiley & Sons, New York.

8. Hausinger, R. P., Colpas, G. J., and Soriano, A. (2001) Urease: a paradigm for protein-assisted metallocenter assembly, *ASM News 67*, 78-84.

9. Lee, M. H., Mulrooney, S. B., and Hausinger, R. P. (1990) Purification, characterization, and in vivo reconstitution of *Klebsiella aerogenes* urease apoenzyme, *J. Bacteriol. 172*, 4427-4431.

10. Jabri, E., and Karplus, P. A. (1996) Structures of the *Klebsiella aerogenes* urease apoprotein and two active-site mutants, *Biochemistry 35*, 10616-10626.

11. Park, I.-S., Carr, M. B., and Hausinger, R. P. (1994) *In vitro* activation of urease apoprotein and role of UreD as a chaperone required for nickel metallocenter assembly, *Proc. Natl. Acad. Sci. USA 91*, 3233-3237.

12. Moncrief, M. B. C., and Hausinger, R. P. (1996) Purification and activation properties of UreD-UreF-urease apoprotein complexes, *J. Bacteriol. 178*, 5417-5421.

13. Moncrief, M. B. C., and Hausinger, R. P. (1997) Characterization of UreG, identification of a UreD-UreF-UreG complex, and evidence suggesting that a nucleotide-binding site in UreG is required for in vivo metallocenter assembly of *Klebsiella aerogenes* urease, *J. Bacteriol. 179*, 4081-4086.

14. Soriano, A., Colpas, G. J., and Hausinger, R. P. (2000) UreE stimulation of GTP-dependent urease activation in the UreD-UreF-UreG-urease apoprotein complex, *Biochemistry 39*, 12435-12440.

15. Soriano, A., and Hausinger, R. P. (1999) GTP-dependent activation of urease apoprotein in complex with the UreD, UreF, and UreG accessory proteins, *Proc. Natl. Acad. Sci. USA 96*, 11140-11144.

16. Kim, J. K., and Hausinger, R. P. (2006) The UreEF fusion protein provides a soluble and functional form of the UreF urease accessory protein, *J. Bacteriol. 188*, 8413-8420.

17. Salomone-Stagni, M., Zambelli, B., Musiani, F., and Ciurli, S. (2007) A model-based proposal for the role of UreF as a GTPase-activating protein in the urease active site biosynthesis, *Proteins 68*, 749-761.

18. Zambelli, B., Stola, M., Musiani, F., De Vriendt, K., Samyn, B., Devreese, B., Van Beeumen, J., Dikiy, A., Bryant, D. A., and Ciurli, S. (2005) UreG, a chaperone in the urease assembly process, is an intrinsically unstructured GTPase that specifically binds Zn^{2+}, *J. Biol. Chem. 280*, 4684-4695.

19. Zambelli, B., Musiani, F., Savini, M., Tucker, P., and Ciurli, S. (2007) Biochemical studies on *Mycobacterium tuberculosis* UreG and comparative modeling reveal structural and functional conservation among the bacterial UreG family, *Biochemistry in press*.

20. Neyroz, P., Zambelli, B., and Ciurli, S. (2006) Intrinsically disordered structure of *Bacillus pasteurii* UreG as revealed by steady-state and time-resolved fluorescence spectroscopy, *Biochemistry 45*, 8918-8930.

21. Gasper, R., Scrima, A., and Wittinghofer, A. (2006) Structural insights into HypB, a GTP-binding protein that regulates metal binding, *J. Biol. Chem. 281*, 27492-27502.

22. Leach, M. R., and Zamble, D. B. (2007) Metallocenter assembly of the hydrogenase enzymes, *Curr. Opinion Chem. Biol. 11*, 159-165.

23. Leach, M. R., Sandal, S., Sun, H., and Zamble, D. B. (2005) Metal binding activity of the *Escherichia coli* hydrogenase maturation factor HypB, *Biochemistry 44*, 12229-12238.

24. Colpas, G. J., Brayman, T. G., Ming, L.-J., and Hausinger, R. P. (1999) Identification of metal-binding residues in the *Klebsiella aerogenes* urease nickel metallochaperone, UreE, *Biochemistry 38*, 4078-4088.

25. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature (London) 227*, 680-685.

26. Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucl. Acids Res. 32*, W668-673.

27. Weatherburn, M. W. (1967) Phenol-hypochlorite reaction for determination of ammonia, *Anal. Chem. 39*, 971-974.

28. Maier, T., Drapal, N., Thanbichler, M., and Bock, A. (1998) Strep-tag II affinity purification: an approach to study intermediates of metalloenzyme biosynthesis, *Anal Biochem 259*, 68-73.

29. Leipe, D. D., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases, *J Mol Biol 317*, 41-72.

30. Brayman, T. G., and Hausinger, R. P. (1996) Purification, characterization, and functional analysis of a truncated *Klebsiella aerogenes* UreE urease accessory protein lacking the histidine-rich carboxyl terminus, *J. Bacteriol. 178*, 5410-5416.

31. Lee, M. H., Mulrooney, S. B., Renner, M. J., Markowicz, Y., and Hausinger, R. P. (1992) *Klebsiella aerogenes* urease gene cluster: sequence of *ureD* and demonstration that four accessory genes (*ureD*, *ureE*, *ureF*, and *ureG*) are involved in nickel metallocenter biosynthesis, *J. Bacteriol. 174*, 4324-4330.

32. Rain, J.-C., Selig, L., de Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schächter, V., Chemama, Y., Labigne, A., and Legrain, P. (2001) The protein-protein interaction map of *Helicobacter pylori*, *Nature 409*, 211--215.

33. Voland, P., Weeks, D. L., Marcus, E. A., Prinz, C., Sachs, G., and Scott, D. (2003) Interactions among the seven *Helicobacter pylori* proteins encoded by the urease gene cluster, *Liver Physiol. 284*, G96-G106.

34. Lu, J., Zheng, Y., Yamagishi, H., Odaka, M., Tsujimura, M., Maeda, M., and Endo, I. (2003) Motif CXCC in nitrile hydratase activator is critical for NHase biogenesis in vivo, *FEBS Lett 553*, 391-396.

Chapter 4

Additional studies, conclusions and remaining questions

ADDITIONAL STUDIES

1. Maltose binding protein-UreF (MBP-UreF) fusion protein crystallization attempts.

Previous efforts to isolate UreF for characterization demonstrated that this protein is soluble only as a fusion protein with large tags such as thioredoxin or Maltose binding protein (MBP)(*1, 2*). I decided to use the MBP-UreF fusion protein for structural characterization through crystallization. The use of the MBP to express and purify UreF has many advantages: it enhances solubility and expression in the host; it allows purification with an affinity column giving high yield and purification fold in a single step. Additionally, upon crystallization, molecular replacement can be used to solve the structure, which is a simpler and less time-consuming method.

The use of fusion proteins also has some disadvantages such as impairment of the protein function. The possibility of obtaining crystals is also reduced in the case of the fusion protein. Multidomain proteins are usually less conductive to forming well-ordered, diffracting crystals, presumably due to the conformational heterogeneity allowed by the flexible linker region.

a. Expression and purification of MBP-UreF

Isolated colonies of *E. coli* DH5α containing the plasmid pMal-UreF (*1*) were grown in 250 mL of Terrific broth supplemented with 100 ug/mL ampicillin at 37°C for 5 hours and then induced overnight with 10 or 100 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The cells were harvested by centrifugation and

resuspended in 20 mL of buffer PED (100 mM potassium phosphate pH 7.2, 1 mM ethylenediaminetetraacetic acid (EDTA) and 1 mM dithiothreitol (DTT)) containing 0.5 mM phenylmethylsulfonyl fluoride (PMSF). The suspension was sonicated and centrifugated at 100,000 x *g* for 45 minutes at 4°C. The supernatant was loaded onto an amylose column (Bio-Rad) previously equilibrated in PED buffer. The fusion protein was eluted with PED buffer containing 10 mM maltose. The collected fractions were analyzed by 12% or 15% SDS-PAGE stained with coomasie blue. Final protein quantification was done with Bio-Rad Protein assay method using BSA as standard.

The fusion protein MBP-UreF was produced in higher amount when cultures were induced with 10 mM IPTG (Figure 4.1, lane 1). More than 95% of the protein was in the soluble fraction. After elution from the amylose resin, MBP-UreF was more than 95% pure as judged by inspection of SDS-PAGE (Figure 4.2). A culture of 250 mL produced 7.5 mg of fusion protein. Protein UreF could be separated from MBP by Factor Xa proteolysis but could not be recovered from the following DEAE-Sepharose chormatography. This is in agreement with a previous report (*1*).

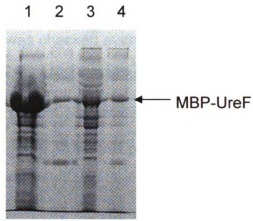


Figure 4.1. SDS-PAGE of MBP-UreF expression on *E. coli* after induction. Lane 1, soluble fraction of cultures induced with 10 mM IPTG; lane 2, insoluble fraction of cultures induced with 10 mM IPTG; lane 3, soluble fraction of cultures induced with 100 mM IPTG; lane 4, insoluble fraction of cultures induced with 100 mM IPTG.

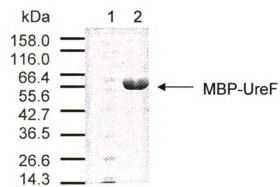


Figure 4.2. SDS-PAGE of purified MBP-UreF. Lane 1, molecular weight markers; lane 2, purified MBP-UreF.

b. Crystallization of MBP-UreF.

Crystallization screening experiments were set up in Dr. Michael Garavito's lab using microbatch sitting drop vapor diffusion method. Each experiment consisted of mixing 2 uL of MBP-UreF at either 15, 20 or 25 mg/mL with 2 uL of 138 crystallization screen solutions obtained from a combination of commercially available crystallization kits. Protein drops were covered with 50/50

paraffin/silicon mix and stored at room temperature or 4°C for two months. The status of each plate was checked every week.

After eight weeks, approximately 50% of the drops remained clear. No crystal was found although protein precipitates were observed on the drops. The same results were obtained with plates incubated at either room temperature or 4°C.

2. Crystallizations attempts for UreG

The accessory protein UreG was purified as described in chapter 3 to >90% purity as determined by SDS-PAGE. A 2µl aliquot of a 10 mg/ml solution of the protein in buffer Tris 100 mM, pH 8.0 containing 150 mM NaCl and 1 mM DTT was mixed with 2µl of 168 crystallization solutions in Dr. Garavito's lab using a microbatch sitting drop vapor diffusion method . Protein drops were covered with 50/50 paraffin/silicon mix and stored at room temperature for two weeks. The status of the plates was checked every 5 days.

When the drops were set up, most of them presented precipitation, however when the plates were checked after five days, crystals or promising precipitates were observed under several condition detailed in table 4.1. No follow up was possible for lack of time.

Table 4.1. Conditions that showed UreG crystals or promising precipitantes.

Precipitant	**Buffer**	**Salt/additive**
1.5 M (NH4)2SO4	0.1M Tris pH 8.5	12% glycerol
1.0 M (NH4)2PO4	0.1M citrate pH 5.5	200 mM NaCl
1.0 M (NH4)2PO4	0.1M Imidazole pH 8.0	200 mM NaCl
0.4 M (NH4)2PO4	None	None
2.0 M (NH4)2SO4	0.1M Tris pH 8.5	None
2.0 NaFormate	0.1 M Na acetate pH 4.6	None
1.6 M NaH2PO4/ 0.4 M K2HPO4	0.1 M Phosphate-citrate pH 4.2	None

3. UreG homology model

Before the publication of the HypB crystal structure, I created a homology model for UreG with the assistance of Dr. Michael Feig. The sequences of *K. aerogenes* UreG was submitted to the Meta server Bioinfo (http://bioinfo.pl/Meta/) (*3*) for a primary structure analysis. The Meta server provides access to several fold recognition and local structure prediction methods. Each method processes the amino acid sequence separately with a specific algorithm. The fold recognition servers search for structures that can be adopted by the submitted amino acid sequence. The local structure prediction methods propose a secondary structure for the protein. The results are collected and then translated into uniform formats to evaluate them for structural similarity. When different methods predict the same structure, the result is considered more reliable and a high score is assigned. The Meta server also gathers primary structure search results for sequence homologs. A more complete structure prediction for UreG was done using the Multiscale Modeling Tools for Structural Biology (MMTSB) tool set (*4*). The crystal structure of the signal sequence binding protein Ffh from *Thermus aquaticus* (Protein Data Bank code: 2Ffh) was used as template. The alignment of secondary structure elements, taken from the Bioinfo Meta server, was used to generate a structure scaffold where side chains in the template structure were replaced with the corresponding amino acids in *K. aerogenes* UreG. Connecting loops and other missing fragments were then added with the MMTSB Tool Set by sampling different conformations and selecting the most favorable based on similarity and a force-field scoring function.

The PDB-Blast search for sequence homologs of UreG gives high statistical confidence values to several GTPases. The four proteins with the highest similarity scores (all GTPases) were: the GTP-binding protein EngA from *E. coli*, Era GTPase from *E. coli*, Yjia protein from *E. coli*, and the large γ subunit of initiation factor eif2 from *Pyrococcus abyssi*. None of the mentioned proteins have more than 20% identity with UreG; therefore, none can be used to build a homology model because this method requires at least a 30% sequence identity to be reliable.

All three secondary structure prediction methods used in the Bioinfo Meta server predict a protein with alpha helices and beta strands. The Meta server evaluation of the structural models proposed by several fold recognition servers gives a high score to structures based on Yjia protein from *E. coli* (mentioned before within the high ranked proteins in the PDB-Blast search), the signal recognition particle receptor from *E. coli* and the signal sequence binding protein from *Thermus aquaticus* (Ffh). Since the secondary structure prediction was in better agreement with Ffh, we decided to use this protein as the template for the model.

The final model of the structure (Figure 4.3) shows a core of parallel beta strands surrounded in both sides by alpha helices. The methodology used to build the structural model of UreG cannot give atomic details of the structure, but the overall topology is considered very reliable. In agreement with this affirmation, when the model is compared to the structures in the Protein Data Bank (PDB) using the Dali Server (http://www.ebi.ac.uk/dali/) only P-loop containing

nucleotide binding proteins appear to be the most similar in structure with the UreG model.

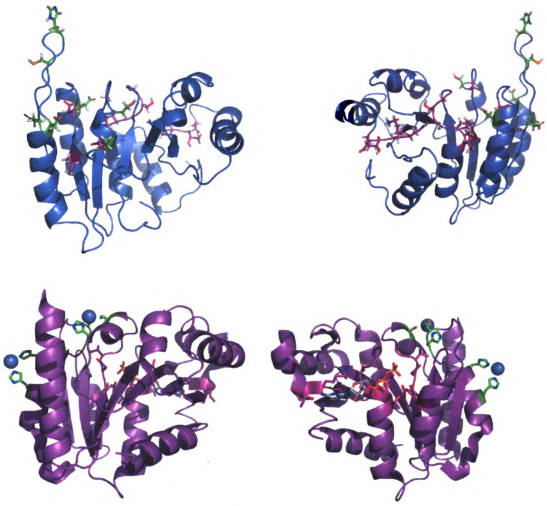


Figure 4.3. UreG model and HypB structure. The top panels show the UreG model in light blue, HypB is shown in purple in the lower panels. Both structures were turned in 180 degrees for the right side panels. Metal binding residues are shown in green and GTP binding residues are shown in pink. The hypB structure also shows the GTP analog in orange.

There are three highly conserved motifs within GTPases: the Walker A motif or P-loop, the Walker B motif and the NDxK motif. The P-loop binds the phosphate groups of the nucleotide. The Walker B motif (involved in Mg^{+2} binding) was modified in UreG as ExxG instead of the more common DxxG

sequence. The NKxD motif confers specificity to GTP by direct interaction with the nucleotide reassuring that GTP is the natural substrate of UreG.

The conserved motifs between Ffh and UreG are the GTPase-related P-loop, Walker B and NKxD motifs. All were localized near from each other in the protein model, in an adequate position to interact with GTP. Since there is no obvious obstacle for GTP binding in the model, the proposed conformation is probably adopted by UreG when is in the apoDFG complex and not isolated when is unable to bind GTP (*5*).

The metal binding residues are not very well aligned as shown in figure 4.4. However, Cys72 and His74 are in a flexible loop that could be easily rearranged. Ser111 is not far from Cys127 position.

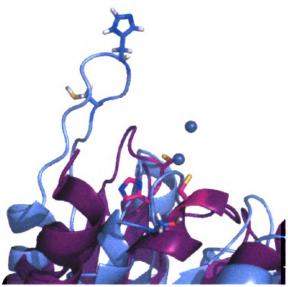


Figure 4.4. Partial superimposition of HypB crystal structure and UreG model. UreG is colored light blue and HypB is in purple. The metal binding residues are shown as sticks and the two spheres are the zinc ions found in the crystal.

CONCLUSIONS AND REMAINING QUESTIONS

The studies presented in this thesis resulted in numerous findings that allow us to propose a more detailed model for urease metallocenter assembly.

In chapter 2, *in silico* flexibility analysis of the urease crystal structure by our collaborators suggested that a flexible hinge in UreB may be involved in urease activation by allowing the subunit to change its conformation and open up the active site. This proposed conformational change agrees with an earlier proposal from our lab that was based on results from chemical cross-linking studies that showed two residues were in close contact, but only when facilitated by the addition of UreF to the $(UreABC\text{-}UreD)_3$ complex. My exchange of Gly11 and Gly18 for proline residues in UreB had mixed effects. In the case of the G18P UreB variant within the context of the other urease proteins, I observed no changes in urease activity. In contrast, the G11P UreB mutation caused a significant reduction in urease activity, a decrease in nickel content, and the recovery of a new activation complex $(UreABC\text{-}UreDFG)_3UreE_n$ from cell cultures. I interpret these results as suggesting that even though the accessory proteins can bind urease, the activation process is somehow hindered in this mutant. The SAXS analyses carried out by our collaborator at Oak Ridge National Laboratories provided data that positioned the UreD and UreF accessory proteins at the vertices of the urease trimer, very close to each other and to UreB. All together, these analyses are consistent with our initial hypothesis and previous cross-linking data.

Chapter 3 describes an improved purification method for UreG that adds a short amino acid sequence at the C-terminus of the protein which becomes biotinylated in the cell. The addition of the tag did not modify the tertiary structure of the protein nor did it interfere with urease activation. Based on sequence alignments and comparison to the crystal structure of the related protein HypB, I mutated selected conserved amino acids in UreG and determined the effects on urease activation. The biotin tag also allowed me to test the interaction of these mutant proteins with urease apoprotein and the other accessory proteins. Finally, I measured the nickel-binding abilities of selected mutants. I found that the

majority of the mutations affecting UreG nearly abolished the ability of this protein to facilitate urease activation, but all of the purified UreG mutants tested (including double mutants) retained the ability to bind at least 1.6 nickel ions per molecule (compared to 2 nickel ions per molecule for the non-mutated protein). Therefore, these results suggest that these residues may not be the ligands used for nickel binding, and the metal binding sites remain to be clearly identified. On the other hand, I showed that mutations in the GTP binding amino acids are key to urease activation. In one case (the D80A UreG mutant) I observed a lower affinity for the complete activation complex, but continued interaction with UreE.

I've made substantial progress in characterizing several urease apoprotein complexes and some properties of wild-type and mutant UreG species; however, numerous questions remain to be answered about the urease activation process. Here I propose some experiments that could help others to a better understand the fascinating process of the urease activation, which may have relevance to the activation of other metalloproteins:

a. The G11P mutant of UreB allows the isolation of a new activation complex that contains all the structural subunits and accessory proteins. If this complex can be characterized further (e.g., analysis by SAXS, chemical modification studies, or even crystallization) it would give very valuable information. Thus, I have created a potentially very useful construct for other researchers to investigate.
b. The urease SAXS studies that I've pioneered can be extended in new directions to obtain additional insights. A soluble form of UreF (a fusion of the UreE and UreF proteins) is known to bind $(UreABC\text{-}UreD)_3$ to form the $(UreABC\text{-}UreD(E)F)_3$ complex. The extra domain provided by UreE could be used to better localize UreF with the $(UreABC\text{-}UreDF)_3$ complex. Of particular interest might be the isolation of UreEF from cells grown in deuterated buffer, formation of the larger complex with $(UreABC\text{-}UreD)_3$ from normal buffer, and analysis using neutron scattering methods. This approach has the potential to more clearly define the position of UreEF in the large complex. Similarly,

other fusion proteins (such as maltose binding protein-UreD) could be use in a similar approach.

c. Two nickel ions are bound to each UreG molecule according to my membrane centrifugation/PAR studies. The crystal structure of HypB hints at the location of one set of metal ligands, but the mutation studies I performed do not provide evidence to confirm that those targeted residues coordinate the metal. More mutational, spectroscopic, and structural analyses are required to clarify this point. For example, mutations could be constructed to alter the second Cys in the protein, and several other potential nickel ion-binding residues could be targeted. A type of spectroscopy known as X-ray absorption spectroscopy can provide information about ligands to metal ions, and this might be a particularly useful approach to identify whether the metal is bound by imidazole ligands (the method cannot distinguish among most N and O ligands, but the secondary scattering from the ring carbons allows one to identify and quantify His ligands). My CD studies suggest that, unlike the highly unstructured situations for other UreG proteins, the *K. aerogenes* UreG might be amenable to crystallographic analysis. Indeed, I obtained preliminary results suggesting that a particular crystallization buffer held promise (data not shown). The structure of apoprotein or nickel ion-bound protein would be valuable.

d. The interaction between UreG and UreE now can be further characterized thanks to the D80A variant of UreG. These two accessory proteins are likely to be at the core of the activation process because they allow for the coupling of nickel ion delivery to urease apoprotein and GTP hydrolysis. Crystallization of these isolated proteins or co-crystallization of the UreG-UreE complex could give better insight into the process of nickel introduction into urease. It is possible that the inclusion of GTP could trigger conformational changes of UreG within the UreG-UreE complex that could provide enhanced understanding of the function of GTP in nickel insertion into the apoprotein.

e. The D80A mutant can also be used to better understand the interaction between UreG and the $(UreABC\text{-}UreDF)_3$ activation complex, again making

use of several fusion proteins available in our laboratory (i.e. UreEF, MBP-UreF, MBP-UreD) to determine if the interaction is with UreD, UreF or the $(UreABC\text{-}UreDF)_3$ complex.

f. The approach I used to target amino acids for site-directed mutagenesis of UreG could be used to analyze UreD and UreF and possibly assess their function. Despite of the low level of identity between UreD and UreF orthologs, it is possible to identify some highly conserved residues that can be mutated. It is important to note that no proteins of related sequence have yet been crystallized for these two proteins and the insoluble nature of the proteins prevents more structural analyses.

In conclusion, the information provided by my studies has improved our understanding of urease metallocenter assembly. In addition, several mutants and the new purification system for UreG could be used as tools for future research of the system.

REFERENCES

1. Kim, K. Y., Yang, C. H., and Lee, M. H. (1999) "Expression of the recombinant *Klebsiella aerogenes* UreF protein as a MalE fusion" *Arch Pharm Res 22*, 274-278.

2. Moncrief, M. B. C. "Unpublished results".

3. Ginalski, K., Elofsson, A., Fisher, D., and Rychlewski, L. (2003) "3D-Jury: a simple approach to improve protein structure predictions" *Bioinformatics 19*, 1015-1018.

4. Feig, M., Karanicolas, J., and Brooks, C. L., III. (2004) "MMTSB tool set: enhanced sampling and multiscale modeling methods for applications in structural biology" *J Molec Graphics Modeling 22*, 377-395.

5. Moncrief, M. B. C., and Hausinger, R. P. (1997) "Characterization of UreG, identification of a UreD-UreF-UreG complex, and evidence suggesting that a nucleotide-binding site in UreG is required for in vivo metallocenter assembly of *Klebsiella aerogenes* urease" *J Bacteriol 179*, 4081-4086.