Characterization of Small Metal-binding Protein (SmbP)

From *Nitrosomonas Europaea*

by

Qin Yan

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2010 by the
Graduate Supervisory Committee:

Wilson Francisco, Chair
James Allen
Giovanna Ghirlanda

ARIZONA STATE UNIVERSITY

December 2010

ABSTRACT

A novel small metal-binding protein (SmbP), with only 93 residues and no similarity to other known proteins, has been isolated from the periplasm of *Nitrosomonas europaea*. It is characterized by its high percentage (17%) of histidines, a motif of ten repeats of seven residues, a four α-helix bundle structure, and a high binding affinity to about six equivalents of Cu^{2+}. The goal of this study is to investigate the Cu^{2+} binding sites in SmbP and to understand how Cu^{2+} stabilizes the protein.

Preliminary folding experiments indicated that Cu^{2+} greatly stabilizes SmbP. In this study, protein folding data from circular dichroism (CD) spectroscopy was used to elucidate the role of Cu^{2+} in stabilizing SmbP structure against unfolding induced by decreased pH, increased temperature, and chemical denaturants. The significant stabilization effects of Cu^{2+} were demonstrated by the observation that Cu^{2+}-SmbP remained fully folded under extreme environmental conditions, such as acidic pH, 96 °C, and 8 M urea. Also, it was shown that Cu^{2+} is able to induce the refolding of unfolded SmbP in acidic solutions. These findings imply that the coordination of Cu^{2+} to histidine residues is responsible for the stabilization effects.

The crystal structure of SmbP without Cu^{2+} has been determined. However, attempts to crystallize Cu^{2+}-SmbP have not been successful. In this study, multidimensional NMR experiments were conducted in order to gain additional information regarding the Cu^{2+}-SmbP structure, in particular its metal binding sites. Unambiguous resonance assignments were successfully made. Cα

secondary chemical shifts confirmed that SmbP has a four α-helical structure. A Cu^{2+}-protein titration experiment monitored by NMR indicated a top-to-bottom, sequential metal binding pattern for SmbP.

In addition, several bioinformatics tools were used to complement the experimental approach and identity of the ligands in Cu^{2+}-binding sites in SmbP is proposed.

DEDICATION

This is dedicated to my lovely wife Yinan, my son Siyuan, and my parents.

谨以此献给我的父母，妻子和儿子。

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Wilson Francisco for all his patience and advice. I would also like to thank all of my lab mates, especially Sara Bowen. Thank you to my committee members, Dr. James Allen and Dr. Giovanna Ghirlanda. And thanks to my friends.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Chapter 1

GENERAL INTRODUCTION OF SMALL METAL BINDING PROTEIN (SMBP)

Summary

A novel small metal-binding protein (SmbP), with only 93 residues and no similarity to other known proteins, has been isolated from the periplasm of *Nitrosomonas europaea*. In this chapter, some general background of the protein will be introduced, including its unique sequence characterization, its binding affinity to multiple equivalents of metal ions such as Cu^{2+}, and its potential function in the autotrophic nitrifying bacterium, *N. europaea*.

Introduction

Copper has a complex role in living cells. It is an essential element as a cofactor in a large number of enzymes, and also plays an important role in the structural stabilization of proteins. However, at elevated concentrations, it can be toxic. Several systems to regulate copper concentrations and conferred copper resistance in a variety of organisms has been identified [1, 2].

Nitrosomonas europaea is an Gram-negative obligate chemolithoautotroph that derive its energy and reductant for growth from the oxidation of ammonia to nitrite [3]. It lives in soil, sewage, freshwater, and the walls of buildings, and it has a slow growth rate (several days for cell division). Its optimum growth conditions are within a pH range of 6.0-9.0 and a temperature range of 20-30 ^{o}C. *N. europaea* expresses two key enzymes, which are involved in the ammonia

oxidation reactions: Ammonia monooxygenase (AMO) and hydroxylamine oxidoreductase (HAO) (Figure 1.1). AMO catalyzes the oxidation of ammonia to hydroxylamine, which is further oxidized to nitrite, a reaction catalyzed by HAO. HAO has been well studied. It is a soluble periplasmic enzyme with heme as its cofactor. AMO, by contrast, is a membrane-bound protein with copper and iron as its proposed cofactors [3]. As a chemoautolithotroph, *N. europaea* has been chosen for complete genome sequencing [4].

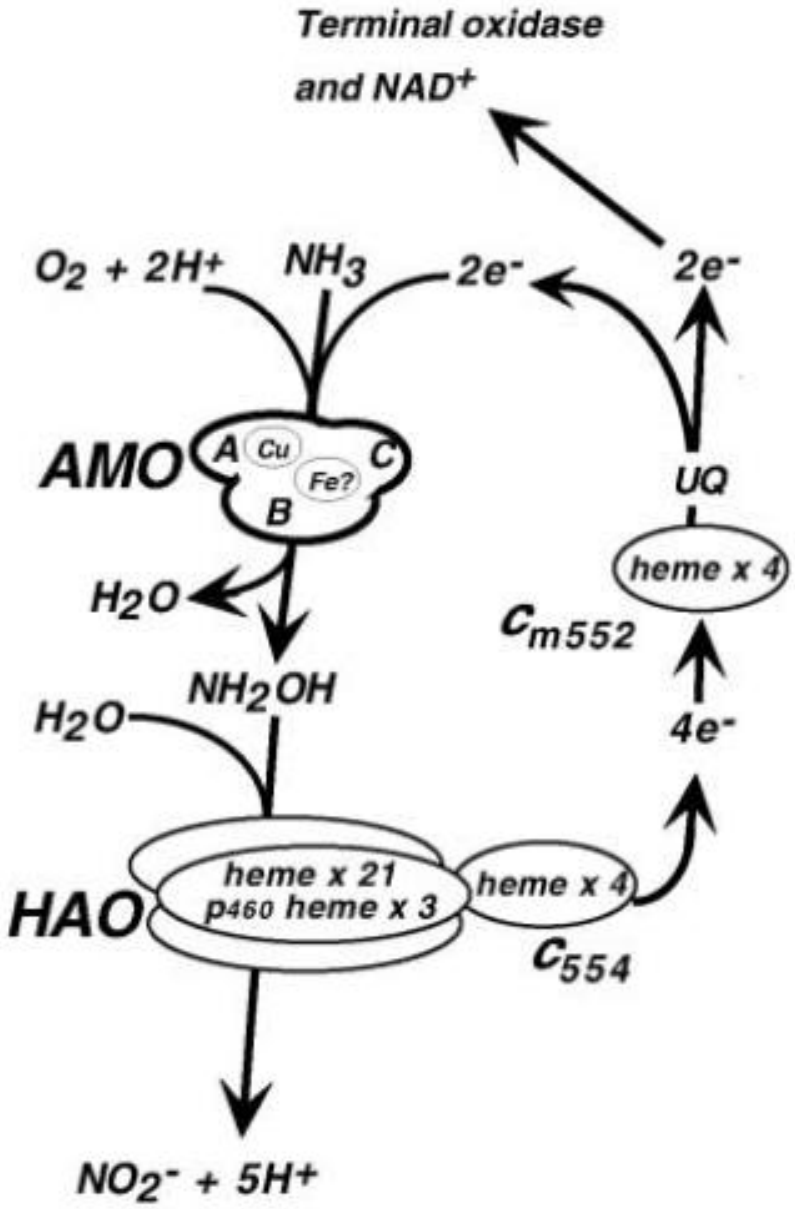

Figure 1.1: Ammonia-oxidizing pathway in *Nitrosomonas europaea* shows two key enzymes ammonia monooxygenase (AMO) and hydroxylamine oxidoreductase (HAO), and a few electron transfer proteins. (From Arp, D.J. et al. 2002 [3])

A novel small metal-binding protein (SmbP), has been purified from the periplasm of *N. europaea* [5]. The protein was first isolated from *N. europaea* with Cu^{2+}-bound, and was later found to be capable of binding multiple equivalents of a variety of divalent and trivalent metals [5]. SmbP has been proposed to have a role in cellular copper management in *N. europaea.*

The Primary Sequence of SmbP

As shown in Figure 1.2, the precursor protein of SmbP contains a leader sequence (shown in green), which directs the protein to the periplasm of *N. europaea*. The mature protein (in red) has 93 residues and a molecular weight of 9.9 kDa. This protein shares no sequence similarity to any other known proteins.

MKTTLIKVIAASVTALFLSMQVYA**SGHTAHVDEAVKHAEE**
AVAHGKEGHTDQLLEHAKESLTHAKAASEAGGNTHVGH
GIKHLEDAIKHGEEGHVGVATKHAQEAIEHLRASEHKSH

Figure 1.2: Full sequence of SmbP with a cleaved leader sequence shown in green and the mature protein shown in red and bold.

The protein contains an unusually high percentage of histidine residues (17%, compared to an average of 2% in 1021 unrelated proteins [6, 7]). Besides histidine, the protein is primarily an assemble of alanine (16%), glutamate (14%), glycine (11%) and lysine (9%), and these five residues account for 67% of the amino acid composition of SmbP. The mature protein lacks aromatic and sulfur-containing residues.

Ten Sequential Repeats of a Seven Amino Acid Motif (abcdefg)

A notable feature of this protein is the presence of 10 repeats of a seven amino acid motif (*abcdefg*), as shown in Figure 1.3. The residues at positions *a*, *d* and *g* are mostly hydrophobic, while the *b*, *c*, *e* and *f* positions are generally hydrophilic residues.

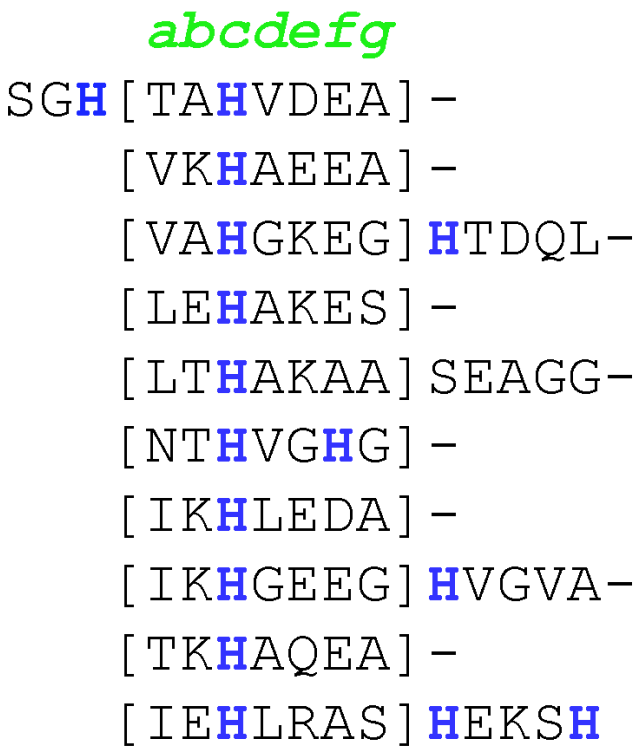

Figure 1.3: The sequence of SmbP contains 10 repeats of a seven amino acid motif. All histidines are highlighted in blue and bold. The positions of *abcdefg* are indicated on the top of the seven amino acid motif.

Based on this sequence motif, the protein was predicted to be a four-alpha-helix bundle. The hydrophobic residues likely contribute to the stabilization of a hydrophobic core, while the hydrophilic residues are responsible for the increased

solubility of SmbP in aqueous solutions. Another unique characteristic of the protein is the conserved histidine residue at the third position. Also, position *f* is highly conserved, and it is occupied by acidic amino acids (Glu or Asp) [5].

The Crystal Structure of SmbP

The structure of SmbP has been determined to a 1.9 Å resolution limit using X-ray protein crystallography. The crystal structure has shown that SmbP is a homodimer with each subunit forming a classical four-helical bundle with approximate dimensions of 30 Å x 20 Å x 20 Å (Figure 1.4). The four helices are arranged in an antiparallel fashion. The first two helices are approximately parallel to each other and are tilted approximately 20° relative to the third and fourth helices, which are also approximately parallel to each other. The helices range from 17 to 22 amino acid residues in length. And the primary sequence is divided into helices as following: residues 2-23 (helix 1), 26-43 (helix 2), 51-68 (helix 3), and 71-87 (helix 4). These four helices are connected by two short loops of two residues (^{24}GH^{25} between helix 1 and 2, and ^{69}GH^{70} between helix 3 and 4) or one longer loop of seven residues (^{44}SEAGGNT^{50} between helix 2 and 3). Two segments ^{45}EAGGN^{49} and ^{89}EHKSH^{93} are missing in the crystal structure due to disorder.

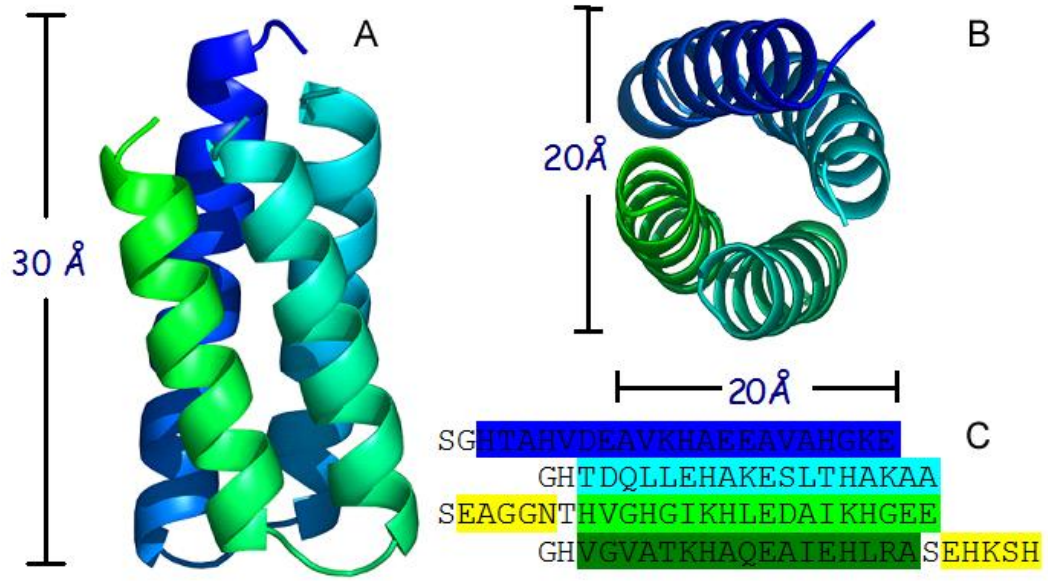

Figure 1.4: SmbP structure shown with the alpha helices as ribbons viewed along a direction (A) perpendicular and (B) parallel to the pseudo-symmetrical axis aligning the four helices. (C) shows the sequence of SmbP. The highlighting color in (C) corresponds to the color in (A) and (B). Yellow highlights the two disordered segments.

The four helices pack together mainly through hydrophobic interactions. The internal surface of the bundle is occupied mainly by hydrophobic residues, Leu, Ile, Ala and Val (Figure 1.5). Analysis of the X-ray crystal structure indicates that most of the larger hydrophobic residues (Leu, Ile and Val) are clustered at the interface between helices 1 and 4 and between helices 2 and 3. On the other hands, most of the residues at the interface between helices 1 and 2 or between helices 3 and 4 are small residues (e.g. Ala and Gly) (Figure 1.5). Larger residues provide

stronger hydrophobic interactions between helices 1 and 4 and between helices 2 and 3 than at the other two interfaces.

In additional, the four-helix bundle is further stabilized by three inter-stand ionic bonds between residues E_{15}–R_{86}, K_{34}–E_{67}, and K_{41}–E_{60} (Figure 1.5), which strengthen the already stronger interactions between helices 1 and 4 and between helices 2 and 3 respectively. In contrast, there is no ionic bond at the other two interfaces to strengthen the hydrophobic interactions between helices 1 and 2 and between helices 3 and 4. Thus the interfaces between helices 1 and 2 and between helices 3 and 4 will constitute the weakest points of the four-helical bundle global structure. Besides the inter-stand ionic bonds, three intra-stand ionic bonds, between residues E_9-K_{12} in the first helix, K_{64}-E_{68} in the third helix, and K_{76}-E_{80} in the fourth helix, also contribute to stabilize the helical structure.

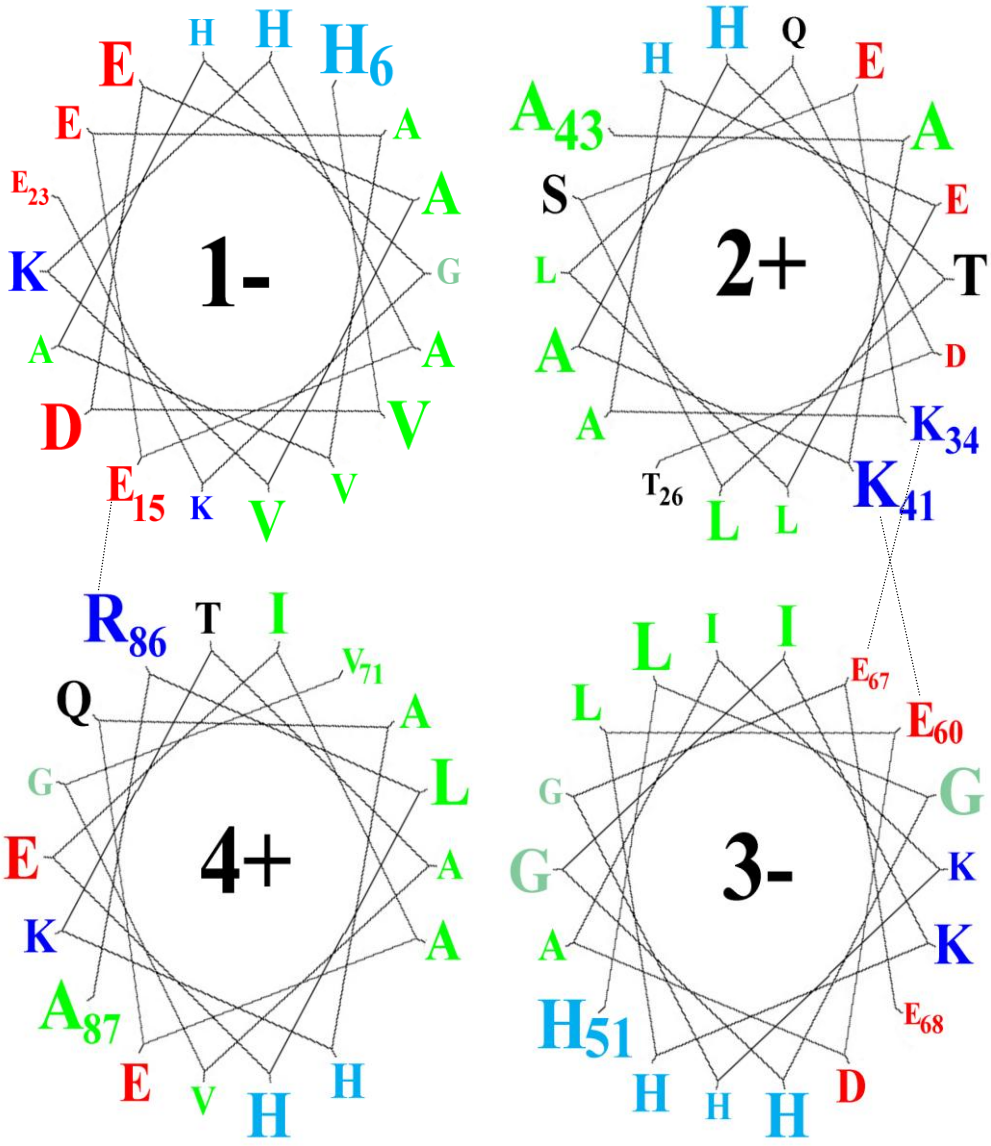

Figure 1.5: Helical-wheel diagram for the four-helices of SmbP (same view as Figure 1.4B). The helices are numbered with the sign indicating direction and the size of the one-letter code amino acid denotes projection. The first and last residues of each wheel are numbered. The amino acids are colored based on their type: hydrophobic (A, L, V) (green), basic (K, R) (blue), acidic (D, E) (red), H (light cyan), G (pastel green), and others (Q, S, T) (black). Three interstrand ionic bonds between residues E_{15}–R_{86}, K_{34}–E_{67}, and K_{41}–E_{60} are depicted by a solid line.

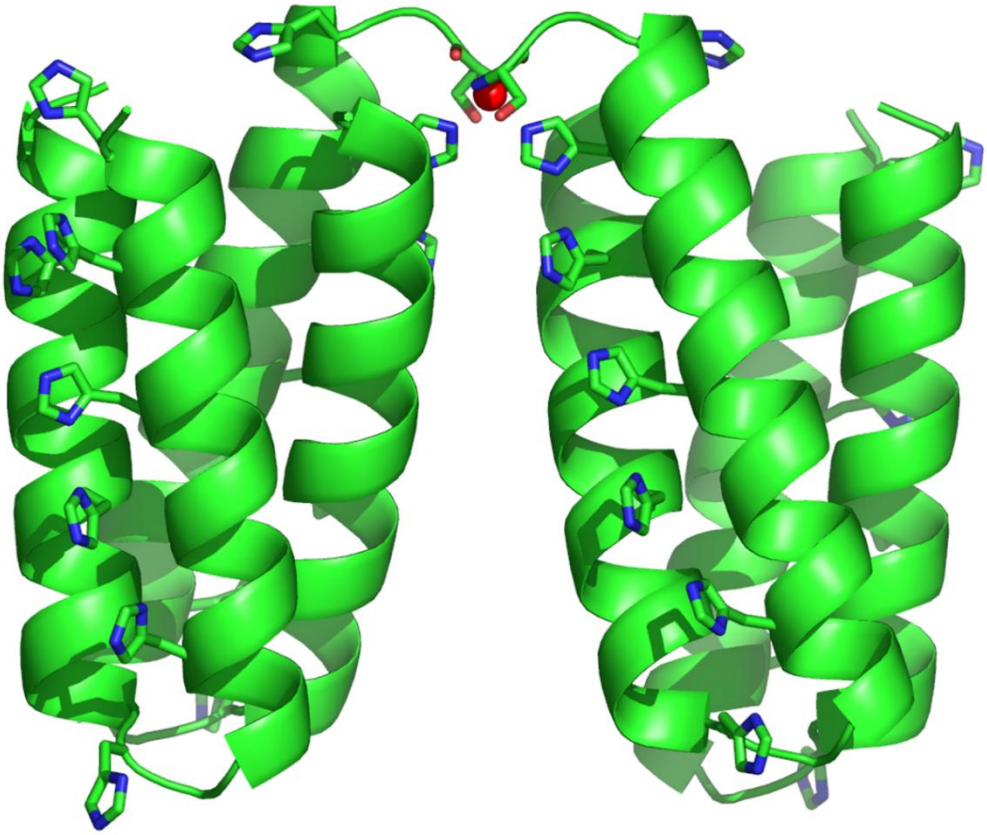

Figure 1.6: Dimeric structure of SmbP in the asymmetric unit of the crystals. The two proteins are related by a pseudo two-fold symmetry axis about the center of the two proteins in the plane of the paper. The shared nickel atom (red sphere) is coordinated by the amino-terminal residues Ser 1 and His 6 of each protein (color coded by atom type). Also shown are all of the histidine residues (colored by atom type) that appeared to be aligned along the two opposite sides of the four-helix bundle.

SmbP contains an unusually large percentage of histidine residues (17%) and most of them are highly conserved, appearing regularly at the *c* position of the repeating motif (Figure 1.3). In three-dimensional space, these ten histidines and two additional histidines located at the first and third loops, line up at uniform

spacings along the helices (Figure 1.5 and Figure 1.6). The distances between two adjacent histidines are range from 2.8 Å to 4.8 Å, with an average of 3.6 Å (Figure 1. 7). The only exception is that the distance between His 51 and His 84 in chain A is 6.2 Å. The corresponding distance in chain B is 2.9 Å. These distances are close to the average hydrogen bond lengths of N-H···N (3.1 Å) [8]. The perfect alignments of the twelve histidines allow for the formation of hydrogen bonds at the interfaces between helices 1 and 2 and between helices 3 and 4. Furthermore, six glutamates (E9, E16, E23, E35, E68 and E80), one aspartate (D61), one glutamine (Q28) and another histidine (H54) on position *f*, spatially close to position *c*, form additional hydrogen bonds with the above histidines directly or mediated by water molecules. These hydrogen bonds stabilize the four-helical bundle at the two interfaces between helices 1 and 2 and between helices 3 and 4, which have weaker hydrophobic interactions and without ionic bonds.

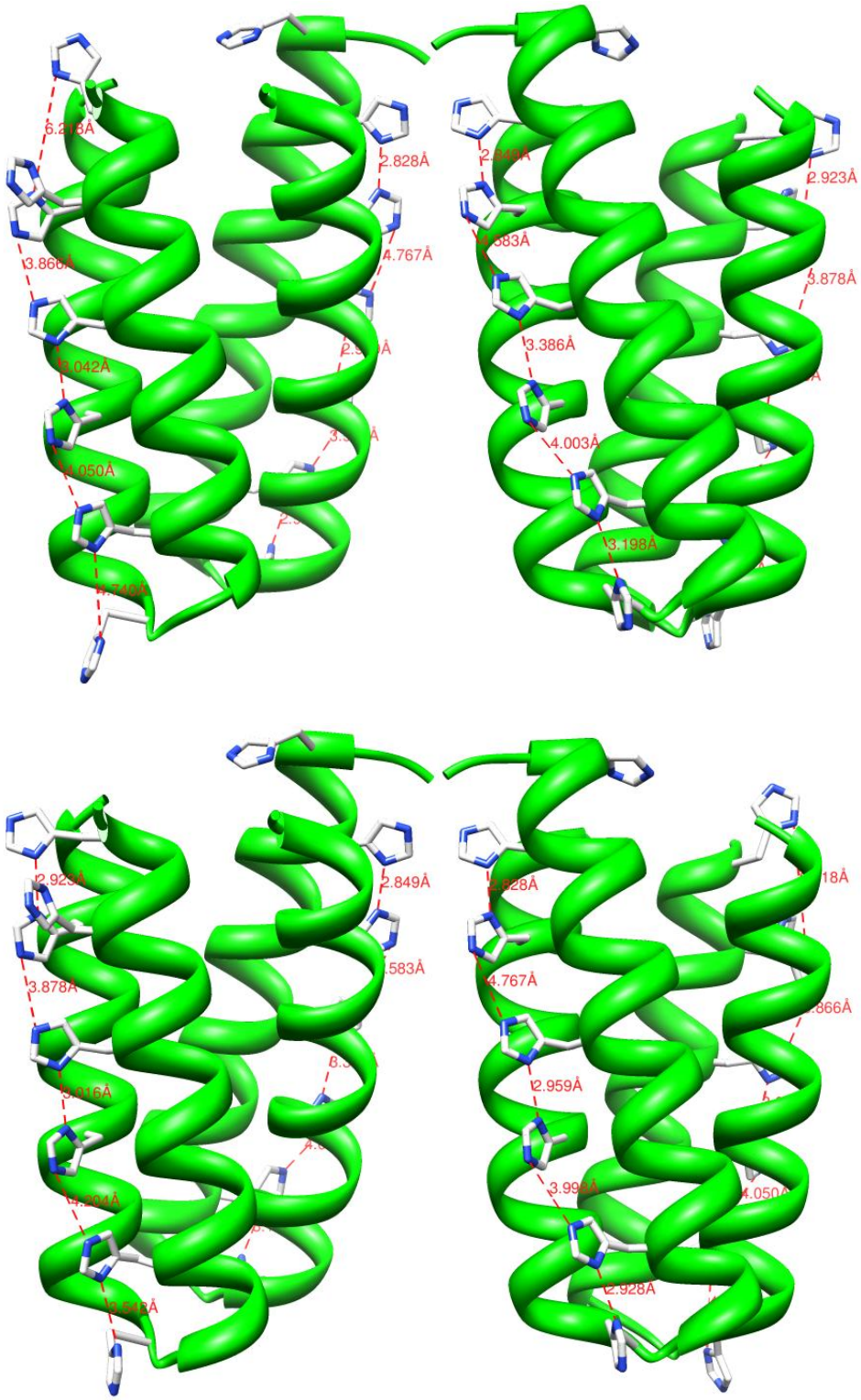

Figure 1. 7: The distances between two adjacent histidines.

Overall, the four-helical bundle is stabilized by three important forces: hydrophobic interactions, ionic bonds, and hydrogen bonds. While stronger hydrophobic interactions and ionic bonds present at the interfaces between helices 2 and 3 and between helices 1 and 4, more hydrogen bonds form at interfaces between helices 1 and 2 and between helices 3 and 4 mediated by adjacent histidines. These interactions are responsible for the structural stability of SmbP.

While the protein is preferentially a monomer in solution under normal biological conditions, in the crystal structure two proteins are closely associated forming a homodimer (Figure 1.6). The dimerization appears to be stabilized by a nickel bridge between the two monomers. A single nickel atom, whose identity was unambiguously confirmed by the multiple-wavelength anomalous dispersion (MAD) results, is found at the interface between the two monomers. The nickel ion is coordinated by the N-terminal residues Ser 1 and His 6 from each monomer (Figure 1.8).

This nickel binding site appears to be an artifact of the purification process, as our previous biochemical studies [5] have indicated that the protein is monomeric in solution and the nickel does not appear to play a direct role in the function of the protein. Most likely, it is present due to the use of a nickel affinity column during the purification of the protein. The argument that SmbP is not natively dimerized with a nickel bridge is supported by sedimentation equilibrium analytical ultracentrifugation experiments, which have confirmed that both apo-SmbP and Ni^{2+}- or Cu^{2+}- bound SmbP are monomeric in solution. Fitting the sedimentation data to a single ideal species yielded an apparent molecular weight

of 9936 ± 66 g/mol, compared to the calculated molecular weight of 9904 g/mol. In the presence of Cu^{2+} and Ni^{2+}, the apparent molecular weight was found to be 10619 ± 50 g/mol and 10950 ± 110 g/mol, respectively. The increase in the molecular weight for SmbP in the presence of metal ions, along with previous data, suggests metal ions are bound to the protein, but the protein remains in monomeric form.

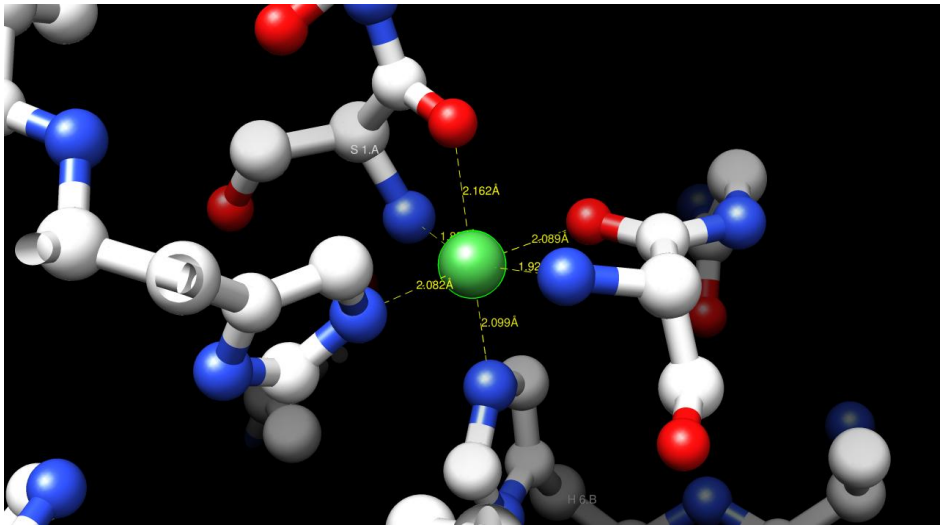

Figure 1.8: The nickel bridge between the two proteins. This nickel ion is coordinated by N and O from the peptide backbone of Ser 1 and ε-N in His 6 from each protein.

Superposition of SmbP Structure with Other Proteins

The Dali server [9][10] is a web based service, which compares newly solved structure against solved protein structures in the Protein Data Bank (PDB). Instead of comparing the primary sequence, it compares the protein structure in 3D. By searching structural homologs of SmbP, excellent alignment has been

found to a number of proteins that contain a four helix bundle motif with up-and-down topology, including cytochrome b_{562} (256B, [11, 12]), PsbQ protein of photosystem II (1NZE, [13]) and Ni-containing superoxide dismutase (1QOD, [14]). Each could be superimposed with SmbP at rms deviations of 1.8, 1.9, and 2.2 Å, respectively, despite a lack of any sequence homology with SmbP (Figure 1.9 and Table 1.1). A list of the top ten structural homologues of SmbP is included in Table 1.1.

. The four helices of SmbP could be closely aligned with those of cytochrome, even though cytochrome b_{562} contains a heme cofactor. As expected, the largest deviations are in the region involved in heme coordination, such as His 102 at the carboxyl terminal region of cytochrome. While the loops of the two structures generally agree, there is no correspondence in SmbP to the extended loop formed by residues Ala 40 to Pro 56 that partially covers the heme cofactor in cytochrome. Although the four helices of superoxide dismutase are slightly longer than those of SmbP, they could be superimposed. The largest deviations occur at the ends of the helices and the more extended loops of superoxide dismutase. The catalytic center of superoxide dismutase is a Ni^{2+} ion that is coordinated by several residues of the amino terminal region, including the amino group of His 1. The nickel atom of SmbP is located at an homologous location and is also coordinated by the backbone of the N-terminal residue, Ser 1. Unlike the nickel atom in superoxide dismutase, the nickel atom of SmbP has no known redox or catalytic activity.

The goal of this study is to investigate the Cu^{2+} binding sites in SmbP, and understand how Cu^{2+} stabilizes the protein.

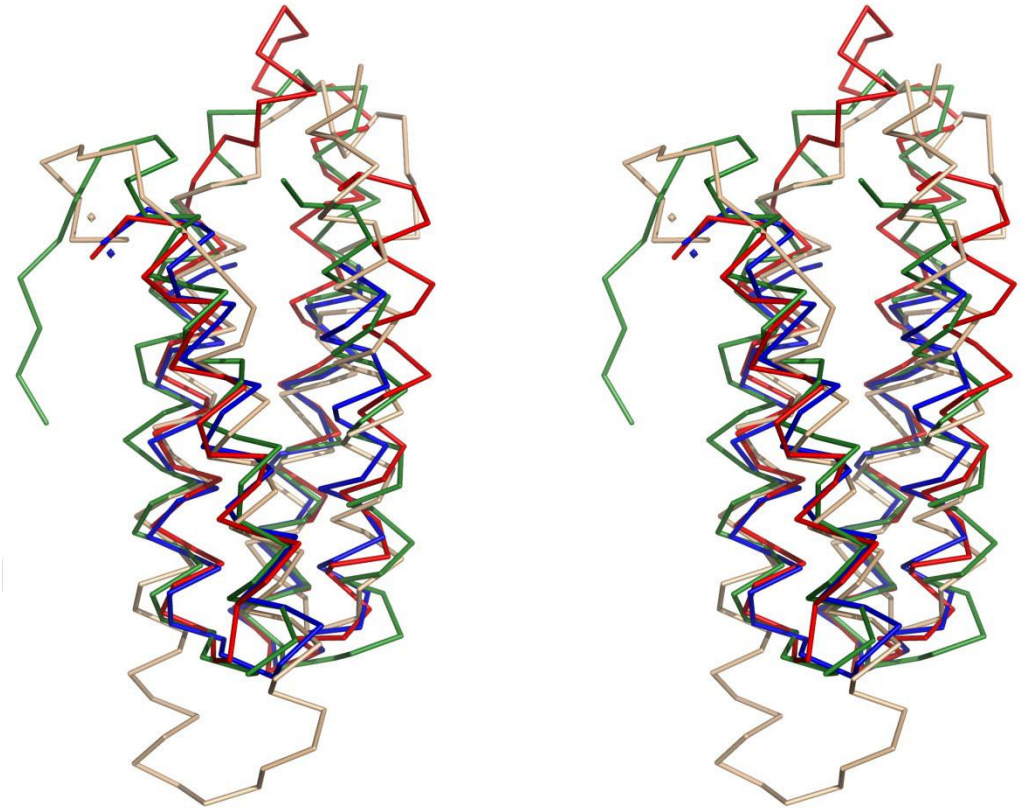

Figure 1.9: Superposition of the Cα trace of SmbP (green) with cytochrome b562 (red, PDB file 256B) (without the heme), nickel-containing superoxide dismutase (blue, PDB 1Q0D), and PsbQ protein of photosystem II- oxygen (wheat, PDB 1NZE). The nickel atoms of SmbP and nickel-containing superoxide dismutase are shown as small spheres in blue and wheat color, respectively. The positions of the nickel atoms in these two proteins appear to be closely aligned.

Table 1.1 SmbP DALI Server Results

N	ST	Z	R	L	L	%	NF	PROTEIN
1	256	1	1.	7	10	1	4	Cytochrome b562 (oxidized)
2	1nz	9	1.	8	11	5	3	PsbQ protein of photosystem II- oxygen-evolving
3	1w	8	1.	8	13	7	3	Histidine-containing phosphotransfer protein ZmHP2
4	1sr	8	2.	8	11	1	2	YojN Histidine-phosphotransferase (HPt) domain
5	1q0	8	2.	7	11	1	5	Ni-containing superoxide dismutase
6	1xz	8	2.	8	45	2	4	GTP-binding protein TrmE from *Thermotoga*
7	1wf	8	1.	6	93	1	2	Mouse MIT domain
8	1kt	8	1.	6	35	8	2	Large FKBP-like Protein, FKBP51
9	1x8	7	1.	6	15	1	2	Pectin methylesterase inhibitor from *Arabidopsis*
1	1zu	7	2.	6	15	1	4	Plant Tom20 mitochondrial import receptor from

The following notation is used for data columns in Table 1.1:

STRID1/STRID2	PDB identifiers of search structure and aligned structure with chain identifier
Z	Z-score, i.e., strength of structural similarity in standard deviations above expected. The matched structures are sorted by Z-score. Only matches above a threshold of Z=2 are reported.
RMSD	Positional root mean square deviation of superimposed Cα atoms in Angstroms
LALI	Total number of equivalent residues
LSEQ2	Length of the entire chain of the equivalent structure
%IDE	Percentage of sequence identity over equivalent positions
NFRAG	Total number of equivalent fragments
PROTEIN	COMPND record from the PDB file of the aligned structure

EPR Characterization of Metal Binding Sites

The EPR spectra of SmbP versus equivalents of Cu^{2+} added (Figure 1.10) indicated SmbP can bind approximately 6-8 equivalents of Cu^{2+} [5]. Simulation of the EPR spectrum of SmbP with one equivalent of Cu^{2+} indicated that the first binding site has a square-planar geometry with four nitrogen ligands (Figure 1.11). By comparison, the simulation of the difference spectra between four and two

equivalents of Cu^{2+} indicated a tetragonal symmetry with three nitrogen/one oxygen or two nitrogen/two oxygen ligands [5].

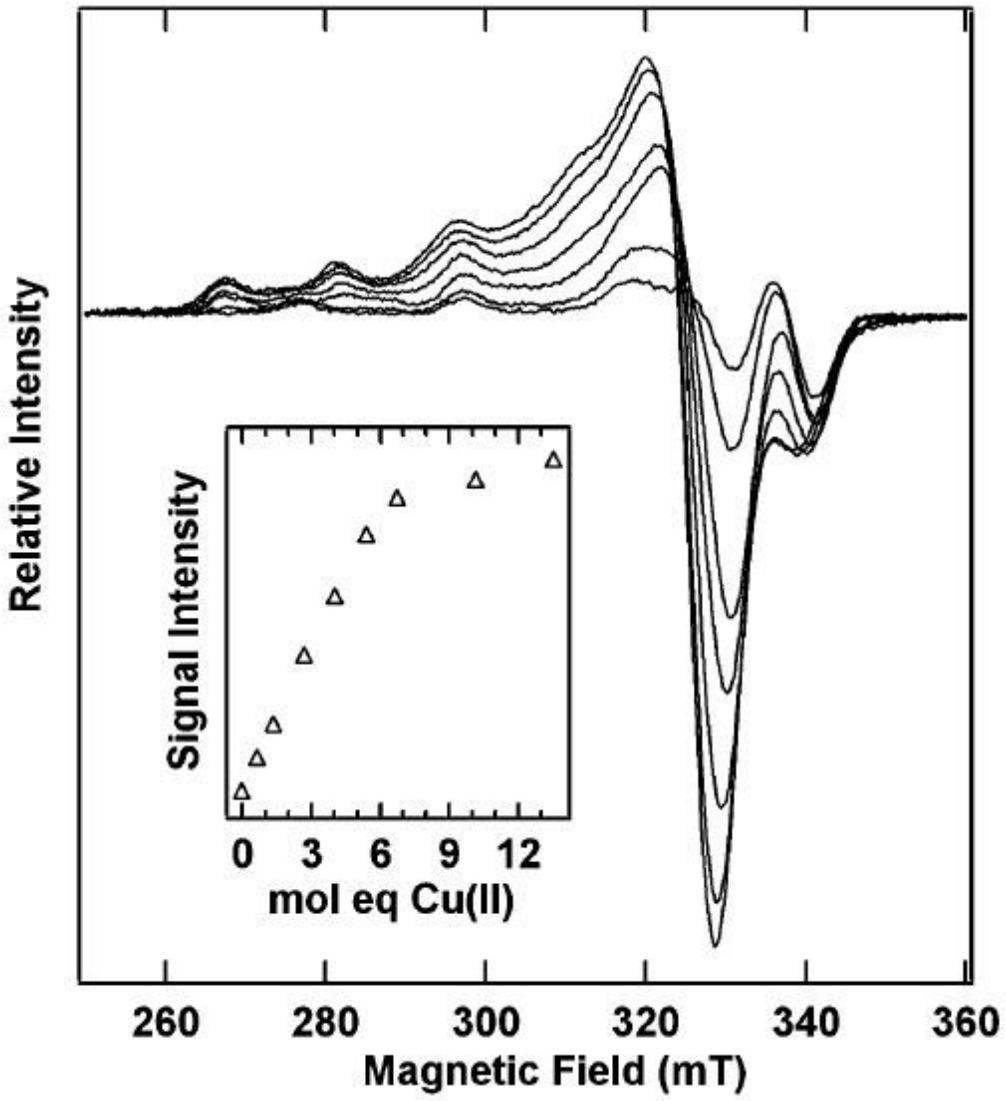

Figure 1.10: EPR spectra of SmbP with increasing equivalents of Cu^{2+} (0.7, 1.4, 2.7, 4.0, 5.4, 6.8, 10.2, and 13.6) (From Barney, B.M. et al. 2004 [5]).

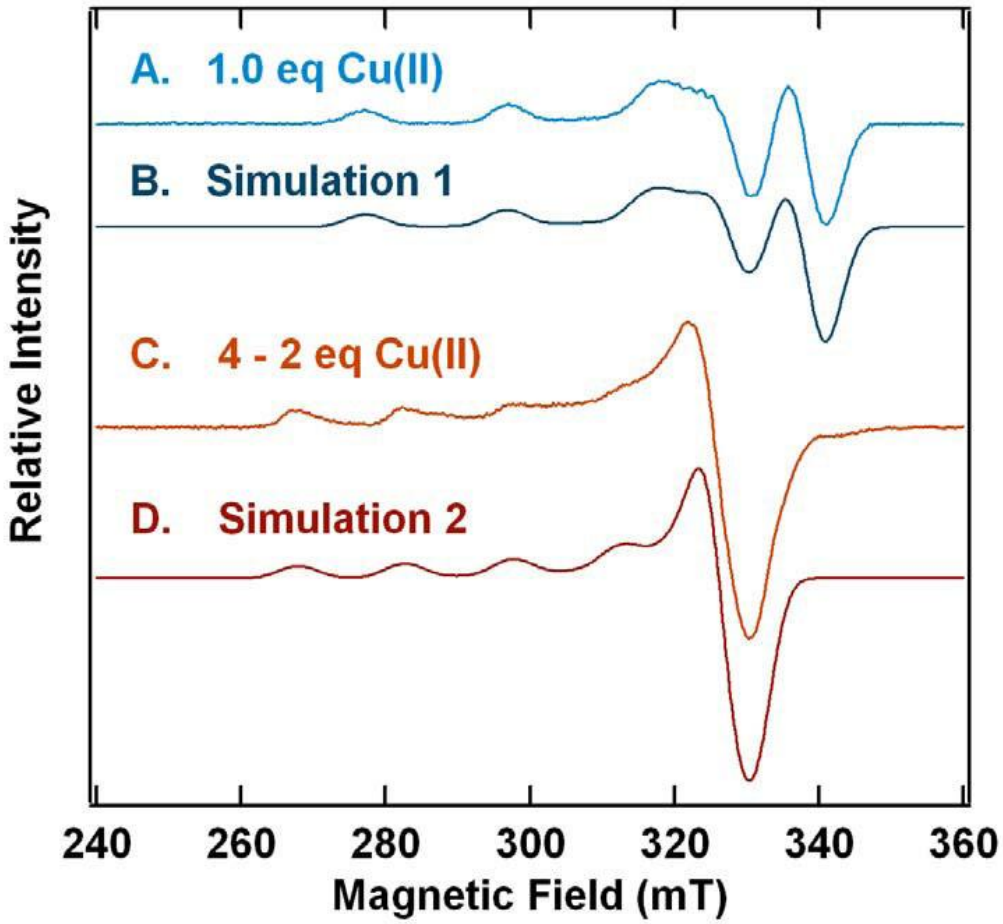

Figure 1.11: Simulations of EPR spectra of SmbP with different equivalents of Cu^{2+}. Simulation 1 (B) is the simulation of EPR spectrum of SmbP with one equivalent of Cu^{2+} (A); while Simulation 2 (D) is the simulation of difference spectra between four and two equivalents of Cu^{2+}, by subtracting the spectrum of two equivalents of Cu^{2+} from the spectrum of four equivalents of Cu^{2+} (C) (From Barney, B.M. et al. 2004 [5]).

Chapter 2

SEARCHING HOMOLOG GENES OF SMALL METAL BINDING PROTEIN (SMBP) & PREDICTION OF METAL BINDING SITES

Summary

Several homologs of the small metal binding protein (SmbP) have been identified by searching the BioCyc Database and the Orthologuous MAtrix (OMA) browser. None of the identified homologs have yet been characterized. They are all in isolated open reading frames, which therefore provides little insight into the possible function of SmbP or these homologs. However, all of these homologs are from a few specific chemolithoautotrophs, which live on ammonia or methane oxidation as their sole energy and electron sources. This supports our proposal that SmbP and its homologs have a role in the cellular management of copper, because copper is a cofactor in some essential enzymes, including ammonia monooxygenase (AMO) and particulate methane monooxygenase (pMMO), involved in the oxidation of ammonia or methane in these organisms. Also, a sequence alignment of SmbP homologs has identified a few highly conserved residues, including glycine, alanine, leucine, isoleucine and valine, which are responsible for the hydrophobic core of the four-helix bundle structure of SmbP. Some other highly conserved residues are several histidines, which, along with nearby acidic residues are predicted by the CHED server to coordinate the six copper atoms that can tightly bind to these proteins.

Introduction

Search and Analysis of Homologs of SmbP

Searching for homolog genes of a protein and analyzing the functions of their protein products is a practical method to infer the function of a novel protein. There are a few different bioinformatics programs that are able to perform these type of analysis. Among them, the BioCyc Database [15][16] and the OMA browser [17][18] were used to search and analyze homologs of SmbP in this study, as these servers can easily search homologs from hundreds of organisms with known genome sequences and generate accurate results.

After identifying the homologs, sequence alignment can provide useful information such as conserved residues. Such residues normally play important roles in the protein, such as maintaining the protein's structure and stability, binding cofactors, or interfacing with other macrobiomolecules (proteins, DNA, or RNA).

Prediction of Cu^{2+} Binding Sites

Two types of metal binding site prediction programs, the FINDSITE-metal server and the CHED server, have been used to predict the binding sites of Cu^{2+} in SmbP.

The FINDSITE-metal server [19] predicts metal-binding sites, ligand residues and binding metal preferences from evolutionarily related templates detected by threading [20].

The CHED server [21] predicts 3D intra-chain protein binding sites for transition metals by using the "CHED" algorithm, which searches for a triad of amino acids composed of 4 types of residues, including Cys (C), His (H), Glu (E), and Asp (D), within specific distances [22].

Methods

Search and Analysis of Homologs of SmbP

All searching and analysis on the BioCyc Database [15] and the OMA browser [17] were done by following their published instructions. The sequence alignments of homologs was generated by using the BioEdit software [23].

Prediction of Cu(II) Binding Sites

To predict the metal binding sites, the pdb file of SmbP was uploaded and analyzed by both the FINDSITE-metal server [20] and the CHED server [22]. The resulting graphics images were produced using the UCSF Chimera package [24] or the YASARA program [25] with the FoldX plugin [26].

Results

Homologs of SmbP Identified from BioCyc Database

The gene number for *smbp* (NE2461) was used to search for homolog genes from other organism on the BioCyc Database [15]. Out of more than 500 organisms on the database as of version 14.0 (March 2010), only 2 organisms (*Nitrosospira multiformis* ATCC 25196 and *Methylococcus capsulatus* Bath)

contain homologous genes to *smbp* (blue shadow arrow boxes in Figure 2.1). Neither of the two homologs has been characterized yet, which confirms that SmbP is a novel protein.

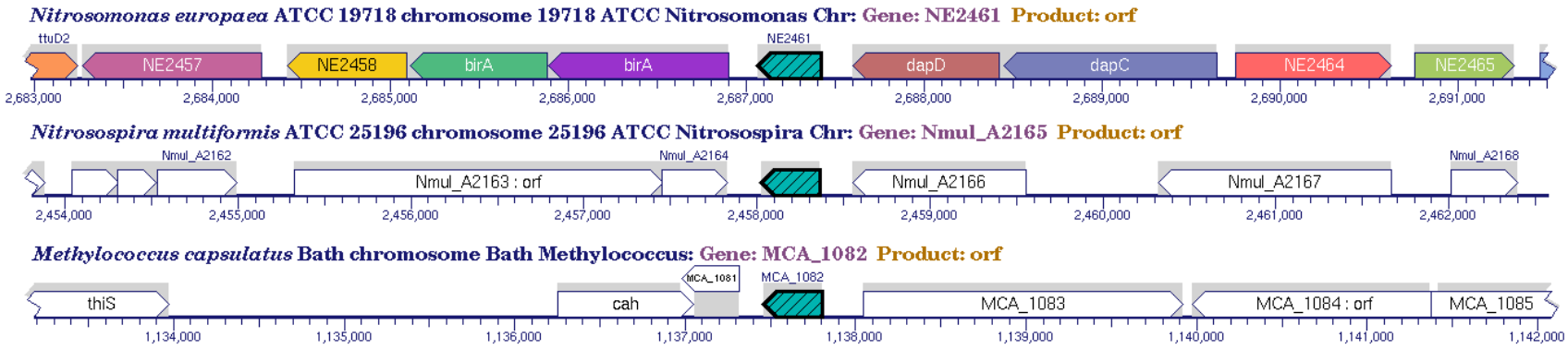

Figure 2.1: Search results of homologous genes of *smbp* (NE2461) from *Nitrosomonas europaea ATCC 19718*. Nmul_A2165 from *Nitrosospira multiformis* ATCC 25196 and MCA_1082 from *Methylococcus capsulatus* Bath are shown as blue shadow arrow boxes in the center of the chromosome pieces.

Smbp and its homologs are located in an isolated open reading frame (ORF), which contains *smbp* or its homologs as the only gene. Generally, the functions of unknown gene can be predicted by analyzing the function of other genes in a given ORF, but *smbp* and its homologous genes are all the sole gene in their respective ORF. Without association with other genes of known function in the same reading frame, no information of the possible functions of SmbP could be obtained from this ORF analysis.

No homolog could be found in more than 500 other organisms in the database, even in some genetically closely related organisms of the above three. Such organisms include: *Nitrobacter winogradskyi* Nb-255, *Nitrobacter hamburgensis* X14, *Methylobacillus flagellatus* KT, *Methylibium petroleiphilum* PM1, *Methanothermobacter thermautotrophicus* Delta H, *Methanosphaera stadtmanae* DSM 3091, *Methanosarcina mazei* Go1, *Methanosarcina barkeri* Fusaro, *Methanosarcina acetivorans* C2A, *Methanosaeta thermophila* PT, *Methanopyrus kandleri* AV19, *Methanocorpusculum labreanum* Z, *Methanococcus maripaludis* S2, *Methanococcus maripaludis* C7, *Methanococcus maripaludis* C5, and *Methanococcoides burtonii* DSM 6242.

Homologs of SmbP Identified from the OMA Browser

By searching the OMA browser [17] for homologs of *smbp* (NE2461), five homologs were identified from four organisms (Table 2.1).

Table 2.1: The search results of *smbp*'s homologs by the OMA browser

Homologs	Organisms	Entry No.
smbp	*Nitrosomonas europaea*	NITEU02359
1:1 homologs	*Nitrosospira multiformis* (strain ATCC 25196 / NCIMB 11849)	NITMU02078
	Nitrosococcus oceani (strain ATCC 19707 / NCIMB 11848)	NITOC01647
	Methylococcus capsulatus	METCA01005
1:2 homologs	*Nitrosomonas eutropha* (strain C71)	NITEC00219, NITEC00220

While *Nitrosomonas europaea*, *Nitrosospira multiformis*, *Nitrosococcus oceani*, and *Methylococcus capsulatus* contain only a single copy of a homolog, *Nitrosomonas eutropha* contains two copies, which are clustered together with a distance of only 86 base pairs in the genome. It implies that possible duplication happened after the speciation event of *Nitrosomonas eutropha*, which is confirmed by the calculated distance tree in Figure 2.2. The reason of why there are two copies of this gene in *Nitrosomonas eutropha* is not yet known; however, it could be an indication of an important function of the protein in these organisms.

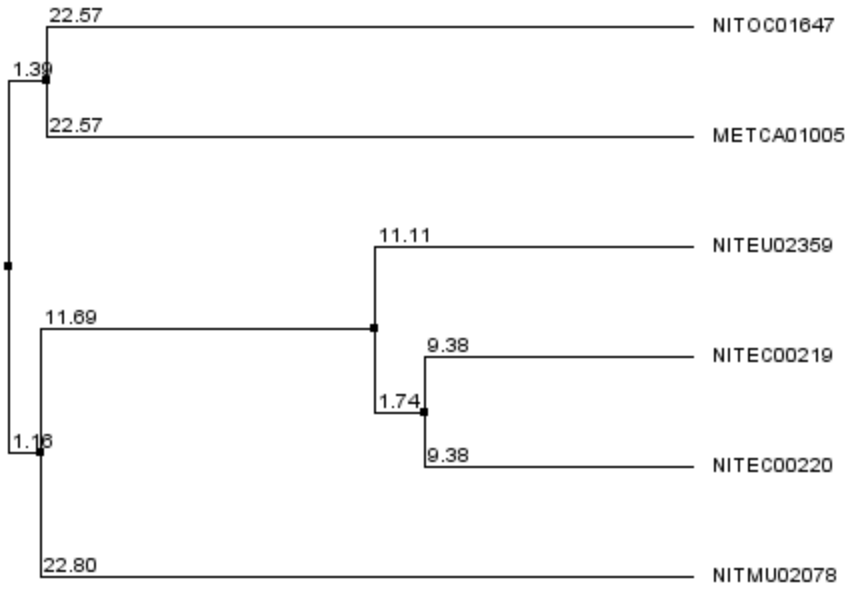

Figure 2.2: Calculated average distance tree of SmbP and its homologs using percentage of identity (PID) by jalview[24]. From top to down, NITOC01647 is from *Nitrosococcus oceani*, METCA01005 is from *Methylococcus capsulatus*, NITEU02359 (*smbp*) is from *Nitrosomonas europaea*, NITEC00219 and NITEC00220 are from *Nitrosomonas eutropha*, and NITMU02078 is from *Nitrosospira multiformis*

Protein Sequence Alignment of SmbP and its Homologs

The protein sequences of SmbP and its homologs are aligned in Figure 2.3. Most histidines are highly conserved across the 5 organisms. Six out of 17 histidines in SmbP are invariant in its homologs, while another four histidines are highly (more than 66%) conserved. Besides histidines, the other highly conserved residues are glycines, alanines, leucines, isoleucines and valines.

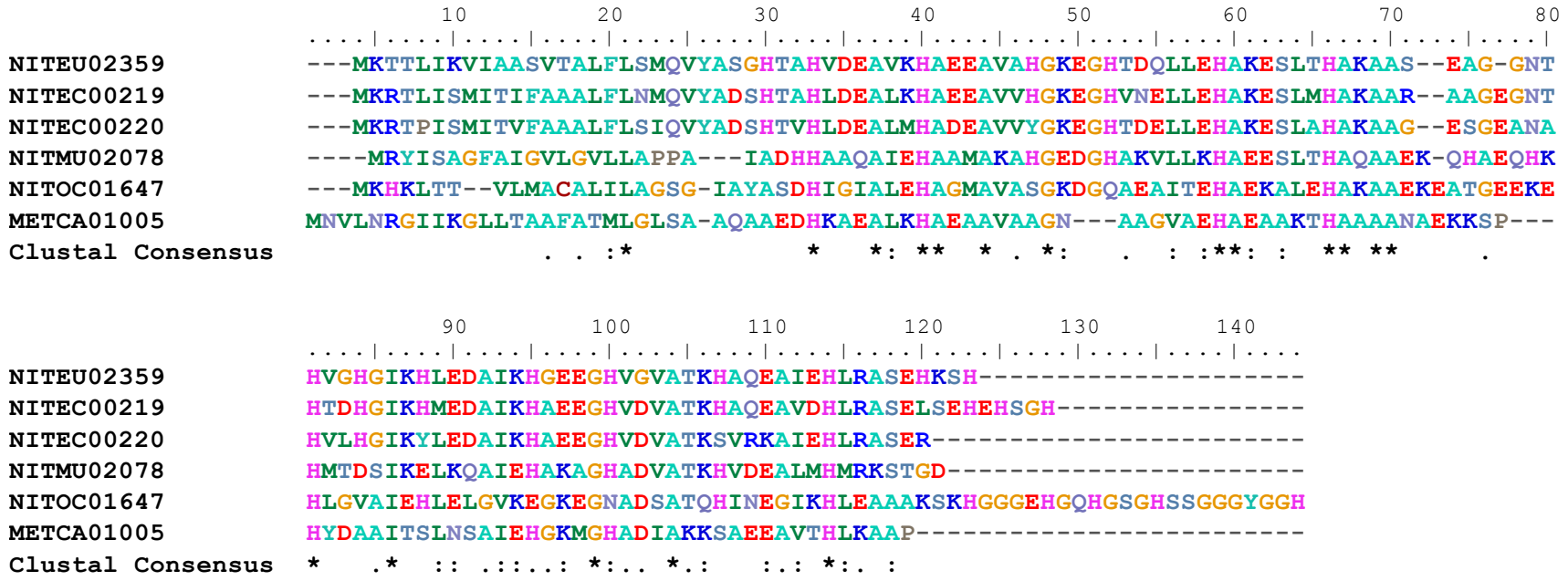

Figure 2.3: The protein sequence alignment of SmbP and its homologs shows highly conserved residues of histidines, glycines, alanines, leucines, isoleucines and valines.

Prediction of Cu(II) Binding Sites

The FINDSITE-metal server (http://cssb.biology.gatech.edu/findsite-metal) predicts metal-binding sites and binding metal preferences from evolutionarily related templates detected in PDB database [20]. However, because there is no homolog of SmbP in the PDB database, this method was not practical for providing a good prediction of the metal binding sites of SmbP.

The CHED server (http://ligin.weizmann.ac.il/~lpgerzon/mbs4/mbs.cgi) predicts 3D intra-chain protein binding sites for transition metals by using the "CHED" algorithm, which searches for a triad of amino acids composed of four types of residues, Cys (C), His (H), Glu (E), and Asp (D), within specific distances [22]. Considering the possible structural rearrangement upon metal binding, it allows one target residue to rotate in rotamer space. A binding site is considered to be true if one or more correct amino acid ligands have been predicted. It predicted a total of 18 residues from SmbP as possible ligands to Cu^{2+} (Table 2.2). The 18 residues were divided into two sets. One contains 8 residues, and other contains 10 residues. They include most of the His residues that line up at uniform spaces along the helices, and also some Glu/Asp/His residues around them (Figure 2.4).

Table 2.2: Predicted residues in Cu^{2+} binding sites by the CHED server

Set No.	Number of Residues	Predicted Set of Residues
1	8	E9, H13, H39, H32, E35, E16, H20, E31
2	10	H51, H54, H84, H58, E80, D61, H77, H65, E68, H70

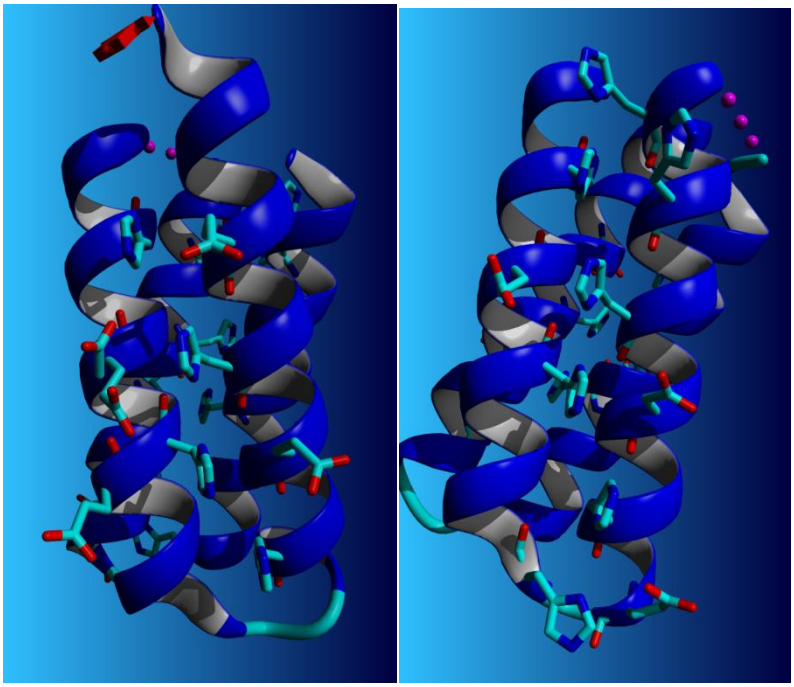

Figure 2.4: Structure of SmbP with predicted ligands residues shown as sticks.

One of the weaknesses of this method is that it cannot separate binding sites when they are considerably closed to each other. In the case of SmbP, it predicted the ligand residues of six Cu^{2+} binding sites as two sets, and each set contains three Cu^{2+} binding sites.

Discussion

The fact that SmbP and its homologs exist in only a few specific chemolithoautotrophs, which live by ammonia or methane oxidation as their sole energy and electron source, indicates that they may fulfill some unique function(s) in these organisms. This is in agreement with our proposal that SmbP may have a role in the cellular management of copper [5], which is a cofactor for a number of essential enzymes in the oxidation of ammonia or methane, including ammonia monooxygenase (AMO) and particulate methane monooxygenase (pMMO)[3, 27].

The protein sequence alignment of SmbP and its homologs shows highly conserved residues of histidines, glycines, alanines, leucines, isoleucines and valines (Figure 2.3).

The conserved histidine residues in SmbP and its homologs line up at uniform spacings along the helices in crystal structure of SmbP, which indicates the importance of histidine at specific positions in the proteins for proper functions. Data will be presented to provide additional evidence that these histidines are responsible for structural stability of SmbP and the binding affinity of SmbP to Cu^{2+} and other metal ions.

According to our X-ray crystal structure, two invariant glycines at positions 51 and 99 in Figure 2.3 are in the turn regions connecting two helices, and they serve as terminators of the previous helix. The other highly conserved residues are hydrophobic residues, such as alanine, leucine, isoleucine and valine. They form the hydrophobic core of the four α-helix bundle structure of SmbP.

It's interesting that all of the predicted Cu^{2+} binding sites are located on the interfaces, between helices 1 and 2 and between helices 3 and 4, with weaker hydrophobic interactions and ionic bonds. As a result of the stronger metal-ligands interactions forming between Cu^{2+} and residues of histidine, glutamate acid and aspartate acid on the two interfaces, Cu^{2+} can greatly increase the structural stability of SmbP, which will be discussed in details in next chapter.

Chapter 3

COPPER (II) CAN INCREASE THE STRUCTURAL STABILITY OF SMBP

Summary

Preliminary folding experiments indicated that Cu^{2+} greatly stabilizes SmbP. In this study, experiments designed to elucidate the role of Cu^{2+} in increasing the stability of SmbP structure were conducted by circular dichroism (CD) spectroscopy. The protein folding data was used to elucidate the role of Cu^{2+} in stabilizing SmbP structure against unfolding induced by decreased pH, increased temperature, and chemical denaturants. The significant stabilization effects of Cu^{2+} were demonstrated by the observation that Cu^{2+}-bound SmbP remained fully folded under extreme environment conditions, such as acidic pH, 96 ℃ , and 8 M urea. Also, it was shown that Cu^{2+} is able to induce the refolding of denatured SmbP in acidic solutions. However, reduced stabilization effects of Cu^{2+} were observed in basic solution, such as pH 8, and higher pH than 9 even results in destabilization effects of Cu^{2+} on SmbP structures. These findings imply that coordination of Cu^{2+} to histidine and acidic residues is responsible for the stabilization effects, while coordination of Cu^{2+} to lysine residues in basic pH counteracts the stabilization.

Introduction

Circular Dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light due to structural asymmetry. It is useful in determining the optical isomerism and

secondary structure of molecules, such as nucleic acids and proteins (Figure 3.1 and Table 3.1).

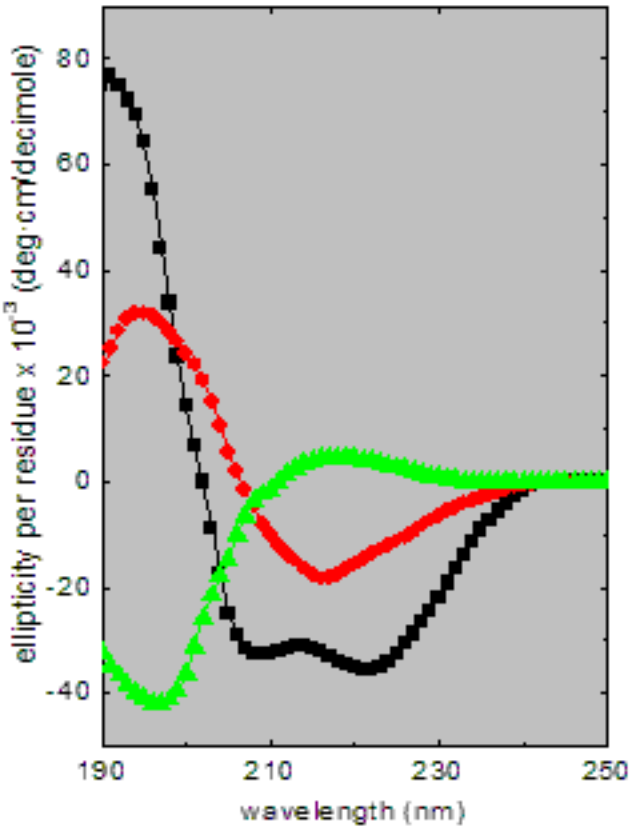

Figure 3.1: The circular dichroism (CD) spectra of three typical secondary structures of proteins: α-helix in black, β-sheet in red and random coil in green [28].

Table 3.1: The negative (-) and positive (+) peaks of each secondary structure of proteins in CD spectroscopy.

Secondary Structure	- peak (nm)	+ peak (nm)
α-Helix	**222 & 208**	**192**
β-Sheet	**216**	**195**
Random Coil	**200**	**212**

The CD spectrum of folded SmbP is the classical example of an α-helical protein displaying with two negative peaks at 222 nm and 208 nm and one positive peak at 192 nm [5]. Here, we report that the secondary structure of SmbP is effectively stabilized by Cu^{2+} ions against thermal denaturation, pH change, and chemical denaturants. In addition it has been observed that the refolding of SmbP in acidic solutions can be induced by Cu^{2+}.

Materials and Methods

Materials

Luria-Bertani (LB) broth was obtained from Fisher Biotech. Ni-NTA His·Bind resin was obtained from Novagen. SDS-PAGE was performed using equipment and reagents from BioRad Laboratories, Inc. Proteins were visualized with Coomassie Brilliant Blue R-250. Isopropyl-1-thio-D-galactopyranoside (IPTG) was obtained from Calbiochem. All other chemicals and biochemicals were purchased from Sigma-Aldrich of the highest purity available.

Purification of SmbP

Recombinant SmbP was purified as previously described [5], with a few modifications. Briefly, *E. coli* BL21(DE3) cells harboring the plasmid pSMBP were grown on LB media containing kanamycin (30 mg/L) at 37 °C to A_{600} of 0.60. They were then induced with IPTG (50 mg/L), grown for an additional 4 h, and harvested by centrifuging at 7000g for 10 min. The periplasmic protein fraction was released by osmotic shock and the supernatant containing the periplasmic fraction was purified by chromatography on a Ni-NTA His·Bind column. SmbP was eluted as a pure protein with 250 mM imidazole. The purified SmbP was dialyzed against 5mM EDTA for 24 hours and against water for several days. The Protein was concentrated to about 50 mg/ml using a Vivaspin 15R (2,000 MWCO) centrifugal concentrator.

The purity of SmbP protein was determined by 17% SDS-PAGE. The protein concentration was determined spectrophotometrically, as previously described [5]. Briefly, due to the lack of aromatic residues in SmbP, the concentration of protein was determined by the simple ultraviolet (UV) spectrophotometric analysis of Waddell [29], based on the absorbance difference between 215 and 225 nm. The modified equation $100*(A_{215}-A_{225})$ was routinely used for calculation of concentration in micrograms per milliliter of SmbP in solution. The molecular weight of the protein was determined by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS).

Folding Analysis by Circular Dichroism

All CD spectra were recorded on a Jasco J-710 spectropolarimeter equipped with a Peltier temperature control unit. The samples were prepared to a final concentration of SmbP of 30μM, in 20 mM pH buffer in the absence or presence of 300 μM $CuSO_4$, unless stated otherwise. CD cuvettes with a path length of 0.1 cm were used for most spectra recording. Far-UV CD spectra were recorded from 260 nm to 190 nm at 25 °C with the following settings: data pitch of 1 nm, scanning speed of 50 nm/min, response time of 4 s and band-width of 1.0 nm. Three spectra were collected and averaged for each sample, and the baseline of the corresponding buffer was subtracted from each spectrum. Experiments conducted at 3 and 30 μM SmbP provided identical results, except in the intensity of signals, indicating no concentration dependence within the range tested. Therefore, a concentration of 30 μM was used in all the following experiments to improve the signal-to-noise ratio.

CD spectra were recorded in θ (mdeg), which are then were converted to the mean residue ellipticity, [θ], ($deg.cm^2.dmol^{-1}.res^{-1}$) using equation 1:

$$[\theta] = \theta / [10*l*c*(N-1)] \quad (1)$$

where l is the pathlength of the cell in centimeters, c is the protein concentration in mol/L, and N is the number of residues per protein (93 for SmbP).

The converted CD data were calculated by using Selcon3 algorithms[30, 31] on the DichroWeb server [32] [33-37] to predict the percentage of each secondary structure.

Thermal Denaturation

Thermal denaturation was monitored over a temperature range of 4–96 °C in 20 mM 3-(N-morpholino)-propanesulfonic acid (MOPS) pH 6.4 buffer or 20 mM Tricine pH 8.0 buffer. Instrument settings were as follows: temperature slope of 30 °C/h, data pitch of 1.0 °C, delay time of 60 s, band width of 1.0 nm, sensitivity of 100 mdeg, and response time of 2 s. CD spectra were recorded before and after thermal denaturation to determine reversibility.

The temperature dependence of the θ_{222} was fitted to obtain the thermodynamic parameters (T_m, ΔH and ΔS) using Jasco's denatured analysis program assuming a two-state unfolding reaction with pre-transitional (folded state, θ_F) and post-transitional (unfolded state, θ_U) baseline corrections. The apparent equilibrium constant of unfolding (K_{eq}) and free-energy change (ΔG) in the transition region were calculated by using equation 2 and 3:

$$K_{eq} = (\theta(T) - \theta_F(T))/(\theta_U(T) - \theta(T)) \quad (2)$$

$$\Delta G = -RT\ln(K_{eq}) = \Delta H - T\Delta S \quad (3)$$

where θ is the ellipticity observed at the given temperature, and θ_F and θ_U are the ellipticities characteristic of the folded and unfolded protein, respectively, and extrapolated from the pre- and post-transition baselines. The van't Hoff plot of $\ln(K_{eq})$ against 1/T was linearly fitted to

$$\ln(K_{eq}) = \Delta S_m/R - \Delta H_m/R(1/T) \quad (4)$$

where ΔH_m and ΔS_m are the enthalpy and entropy changes upon unfolding at the half-denaturation temperature, T_m, respectively. The fit of the $[\theta]_{222}$ versus temperature was generated using equation 5 [38]:

$$[\theta](T) = ([\theta]_F(T)+K_{eq}[\theta_U](T))/(1+K_{eq}) \quad (5)$$

pH Stability

Determination of the pH stability of apo-SmbP in the presence or absence of 10 times [Cu^{2+}] was conducted by diluting a concentrated apo-SmbP stock solution to a final concentration of 30 μM in 20 mM various pH buffers (Table 3.2) with pH range from 2.0-12.8 with or without 300 μM $CuSO_4$.

CD spectra were recorded at 25°C after equilibrating at 4°C for more than 24 hours. The plot of $[\theta]_{222}$ vs pH was fitted by non-linear least-square analysis to equation 5, which is a transformation of the Hill equation:

$$[\theta] = \frac{[\theta]_U + [\theta]_F 10^{n(pH-pK_a)}}{1+10^{n(pH-pK_a)}} \quad (6)$$

where $[\theta]_F$ and $[\theta]_U$ are the ellipticities characteristic of the folded and unfolded protein, respectively, n is the Hill coefficient and pK_a is the half protonated pH.

Table 3.2: pH buffers and their pH ranges used for pH stability study of SmbP

Buffer Name	Conjugate Acid	Conjugate Base	pKa	pH range
Phosphoric acid	H_3PO_4	KH_2PO_4	2.12	2.0-3.0
Formic acid	HCO_2H	HCO_2K	3.75	2.9-4.7
Acetic acid	CH_3CO_2H	CH_3CO_2K	4.76	3.7-5.5
MES [1]	MES	MES-K	6.10	5.3-7.0
Potassium phosphate	KH_2PO_4	K_2HPO_4	7.20	6.4-8.2
MOPS [2]	MOPS	K-MOPS	7.20	6.4-7.5
Tricine [3]	Tricine	K-Tricine	8.15	7.3-8.7
Boric Acid	H_3BO_3	$KB(OH)_4$	9.14	8.2-9.8
Sodium bicarbonate	$NaHCO_3$	Na_2CO_3	10.25	9.4-11.0
CABS [4]	CABS	K-CABS	10.70	10.0-11.6
Potassium hydroxide [5]	H_2O	KOH	-	11.5-12.8

[1] MES: 2-(N-morpholino) ethanesulphonic acid

[2] MOPS: 3-(N-morpholino) propanesulphonic acid

[3] Tricine: N-[tris (hydroxymethyl) methy] glycine

[4] CABS: 4-(Cyclohexylamino)-1-butanesulfonic acid

[5] Potassium hydroxide: concentration ranges from 4 mM to 100 mM

Refolding Induced by Cu^{2+} in Acidic Solution

The Cu^{2+} induced refolding of SmbP was conducted in 20 mM sodium acetate buffer at pH 5.0 in a 3.5 mL cuvette with pathlength of 1.0 cm. Unfolded

SmbP protein, with an initial concentration of 3.1 μM, was titrated with $CuSO_4$ or $NiSO_4$ solutions. CD spectra were recorded after adding metal solution and reaching equilibrium with vigorous stirring for about 30 min at 25 °C. The volume change after the addition of metal solution was corrected for each recorded spectrum. $[\theta]_{222}$ was monitored as a function of metal ion concentration, ranging from 0 to 10 equivalents.

Unfolding Induced by Chemical Denaturants

Chemical induced denaturation was carried out by diluting a concentrated apo-SmbP stock solution to a final concentration of 30 μM in 20 mM of Tricine buffer at pH 8.0 with various concentrations of guanidine hydrochloride (0-6 M) or urea (0-8 M), in the presence or absence of 300 μM $CuSO_4$. After equilibrating overnight at 4°C, CD spectra were recorded at 25°C.

The plot of $[\theta]_{222}$ vs concentration of denaturant [D] was fitted by non-linear least-square analysis to equation 7:

$$[\theta] = \frac{[\theta]_F + [\theta]_U e^{-(\Delta G^{H_2O} - m[D])/RT}}{1 + e^{-(\Delta G^{H_2O} - m[D])/RT}} \qquad (7)$$

where $[\theta]_F$ and $[\theta]_U$ are the ellipticities characteristic of the folded and unfolded protein, respectively, [D] is the concentration of denaturant, ΔG^{H2O} is the Gibbs free energy change in water solution when [D] = 0, *m* is denaturant *m* values which is correlates very strongly with the amount of protein surface exposed to solvent upon unfolding [39].

Results

SmbP Has a Predominantly *α-Helical Structure*

At pH 7.0 and room temperature (25 ℃), SmbP in the absence or presence of Cu^{2+} has a classical α-helical CD spectrum with two negative peaks at 208 and 222 nm (Figure 3.2). This observation agrees with the predominantly α-helical structure prediction [5] and the X-ray crystal structure of SmbP. The calculation results using Selcon3 algorithms [30, 31] on the DichroWeb server [33-37] predicts about 70% α-helical structures, agreeing with the observation of 78% in X-ray crystal structure.

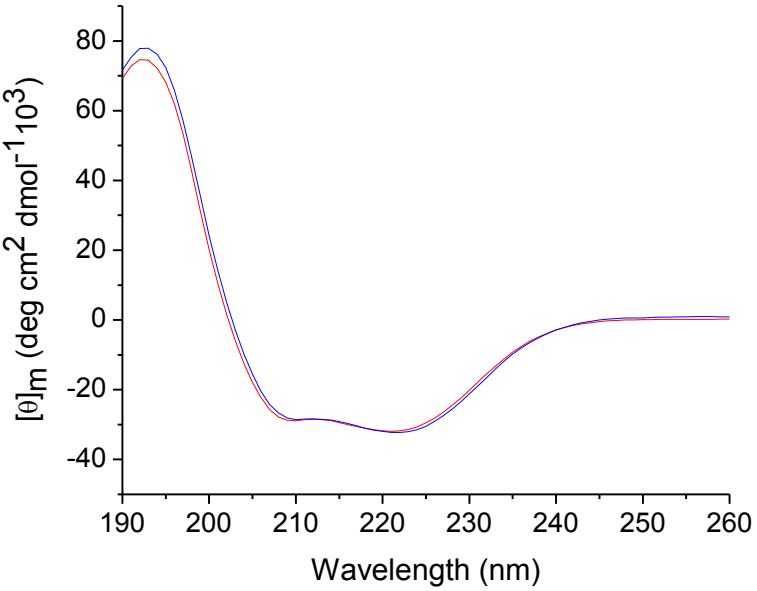

Figure 3.2: CD spectra of SmbP without (in red) and with (in blue) Cu^{2+} at pH 7 show the proteins are completely folded into α-helix structures.

Cu^{2+} Stabilizes SmbP against Thermal Denaturation

At pH 6.4 and room temperature, SmbP also has a classical α-helical CD spectrum. As the temperature increases, apo-SmbP converted from a folded α-helical structure to a unfolded random coil structure, as demonstrated by the spectra changes in Figure 3.3A. An isodichroic point of about 204 nm was observed, which implies a two-state transition from an α-helical structure to a random coil structure [40]. Because $[\theta]_{222}$ is most sensitive to the change of secondary structure from α-helices to random coils, it was plotted as a function of temperature in range 4 - 96 °C (Figure 3.3B) to monitor the unfolding process,. The plot shows a typical thermal denaturation curve in absence of Cu^{2+}, with three phases: a pre-melting phase, a melting transition phase and a post-melting phase. The pre-melting phase is lower than 27 °C, and has a slightly positive slope, which is a feature of an α-helical structure [41]. The melting transition phase is between 27 °C and 72 °C, and has a dramatic change in $[\theta]_{222}$. The post-melting phase is higher than 72 °C, and has a slightly negative slope which is a feature of a random coil structure [41]. The three phases indicate that the thermal denaturation of SmbP is a secondary structural change from an α-helical structure to a random coil structure. By fitting the denaturation curve, a melting temperature of 48.5 °C was extracted.

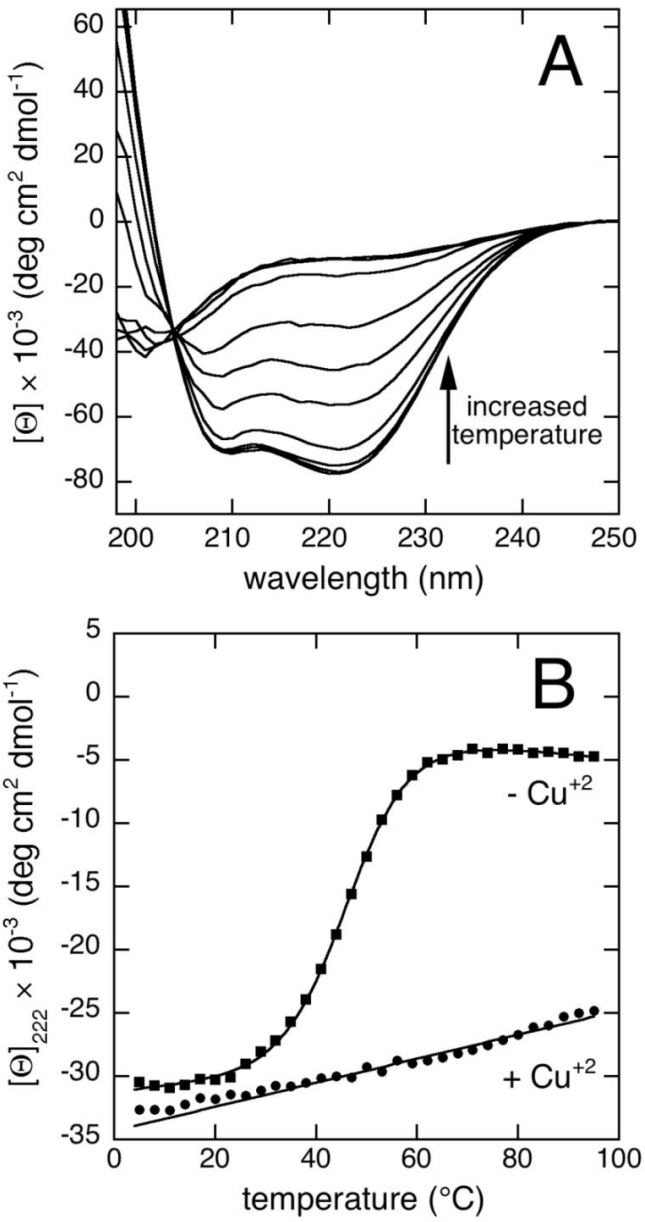

Figure 3.3: Temperature stability of SmbP at pH 6.4. (A) Far–UV CD spectra of apo-SmbP as a function of temperature. (B) Unfolding of SmbP (30 μM) in the absence (- Cu^{2+}) (solid square) and presence of 300 μM Cu^{2+} (+ Cu^{2+}) (solid cycle). The solid lines are fitting results of the data by equation 5. While apo-SmbP melts at 48.5 °C, the protein in the presence of Cu^{2+} does not denature up to 96 °C.

When the temperature is reduced after thermal denaturation, the protein is observed to regain its α-helical structure. In fact, CD spectra, at the same temperature during the unfolding process (increasing temperature) and the refolding process (decreasing temperature), are identical (Figure 3.4). It indicates that the thermal denaturation of SmbP is a completely reversible process.

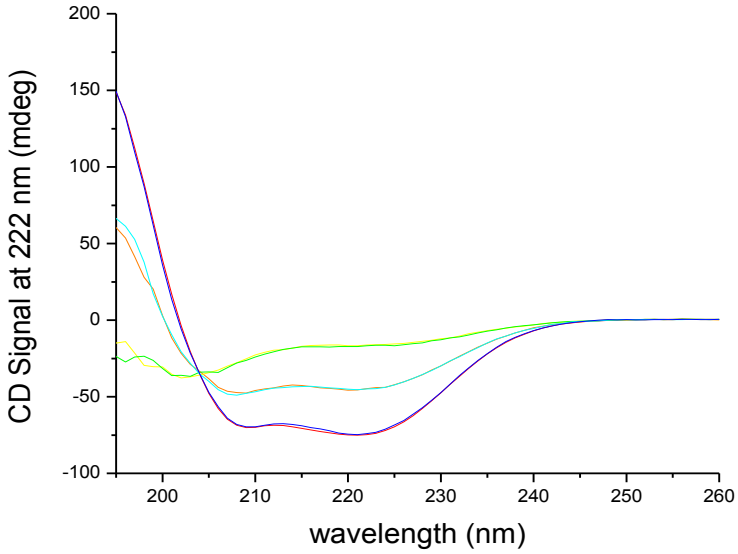

Figure 3.4: CD Spectra of apo-SmbP at different temperatures. As the temperature is increased: 20 °C (red), 45 °C (golden), and 60 °C (yellow), and after complete unfolding at 95 °C as the temperature is decreased: 60 °C (green), 45 °C (cyan), and 20 °C (blue).

In the presence of 10 equivalents of Cu^{2+}, only the pre-melting phase with a positive slope baseline was observed, and no melting transition or post-melting phases were observed up to a temperature of 96 °C (Figure 3.3B), which is very

close to the boiling point of water. The small change in ellipticity at 222 nm observed for the protein in the presence of Cu^{2+} correlates with the linear relationship of the ellipticity of an α-helical structure with temperature, and it is the same as the positive-sloped baseline observed in the pre-melting phase of SmbP in the absence of Cu^{2+} [41]. The results indicate that Cu^{2+}-SmbP still forms a predominantly α-helical structure even at 96 °C. In other words, Cu^{2+} increases the stability of SmbP so significantly that its melting temperature increases over 50 °C. As a result, we are not able to characterize of the thermodynamic folding parameters of Cu^{2+}-SmbP at this pH. It is, therefore, of great interest to quantitively measure the effects of Cu^{2+} on the structure stability of SmbP protein against increases on temperature.

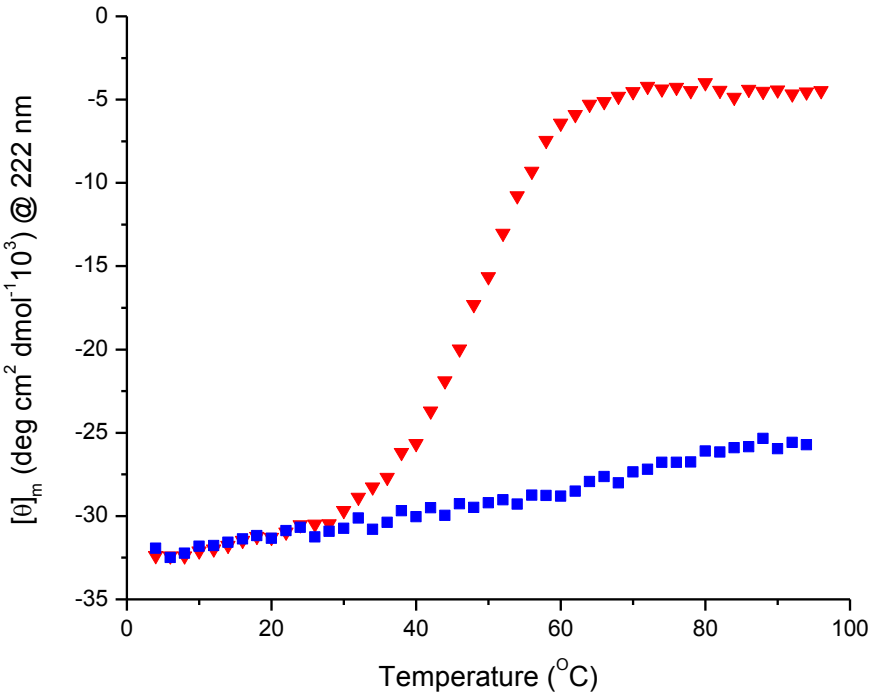

Figure 3. 5: Temperature stability of SmbP at pH 7.0. Unfolding of SmbP (30 μM) in the absence (- Cu^{2+}) (red) and presence of 300 μM Cu^{2+} (+ Cu^{2+}) (blue).

To overcome this difficulty, thermal denaturation experiments were conducted in several different pH conditions. At pH 7.0, SmbP in presence of Cu^{2+} remains completely folded at 96 °C (Figure 3. 5), which is similar to the observation at pH 6.4 (Figure 3.3). However, at pH 8.0, a normal denaturation process is observed, with a pre-melting phase, a melting transition phase and a post-melting phase (Figure 3.6). At this pH, the melting temperature of SmbP shifts to a higher temperature, from 45.1 °C in the absence of Cu^{2+} to 58.3 °C in the presence of 10 equivalents of Cu^{2+} (Figure 3.6 and Figure 3.7). The first seven

equivalents of Cu^{2+} have more drastic effects on the melting temperature, with an increase of more than 1.0 °C for each equivalent of Cu^{2+}. However, the last three equivalents have less effects, with an increase of less than 1.0 °C for each equivalent of Cu^{2+} (Figure 3.7 and Table 3.3). The enthalpy change upon unfolding at the half-denaturation temperature also increases from 143 kJ/mol to 187 kJ/mol in the presence of 10 equivalents of Cu^{2+}, while the respective entropy change increases from 448 J/mol/K to 565 J/mol/K in presence of 10 equivalents of Cu^{2+} (Table 3.3).

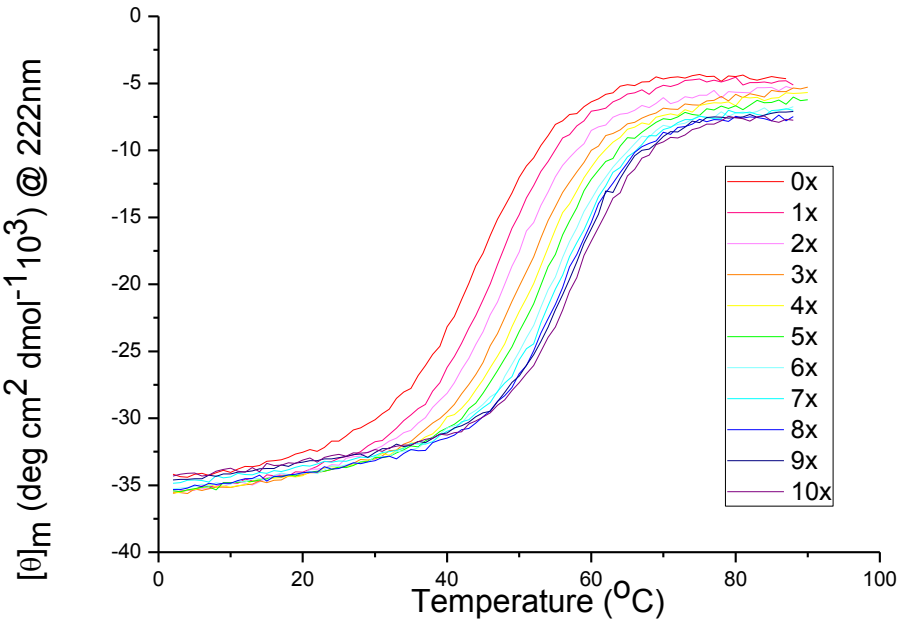

Fig. 5. Thermal Denat

Figure 3.6: Thermal denaturation curve of SmbP shifts to higher temperature in the presence of increasing equivalents of Cu^{2+} at pH 8.0. Each line is the addition of one equivalent of Cu^{2+} from 0 on the left (red), and a maximum of 10 equivalents on the right (purple).

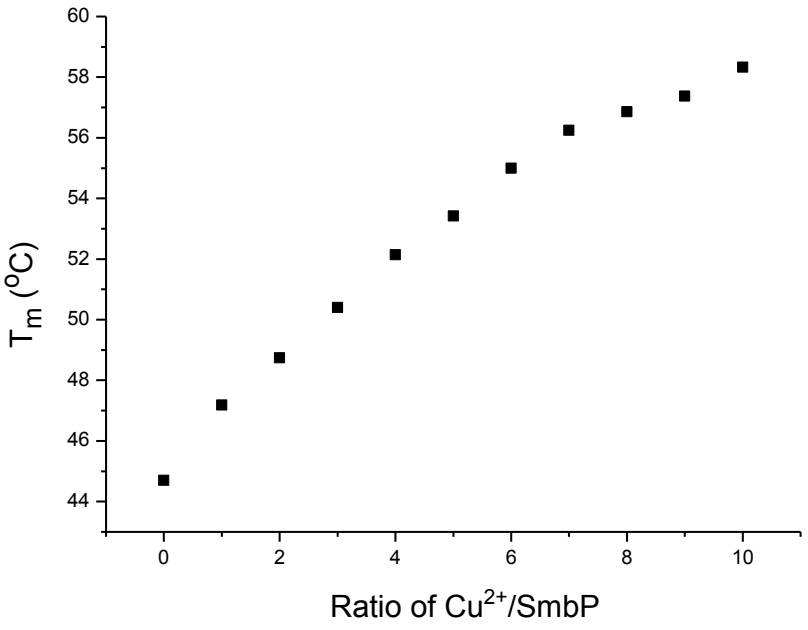

Figure 3.7: The plot of the melting temperature versus increasing equivalents of Cu^{2+} at pH 8.0.

Table 3.3: Effects of Cu^{2+} on the unfolding Tm, ΔH and ΔS of SmbP at pH 8.0

Cu/SmbP	Tm [^{o}C]	Stepwise increase of Tm [^{o}C]	ΔH [kJ/mol]	Stepwise increase of ΔH [kJ/mol]	ΔS [J/mol/K]	Stepwise increase of ΔS [J/mol/K]
0	45.1	-	143	-	448	-
1	47.2	2.1	152	9	475	27
2	48.7	1.5	161	9	500	25
3	50.4	1.7	166	5	512	12
4	52.1	1.7	171	5	526	14
5	53.4	1.3	173	2	530	4
6	55.0	1.6	177	4	539	9
7	56.2	1.2	177	0	537	-2
8	56.9	0.7	183	6	556	19
9	57.4	0.5	190	7	575	19
10	58.3	0.9	187	-3	565	10

Cu^{2+} Stabilized SmbP against pH Changes

The CD spectra of SmbP in the absence and presence of 10 equivalents of Cu^{2+} in a series of pH buffers were recorded. A decrease in pH results in the unfolding of SmbP from an α-helical structure to a random coil, as shown in Figure 3.8A. In the absence of Cu^{2+}, the unfolding transition occurs between pH

6.4 and 5.4, as indicated by a rapid and marked reduction of $-[\theta]_{222}$ (Figure 3.8B). By fitting these data to equation 5, a Hill coefficient (n) of 2.6, and a pH of half unfolded protein (pK_a) of 5.9 were observed (Table 3.4). After the complete unfolding of SmbP has been induced by acidic pH, the CD spectrum shows less than 10% residual helical content, based on the analysis results using the Selcon3 algorithms [30, 31] on the DichroWeb server [33-37], By comparison, fully folded SmbP at neutral pH consists of about 70% α-helical secondary structure.

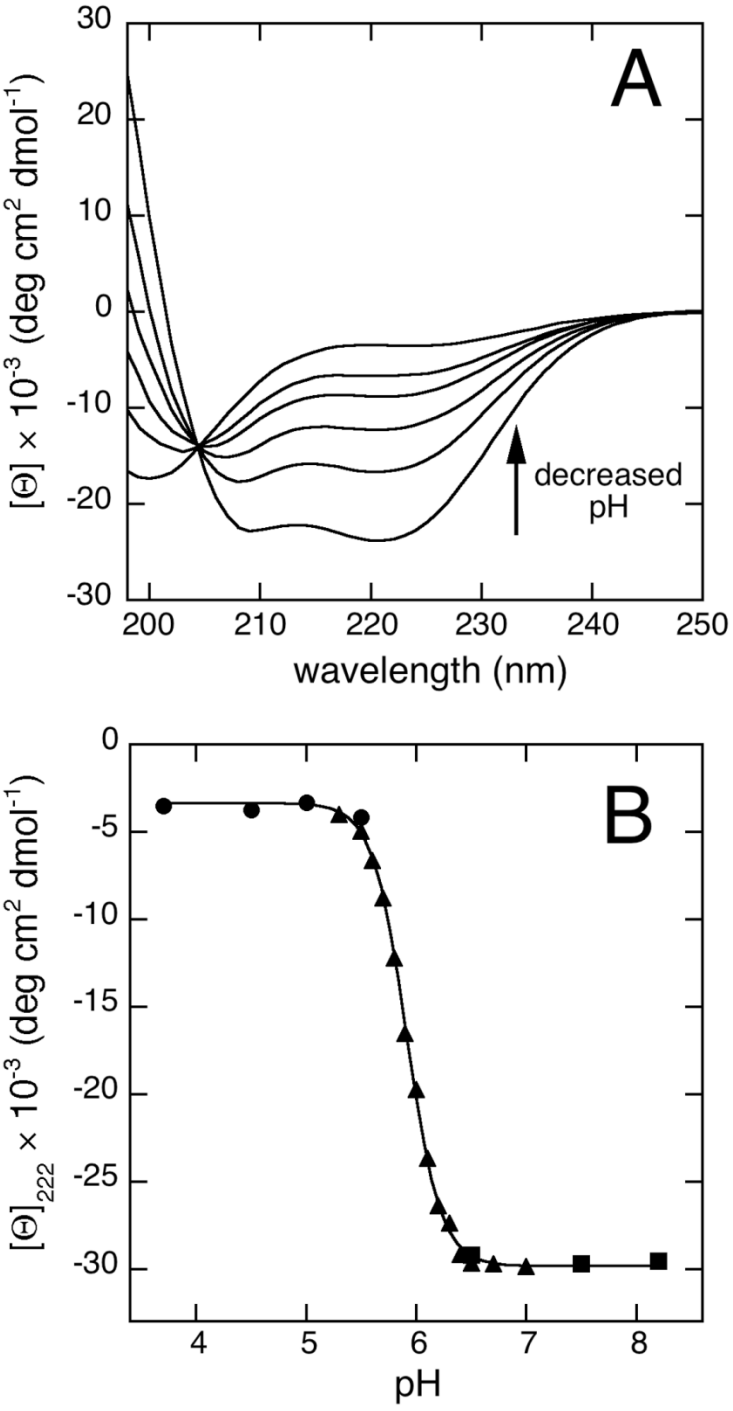

Figure 3.8: Acid (H^+) induced unfolding of SmbP in the absence of Cu^{2+}.

Ni^{2+} shows a weaker effect on the stability of SmbP structure over this pH range, by shifting the pK_a to 5.5 (0.4 pH unites only), and increasing n to 3.8 (Figure 3.9 and Table 3.4).

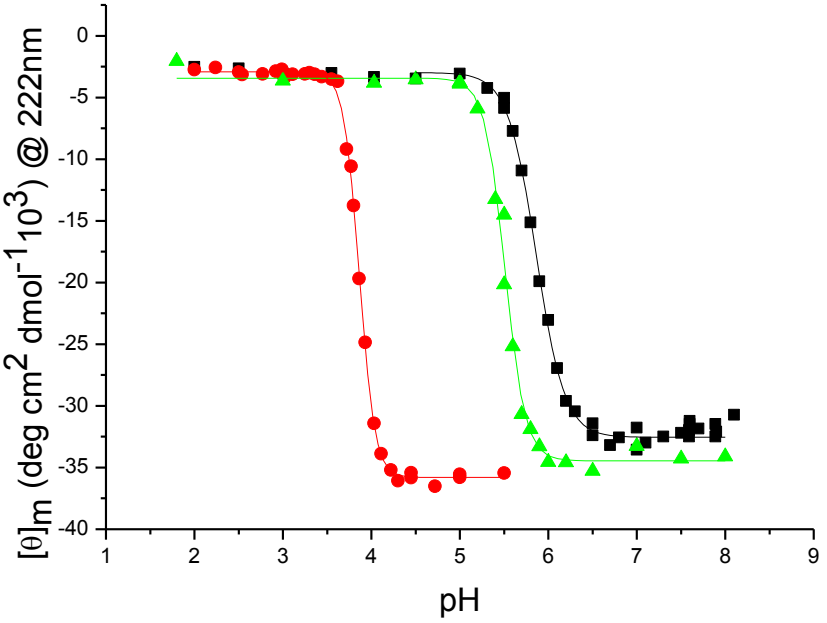

Figure 3.9: Acid (H^+) induced unfolding of SmbP in the presence of no metal (black), 10 molar equivalents of Ni^{2+} (green), or 10 molar equivalents of Cu^{2+} (red).

Table 3.4: Data fitting results for the acid induced unfolding of SmbP in the presence of no metal, 10 molar equivalents of Ni^{2+}, or 10 molar equivalents of Cu^{2+}.

Parameters	Apo-SmbP	Ni-SmbP	Cu-SmbP
U	-3	-3.4	-2.9
F	-32.6	-34.5	-35.8
n	2.6	3.8	4.9
pK_a	5.86	5.51	3.86

In presence of 10 molar equivalents of Cu^{2+}, SmbP is stabilized against the acidic pH change, as the unfolding of SmbP is observed at much lower pH in the presence of Cu^{2+} than without metal, shown in Figure 3C. In fact, the pH at which half SmbP is unfolded is lowered by 2.0 pH units to a pK_a = 3.9 in presence of Cu^{2+}, and the Hill coefficient is almost doubled to n = 5.0. An additional transition of SmbP in presence of Cu^{2+} was observed around pH 5.5. It is interesting that acidic conditions induce the increase of α-helicity folding of SmbP in presence of Cu^{2+}, but such conditions induce the unfolding of apo-SmbP.

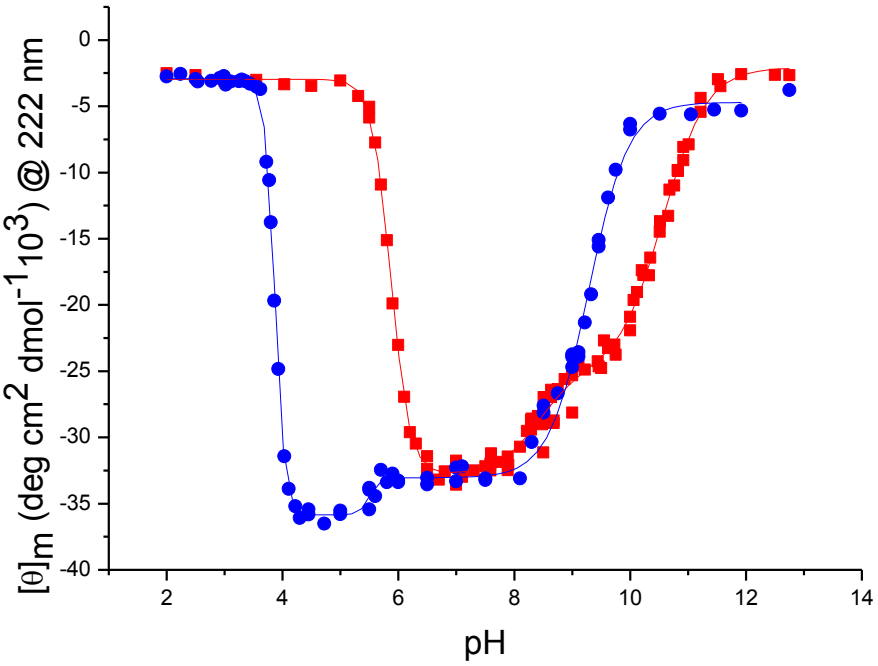

Figure 3.10: The entire pH profile of SmbP in the presence of no metal or 10 molar equivalent of Cu^{2+}

Table 3.5: Data fitting results of the entire pH profile of SmbP in the presence of no metal or 10 molar equivalent of Cu^{2+}

Parameters	SmbP	SmbP +10x Cu^{2+}
A	-2.96	-2.92
B	-32.82	-35.86
a	**2.65**	**4.95**
pK_1	**5.87**	**3.86**
C	-24.79	-33.05
b	1.22	4.95
pK_2	8.42	5.50
D	-2.08	-4.70
c	1.16	1.27

Data fitting were based on the following modified Henderson-Hasselbach equations:

$A \rightleftharpoons B + a\ H^+$ $\quad K_1{}^a = [B][H^+]^a/[A]$ $\quad pK_1 = pH + (1/a)^* \log[A]/[B]$

$B \rightleftharpoons C + b\ H^+$ $\quad K_2{}^b = [C][H^+]^b/[B]$ $\quad pK_2 = pH + (1/b)^* \log[B]/[C]$

$C \rightleftharpoons D + c\ H^+$ $\quad K_3{}^c = [D][H^+]^c/[C]$ $\quad pK_3 = pH + (1/c)^* \log[C]/[D]$

Cu^{2+} Induces the Refolding of SmbP in Acidic Solution

As shown in the above pH profile (Figure 3.10), SmbP is unfolded in the absence of Cu^{2+} at pH 5.0, but it is completely folded in the presence of Cu^{2+}. This observation led us to test if the addition of Cu^{2+} can induce the refolding of denatured apo-SmbP in acidic solutions. To do this, the binding of each metal ion to SmbP was analyzed by titration of unfolded apo-SmbP at pH 5.0 with Cu^{2+} or Ni^{2+}. The α-helicity was increased by the addition of Cu^{2+}, with dramatic spectral changes indicating a random coil to α-helix transition (Figure 3.11A). The CD spectra had an isodichroic point at ~204 nm, which indicates a two-state refolding transition model from a random coil to an α-helical structure corresponding to Cu^{2+} binding to the unfolded protein [40]. No significant change in the CD spectra was observed after the addition of six equivalents of Cu^{2+}, which indicates full metal occupancy and complete refolding of SmbP has been achieved (Figure 3.11B). The same effects were not observed with the addition of Ni^{2+} (data not shown).

A very similar refolding process induced by Cu^{2+} was observed at pH 4.5, but the refolding rate was much slower, as a much longer equilibration time (hours at pH 4.5 comparing to minutes at pH 5.0) is required to reach the final results.

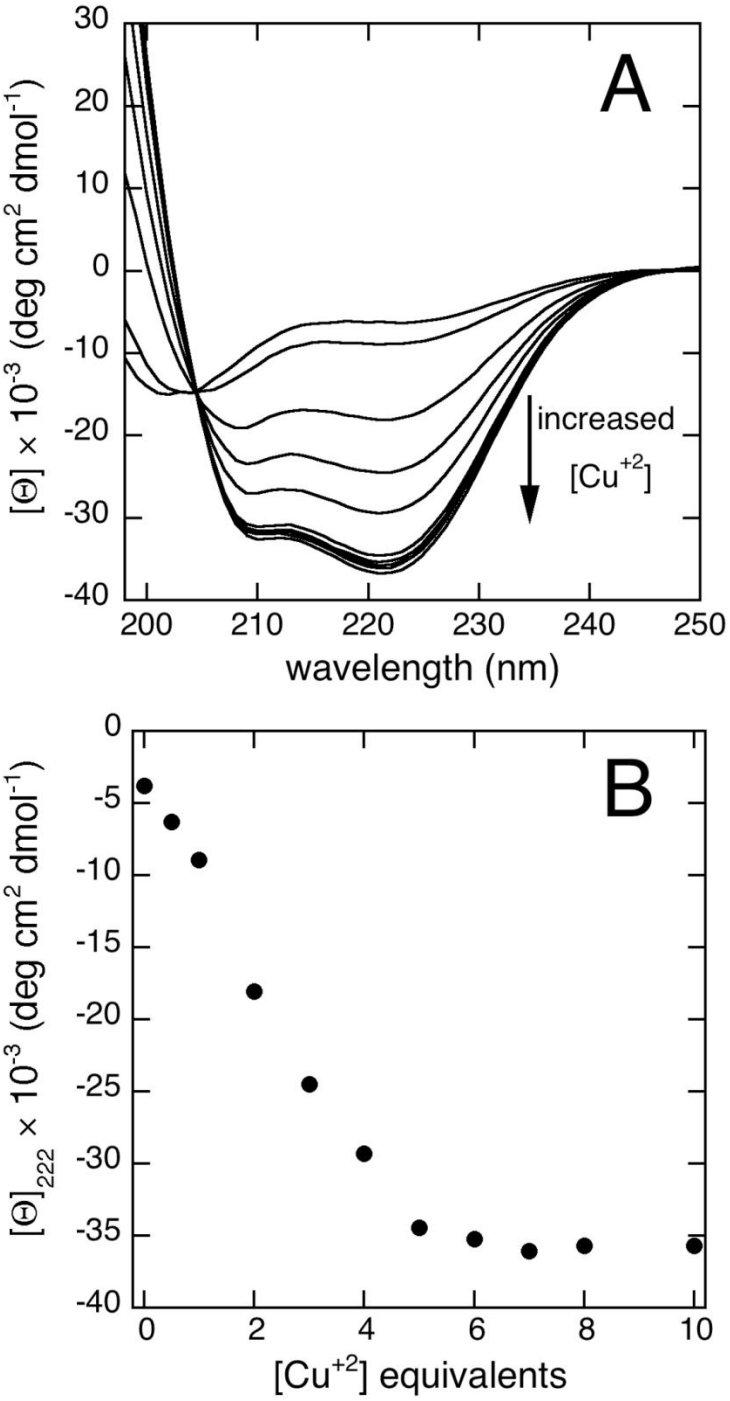

Figure 3.11: Cu^{2+} Assisted Folding of SmbP at pH 5.0

Cu^{2+} Stabilized SmbP against Chemical Denaturation

Chemical denaturants (e.g. urea and guanidinium chloride) are frequently used to disrupt nature structure in studies of protein folding. The unfolding of SmbP induced by urea in the absence of Cu^{2+} at pH 8.0 was followed by measuring the change in ellipticity at 222 nm as a function of urea concentration (Figure 3.12). The presence of 10 molar equivalents of Cu^{2+} greatly stabilized SmbP against chemical denaturation by increasing the free energy change of unfolding ΔG^{H_2O} four-fold, from 3.1 kJ/mol in the absence of Cu^{2+} to 11.1 kJ/mol in the presence of 10 molar equivalents of Cu^{2+} (Table 3.6). However, the denaturant value *m* was not affected by the presence of Cu^{2+} (3.0 kJ/mol/M in the absence of Cu^{2+} compared to 3.1 kJ/mol/M in the presence of Cu^{2+}). Since the *m* value correlates strongly with the amount of protein surface exposed to solvent upon unfolding [39], the comparable *m* values indicate Cu^{2+} greatly increases the free energy change associated with unfolding of SmbP without significantly changing the overall structure of the protein.

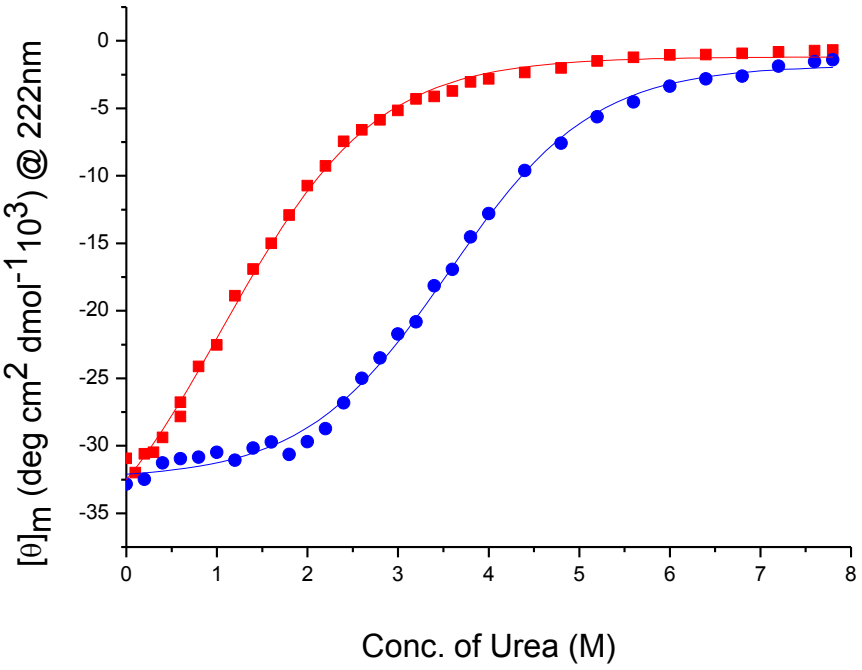

Figure 3.12: Denaturation of SmbP by urea at pH 8.0 in the absence of Cu^{2+} (red) and in the presence of 10 molar equivalent of Cu^{2+} (blue).

Table 3.6: Data fitting results of chemical denaturation of SmbP by urea

	apo-SmbP	SmbP + 10x Cu^{2+}
ΔG^{H_2O} (kJ/mol)	3.14	11.06
m (kJ/mol/M)	2.95	3.11

At pH 7.0, apo-SmbP has a similar unfolding behavior induced by urea, as described above at pH 8.0. However, Cu^{2+}-SmbP remains fully folded even in 8 M urea after more than 50 hours incubation (Figure 3. 13).

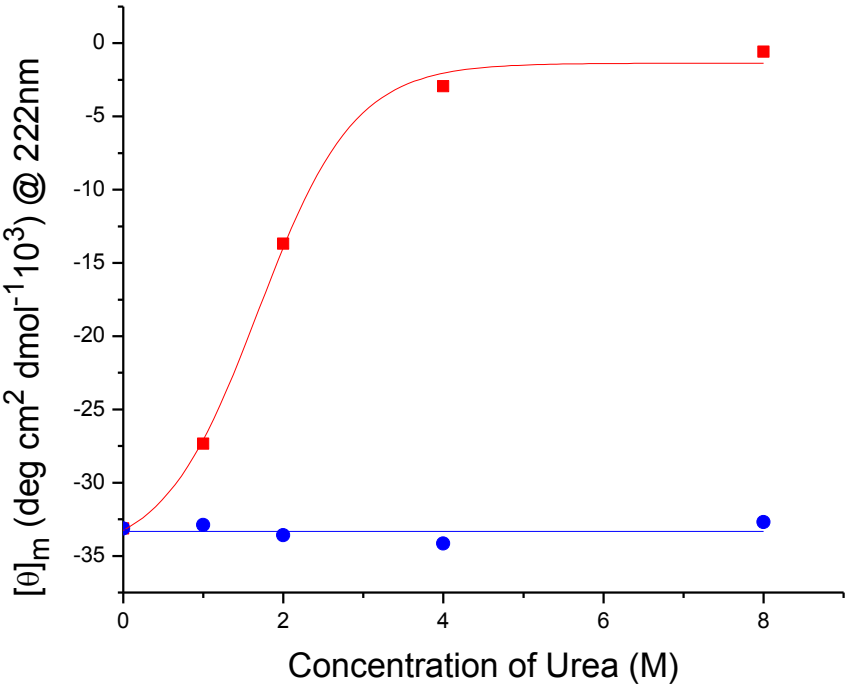

Figure 3. 13: Denaturation of SmbP by urea at pH 7.0 in the absence of Cu^{2+} (red) and in the presence of 10 molar equivalent of Cu^{2+} (blue).

A similar stabilization effect of Cu^{2+} on stability of SmbP was observed against chemical denaturation induced by guanidine chloride at pH 8.0 (Figure 3. 14) and pH 7.0 (data not shown).

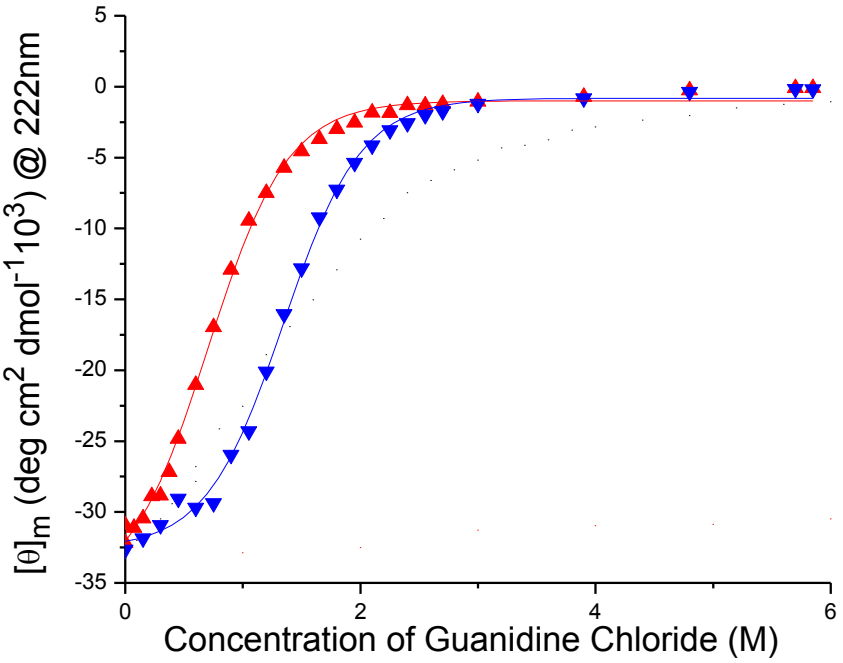

Figure 3. 14: Denaturation of SmbP by guanidine chloride at pH 8.0 in the absence of Cu^{2+} (red) and in the presence of 10 molar equivalent of Cu^{2+} (blue).

Discussion

SmbP was first isolated and characterized as a metal binding protein from the periplasm of the ammonia-oxidizing bacterium *N. europaea* [5]. The mature protein is only 93 residues long, but it can bind up to six equivalents of Cu^{2+}. This high affinity for Cu^{2+} is likely to be a consequence of the unusually high percentage of histidine residues (17%, versus an average of 2% from 1021 unrelated proteins [42]). The protein has a unique sequence with no homology to any known protein. It contains ten repeats of a seven amino acid motif (*abcdefg)*, which forms an α-helical structure, and histidine is the absolutedly conserved residue at position *c*.

The far-UV CD spectrum of apo-SmbP at room temperature is consistent with a secondary structure of high α-helical content (~70%). The protein experiences a reversible and cooperative α-helix to random-coil transition as the temperature is increased, with a melting temperature of 48.5 °C at pH 6.4. A similar conformational change can also be induced by decreasing the pH, with an apparent pK_a of 5.9. Recovery of the native structure by cooling or increasing the pH suggests that this process is a two-step process. Chemical denaturants also induce the unfolding of SmbP, and analysis of these processes in terms of a two-state transition has allowed expression of the stability of SmbP in terms of Gibbs free energy changes. A ΔG^{H_2O} value of 3.14 kJ/mol for unfolding process shows that the folded structure of SmbP is moderately stable.

Cu^{2+} has been shown to play a substantial role in the stabilization of SmbP protein structure against various suboptimal environments including acidic pH, high temperature, and the presence of chemical denaturants. In fact, Cu^{2+}

stabilizes the protein structure in boiling water at pH 6.4 or 7.0. Cu^{2+} also extends the pH range of folded SmbP by 2.0 pH units in acidic solution. And Cu^{2+} induces the refolding of SmbP protein at pH 5.0. Cu^{2+} also increases the Gibbs free energy change of unfolding process by 7.9 kJ/mol from data obtained at pH 8.0.

In acidic solutions, the protonation of side-chains of histidine residues is most likely responsible for the denaturation of SmbP. Because apo-SmbP is disordered around pH 5.9, which is very close to the pKa of imidazole group of histidine exposed to water solvent at the surface of protein. The X-ray crystal structure of SmbP in the absence of Cu^{2+}, shows that SmbP has a four-α-helix bundle structure with histidines aligned on two opposite faces of the bundle. When these histidines are deprotonated at neutral pH, they form hydrogen bonds with each other and stabilize the protein structure. They are functioning like two hydrogen-bond-zippers locking the bundle structure from two opposite sides. In acidic solution, however, these histidine residues become protonated. This breaks the two zippers of hydrogen bonds, and the repulsive force between the positive charges of protonated histidines further dissociates the bundle structure.

At pH 5.0, Cu^{2+} binding to histidines and glutamate acids can re-bridge the gaps of hydrogen bonds breaking and form two new copper-bond-zippers, thus reordering the protein. This explains the observation that the addition of Cu^{2+} can induce the refolding of SmbP in the acidic solutions. The binding of Cu^{2+} to histidine out competes the binding of H^{+}. This shifts the pK_a of the histidines to 3.9, which is 2.0 pH units lower than the pK_a of the histidines in apo-SmbP. At pH3.9, protonation of glutamate results in the disruption of these copper bonds

and three salt bridges between glutamate and lysine or arginine on the other two faces of the four-helix bundle. This is likely another explanation why the Cu^{2+}-SmbP unfolds at this pH. Ni^{2+} has similar, but much weaker effects on the stability of SmbP structure than Cu^{2+}, possibly due to its lower binding affinity than Cu^{2+} to SmbP.

Thermal unfolding of SmbP appears to be a single cooperative transition, which is a characteristic of small, single domain globular proteins. These proteins typically have unfolding enthalpy change (ΔH) on the order of 40 kcal/mol (about 167 kJ/mol) [43, 44]. The values of ΔH we observed for both apo-SmbP (143 kJ/mol) and Cu^{2+}-SmbP (187 kJ/mol) at pH 8.0 are closed to the average value above. The presence of Cu^{2+} increases the ΔH by 44 kJ/mol. It probably due to the formation of coordination bonds between Cu^{2+} and protein ligands (such as histidines) and some other consequent structural changes (such as conformation and solvation changes). On the other hand, the binding of Cu^{2+} decreases the degree of freedom of ligand residues, which results in a larger entropy change (ΔS) upon unfolding. The 117 J/mol/K increase of ΔS is unfavorable for the stabilization of SmbP by Cu^{2+}, but it is counteracted by the large favorable increase of ΔH of 44 kJ/mol. Overall, unfolding Gibbs free energy change (ΔG) is the combination of ΔH, ΔS and temperature (T). Assuming Cu^{2+}-SmbP has similar ΔH and ΔS dependences on temperatures to apo-SmbP, the increase of ΔG of Cu^{2+}-SmbP than apo-SmbP is calculated to be about 9 kJ/mol at room temperature and pH 8.0. This increase of ΔG significantly increases the thermal stability of SmbP in the presence of Cu^{2+}. At pH 6.4 and pH 7.0, it is reasonable

to believe that the larger increases of ΔG are responsible for the extreme stabilities of Cu^{2+}-SmbP at high temperatures.

At pH 8.0, the temperature dependences of the ellipticities, the baseline slopes determined from both the pre-melting phases of apo-SmbP and Cu^{2+}-SmbP, correspond closely to the temperature dependence for ellipticity of an α-helical structure [41]. It indicates that both of apo-SmbP and Cu^{2+}-SmbP at low temperatures have a folded α-helical structure. The baselines in post-melting phase are shifted to a more negative ellipticity with increasing equivalents of Cu^{2+} at pH 8.0. This is an indication of a higher percentage of α-helical structure in the presence of higher equivalents of Cu^{2+} at high temperatures. In other words, Cu^{2+} helps SmbP maintain partial α-helical structure, resulting in incomplete unfolding of SmbP, even at the high temperatures and unfavorable pH condition. At a more favorable pH (6.4 or 7.0), the elliticities of SmbP in the presence of Cu^{2+} show no melting transition but match closely with the baselines of apo-SmbP in pre-melting phases, respectively. It indicates that the observed slope is due to the intrinsic temperature dependence of $[\theta]_{222}$ for an α-helix rather than to the unfolding of the protein. It is very rare for a protein to have a fully folded structure at such high temperatures, but it appears that Cu^{2+} allows SmbP to remain fully folded at these temperatures close to the boiling point of water.

These observations also expose another interesting question: while Cu^{2+} increases the melting temperature (T_m) of SmbP by more than 50 °C at pH 6.4 or 7.0, why does it increase T_m by only 13.2 °C at pH 8.0? A similar question rises again in the case of unfolding induced by chemical denaturants: the Cu^{2+}-SmbP is

still fully folded in 8M urea at pH 7.0, however, why is it completely unfolded in about 6M urea at pH 8.0? In other words, what factor(s) decrease the stabilization effects of Cu^{2+} on SmbP at pH 8.0 compared to pH 6.4 or 7.0? One possible factor is the deprotonation of the side chains of lysines, upon Cu^{2+} binding at pH 8.0. The protonated lysines are essential to form the inter-strand salt bridges between them and glutamates, which stabilize the structure of SmbP. Another possible factor is that the deprotonated lysine residues compete with Cu^{2+} ligation sites, which has been observed in other proteins [45]. But these perturbations do not happen at lower pH because higher concentrations of H^{+} compete with the binding of Cu^{2+} to lysine, making protonated lysine the predominate state. This argument is supported by the increasing equivalent of bound Cu^{2+} at high pH (seven at pH 8.0 versus six at pH 5.0 to pH 7.0). It is reasonable to extrapolate that the stabilization effects of Cu^{2+} on the structure of SmbP will decrease further at higher pH than 8. Indeed, a destabilization effect of Cu^{2+} has been observed in pH 9-11(Figure 3.10).

A similar stabilization effect of Cu^{2+} on the structure of SmbP has been observed in unfolding induced by chemical denaturants. At pH 8.0, binding of Cu^{2+} results in a more negative Gibbs free energy (-8 kJ/mol more in presence of 10 equivalents of Cu^{2+}, which is close to the value of -9 kJ/mol evaluated from thermal denaturation). It gives Cu^{2+}-SmbP more resistance to unfolding induced by chemical denaturation, and the addition of Cu^{2+} results in an increase of the concentration of urea, at which half of SmbP is unfolded, from about 1.5 M to about 4 M. However, the presence of Cu^{2+} does not affect the denaturant value *m*,

which is an indication of the amount of protein surface exposed to solvent upon unfolding [39]. The similar *m* values imply that the hydrophobic core of SmbP is not affected by the presence of Cu^{2+}, and Cu^{2+} is binding to residues on the surface of the protein, such as histidines and lysines. At pH 7.0, no unfolding event was observed in the presence 10 equivalents of Cu^{2+}, even in 8 M urea. The difference in the stabilization effects of Cu^{2+} at different pH is possibly a result of Cu^{2+} ion's binding affinity to side chains of lysine residues at high pH as discussed above.

In conclusion, SmbP, a small protein with only 93 residues, can bind up to about six equivalents of Cu^{2+}, and the binding of Cu^{2+} can significantly increase the stability of α-helical structure of SmbP against pH change, high temperature, and chemical denaturants. These exceptional properties make its structure information, specifically the binding sites of Cu^{2+}, critical to elucidating its function in the host organism, to revealing the roles of Cu^{2+} in protein structure and stability, and for designing *de novo* proteins with metal binding capabilities.

Chapter 4

NMR STUDY TO IDENTIFY COPPER BINDING SITE OF SMBP

Summary

The crystal structure of SmbP without Cu^{2+} has been determined. However, attempts to crystallize of SmbP with Cu^{2+} bound have not been successful. In this chapter, multidimensional NMR experiments were conducted in order to gain additional information regarding the Cu^{2+}-SmbP structure, in particular its metal binding sites. Unambiguous resonance assignments were successfully made. Cα secondary chemical shifts confirmed that SmbP has a four α-helical structure. Cu^{2+}-protein titration experiments monitored by NMR have indicated a top-to-bottom, sequential metal binding pattern for SmbP. NMR titration results combined with results presented in Chapters 2 and 3, confirm that SmbP contains six copper binding sites and the identity of the Cu^{2+} ligands is proposed.

Introduction

Since the ground breaking report on NMR spectroscopy studies on ribonuclease in 1957 [46], nuclear magnetic resonance (NMR) spectroscopy has been widely used for structural studies of biological macromolecules. NMR and one other technique, X-ray crystallography, are currently the only techniques able to determine the three dimensional structures of macromolecules at atomic resolution. In addition, NMR spectroscopy is a powerful tool for the study of protein conformational changes, denaturation and internal mobility, pH titration

of individual ionizable amino acid side chains in enzyme active sites, and for the investigation of paramagnetic centers in metalloproteins [47, 48].

The crystal structure of SmbP without Cu^{2+} has been determined, however, attempts to crystallize the Cu^{2+}- SmbP have not been successful. In order to obtain information regarding the Cu^{2+} binding sites, Cu^{2+}-protein titration experiments monitored by NMR were conducted. Paramagnetic Cu^{2+} ion causes acceleration of T_2 relaxation of nuclei, and therefore, broadens NMR signals or reduces resonances intensity of residues in the vicinity of Cu^{2+} binding sites in the protein. Monitoring the NMR signal intensities during Cu^{2+}-protein titration is a site specific probe to identify the residues close to the Cu^{2+} ion, which has been used to provide information on Cu^{2+} binding sites in several Cu^{2+}-protein complexes [49-51].

Materials and Methods

Isotopic Labeling of Recombinant SmbP Protein

In order to generate uniformly ^{13}C, ^{15}N-double labeled SmbP, *Escherichia coli* BL21(DE3) cells harboring the plasmid pSMBP were grown on M9 minimal medium prepared with $^{15}NH_4Cl$ (1 g/L) and ^{13}C glucose (4 g/L). The , the ^{15}N-labeled SmbP was generated by growing the same *E. coli* cells in M9 minimal medium containing $^{15}NH_4Cl$ (1 g/L) [52]. The purification was following the same procedures as presented in Chapter 3.

NMR Spectroscopy

NMR experiments were conducted on samples of 0.9 mM ^{13}C, ^{15}N-double labeled SmbP in 90% H_2O, 10% D_2O, 1 mM NaPi, pH 7.0. 0.1% NaN_3 was included to inhibit bacterial growth. A series of heteronuclear 2D and 3D NMR spectra (gNhsqc_trosy, CBCANH, CBCAcoNH, HNCA, HNcoCA, HNCO, and HNcaCO) were recorded at 295K on a Bruker 800 MHz NMR spectrometer. All spectra were processed using NMRPipe and analyzed using the program Sparky.

Secondary chemical shift values, which are the differences between the measured Cα chemical shifts and the empirical random coil value of the corresponding amino acid at pH 2.3 and 293K in 8M urea, were plotted as a function of the protein sequence.

Cu^{2+}-Protein Titration

Cu^{2+}-protein titrations were performed on ^{15}N-labeled protein samples containing 0.2 mM SmbP in 90% H_2O, 10% D_2O, 20 mM MOPS buffer, pH 7.0. Aliquots of a copper sulfate ($CuSO_4$) solution in the same buffer were added incrementally to increase the Cu^{2+} concentration. 2D ^{1}H-^{15}N HSQC spectra were recorded for each step of the titration at 295K, with ratios (Cu^{2+}/SmbP) of 0.005, 0.01, 0.02, 0.04, 0.1, 0.2, 0.4, 1, 1.5, 2, 4 and 6.

To identify the residues around the binding sites of Cu^{2+}, I/I_0 ratio was plotted as a function of the protein sequence, where I_0 is the intensity of amide backbone cross peak in the absence of Cu^{2+}, and I is the peak intensity of the

same residue after Cu^{2+} addition, corrected for the corresponding volume change after the addition of $CuSO_4$.

Results

Assignment of NMR Spectra of SmbP

2D NMR spectra of ^{1}H-^{15}N heteronuclear single quantum correlation (HSQCs) confirmed that SmbP has a well-folded structure. SmbP's moderate wide dispersion of resonances in the proton dimension (around 3 ppm) (Figure 4.1), is significantly different from that of a disordered protein, which is only 1 ppm dispersion.

The structure of SmbP, dominated by α-helices and lacking β-sheets, narrowed the dispersion of resonances, which made the backbone assignment challenging. Furthermore, the fact that 67% of the protein is composed of only five amino acids: histidine (17%), alanine (16%), glutamate (14%), glycine (11%) and lysine (9%), results in further resonance overlap in the HSQC spectra. However, unambiguous assignment was completed using CBCANH, CBCAcoNH, HNCA and HNcoCA spectra. The resonances of only 8 residues were missing in the NMR spectra.

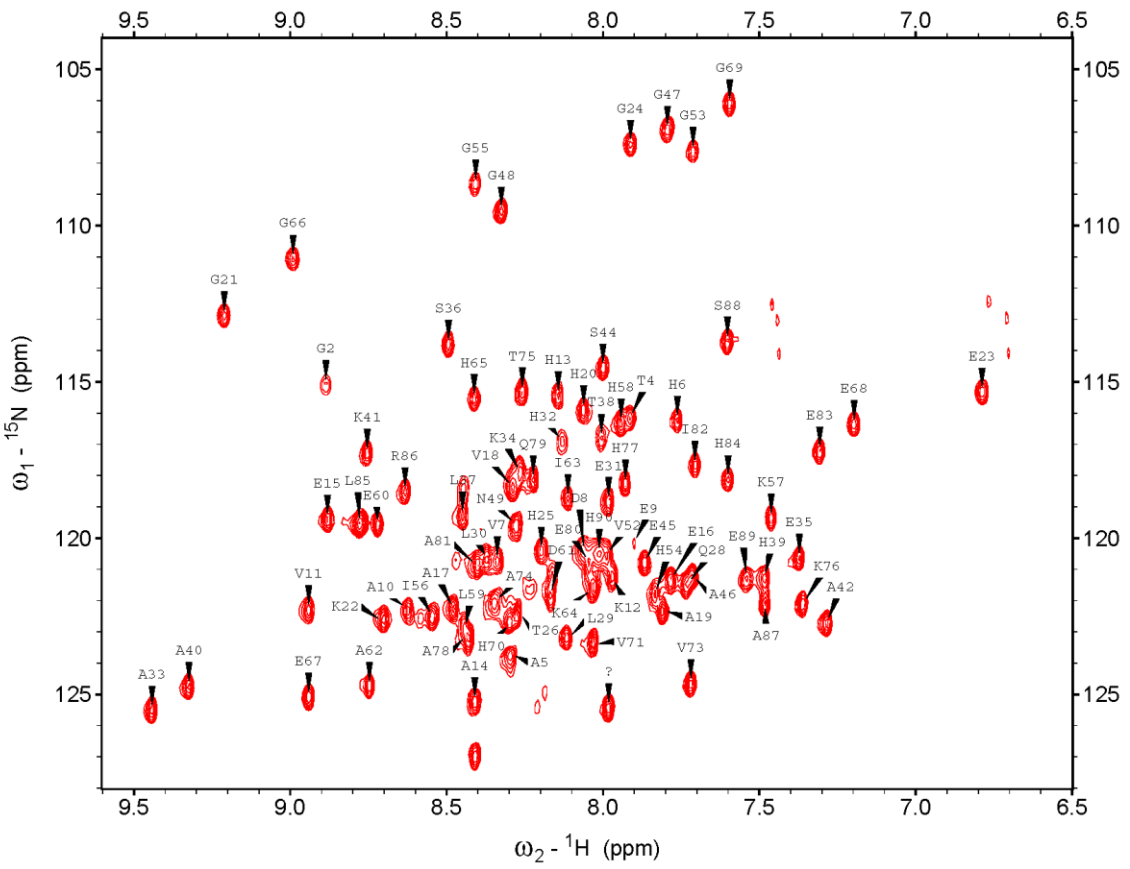

Figure 4.1: HSQC assignments of SmbP.

NMR Spectra Demonstrate that SmbP Has a Well-folded, Four α-Helical Structure

NMR chemical shifts of Cα atoms, which are sensitive to the secondary structure of proteins, were used to investigate the secondary structure of SmbP. T4-K22, T26-E45, H51-E67 and V71-S88 were identified as α-helixes (Figure 4.2). They are essentially identical to the structure obtained from X-ray crystallization. E23-H25, A46-T50, E68-H70 and E89-H93 did not show any structural propensity. These correspond to the loops or turns connecting two α-helices or the tail located at the C-terminus. E45-N49, which were absent in the X-ray crystal structure, had strong resonances in the NMR spectrum and are found in the loop connecting α-helices 2 and 3. E45 was identified as the last residue of α-helix 2.

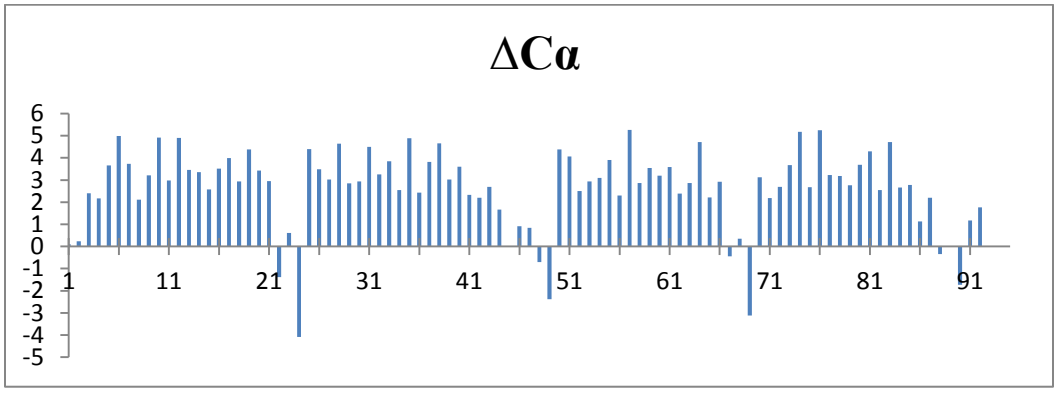

Figure 4.2: Cα secondary chemical shifts observed in SmbP as a function of residue number. Positive secondary chemical shifts spanning four regions of SmbP sequence indicate four α-helixes.

Mapping Cu^{2+} Binding Sites by NMR Titration

The sequence-specific assignment of apo-SmbP resonances was first determined in sodium phosphate buffer at pH 7.0. Because phosphate buffer is not compatible with Cu^{2+}, a Cu^{2+} compatible buffer, MOPS, was used in the Cu^{2+} titration experiments. Fortunately, there are no observed differences between the spectra of apo-SmbP in phosphate and MOPS buffer.

By titrating incrementally increasing concentrations of $CuSO_4$ solutions into a solution of SmbP protein sample, the HSQC spectra changes resulting from the binding of Cu^{2+} to SmbP were monitored. Analysis of the obtained information revealed the residues involved in Cu^{2+}-binding as well as residues located close to the Cu^{2+}-binding sites at residue resolution. The addition of Cu^{2+} induced very small chemical shift changes for only a few residues, which indicating that Cu^{2+} binding does not dramatically change the overall structure of SmbP. This is consistent with the observations from the CD data at pH7.0. Reduction in the signal intensities were observed while increasing the concentration of Cu^{2+}. At low concentrations of Cu^{2+} (Cu^{2+}:SmbP equal to 1.0), Cu^{2+} availability is limited. Most of the Cu^{2+} binds to the first two binding sites, whose binding affinities are much higher than the remaining four. Affected by the paramagnetic Cu^{2+} cation, the signal intensities of residues liganding directly to Cu^{2+} or in the inner coordination shell reduce dramatically. The signal intensities of residues farther from Cu^{2+} binding sites are reduced to a less extent or are not affected at all (Figure 4.3).

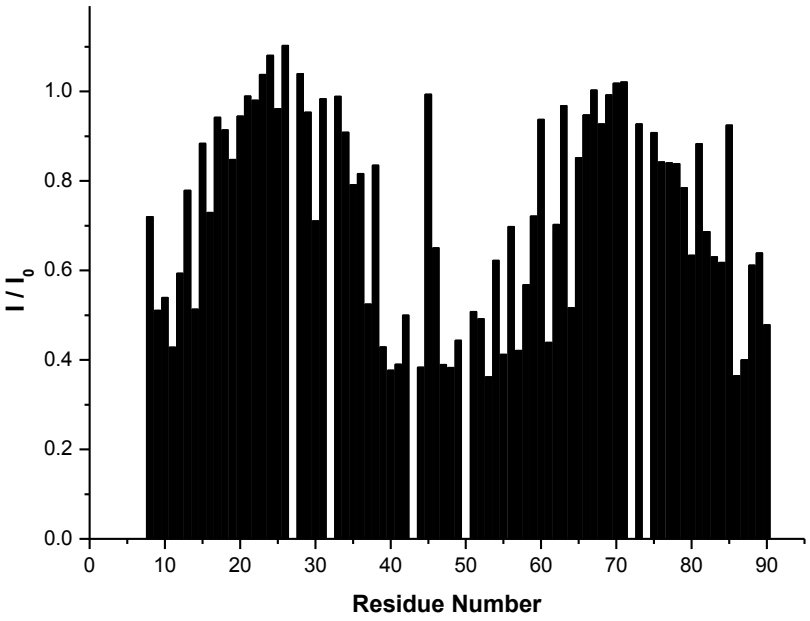

Figure 4.3: Mapping SmbP:Cu^{2+} Binding Sites at single residue resolution by NMR titration.

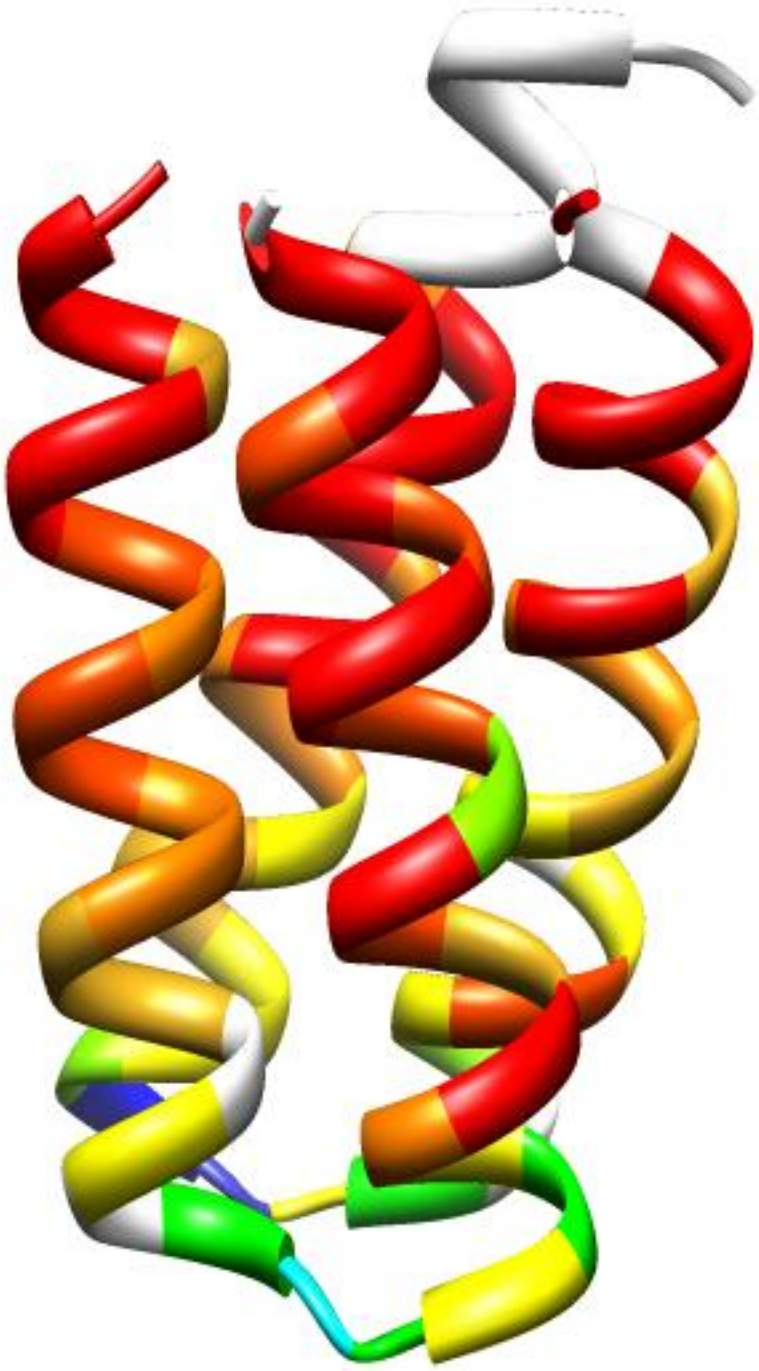

Figure 4.4: Color coded SmbP residues, from strongest (red) to weakest (blue) affected by addtion of Cu^{2+}.

The NMR signal intensities of four peptides were affected dramatically by the addition of one equivalent of Cu^{2+}: ^{9}***E***AVK[H]A^{14} from α-helix 1, ^{37}L[T]***H***AKA(A)S^{44} from α-helix 2, ^{47}GGN(T)***H***VG***H***G[I]K***H***^{58} from α-helix 3 and the loop before it, and ^{80}E[A]IE***H***[L]RASEH ^{90} from α-helix 4 and the C-terminal tail after it (Figure 4.4). In this notation, the residues in red were proposed as Cu^{2+}-binding ligands. The residues in parenthesis had missing resonances in the NMR spectra, and the residues in brackets showed only slight decrease of the signal intensities in NMR spectra upon addition of one equivalent of Cu^{2+}. Two additional residues, D61 and K64, which are from α-helix 3 and are closed to the third peptide mentioned above, were affected strongly by the addition of one equivalent of Cu^{2+}.

The greatest contrast of residues surround the first two binding sites compared to the rest of the residues was observed after the addition of one equivalent of Cu^{2+}, as shown in Figure 4.3. Upon further additions of Cu^{2+}, the Cu^{2+} ions binds not only to the first two high affinity binding sites, but also to the remaining four lower affinity binding sites. It results in further decreases of the NMR signal of residues surround the first two binding sites discussed above, but also dramatically decreases the intensity of the signals of other residues spanning the middle and part of the bottom end of the four α-helixes bundle. This indicates with high probability that the third and fourth binding sites are located in the middle of the bundle. After adding six equivalents of Cu^{2+}, signals of all residues, including a few strongest resistant residues located at the bottom end of the four

α-helical bundle, retain barely more than 20% of original intensities. This indicates that the protein has been fully loaded with Cu^{2+}.

After fully loading of Cu^{2+}, 80 molar equivalents of EDTA in term of the concentration of SmbP , which is 12 times excess in term of the concentration of Cu^{2+}, was added to test the reversibility of the NMR spectral changes induced by Cu^{2+} binding. The intensities of most peaks were recovered to about 80% of their original intensities (Figure 4.5), This indicates that the decrease in signal intensities was due to the binding of Cu^{2+}, as excess EDTA can strip Cu^{2+} from SmbP to reverse its effects on the NMR spectra.

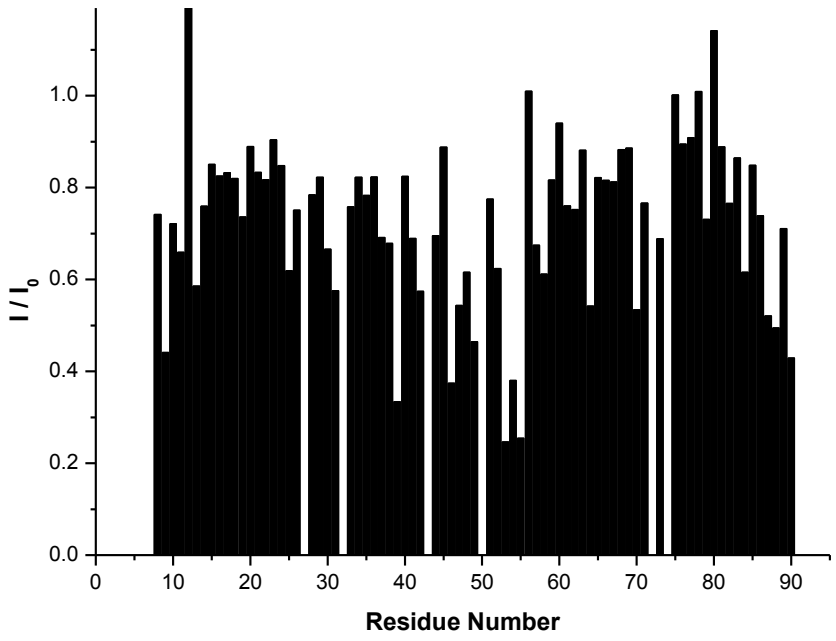

Figure 4.5: The intensities of most peaks are recovered by addition of excess EDTA to strip Cu^{2+} from SmbP

Discussion

By locating the four peptides, that affected dramatically by the addition of one equivalent of Cu^{2+}, on the X-ray crystal structure of SmbP, it was found that all four peptides are located at the top end of the four α-helical bundle, where the N-terminus, the C-terminus and the long loop between helices 2 and 3 are located. The influence of lower concentrations of Cu^{2+} (as low as 0.2 equivalents) on these residues were observed, and this influence became more obvious as the concentration of Cu^{2+} was increased. The remaining residues, which are more resistant to one equivalent of Cu^{2+}, are mostly located in the middle and at the bottom end of the four α-helical bundle, where the short loops between helices 1 and 2 and between helices 3 and 4 are located.

By investigating the ligand properties of the residues most affected, two binding sites of Cu^{2+} have been identified (Figure 4.6). The first binding site is located at the interface between α-helix 3 and α-helix 4, with four possible ligand residues: H51, H54, H58 and H84. This is supported by the simulation result of EPR spectrum of SmbP with one equivalent of Cu^{2+}, which has indicated the first binding site has a square-planar geometry with four nitrogen ligands. Another two nearby residues, K57 and E80, may be playing indirect roles in Cu^{2+} binding by providing the hydrogen bonds network or balancing the change in charge upon Cu^{2+} binding in the second binding site.

The second binding site is formed by residues E9 and H39 from α-helix 1 and α-helix 2, respectively. The position of residue H13 makes it a good candidate as one of the ligands in the second Cu^{2+} binding site, but its NMR signal intensity was only moderately affected by the addition of one equivalent of Cu^{2+}. Another

possible ligand is H6, which is located at a suitable position for Cu^{2+} binding, but because its resonance was missing in the NMR spectra, it was not identified as one of the highly influenced residues. The two binding sites look very similar to each other, except that H54 in the first binding site is replaced by E9 in the second site. Corresponding to K57 and E80 in the first binding site are K12 and E35, which may be involved indirectly by providing the hydrogen bonds network and balancing the change in charge upon Cu^{2+} binding.

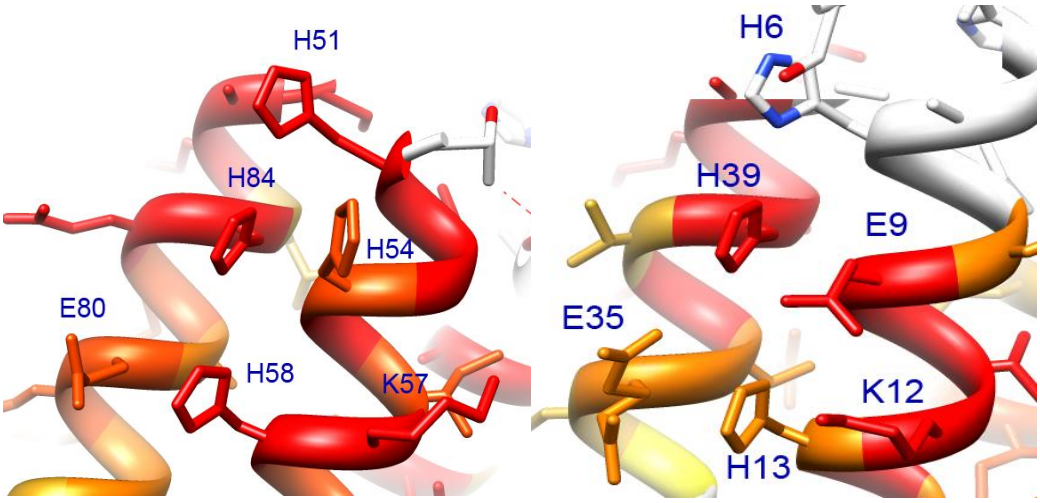

Figure 4.6: First two binding sites of Cu^{2+} indicated by NMR tritration experiment

After the addition of excess EDTA to remove Cu^{2+} from the Cu^{2+} loaded SmbP sample, the intensities of most peaks were recovered to about 80% of their original intensity (Figure 4.5). However, for a few residues close to the first two binding sites, the NMR signals were recovered to less than 60% (H13, E31, A42, G47, N49, I63, H70, R86, and A87) or even less than 45% (E9, H39, A46, G53,

H54, G55, and H90) of the original intensity. Their close positions further support that the correct binding sites have been identified.

When the sequence of SmbP from *N. europaea* and the homologs proteins from its relatives are compared (Figure 2.3), all residues identified as ligands in SmbP are well-conserved. Residues H6, H13, H39, H51, H84 are absolutely conserved (100%), while the other residues E9 (67%), H54 (50%), H58 (50%), are all highly conserved. Four other residues mentioned above, K12 (50%), E35 (67%), K57 (67%), E80 (83%), are also highly conserved. This indicates that the two identified binding sites are most likely the authentic sites of Cu^{2+} binding *in vivo*.

Based on the above information from the Cu^{2+}-protein titration, the sequence alignment of SmbP and its homologs, the predicted metal binding residues by the CHED server, and the distances of the participant residues, six Cu^{2+} binding sites are proposed (Figure 4.7). The first two sites (at the top) are the highest affinity sites, while the other four sites located in the middle and at the bottom of the four helix bundle are the lower binding affinity sites.

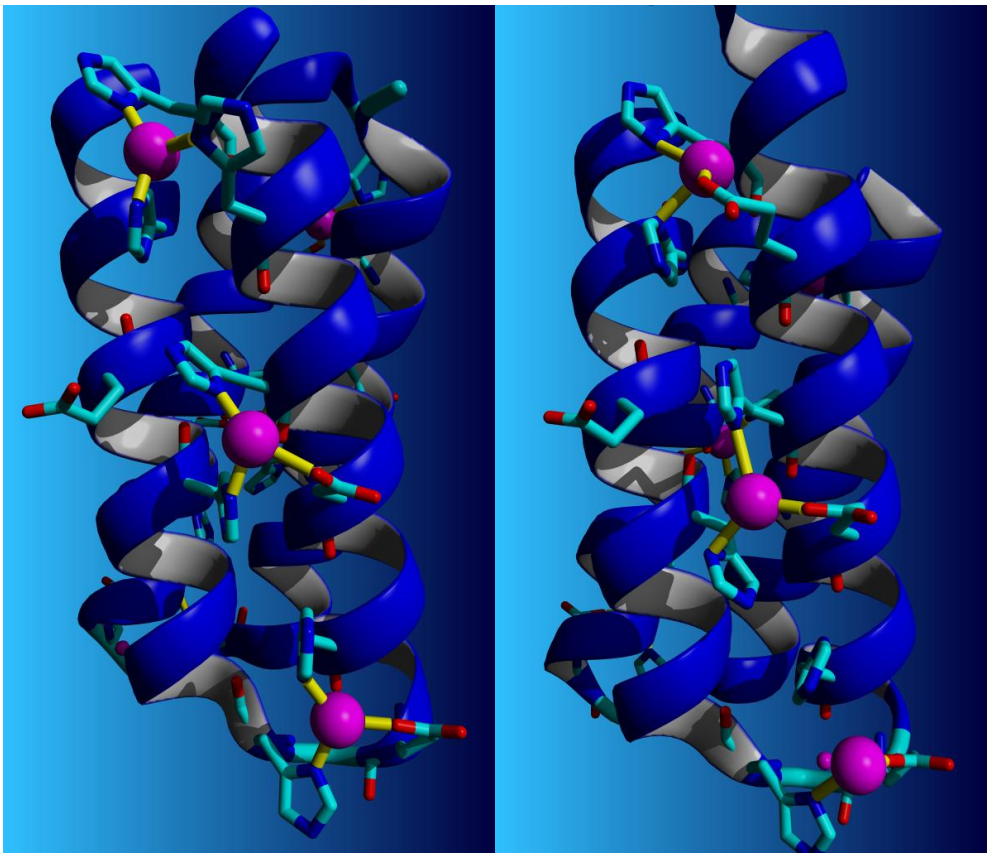

Figure 4.7: Proposed six Cu^{2+} binding sites. Six Cu atoms were added artificially into the crystal structure of SmbP, and side chains were rotated to allow best binding conformation while the backbones were fixed, using the YASARA program [25] with the FoldX plugin [26].

REFERENCES

1. Outten, F.W., et al., *The independent cue and cus systems confer copper tolerance during aerobic and anaerobic growth in Escherichia coli.* Journal of Biological Chemistry, 2001. **276**(33): p. 30670-30677.

2. Huffman, D.L., et al., *Spectroscopy of Cu(II)-PcoC and the multicopper oxidase function of PcoA, two essential components of Escherichia coli pco copper resistance operon.* Biochemistry, 2002. **41**(31): p. 10046-10055.

3. Arp, D.J., L.A. Sayavedra-Soto, and N.G. Hommes, *Molecular biology and biochemistry of ammonia oxidation by Nitrosomonas europaea.* Archives of Microbiology, 2002. **178**(4): p. 250-255.

4. Chain, P., et al., *Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph Nitrosomonas europaea.* Journal of Bacteriology, 2003. **185**(9): p. 2759-2773.

5. Barney, B.M., R. LoBrutto, and W.A. Francisco, *Characterization of a small metal binding protein from Nitrosomonas europaea.* Biochemistry, 2004. **43**(35): p. 11206-11213.

6. McCaldon, P. and P. Argos, *Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences.* Proteins: Structure, Function, and Bioinformatics, 1988. **4**(2): p. 99-122.

7. Nilsson, B.L., M.B. Soellner, and R.T. Raines, *CHEMICAL SYNTHESIS OF PROTEINS*, in *Annual Review of Biophysics and Biomolecular Structure*. 2005. p. 91-118.

8. Wallwork, S., *Hydrogen-bond radii*, in *Acta Crystallographica*. 1962. p. 758-759.

9. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D.* Nucl. Acids Res., 2010. **38**(suppl 2): p. W545-W549.

10. http://ekhidna.biocenter.helsinki.fi/dali_server

11. Lederer, F., et al., *Improvement of the 2.5 Å resolution model of cytochrome b562 by redetermining the primary structure and using molecular graphics.* Journal of Molecular Biology, 1981. **148**(4): p. 427-448.

12. Mathews, F.S., P.H. Bethge, and E.W. Czerwinski, *Structure of cytochrome b562 from Escherichia coli at 2.5 Å resolution.* Journal of Biological Chemistry, 1979. **254**(5): p. 1699-1706.

13. Calderone, V., et al., *Crystal structure of the PsbQ protein of photosystem II from higher plants.* EMBO Reports, 2003. **4**(9): p. 900-905.

14. Wuerges, J., et al., *Crystal structure of nickel-containing superoxide dismutase reveals another type of active site.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(23): p. 8569-8574.

15. Karp, P.D., et al., *Expansion of the BioCyc collection of pathway/genome databases to 160 genomes*. p. 6083-6089.

16. BioCyc Database: http://biocyc.org/

17. Schneider, A., C. Dessimoz, and G.H. Gonnet, *OMA Browser-Exploring orthologous relations across 352 complete genomes*, in *Bioinformatics*. 2007. p. 2180-2182.

18. OMA BrowserWebsite: http://omabrowser.org/

19. FINDSITE-metal web server: http://cssb.biology.gatech.edu/findsite-metal

20. Brylinski, M. and J. Skolnick, *FINDSITE-metal: integrating evolutionary information and machine learning for structure-based metal binding site prediction at the proteome level.* Proteins: Structure, Function, and Bioinformatics, 2010. **in press**.

21. CHED web server: http://ligin.weizmann.ac.il/~lpgerzon/mbs4/mbs.cgi

22. Babor, M., et al., *Prediction of transition metal-binding sites from apo protein structures.* Proteins: Structure, Function, and Bioinformatics, 2008. **70**(1): p. 208-217.

23. Hall, T.A., *BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.* Nucl. Acids. Symp. Ser., 1999. **41**: p. 95-98.

24. Pettersen, E.F., et al., *UCSF Chimera—A visualization system for exploratory research and analysis.* Journal of Computational Chemistry, 2004. **25**(13): p. 1605-1612.

25. Krieger, E., G. Koraimann, and G. Vriend, *Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force*

field. Proteins: Structure, Function, and Bioinformatics, 2002. **47**(3): p. 393-402.

26. Schymkowitz, J., et al., *The FoldX web server: an online force field*, in *Nucleic Acids Res.* 2005 p. W382-W388.

27. Lieberman, R.L. and A.C. Rosenzweig, *Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane.* Nature, 2005. **434**(7030): p. 177-182.

28. http://www.ap-lab.com/circular_dichroism.htm

29. Waddell, W.J., *A Simple Ultraviolet Spectrophotometric Method for the Determination of Protein.* Journal of Laboratory and Clinical Medicine, 1956. **48**(2): p. 311-314.

30. Sreerama, N. and R.W. Woody, *A Self-Consistent Method for the Analysis of Protein Secondary Structure from Circular Dichroism.* Analytical Biochemistry, 1993. **209**(1): p. 32-44.

31. Sreerama, N., S.Y.U. Venyaminov, and R.W. Woody, *Estimation of the number of α-helical and β-strand segments in proteins using circular dichroism spectroscopy.* Protein Science, 1999. **8**(2): p. 370-380.

32. The DichroWeb Server: http://dichroweb.cryst.bbk.ac.uk/html/home.shtml

33. Lobley, A., L. Whitmore, and B.A. Wallace, *DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra.* Bioinformatics, 2002. **18**(1): p. 211-212.

34. Whitmore, L. and B.A. Wallace, *DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data.* Nucleic Acids Research 2004. **32**(suppl 2): p. W668-W673.

35. Whitmore, L. and B.A. Wallace, *Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases.* Biopolymers, 2008. **89**(5): p. 392-400.

36. Rost, B. and C. Sander, *Combining evolutionary information and neural networks to predict protein secondary structure.* Proteins: Structure, Function, and Genetics, 1994. **19**(1): p. 55-72.

37. Rost, B. and C. Sander, *Prediction of Protein Secondary Structure at Better than 70% Accuracy.* Journal of Molecular Biology, 1993. **232**(2): p. 584-599.

38. Collaborative Computational Project, N., *The Ccp4 Suite - Programs for protein crystallography.* Acta Crystallogr. D, 1994. **50**: p. 760-763.

39. Myers, J.K., C. Nick Pace, and J. Martin Scholtz, *Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding.* Protein Science, 1995. **4**(10): p. 2138-2148.

40. Holtzer, M.E. and A. Holtzer, *α-helix to random coil transitions: Determination of peptide concentration from the CD at the isodichroic point.* Biopolymers, 1992. **32**(12): p. 1675-1677.

41. Rohl, C.A. and R.L. Baldwin, *Comparison of NH Exchange and Circular Dichroism as Techniques for Measuring the Parameters of the Helix-Coil Transition in Peptides.* Biochemistry, 1997. **36**(28): p. 8435-8442.

42. McCaldon, P. and P. Argos, *Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences.* Proteins: Structure, Function, and Genetics, 1988. **4**(2): p. 99-122.

43. DeGrado, W.F., D.P. Raleigh, and T. Handel, *De novo protein design: what are we learning?* Current Opinion in Structural Biology, 1991. **1**(6): p. 984-993.

44. Betz, S.F., D.P. Raleigh, and W.F. DeGrado, *De novo protein design: from molten globules to native-like states: Current opinion in structural biology 1993, 3:601-610.* Current Opinion in Structural Biology, 1993. **3**(4): p. 601-610.

45. Weinkam, P., et al., *The Folding Energy Landscape and Free Energy Excitations of Cytochrome c.* Accounts of Chemical Research, 2010. **ASAP**(ASAP): p. ASAP.

46. Saunders, M., A. Wishnia, and J.G. Kirkwood, *THE NUCLEAR MAGNETIC RESONANCE SPECTRUM OF RIBONUCLEASE1.* Journal of the American Chemical Society, 1957. **79**(12): p. 3289-3290.

47. Wuthrich, K., *Protein structure determination in solution by NMR spectroscopy*. 1990. p. 22059-22062.

48. John Cavanagh, W.J.F., Arthur G. Palmer III, Mark Rance and Nicholas J. Skelton, *Protein NMR Spectroscopy: Principles and Practice* 2nd ed. 2007 Elsevier Inc.

49. Rajalingam, D., et al., *The C2A domain of synaptotagmin exhibits a high binding affinity for copper: Implications in the formation of the multiprotein FGF release complex.* Biochemistry, 2005. **44**(44): p. 14431-14442.

50. Chong, L.X., et al., *Unprecedented Binding Cooperativity between CuI and CuII in the Copper Resistance Protein CopK from Cupriavidus metallidurans CH34: Implications from Structural Studies by NMR Spectroscopy and X-Ray Crystallography.* Journal of the American Chemical Society, 2009. **131**(10): p. 3549-3564.

51. Zambelli, B., et al., *Structural Characterization of Binding of Cu(II) to Tau Protein.* Biochemistry, 2008. **47**(41): p. 10841-10851.

52. Marley, J., M. Lu, and C. Bracken, *A method for efficient isotopic labeling of recombinant proteins.* Journal of Biomolecular NMR, 2001. **20**(1): p. 71-75.