NORTHWESTERN UNIVERSITY

Structures and Functions of New Copper Receptor Proteins Controlling

Gene Regulation and Metal Trafficking

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Chemistry

By

Yi Xue

EVANSTON, ILLINOIS

June 2008

ABSTRACT

Structures and functions of new copper receptor proteins controlling gene regulation and metal trafficking

Yi Xue

Copper serves as a cofactor for many proteins and enzymes involved in important biological processes, but at the same time, excessive copper is toxic to cell. Thus, the intracellular copper concentration must be tightly controlled such that copper is provided to essential enzymes but does not accumulate to toxic levels. As a result, organisms have evolved various mechanisms to obtain and distribute copper safely.

The thesis represents a structural-based investigation into copper receptor proteins involved in copper homeostasis, including copper regulator CueR, copper chaperones CusF and Atx1. The major technique is X-ray crystallography.

CueR is one member of MerR family metalloregulatory proteins and controls the transcription of genes involved in copper detoxification in *E.coli*. The structures of metal bound forms of CueR, combined with the zinc regulator ZntR of the same family, elucidated the unusual metal coordination environment accounting for the extraordinary metal responsiveness and metal selectivity. The mutational study was consistent with the observation from structural models. Biochemical and biophysical investigations of the binding of CueR and its responsive promoters revealed important minor groove interaction crucial for protein-DNA recognition. Extensive trials to cocrystallize CueR-DNA complex finally led to the very recent intriguing breakthrough, which will eventually produce the structural model of CueR/DNA complex.

CusF is a small periplasmic protein thought to function as part of the *cus* copper tolerance system. The structure characterized a Cu(I) center coordinated with Met-rich motif. It also revealed a Trp close to the metal center at a remarkable distance, suggesting the existence of cation-π interaction, which has been recently confirmed the Resonance Raman spectroscopic data from Dr. Anna Davis.

Atx1 is a small cytosolic copper chaperone in yeast, and has been characterized extensively. It was utilized to investigate the interaction between tetrathiomolybdate (TM) and copper proteins. The structure was solved by MR and Cu-MAD. The model presented the first structural demonstration of the protein-Cu-TM complex. It elucidated the binding mechanism between TM and Cu-Atx1, and revealed how the versatile inorganic Mo-Cu cluster adapted to a biological environment. It greatly deepens the understanding of the classical Mo-Cu antagonism and provides new insights into TM as a potential anti-cancer drug.

_____

Thomas V O'Halloran

Dissertation advisor

ACKNOWLEDGMENTS

TABLE OF CONTENTS

CHAPTER 1   Introduction and Scope of Thesis

CHAPTER 2   Metal Selectivity of *E. coli* Metalloregulatory Proteins CueR and ZntR

CHAPTER 3   Biophysical Studies of Interaction between CueR and Its Responsive Promoter

CHAPTER 4   Crystallographic Studies of Metal bound forms of CusF Protein

CHAPTER 5   Crystallographic Studies of Atx1-Cu-Tetrathiomolybdate(TM) Complex

CHAPTER 6 Crystallization of CueR/DNA/Ag(I) Complex

REFERENCES

APPENDIX

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 4

CHAPTER 5

CHAPTER 6

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 1

Introduction and Scope of Thesis

1.1 Biological importance of copper

The biological activities of one-third of all proteins require essential transition metal ions that perform catalytic, structural or regulatory functions[1]. Key transition metal ions include the first row elements zinc, copper, nickel, cobalt, iron and manganese, as well as second and third row elements molybdenum and tungsten.

Copper plays a vital role as catalytic co-factor for a variety of metalloenzymes including superoxide dismutase (for protection against free radicals), cytochrome c oxidase (mitochondrial electron transport chain), tyrosinase (pigmentation), peptidylglycine alpha-amidating mono-oxygenase (PAM) (neuropeptide and peptide hormone processing), and lysyl oxidase (collagen maturation) [4-6], At the same time, copper is toxic to both prokaryotic and eukaryotic cells due to its ability to produce the damaging free radicals. As a result, organisms have evolved elaborate mechanisms to control cellular uptake, distribution, detoxification, and elimination of copper [7, 8].

The imbalance of copper homeostasis causes several human disorders, which include Menkes disease, Wilson disease [9], Indian childhood cirrhosis [10], Endemic Tyrolean infantile cirrhosis [11], and idiopathic copper toxicosis [12]. A deficiency of bioavailable copper in the brain may contribute to the pathogenesis of neurodegenerative disorders, notably Alzheimer's disease (AD) [13]. Abnormalities in copper homeostasis are also associated with spongiform encephalopathies, commonly referred to as prion disease, and probably Parkinson's disease [14, 15]. A cancer connection of copper is suggested by the recent finding that the cuproenzyme lysyl

oxidase is required for hypoxia-induced metastasis [16]. Additionally, the copper nutrition is linked to iron metabolism via copper-containing ferroxidases. Low copper intake can result in anemia, possibly as a result of low activities of copper ferroxidase [17].

1.2 Copper homeostasis

Prokaryotes

In prokaryotes copper homeostasis is maintained by metal-responsive transcriptional regulatory proteins that regulate the transcription of genes encoding proteins responsible for metal detoxification, sequestration and efflux. Two best understood systems include *Escherichia coli* and *Enterococcus hirae*.

Three copper-responsive regulatory have been identified in Gram-negative bacterium *Escherichia coli* and they are induced in a stepwise fashion as the copper concentration in the growth medium is increased (Figure 1-1) [3]. The first are *cue* system, composed of CueO, CopA, and CueR. CueO is a multi-copper oxidase, and is proposed to convert Cu(I) to less toxic Cu(II), thus protecting periplasmic enzymes from copper-induced damage [18]. CopA is the central component of copper homeostasis in *E. coli* and is required for intrinsic copper resistance under both aerobic and anaerobic conditions. CopA is a Cu(I)-translocating P-type ATPase which pumps copper from the cytoplasm into preiplasm [19, 20]. CopA possesses two putative CXXC motifs in the N-terminal domain that are potential metal-binding sites, however, amino acid substitutions by site-directed mutagenesis showed that these motifs are not required for function and do not confer metal specificity [21]. CueR is a member of the MerR family of metal-responsive regulators and regulates *cueO* and *copA* [22]. As Cu concentrations continue to rise, the CusCBFA genes are induced by a two-component system (CusRS), which typically monitors

stress in the cell envelope and is particularly effective in anaerobic Cu stress conditions [19]. The

CusCFBA system is expected to serve as a proton-substrate antiporter to remove excess copper

from the periplasm[23]. The CusA protein is the central component for copper transport, the

putative channel-forming CusC and the clamping CusB are necessary for full Cus-mediated

copper resistance. CusF, a small periplamic protein, is suggested to be a metallochaperone

transporting copper to CusCBA efflux complex and thus facilitating copper detoxification of the

periplasm. If increasing Cu concentrations overwhelm these chromosomally encoded pathways,

cells that harbor either the *pco* (*E. coli*) or *cop* (*Pseudomonas syringae*) Cu resistance plasmids

can invoke one last line of defense. The *pco* operon includes the bacterial multicopperoxidase

PcoA and its putative partner PcoC, both of which are exported to the periplasm [24]. PcoC

exhibits a cuperodoxin fold that binds Cu(I) through two Met sulfur atoms and one nitrogen or

oxygen ligand in a hydrophobic Met-rich loop that is exposed to solvent on the protein surface [25].

Enterococcus hirae, a Gram-positive bacterium, serves another extensive-studied and

well-understood prokaryotic copper homeostasis system. The *cop* operon regulating copper

uptake, availability and export consists of four genes that encode a repressor, CopY, a copper

chaperone, CopZ, and two CPx-type copper ATPases, CopA and CopB [26]. CopA is proposed to

serve in copper uptake, it possesses an N-terminal metal binding site with the consensus motif

CxxC [27]; CopB, exports copper out of cell and features a histidine-reach N-terminus metal

binding domain [28]. Both repressor CopY and chaperone CopZ play a role in the regulation of the

cop operon. CopY binds the DNA and represses the transcription [29]. Cu(I)-CopZ donates copper

to CopY, which then dissociates from the DNA, allowing transcription to proceed [30]. CopY, like

more repressors, acts as homodimer. The C-terminus half of CopY contains the CxCx4CxC

motif, which forms binuclear Cu-thiolate core. The three-coordinate copper is sequestered in

an environment inaccessible to solvent. CopZ, similar to Atx1, a yeast metallochaperone, adopts

the ferredoxin-like fold [31]. The metal binding CxxC motif is located in a loop and the exposed

copper binding site exhibits two- or three- coordination environment. The transfer of copper

between CopZ and CopY is dependent on electrostatic interactions.

Eukaryotes

Eukaryotic organisms from yeast to humans use elaborate systems to regulate copper

homeostasis, consisting of copper importers, copper chaperones, transcription factors, small

metal binding proteins called metallothioneins and copper exporters (Figure 1-2).

From yeast to human, copper is acquired by high affinity, membrane-associated copper

importers exemplified by the copper transporter(Ctr)-family. Cu(I) is the substrate for the Ctr

family members which are relatively small proteins containing three transmembrane domains [32-34]. A conserved feature of some Ctr importers is an N-terminal segment that contains one or

more 'Mets' motif (MxxM or MxM). The Met-rich motifs have been shown by deletion studies

in yeast and human cells to be important for survival under copper starvation.

Due to the toxicity of free intracellular copper, the delivery of copper to the target

proteins relies on copper chaperones. Three functional groups [35] of chaperones are observed. The

Atx1 family proteins include the prototype yeast Atx1 and the human homolog Hah1, which

deliver copper to the Cu-transporting ATPases (Ccc2 in yeast , ATP7A and ATP7B in human) in

the trans-Golgi network [36, 37], where copper is finally incorporated into multi-Cu oxidases such as

Fet3 and ceruloplasmin. A bacterial homolog of Atx1 called CopZ, has also been identified [30].

The copper chaperones for superoxide dismutase (CCS) comprise a distinct family of

metallochaperone, which donate copper to the eukaryotic anioxidant enzyme Cu, Zn-superoxide dismutase (SOD1) [38, 39]. A model for copper transfer from CCS to SOD1 has been developed based on previous studies: the N-terminal Atx1-like domain of CCS is proposed to capture and bind copper and may interact with C-terminal CXC copper binding site; Domain II of CCS exhibits strong homology to SOD1 and is involved in docking with SOD1 target [40]; The small domain III at the C-terminus has essential cysteines proposed to be involved in both disulfide formation with SOD and Cu-insertion into active site [41-43]. Three potential copper trafficking proteins, Cox17, Cox11 and Sco1, play a role in the assembly of mitochondrial cytochrome c oxidase [44]. Cox17 is proposed to facilitate copper uptake by the inner mitochondrial membrane proteins Cox11 and Sco1, which in turn may transfer it to cytochrome c oxidase [45]. Some fungi such as *Podospora anserine* are able to maintain respiration even upon insufficient copper supply, switching from COX to AOX, an iron-containing alternative oxidase [46].

The ATP-dependent transporters of the ATP7 family transport copper to remote organs in multicellular orgainisms. Two isoforms ATP7A and ATP7B are identified [47-51]. At elevated copper levels, ATP7A moves from trans-Golgi network to the plasma membrane and redistributes excess copper by pumping out from the cell [49]. ATP7A returns to the trans-Golgi network as soon as the intracellular copper levels are reduced. ATP7B, a liver-specific copper transporter, trafficks the excessive copper to the biliary canalicular membrane where copper is excreted into the bile duct for disposal from the body via digestive tract [49]. Loss-of-function mutations in the ATP7B gene lead to Wilson disease, an autosomal recessive disease which causes accumulation of copper in the liver and a life-threatening liver cirrhosis.

Even with elaborate control mechanisms of copper export and import, there can be conditions where copper is imported in excess. Therefore, an important aspect of metal homeostasis is the sequestration of toxic heavy metals. Metallothioneins constitute a group of low molecular weight, cysteine-rich proteins with outstanding metal binding capacity [52]. They are found in all eukaryotes as well as in some prokaryotes and mainly function as metal scavenger/metal storage proteins [53-55]. Typically, metallothioneins are expressed at a basal level, but their transcription is strongly induced by the transcription factor upon heavy metal load. In the budding yeast *S. cerevisiae*, the factor is Ace1, which is rich in cysteines and positively charged amino acids, harbors a copper-dependent DNA binding domain [56]. On the converse situation of copper starvation, yeast rely on the transcription factor Mac1 to regulate the expression of the high affinity copper transport system including CTR1, CTR3, FRE1 and FRE7 [57, 58]. From insects to mammals, the best understood transcription regulator for controlling metal homeostasis, including copper excess, is MTF-1 (metal-responsive element binding transcription factor) [59-61]. The DNA binding domain of MTF-1 consists of six zinc fingers. *In vitro* studies have shown that zinc activates human MTF-1 directly, whereas copper activates the transcription indirectly by displacing zinc from zinc-metallothioneins or other zinc binding cellular proteins [62, 63].

1.3 Scope of the thesis and related background

The thesis explores the chemistry and structural biology of copper homeostasis including the metalloregulatory protein CueR, the putative metallochaperone CusF in *E. coli*, and the yeast metallochaperone Atx1-drug complex. The primary questions are addressed in the crystallographic studies of the inorganic chemistry in these metal trafficking and sensing proteins

at atomic level. The results reveal the atypical metal coordination and elucidate the structure-function relationships for these emerging classes of metal receptors.

CueR

CueR, the copper sensor, is one of the metalloregulatory members of the MerR family. It regulates the transcription of CopA, which removes Cu(I) from the cytosol into periplasm, and CueO, which catalyzes the oxidation of Cu(I) to a less toxic form Cu(II) in the periplasm [64].

With the feature of longer spacer between –10 and –35 elements of the responsive promoters [65], the MerR family regulators employ a DNA-distortion mechanism (Figure 1-3) to control the transcription [66, 67]. Crystal structures of BmrR/DNA/TPP and MtaN/DNA [68, 69], two homologues of MerR family proteins bound to their responsive promoters, corroborate several predictions and the DNA-distortion mechanisms. More characterized MerR family members are listed in Table 1-1.

CueR is 28% identical to the archetype Hg-responsive *Tn501* MerR, and contains four cysteine residues, two of which are in the corresponding positions to the Cys117 and Cys126 residues of MerR involved in binding mercury. Overexpressed and purified CueR binds to P$_{copA}$ in gel retardation experiments and the affinity of CueR to the DNA decreases upon addition of Cu(I) [22]. Titration of different metals in the presence of Cu(I)-buffering solution reveals strong responses to Cu(I), Ag(I) and Au(I), no response to Zn(II) and Hg(II), which is consistent with the *in vivo* behaviors of the CueR/promoter system. Using the established Cu^+-CN^- formation constants, the corresponding free Cu(I) can be calculated. Cu(I) titration profile exhibits half-maximal induction at the concentration of 10^{-21} M, corresponding to the zeptomolar sensitivity for Cu(I), which is far less than one atom per cell [70]. The extraordinary copper sensitivity of

CueR suggests that prokaryotic cytoplasm operates under conditions of copper deprivation. Similarly, ZntR, the zinc sensor in *E. coli*, exhibits femtomolar (10^{-15} M) sensitivity to $[Zn(II)]_{free}$ [71], and shows responses to Zn(II), Pb(II) and Cd(II) but not Cu(I) and Ag(I). The ultrasensitivity also indicates there is no free zinc in the cell.

Prior to studies described here, the exact metal coordination environment was unknown for both CueR and ZntR. The mechanism underlying the ultra-sensitivity and selectivity of the metal sensor has not been established yet. Chapter 2 of the thesis describes the X-ray crystallographic studies of metal bound forms of CueR and ZntR at high resolutions. The atomic models reveal the atypical Cu(I) and Zn(II) coordination chemistry and important structural features associated with the extraordinary metal sensitivity and selectivity of both ZntR and CueR. As described later, the rare and unusual metal binding mode is well suited for an ultra-sensitive genetic switch.

Further biochemical and biophysical studies were carried out to investigate the interactions between CueR and P_{copA} (Chapter 3). The protein/DNA binding constant characterized by three different methods agrees well and is in μM range. One crucial residue involved in minor groove interaction is predicted and confirmed by mutagenesis. Initial binding tests with different lengths of DNAs were tried and provided the starting point for crystallization experiments of CueR/DNA complex. Different lengths of DNAs were tested for crystallization, and some of them did produce good crystals with DNA bound, which was confirmed by native gel. The final success was obtained very recently and some initial characterizations were presented (Chapter 6).

CusF

CusF is a small periplasmic protein encoded in the *cus* copper tolerance operon, and CusF is believed to facilitate *cus*-mediated copper export, serving as a periplasmic copper scavenger or copper chaperone and may interact in vivo with other system components [19, 23, 72]. Originally described as forming a pink copper complex, CusF has been reported to be both a Cu(II)- and a Cu(I)-binding protein with recent reports emphasizing the Cu(I) chemistry as being most physiologically relevant [23, 73-75]. The reported coordination chemistry of CusF is complicated by both Cu(I) and Cu(II) metal binding activities. Cu(II)-CusF has been previously described as a pink complex with an absorption maximum around 510 nm, and EPR and ENDOR study has implicated the N-terminal His residues (H3,4,5) as participating in Cu(II) coordination [75]. The apo structure of the protein combined with NMR chemical shift data was used to build a model of the Cu(I) site, involving His36, Met47 and Met49 in metal coordination [73].

To elucidate the properties of the metal binding sites, we carried out the X-ray crystallographic studies on the metal bound forms of CusF protein (Chapter 4). Both Cu(I)-CusF and Ag(I)-CusF were crystallized and the structures were solved. The atomic models reveal a trigonal planar metal coordination for both Cu and Ag, which is formed by two methionine sulfurs, one histidine nitrogen. The Met,Met,His metal binding site represents a recently discovered feature of copper homeostasis proteins and biological copper coordination. Furthermore, one tryptophan is positioned at a remarkably close distance to the metal center (~ 3.3 Å), suggesting strong cation-π interactions.

Atx1 and TM

Atx1, a 72-residue protein, originally isolated as a suppressor of superoxide toxicity [76], is a cytosolic yeast copper chaperone. After copper is transported into the yeast cell via the plasma membrane protein Ctr1[77], Atx1 delivers it to a P-type ATPase, Ccc2, which translocates copper across intracellular membranes to the multicopper oxidase Fet3 [78]. A yeast two-hybrid experiment has revealed a copper-dependent interaction between Atx1 and the N-terminus of Ccc2 in the cell [79]. A similar pathway has been proposed in humans: copper is apparently transported into the cell by plasma-membrane proteins similar to Ctr1 [33] and then shuttled by an Atx1-like protein, Hah1 [80], to the multicopper oxidase ceruloplasmin (a Fet3 analog) [81] via the Menkes and Wilson disease proteins, which are the human counterparts of Ccc2. A striking feature of both the chaperone proteins and target ATPases is the existence of one or more N-terminal MT/HCXXC sequence motifs, which have been proposed to bind copper ions. Atx1 and Hah1 each contain one repeat of this motif. The Menkes/Wilson disease proteins each contain six repeats of the motif [82] and Ccc2 contains two repeats [83]. It has been proposed that facile copper transfer occurs by docking of the chaperone and target proteins with the MT/HCXXC metal binding sites in close proximity and the subsequent formation of two-coordinate and three-coordinate intermediates in which the metal ion is ligated simultaneously by both proteins [84] (Figure 1-4).

The structures of apo- and Hg-Atx1[85] , Cd-, Hg- and Cu-Hah1 [84] reveal the similar overall $\beta\alpha\beta\beta\alpha\beta$ ferredoxin fold. The metal ion is coordinated to the two conserved cysteines located on the loop between the first α helix and the second β strand. The structure of Cu(I)-Atx1 is not available due to the disorder of the metal binding site. X-ray absorption studies suggest Cu(I) is

tri-coordinated by the two conserved cysteines and one exogenous thiol at a longer distance [79].

A variety of bacterial metal transport proteins adopt the similar topology and metal binding mode, such as MerP [86] and CopZ [87].

Tetrathiomolybdate (TM), a copper chelator, has been used as a therapeutic agent for treatment of Wilson's disease, which is caused by the accumulation of copper in the liver due to the dysfunction of Cu-transporting ATPase (ATP7B) [50, 88]. TM has the appropriate qualities of fast action, low toxicity and oral delivery. Recent phase III clinical data show TM is more effective than other treatments at reducing the disease's effects [89]. Furthermore, the dependence of tumor growth on angiogenesis and that copper being critical to angiogenesis result in successful trials of TM as a new and promising anticopper drugs [90, 91].

The mechanism of TM lowering copper levels has been suggested to be the formation of a tripartite complex among TM, Cu and proteins through Mo-S-Cu interactions. It is still unclear how does TM binds Cu and proteins. In order to understand the precise mechanism by which TM binds copper in the body and with the idea that the copper homeostasis always involving the transfer between the chaperone and the target, we started with Cu chaperone of yeast, Atx1, we crystallized the Atx1-Cu-TM complex (by Hamsell) and solved the structures of the complex (Chapter 5). It is the first structural demonstration of the biologically related Mo-S-Cu interactions. The model reveals a very novel Mo-Cu cluster, which includes both tri- and tetra-coordinated Cu atoms. This model can very well explain the EXAFS data of the isolated LEC rat liver lysosomes [92]. The atomic structure of Atx1-Cu-TM complex sheds remarkable insights on understanding the therapeutic mechanisms of TM in lowering copper level and will provide a great model for designing new TM-based drugs.
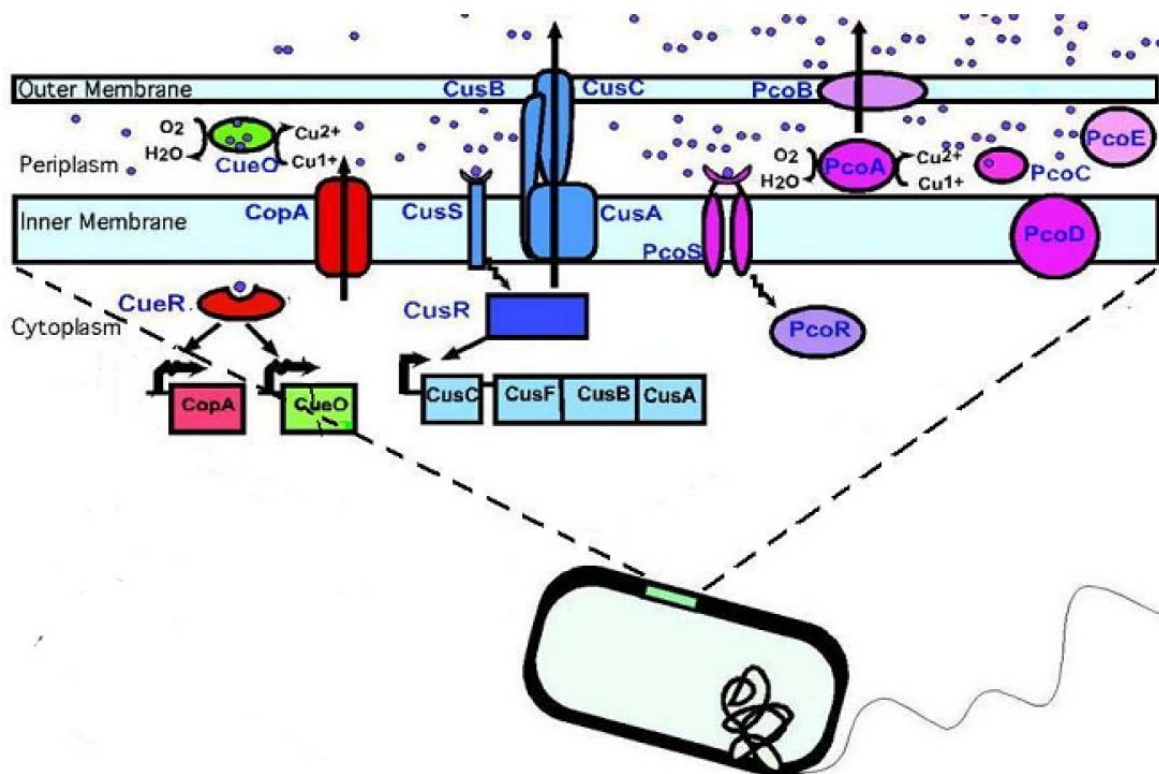
Figure 1-1. Copper trafficking in the periplasm in *E. coli*.
The periplasm, a compartment of the cell envelope of Gram-negative bacteria, is proving to be an important site of Cu trafficking and utilization. Cellular Cu efflux is controlled in *E. coli* by the *cue*, *cus*, and *pco* operons, each of which is induced at different levels of Cu stress by separate metalloregulatory proteins. Recent structural insights for CueO and PcoC are highlighted. The cartoons of Cu ions (purple balls) represent various levels of total Cu content in the periplasm. Picture is taken from reference 3.
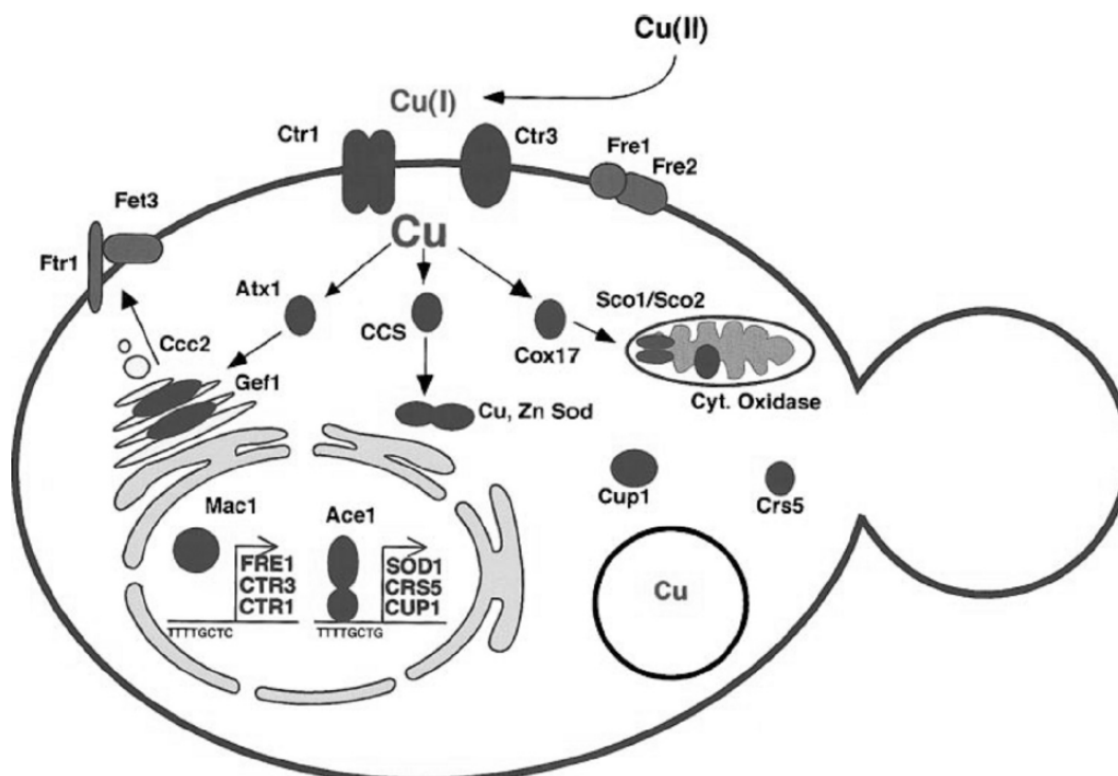
Figure 1-2. Copper transport and distribution in *Saccharomyces cerevisiae* cells.

Copper is first reduced from Cu(II) to Cu(I) by cell surface reductases Fre1/Fre2 prior to uptake. High affinity Cu(I) ion uptake is mediated by the Ctr1 and Ctr3 proteins. Within the cell Cu(I) ions are bound to the copper chaperones Atx1, Cox17, and CCS for specific delivery to Golgi complex, mitochondria, and Cu,Zn_SOD1, respectively. Within the Golgi complex, Cu ATPase (Ccc2) accepts Cu(I) from Atx1, followed by incorporation of Cu(I) into the multicopper ferroxidase, Fet3. Fet3 forms a complex with the iron permease Ftr1, and both proteins are responsible for high affinity iron uptake. In mitochondria Cu(I) ions delivered by Cox17 are incorporated into cytochrome *c* oxidase (COX), a process that requires the integral inner_membrane protein Sco1 and possibly its homolog Sco2. CCS delivers Cu(I) specifically to Cu,Zn-SOD1 in the cytosol. Figure is taken from reference 35.

Table 1-1. MerR family homologues

Protein	Organism	% Identity *	Inducer	Ref.
MerR	*P. aeruginosa (Tn501)*	100	Hg(II)	93-97
ZntR	*E. coli*	34	Zn(II)	71, 98, 99
CueR	*E. coli*	27	Cu(I)	22, 64
SoxR	*E. coli*	29	Superoxide	100-102
CoaR	*Synechocystis PCC 6803*	33	Co(II)	103
CadR	*P. putida*	32	Cd (II)	104
PbrR	*R. metallidurans*	32	Pb (II)	105
HmrR	*R. leguminosarum*	30	Cu (I)	106
PmtR	*P. mirabilis*	30	Zn (II)	107
TipAL	*S. lividans*	32	Antibiotics	108
Mta	*B. subtilis*	32	Multidrug	109
BmrR	*B. subtilis*	28	Multidrug	110, 111
BltR	*B. subtilis*	22	Multidrug	111

*Values were obtained using the BLAST local alignment tool (NCBI)

(The table is taken from Caryn E Outten's Ph.D. thesis)
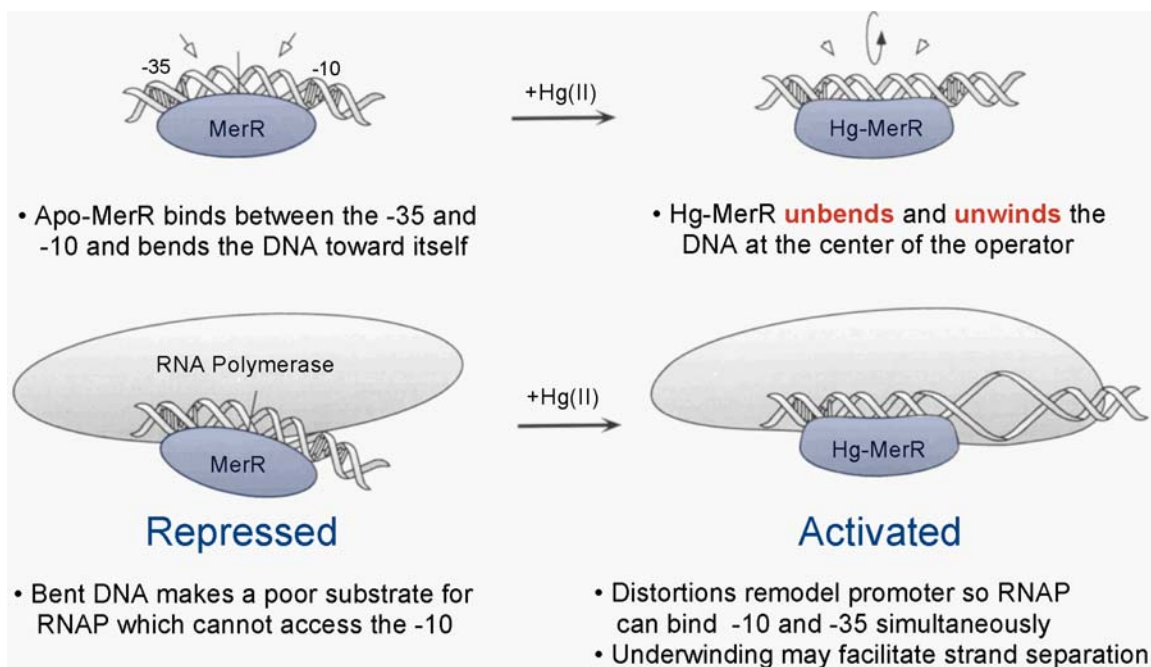
Figure 1-3. MerR DNA distortion model.

Apo-MerR bends the DNA producing two kinks (open arrows) that are symmetrically spaced around the center of the operator (vertical line). Upon binding Hg(II), MerR unbends the DNA and unwinds (circular arrow) the center of the operator. With apo-MerR bound to the operator, RNAP is only accessible to the –35 site and transcription is repressed. The unkinking and untwisting introduced by Hg(II)-MerR allows RNAP to access both the –35 and –10 and transcription is activated. Picture is taken from reference 2.
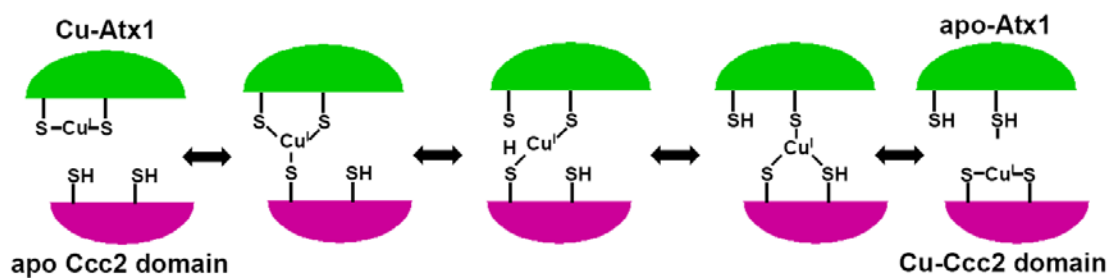
Figure 1-4. Hypothetical mechanism of copper transfer between Atx1/Hah1 and Ccc2/Menkes or Wilson protein.
The formation of a metal-bridged intermediate is supported by the crystal structure of Hah1, which is shown. The picture is adapted from reference 3.

CHAPTER 2

Metal Selectivity of *E. coli* Metalloregulatory Proteins CueR and ZntR

Abstract

The *E. coli* copper sensor CueR and zinc sensor ZntR are metalloregulatory members of MerR-family proteins. They exhibit metal-selective behavior both *in vivo* and *in vitro*. The extraordinary metal selectivity has been demonstrated with *in vitro* transcription assay. In order to elucidate the structural basis for the metal selectivity and sensitivity, we have crystallized the metal bound forms of CueR and ZntR, and solved the structures. The atomic details reveal the atypical metal coordination and the local environment features contributing to the metal selectivity. A residue at the N-terminus of the dimerization domain is suggested to be unique in discriminating +1 and +2 ions. We have evaluated the residue with site-directed mutagenesis combined with metal-ratio analysis and *in vitro* transcription assay. The results are consistent with the structural prediction.

2.1 Introduction

Copper and zinc are essential nutrients. The electron transfer properties of copper are fundamental to processes such as respiration and photosynthesis [1-4]. Zinc forms the catalytic center in numerous enzymes and has an important structural role in a wide range of proteins [5, 6]. However, the metals can be toxic if their levels and distribution are not carefully regulated, as their inappropriate binding may compromise cellular function. Therefore, organisms maintain cytoplasmic metal concentrations at a nontoxic level that is sufficient for growth.

A number of homeostatic mechanisms have been identified which ensure that copper is maintained at a level sufficient for, but not toxic to, cell growth. In mammals, posttranslational mechanisms, such as the intracellular trafficking of copper transporters and the copper-stimulated endocytosis and degradation of proteins involved in copper uptake, play a major role in copper homeostasis [7-10]. Although posttranslational control of transporters exists in fungi, copper homeostasis in these organisms is also mediated by the transcriptional regulation of genes involved in copper acquisition, mobilization and sequestration [11]. Six copper-responsive fungal factors have been characterized in detail. Ace1 [12, 13], Amt1 [13], and Crf1 [14] activate gene expression in response to elevated copper while Mac1 [15], GRISEA [16], and Cuf1 [17] activate gene expression in response to copper deficiency. Homeostasis through copper-responsive transcriptional regulation has been observed in insects and plants, as well as fungi, suggesting that this mechanism of copper control is widespread in eukaryotes [18-20].

To a major extent, zinc homeostasis generally parallels copper homeostasis in that both posttranslational and transcriptional regulatory mechanisms function together to maintain zinc at an optimal level under conditions of either zinc limitation or zinc excess. However, unlike

copper, zinc-responsive transcription factors are found in fungi, mammals, fish and possibly plants, suggesting that the transcriptional control of genes involved in zinc homeostasis is of universal importance [21-25]. Two factors that control gene expression in response to zinc have been characterized in detail. These are Zap1 from *S. cerevisiae* [24], which activates gene expression in response to zinc deficiency, and mammalian MTF-1, which is activated by zinc [26].

Since prokaryotic organisms are devoid of sub-celluar compartments, the major mechanism that maintain celluar copper and zinc level is transcriptional regulatory of genes involved in metal uptake, transport and sequestration.

In *E. coli*, three systems are identified which prevent the overaccumulation of copper. The first one, the plamid-encoded pco copper resistance operon consists of at least seven genes that facilitate copper efflux [27]. The second one, *cusCBFA*, is proposed to form a proton-cation antiporter that exports copper out of the cell. These genes are regulated by *cusRS*, a two-component copper-responsive signal transduction system [28]. The second system includes CopA [29], a P-type ATPase that moves Cu(I) and presumably Ag(I) out of the cell, and CueO, a putative multicopper oxidase targeted for the periplasm [30]. In vivo β-galactosidase assays indicate that the *copA* and *cueO* promoters are activated by copper in a *cueR*-dependent manner [30]. Both *cus* and *cue* systems are chromosomally encoded.

In *E. coli*, the trafficking of Zn(II) is controlled by the metal-sensitive regulatory proteins Zur and ZntR. Zur, a homologue of Fur, controls the transcription of *znuABC* operon, which encodes three proteins involved in Zn(II) uptake [31]. ZntR controls the transcription of the zinc pump ZntA, a cation-translocating ATPase that couples Zn(II), Cd(II), or Pb(II) export with ATP hydrolysis [32].

The sequences of CueR and ZntR proteins exhibit 34% and 28% identity to MerR, respectively. Both proteins are MerR family metalloregulatory proteins that utilize a DNA distortion mechanism to activate metal detoxification or export genes. Two distinct hallmarks of promoters regulated by MerR homologues are a palindromic or pseudo-palindromic sequence and a 19-20 bp spacer between –35 and –10 elements. The unusual spacing (consensus spacing is 17 bp) is a key aspect of the DNA distortion mechanism. RNA polymerase cannot effectively access both promoter elements since the long spacer places them on opposite faces of the DNA. Apo-MerR further prevents RNA polymerase access by binding to the palindrome and bending the DNA towards itself, effectively repressing transcription. Upon binding of Hg(II), MerR unbends and unwinds the DNA, remodel the promoter, and facilitate the RNA polymerase binding and forming open complex [33, 34]. Both ZntR and CueR use a similar DNA distortion mechanism in response to elevated inducer concentrations [35].

One remarkable feature of MerR family metalloregulatory proteins is the metal selectivity and ultrasensitivity. MerR is responsive to nanomolar levels of Hg(II) in presence of excess competing thiols, and, the transcriptional activity increases from 10% to 90% over only a 7-fold increase in Hg(II) concentration [36]. Zn(II) and Cd(II) also activate the transcription of *mer* operon, but at much higher concentrations (micromolar). Both the ultrasensitivity and selectivity phenomena are also observed in vivo for the *mer* operator [27, 37]. Likewise, ZntR exhibits ultrasensitivity both in vitro [35] and in vivo [32]. Femtomolar sensitivity to $[Zn(II)]_{free}$ has been reported by the calibration of strong zinc chelator TPEN [38]. In vivo studies of CueR suggest CueR responds to Cu(II) and Ag(I), and it responds to Ag(I) in an ultrasensitive mode [39].

In order to understand the metal selectivity and sensitivity of the MerR family of metalloregulatory proteins from atomic level, we crystallized Cu(I) and Ag(I) bound forms of CueR and Zn(II)-bound form of ZntR, and solved the structures in collaboration with Professor Mondragon's Lab. The atomic models reveal the structural basis of the selectivity and sensitivity of metal ion recognition, the crucial amino acid discriminating between +1 and +2 metal ions is identified, and, a structure-based sequence alignment of MerR metalloregulatory homologs allows for predictions of metal selectivity in other family members. Mutation of this crucial amino acid of ZntR causes the abolishment of Zn(II) binding ability and *in vitro* transcription assay of the mutant ZntR reveals the total loss of transcription in presence of Zn(II) , which corroborates the prediction from the atomic models.

2.2 Experimental procedures

CueR cloning and purification

The overexpression plasmid pET24aCueR was prepared by Caryn E. Outten [30]. The plasmid was transformed into BL21(DE3) cells (Novagen). The cells were grown in 9L of LB with shaking at 37°C and induced with 400 μM IPTG at OD_{600} = 0.6. After 2.5 hours of IPTG induction, the cells were harvested by centrifugation at 8,000 rpm for 10 minutes, then stored in –80°C. The pelleted cells were frozen by liquid nitrogen and thaw at room temperature three times, then resuspended in 250 mL Tris buffer A (50 mM Tris-Cl, pH 8.0, 2 mM EDTA, 5 mM DTT), sit on ice for 1 hour with occasional mixing. The cells debris was removed by centrifugation at 9,000 rpm for 15 minutes, and the protein in the supernatant was precipitated with 45% $(NH_4)_2SO_4$. The precipitated protein was resuspended in 10 mL of Tris buffer A, and

centrifuged at 15,000 rpm for 15 minutes. The supernatant containing the protein was filtered

with 0.45 μm membrane filter (Milipore) and then desalted on a Sephadex G25 protein. The

crude protein extract was subsequently loaded onto a Heparin column equilibrated with Tris

buffer B (20 mM Tris, pH 8.0, 5 mM DTT). Following elutions with a NaCl gradient, coming off

at 0.25-0.35 M NaCl, the protein fractions were collected and concentrated to 1~2 mL. As a final

purification step, the protein was loaded onto a High Load Superdex 75 gel filtration column

(Pharmacia) equilibrated with Tris buffer C (50 mM Tris, pH 8.0, 250 mM NaCl, 5 mM DTT).

The CueR fractions were concentrated to 1~ 2 mL, and applied to SDS-PAGE gel to inspect

purity. Protein were stored at -80°C in Tris buffer C with 5% glycerol.

A typical 9 L preparation of cells yields ~ 50 mg of protein that is > 95% purity by SDS-

PAGE analysis (Figure 2-1). The molecular weight of CueR (calculated 135 aa for full length

protein: 15236.0 Da) was found by matrix-assisted laser desorption/ionization-time of fight

(MALDI-TOF) (PerSpective Biosystems Voyager-DE) to be 15233.8 Da using sinapinic acid

matrix with myoglobin as the calibration standard (Figure 2-2).

ZntR cloning and purification

The overexpression plasmid pET11cZntR was prepared by Caryn Outten [35]. The

overexpression and purification of ZntR is the same as CueR protein except there is one more

hydrophobic column between Heparin column and gel filtration column. After Heparin column

(protein coming off at 0.1-0.2 M NaCl), the protein fractions were collected and $(NH_4)_2SO_4$ were

added to a final concentration of 1.2 M. The protein was then loaded onto a Phenyl Sepharose

High Performance hydrophobic column (Pharmacia) in Tris buffer D (50 mM Tris, pH 8.0, 1.2

M $(NH_4)_2SO_4$, 5 mM DTT). The protein was eluted with a decreasing salt gradient, coming off

at 0.5-0.3 M $(NH_4)_2SO_4$.

A typical 9 L preparation of cells yields ~ 100 mg of protein that is > 95% purity by

SDS-PAGE analysis (Figure 2-2). The molecular weight of ZntR (calculated 141 aa for full

length protein: 16179.3 Da) was found by MALDI-TOF to be 16181.9 Da using sinapinic acid

matrix with cytochrome C as the calibration standard (Figure 2-3).

Mutangenesis, expression, and purification of ZntRC79S

The ZntRC79S mutant genes were prepared by site-directed mutagenesis (QuikChange®

Site-Directed Mutagensis Kit, Stratagene) of the over-expression vector pET11c inserted with

the wild-type *zntR* gene (prepared by Caryn E. Outten). The sequences of the oligonucleotides

used are shown below, with the modified base shown in bold lower case: 5'-

CGATCCTGAACACCATACCT**ca**CAGGAGTCAAAGGC -3'. The sequence of the mutated

gene was confirmed by DNA sequencing analysis. Similar over-expression and purification

procedures as those described for wild-type ZntR yielded the mutant proteins with >95% purity

(Figure 2-4). The molecular weight (calculated 141 aa for full length protein: 16163.3 Da) was

found by MALDI-TOF to be 16162.2 Da using sinapinic acid matrix with cytochrome C as the

calibration standard (Figure 2-5).

Protein concentration determination

The protein concentration was determined by the Bradford assay (Bio-Rad, with IgG as

the calibration standard) with correction factor derived from quantitative amino acid hydrolysis

(Protein Chemistry Laboratory, Texas A&M University). Comparison of protein concentrations

revealed that the Bradford assay overestimated the concentration of the wild-type CueR by 3.71

fold, ZntR by 2.65 fold. The mutants of CueR and ZntR directly adopted the correction factors

of the wild-type proteins.

DNA design, purification and annealing

Initial attempt was to crystallize protein bound to its responsive promoter and metal ion. Since the responsive promoters of MerR family members are generally palindromic or pseudo-palindromic, one efficient way to design the crystallization candidates is to divide the full sequence into two shorter sequences, the longer sequence has about 4 bp overlap after annealing. Thus, it is easy to modify both sequences (mostly termini) and produce a variety of DNA candidates with different termini and overhangs. Promoters of *copA* and *zntA* were chosen as the design templates (Figure 2-6). The designed annealed DNA covered the full protein-DNA interaction region identified by footprinting experiments [30, 35].

Dry oligos were directly ordered from IDT Corp. and stored in -20°C freezer. The oligos were purified by reverse phase HPLC. Dry oligos were dissolved in 0.1 M TEAAc[*] to a final concentration of ~ 2.4 mg/mL and centrifuged at 15,000 rpm for 3 minutes, the clear supernatant were moved to a clean eppendorf tube. 200~300 μL of sample was injected on a 250×4.1mm (length $\times$ radius) poly(styrene-divinylbenzene) reverse phase HPLC column (Hamilton Corp.). Flowing rate was 1.5 mL/min. 20 mL of 10% CH_3CN/TEAAc was applied to wash off untritylated oligos, followed by 10~20 mL 0.5% TFA to remove the trityl groups of the bound oligos. A 60 mL of 10-30% gradient of CH_3CN/TEAAc eluted the oligos. Trace amount of impurities bound to the column was washed off with 20 mL of 80% CH_3CN/TEAAc. Another

---

[*] 0.1 M TEAAc: 950 mL of water + 5.6 mL of acetic acid + 13.86 mL of TEA, adjust pH to 7.0 with acetic acid, add water to the final volume of 1L.

20~30 mL of 10% CH$_3$CN/TEAAc was applied to re-equilibrate the column. Fractions containing the purified oligos were collected and dried by speed-vac. Dried pure oligos were dissolved in 400 μL 0.3 M sodium acetate, after adding 1200 μL of ice cold ethanol, the solution was put into -80°C freezer and sit overnight. The precipitated oligos were pelleted by centrifugation at 15,000 rpm for 30 minutes, supernatant was romoved. The pellet was washed with 300 μL of 70% ice cold ethanol and dried by speed-vac.

Pure dried oligos were dissolved in TE buffer (Tris: 10 mM , EDTA: 1 mM, pH 8), concentration were decided by absorption coefficient at 260 nm.

$$\textit{Concentration (μg/mL)} = A_{260} \times \textit{33 (μg/mL)} \times \textit{Dilution factor}$$

$$\textit{Concentration (mM)} = \textit{Concentration (μg/mL)} \div \textit{M.W. (g)}$$

Equimolar amounts of purified oligos were mixed thoroughly and heated at 95°C for 3 minutes. After turning off the heat block, the samples were left in the heat block to cool down to room temperature slowly. The concentration of the annealed DNA was decided by absorption coefficient at 260 nm. Annealed DNA was stored in -20°C freezer.

$$\textit{Concentration (μg/mL)} = A_{260} \times \textit{50 (μg/mL)} \times \textit{Dilution factor}$$

$$\textit{Concentration (mM)} = \textit{Concentration (μg/mL)} \div \textit{M.W. (g)}$$

Crystallization of Cu(I)- and Ag(I)-CueR

The protein (1~2mg/ml) was mixed with 10 mM DTT and an equimolar amount of [Cu(CH$_3$CN)$_4$]PF$_6$ (in CH$_3$CN). In an attempt to crystallize the metal-activated protein-DNA ternary complex, a 3-fold molar excess of duplex DNA P$_{copA}$-18/12mer (Figure 2-6) derived from the CopA promoter sequence was included in the crystallization experiment, but turned out not to be present in the crystal. CueR crystals were grown at 18°C by vapor diffusion in hanging

drops (Figure 2-7) equilibrated against 1.0 M sodium citrate and 0.1 M sodium cacodylate (pH 6.5). Crystals of silver-bound CueR (Figure 2-7) were grown in the same manner except that the protein/DNA/metal mixture was incubated for 1 hour at 22°C prior to crystallization and contained a 5-fold molar excess of $AgNO_3$, 5 mM DTT, and a 1.5-fold molar excess of duplex DNA P_{copA}-18/12mer. Crystals appeared overnight, and reached final dimensions of $0.04 \times 0.06 \times 0.2$ mm in approximately one week.

Crystallization of Zn(II)- ZntR

ZntR (4 mg/ml) was mixed with 5 mM DTT, a 5-fold molar excess of $ZnSO_4$, and a 1.5-fold molar excess of duplex DNA P_{zntA}-18/13mer (Figure 2-6) derived from the ZntA promoter sequence. ZntR crystals (Figure 2-7) were grown at 18°C by vapor diffusion in hanging drops equilibrated against 0.1 M Na_2HPO_4, 0.1 M MES (pH 6.5), 20 mM NaCl, and 10 mM $MgCl_2$. Crystals appeared in one week, and reached final dimensions of $0.05 \times 0.05 \times 0.2$ mm in approximately three weeks.

CueR structure determination (by Anita Changela of Prof. Alfonso Mondrgon's lab)

Data collection and structure determination were collaborated with Professor Mondragon's lab (by Anita Changela). The statistics are shown in Table 2-1.

Crystals of derivatized CueR were obtained by soaking Ag(I)-CueR crystals in crystallization solution containing 1 mM $K[Au(CN)_2]$ for 2 days. For data collection, crystals were transferred in a single step to crystallization solution supplemented with 12% ethylene glycol for 1~2 minutes and flash-cooled in liquid nitrogen. All data sets were collected at 100 K using synchrotron radiation. Data were processed using MOSFLM [40] or XDS [41] and scaled with SCALA [42]. CueR crystals belong to space group $P2_12_12_1$ (a = 59 Å, b = 66 Å, c = 82 Å) and

contain one dimer in the asymmetric unit. The structure was determined by SAD methods from a gold soaked CueR crystal. Data were collected just above the theoretical gold LIII absorption edge in 15° segments using inverse beam techniques. One gold site was identified in an anomalous difference Patterson map. Refinement of the first site in SHARP [43] led to the identification of a second site. Heavy atom refinement and SAD phasing with 2 sites in SHARP followed by solvent flattening yielded an excellent electron density map from which the entire structure could be traced. Model building was carried out in O [44] and alternated with cycles of simulated annealing, positional, and individual temperature factor refinement in CNS [45]. The Au-CueR structure was used as a model for subsequent structure determination and refinement of the Ag-CueR and Cu-CueR structures. Metal ions were identified by strong peaks in anomalous difference Fourier and Fo-Fc difference maps. The Cu^+ and Ag^+ atoms were refined at full occupancy whereas the Au^+ atoms were refined at 60% occupancy. For each CueR structure, residues 115-119 of one monomer and residues 128-135 at the C-termini of both monomers are disordered and not included in the final model. All residues in the final model are found within the most favored or allowed regions in the Ramachandran plot.

ZntR structure determination (by Anita Changela of Prof. Alfonso Mondrgon's lab)

A heavy atom derivative was obtained by soaking ZntR crystals in crystallization solution containing 0.5 mM K_2PtCl_4 for 14 hours. Crystals were cryoprotected by a quick transfer to crystallization solution supplemented with 35% glycerol followed by flash cooling in liquid nitrogen. To maximize the anomalous scattering from the platinum, a derivative data set was collected just above the theoretical Pt LIII absorption edge in 15° segments using inverse beam techniques. All data were processed with MOSFLM and scaled using SCALA. ZntR crystal

belongs to space group $P2_12_12_1$ (a = 42 Å, b = 47 Å, c = 119 Å) and contains a dimer in the asymmetric unit. The platinum derivative data set was scaled against a native data set using SCALEIT [42] and two platinum sites were located. Heavy atom refinement and SIRAS phasing with two sites were carried out in SHARP, and a readily interpretable electron density map was obtained after solvent flattening. The electron density map revealed that only a fragment of ZntR had crystallized which was later confirmed by matrix-assisted laser desorption/ionization time of flight mass spectrometry, and was the result of proteolysis occurring in the crystallization drop. Model building was carried out in O and was interspersed with cycles of simulated annealing, positional, and individual temperature factor refinement in CNS. Once most of the structure had been built, the resolution was extended using native data to 1.9 Å. Two Zn^{2+} atoms and a phosphate ion were modeled in at each binding site at full occupancy. The final model contains residues 43-136 for one monomer and residues 45-133 for the other monomer, four Zn^{2+} atoms, two phosphate ions, one hydrated magnesium ion, and 236 water molecules. All residues are within the most favored or allowed regions of the Ramachandran plot. Structure determination of the other two crystal forms was performed by molecular replacement in AMORE [42] by using the ZntR dimer as a search model. All crystal structures reveal identical conformations of the truncated form of ZntR and contain two zinc atoms and a sulfate or phosphate ion (assignment based on crystallization components) at each metal binding site.

Metal-binding assays

Metal binding assay was performed in an Amicon ultrafiltration cell with a YM-10 membrane. 2~4 equivalents of $ZnSO_4$ or $AgNO_3$ were added to 10~20 μM protein (ZntR, ZntRC79S or CueR) in 10 mL of 50 mM Tris, 250 mM NaCl, pH 8, 1~2 mM DTT. The solution

was stirred at 4°C for 30 minutes and washed 3~4 times with buffer to remove excess metal.

Metal concentrations were determined by inductively coupled plasma atomic emission spectroscopy (ICP-AES) using a Varian 700-ES series spectrometer. Protein concentrations were determined by Bradford assay as described in 'Protein concentration determination' section.

In vitro run-off transcription of ZntR and ZntRC79S [#]

Transcription template was prepared using the (-48) and (-47) pUC sequencing primers (New England Biolabs) and the plasmid pUC19Zntfoot. The PCR fragment was then digested with *Bbv*I to make a 171-bp fragment, gel purified, and quantitated by comparison with DNA standards stained with ethidium bromide. Transcription run-off experiments were carried out in 20 μL transcription assay buffer (10 mM Tris-Cl, pH 8.0, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 100 mM K-glutamate, 100 μg/mL BSA, 5% glycerol, 20 mM TPEN and 1 mM fresh DTT). DNA template, RNase inhibitor (Ambion), ZntR/ZntRC79S, $ZnSO_4$/$AgNO_3$, and RNAP (USB) were mixed together and incubated at 37 °C for 20 minutes. Total 10 μL Heparin-NTP solution (0.15 mM ATP, GTP, CTP and 0.0015 mM UTP, 200 μg/mL heparin) combined with [α-^{32}P]-UTP in a 9:1 ratio were added and incubated for another 20 minutes at 37 °C. The reaction was terminated with 30 μL stop buffer (9 M NH_4OAc, 200 μg/mL yeast RNA, 40 mM EDTA) and then was precipitated by sitting at –80 °C freezer for 20 minutes after adding 100 μL EtOH. The pellets were washed with 70% EtOH, dried under vacuum, and then resuspended in loading buffer (0.53 TBE, 80% formamide, 0.05% xylene cyanol FF, 0.05% bromphenol blue). The samples were loaded onto a 7% polyacrylamide sequencing gel containing 7 M urea. The gel was dried for 20~30 minutes and exposed to Kodak film for overnight.

[#] Detailed procedures for in-vitro transcription assay are listed in appendix 1.
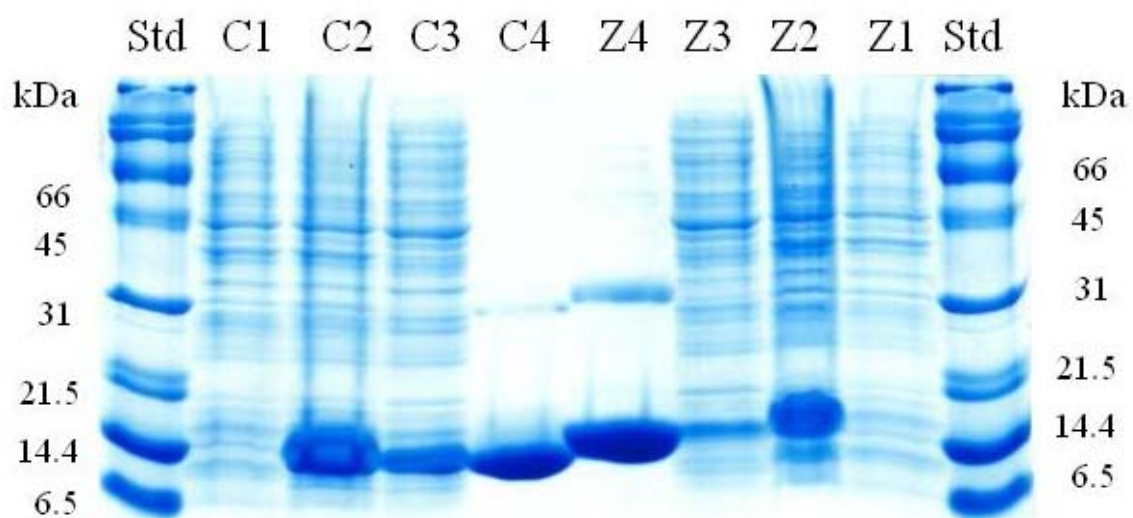
Figure 2-1. CueR and ZntR SDS-PAGE gels. Std: molecular mass standards; C stands for CueR protein. C1: non-induced cells; C2: induced cells; C3: crude protein; C4: purified protein. Z stands for ZntR protein. Z1: non-induced cells; Z2: induced cells; Z3: freeze-thaw supernatant; Z4: purified protein.
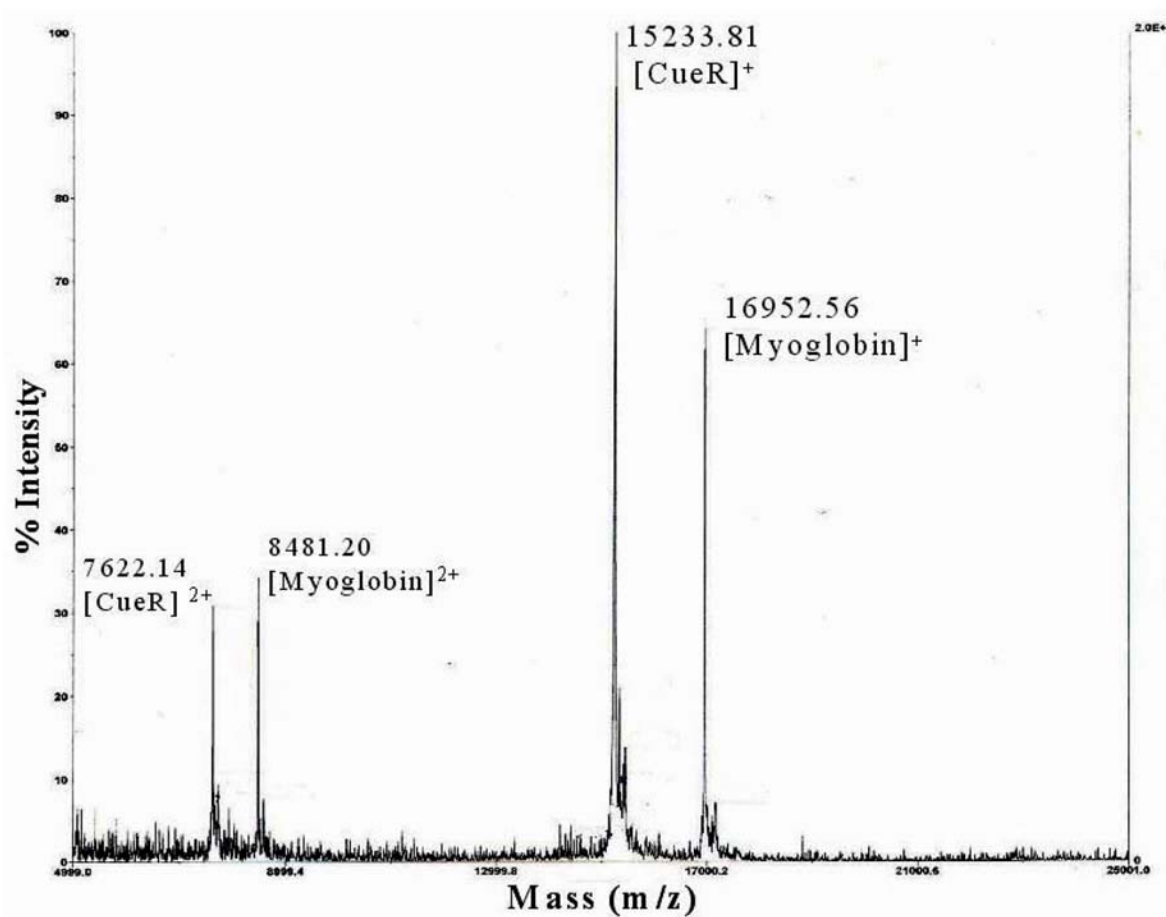
Figure 2-2. CueR MALDI-TOF spectrum
1~2 μg protein sample and ~2 μg myoglobin (internal standard) were
used for the measurement.
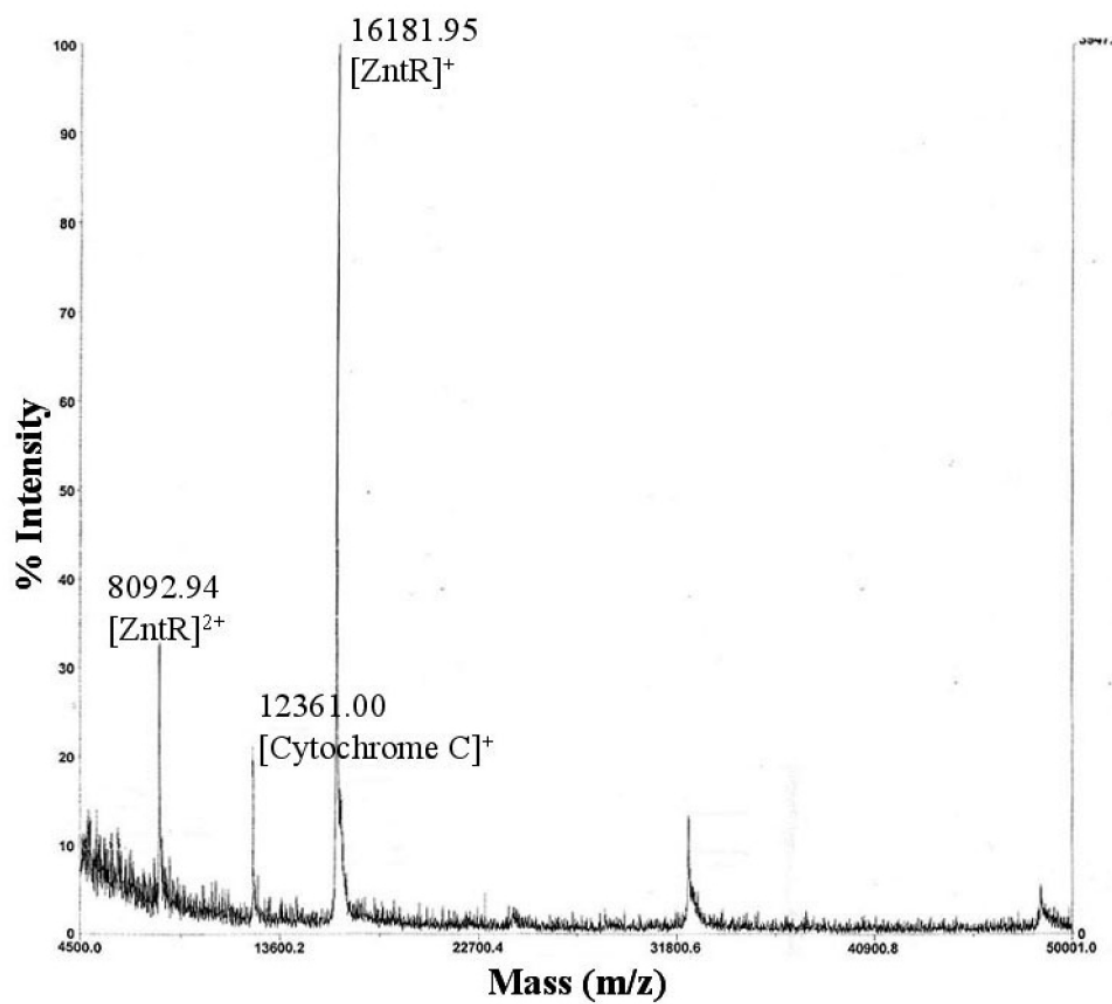
Figure 2-3. ZntR MALDI-TOF spectrum
1~2 μg protein sample and ~2 μg cytochrome C (internal standard) were used for the measurement.

(A) Induction test

Collected fractions

**(B) eluants of final purification column
- gel filtration column**
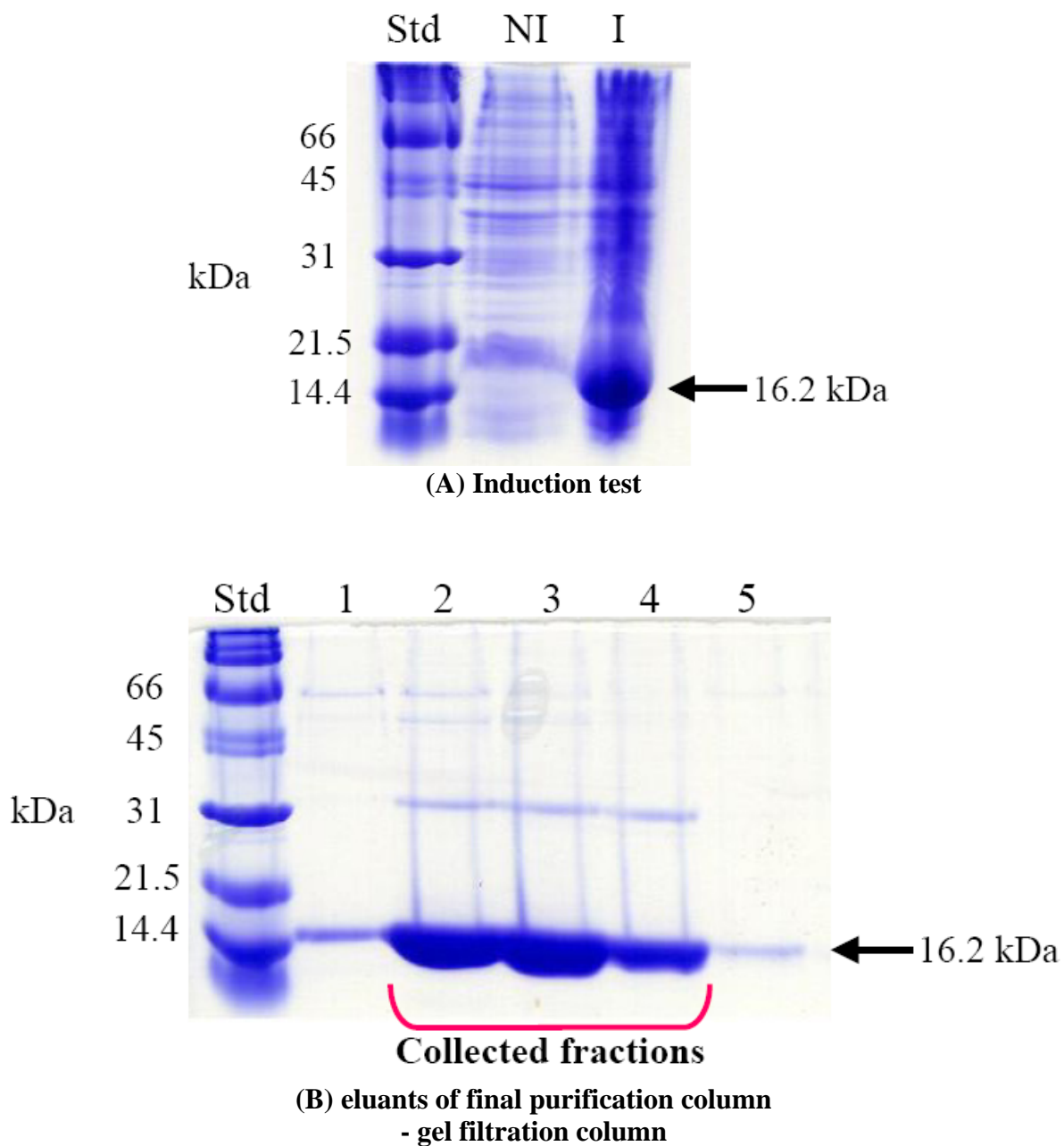
Figure 2-4. SDS-PAGE gels of ZntRC79S
(A) induction test (B) eluants of gel filtration column

16162.21
[ZntRC79S]$^+$

12361.00
[Cytochrome c]$^+$
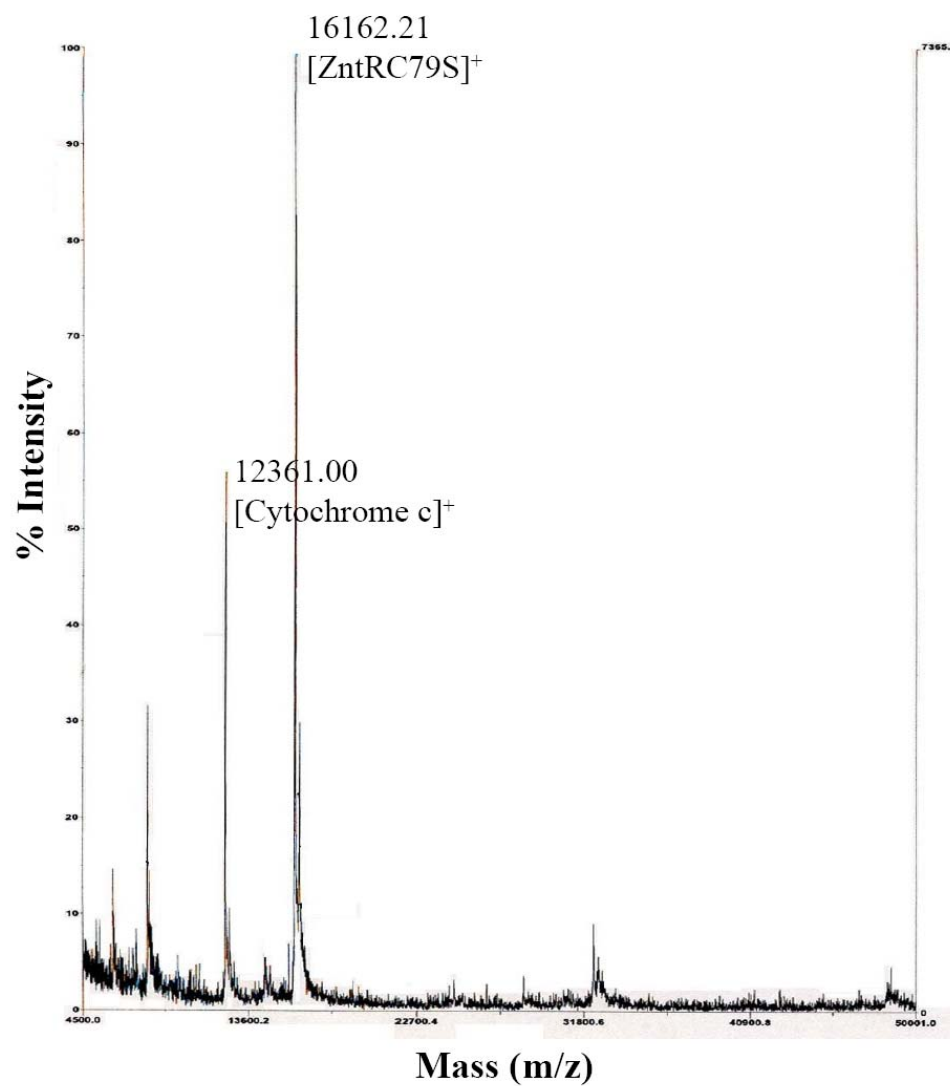
Figure 2-5. ZntRC79S MALDI-TOF spectrum
1~2 μg protein sample and ~2 μg cytochrome C (internal standard)
were used for the measurement.

zntA promoter: 5'– |TTGACT|CTGGAGTCGACTCCAGAGTG|TATCCT| – 3'
                      |AACTGA|GACCTCAGCTGAGGTCTCAC|ATAGGA|
                        -35   -10

18 mer: 5' – GATACACTCTGGAGTCGA – 3'
13 mer: 5' – CTCCAGAGTGTAT – 3'

P_**zntA**-18/13 mer: 5'– GATACACTCTGGAGTCGACTCCAGAGTGTAT – 3'
                         TATGTGAGACCTCAGCTGAGGTCTCACATAG

copA promoter: 5'– |TTGACC|TTCCCCTTGCTGGAAGGTT|TAACCT| – 3'
                      |AACTGG|AAGGGGAACGACCTTCCAA|ATTGGA|
                        -35   -10

18 mer: 5' – GTTAAACCTTCCCCTTAG – 3'
12 mer: 5' – GGGAAGGTTTAA – 3'

P_**copA**-18/12 mer: 5'– GTTAAACCTTCCCCTTAGGGGAAGGTTTAA – 3'
                         AATTTGGAAGGGGATTCCCCTTCCAAATTG

Figure 2-6. Design of crystallization DNA. Arrow: palindromic sequence; Dot: plaindromic symmetry center; Triangle: nicks in DNA; double underline: identical sequences in designed DNA as in wild type promoters
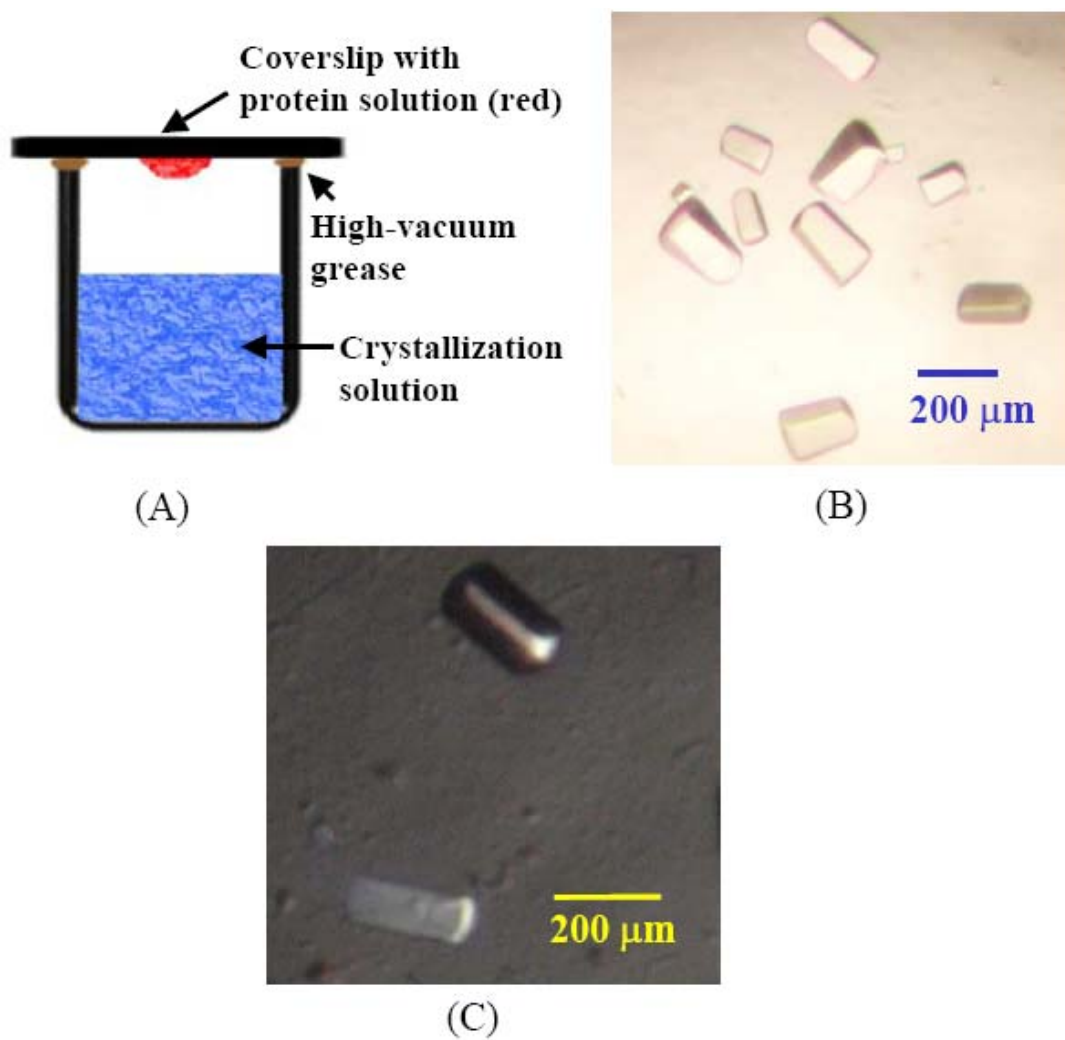
Figure 2-7. Crystallization experiments
(A) Hanging drop vapor diffusion method (B) Ag(I)-CueR crystals
(C) Zn(II)-ZntR crystals

Table 2-1. Data collection and refinement Statistics

	Pt-ZntR	Native ZntR	Au-CueR	Cu-CueR	Ag-CueR
Data collection					
Detector type/source	MAR-CCD/ APS	MAR-345/ SSRL	MAR-CCD/ APS	MAR-CCD/ APS	MAR-CCD/ APS
Wavelength (Å)	1.068	0.97	1.0358	1.000	1.000
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Resolution (Å)	2.7	1.9	2.5	2.2	2.07
Measured reflections	46,624	50,994	76,958	82,696	107,567
Unique reflections	6821	18,897	11,501	15,854	20,248
Completeness (%)[a]	98.0 (99.7)	99.2 (99.0)	98.3 (96.1)	95.8 (97.2)	99.3 (91.9)
R_{sym} (%)[a,b]	4.0 (11.8)	6.6 (28.5)	4.8 (21.4)	5.5 (25.4)	4.2 (23.0)
R_{meas} (%)[a,c]	4.7 (13.8)	8.1 (35.4)	5.7 (28.7)	6.6 (31.2)	5.0 (29.5)
Redundancy	6.8 (6.7)	2.7 (2.7)	6.7 (4.0)	5.2 (4.7)	5.3 (3.8)
$I/\sigma(I)$	12 (6.2)	8.4 (2.6)	12.6 (3.2)	10.9 (3.0)	10.8 (3.1)
Phasing					
Phasing power (acentric)[d]	1.23		1.82		
R_{cullis} (acentric)[e]	0.845		0.737		
Figure of merit	0.37		0.35		
Figure of merit after density modification	0.93		0.77		
Refinement					
Resolution (Å)		29-1.9	27-2.5	24-2.2	24-2.071
Number of reflections:					
working set/test set		917/17,949	780/10,067	1923/26,900	1382/18835
R-factor[f]		18.3	21.0	21.2	21.9
R_{free}[g]		21.4	25.3	25.0	25.9
Protein atoms		1455	1974	1978	1974
Water molecules		236	64	150	135
Other atoms		15	2	2	2
R.m.s.d.					
Bond lengths (Å)		0.004	0.007	0.006	0.005
Bond angles (°)		1.2	1.1	1.1	1.0
Average B-factor (Å^2):					
Main chain		20	43	34	34
Side chain		24	46	38	39
Solvent		36	50	48	44

[a]Numbers in parenthesis represent values in the highest resolution shell.

[b]$R_{sym}=\Sigma|I-\langle I\rangle|/\Sigma I$, where I=observed intensity, and $\langle I\rangle$=average intensity obtained from multiple measurements.

[c]R_{meas} as defined by (9)

[d]Phasing power=r.m.s. ($|Fh|/E$), where $|Fh|$=heavy atom structure factor amplitude and E=residual lack of closure error.

[e]$R_{cullis}=\Sigma||Fh(obs)|-|Fh(calc)||/\Sigma|Fh(obs)|$, where $|Fh(obs)|$=observed heavy atom structure factor amplitude and $|Fh(calc)|$=calculated heavy atom structure factor amplitude.

[f]R-factor=$\Sigma||Fo|-|Fc||/\Sigma|Fo|$, where $|Fo|$=observed structure factor amplitude and $|Fc|$=calculated structure factor amplitude.

[g]R_{free}: R-factor-based on 7% of the data excluded from refinement

2.3 Results and Discussions

Overall structure

The structures of Cu^+-, Ag^+-, and Au^+- bound forms of CueR were determined to 2.2, 2.1, and 2.5Å resolution, respectively, whereas the structure of an N-terminally truncated fragment of ZntR bound to Zn^{2+}- was solved to 1.9 Å resolution. The overall structure of the CueR dimer is identical in all three metal-bound states. Each monomer can be divided into three distinct functional domains: a dimerization domain flanked by a DNA-binding domain and a metal-binding domain (Figure 2-8). The N-terminal DNA-binding domain belongs to the winged helix-turn-helix family [46, 47], which consists of two helix-turn-helix motifs and antiparallel β sheet. The second helix-turn helix motif of the DNA-binding domain is followed by a five-residue loop connecting to a 10-turn α helix. This long α helix links the DNA-binding domain to the metal binding domain and contributes to the bulk of the dimerization interface by forming an antiparallel coiled coil with the equivalent helix of the other monomer. The structure of the CueR DNA-binding and dimerization domains is characteristic of the MerR family of proteins and shares the same topology as the structures of two other MerR-family transcriptional activators, BmrR [48] and MtaN [49, 50], which respond to the presence of organic substrates. The topology arrangement of CueR is α1-α2-β1-β2-α3-α4-α5-α6. The overall structure of the truncated ZntR dimer is similar to that of CueR and consists of one helix-turn-helix motif of the DNA-binding domain, the dimerization helix,and an intact metal-binding domain (Figure 2-8).

DNA binding domain

The DNA binding domain of CueR is stabilized by a hydrophobic core which consists of side chains from α1 (Ile-3, Val-6, Ala-7), α2 (Ile-17, Tyr-20, Glu-21), β2 (Arg-37, Tyr-39),

α3(Leu-44, Leu-47, Leu-50, Ala-53, Val-56), α4 (Leu-60, Ser-63, Leu-66, Val-67), and turns

between α1 and α2 (Leu-12), α2 and β1 (Lys-24, Leu-26, Val-27, Pro-29), α3 and α4 (Phe-58).

Inspecting the structures of MtaN/DNA and BmrR/DNA/TPP, two homologues of MerR

family drug-responsive members bound to their promoters, it is found that the protein/DNA

interactions are throughout the whole bound sequences of the DNA. On each half site, in

addition to the specific recognition of the major groove and minor groove, the non-specific

binding interactions concentrate in two areas: one is near the center area (C region), the other is

near the terminus (T region) (Figure 2-9). The major groove recognition helix $\alpha2$ coordinates

with $\alpha4$ and loop to interact with the C region and T region, respectively (Figure 2-9B). R17 and

Y38 are involved the major groove and minor groove recognition, respectively. In particular, the

minor groove interaction is crucial for protein/DNA binding, which will be addressed in details

in chapter 3. The protein/DNA recognition pattern of MtaN/DNA is similar to that of

BmrR/DNA, and presumably applies to CueR/DNA as well.

Sequence alignments and structure data reveal that the protein residues involved in T

region binding are more conserved from both sequence and structural aspects. Hence the

following discussions are focusing on the T region based on the structural data of protein and

protein/DNA models in the active form.

For clarity, the several important residues are labeled by numbers in the sequence

alignment diagram and are highlighted with stick representation in the cartoon model (Figure 2-

10). Residues **1**(Tyr) and **2**(Arg) are both located in the loop. **1**(Tyr) interacts specifically with

the minor groove and the detail is addressed in chapter 3. **2**(Arg) is next to **1**(Tyr) in sequence

and contacts with two consecutive backbone phosphates (Figure 2-11). Residues **3**(His/Arg),

4(Tyr) and **5**(Asp/Glu) are all located on the major groove recognition helix α2. **3**(His/Arg) forms H-bonds with **2**(Arg) in absence of DNA, while in presence of DNA, it reorient to make contacts with the DNA backbone phosphates (Figure 2-11); **4**(Tyr) interacts with the DNA backbone phosphates in the C region, which could help to break the central basepairs (Figure 2-10); **5**(Asp/Glu) interacts with **2**(Arg) through both H-bonds and electrostatic forces in both absence and presence of DNA (Figure 2-11). It is located at the C-terminus of the helix and next to **4**(Tyr) in sequence. Hence, **2**(Arg), **3**(His/Arg) and **5**(Asp/Glu) form a small network which could communicate or coordinate the minor groove and major groove interactions to accomplish the gene regulation function. In particular, **2**(Arg) not only contacts with the minor groove DNA, it could also affect the positioning of the major groove recognition helix via interaction with **5**(Asp/Glu) and **3**(His/Arg), and as a matter fact, it is absolutely conserved among all MerR family members, whereas **3**(His/Arg) and **5**(Asp/Glu) both have exceptions, suggesting that two other components of the network might be substituted by other nearby residues. It is tempting to testify the idea of the small communicating network by mutating the **2**(Arg) to Lys to break the connection between **2**(Arg) and **5**(Asp/Glu), or by mutating to Ala to further eliminate the DNA binding as well. On the other hand, to avoid disturbing the protein/DNA interface, **5**(Asp/Glu) can be targeted for mutation to break the communication. Furthermore, **2**(Arg) has two different conformations to form the small network, in MtaN, it exhibits a more extended conformation and uses NE and NH2 to contact with DNA, which might facilitate the interaction with **3**(His) with a shorter sidechain in absence of DNA; whereas, in BmrR and CueR, it adopts a different conformation which uses NH1 and NH2 to contact with DNA, and contact with **3**(Arg) using NH1 in absence of DNA.

Previous biophysical studies [33, 51] suggested the gene activation is through both bending and untwisting of the promoter. In the absence of the apo-protein/DNA model, we speculate that, structurally it is realized by the re-adjusting the relative positioning of the two DNA binding domains upon activator binding. However, the exact reconfiguring is still unknown and need further investigation.

Dimerization Domain

The dimers are stabilized primarily by the formation of an antiparallel coiled-coil. Coiled-coils are characterized by the heptad repeat $(abcdefg)_n$ in which the *a* and *d* positions are typically occupied by hydrophobic core of the interface helices [52]. In CueR (Figure 2-6), the hydrophobic core of the interface consists of the side chains of residues Val-88 (*a*2), Ile-95 (*a*3), Ala-109 (*a*5), Ile-91 (*d*2), Leu-98 (*d*3) and Leu-105 (*d*4), and the methylene carbons of Lys-81 (*a*1), Arg-102 (*a*4) and Thr-84 (*d*1). In ZntR (Figure 2-12), the hydrophobic core of the interface consists of the side chains of residues Leu-90 (*a*2), Ile-97 (*a*3), Val-86 (*d*1), Val-93 (*d*2), Leu-100 (*d*3) and Leu-107 (*d*4), and the methylene carbons of Lys-83 (*a*1), Gln-104 (*a*4) and Asn-111 (*a*5). In addition to forming the antiparallel coiled-coil, the helix α3 also interacts with the C-terminal end of α5 and the helix α6, the loop region between α5 and α6 interacts with the N-terminus of α5.

Metal-binding domain

In CueR structure, the metal ion is buried in a solvent-inaccessible site in a loop at the dimer interface and has only two coordinating ligands: the S-atoms of conserved Cys112 and Cys120 (Figure 2-13). These residues define the end points of a 10-residue loop that extends from the C-terminal end of the dimerization helix up to the N-terminus of a short two-turn α-

helix. In the dimer, only one metal-binding loop is fully ordered, even though another metal is bound at the equivalent site in the dimer. Residues 115 to 119 of the second monomer are disordered, perhaps due to differences in crystal packing. This disorder suggests that flexibility in the extended region of the metal-binding loop may allow metal access to the buried binding site. The metal-binding loop of one monomer rests against the N-terminal end of the long dimerization helix and the preceding loop of the other monomer, interacting primarily through backbone contacts, interacting primarily through backbone contacts. The short two-turn α-helix that extends from the metal-binding loop packs against the DNA-binding domain of the other monomer and against both dimerization helices. These extensive hydrophobic interactions effectively form a scaffold that stabilizes the buried metal-binding site.

Coordinate-covalent bonds between the Cu^+ ion and the two sulfur atoms of Cys 112 and Cys 120 exhibit Cu-S bond distances of 2.13 Å with an essentially linear S-Cu-S bond angle of 176 º. Extended x-ray absorption fine structure studies on copper-CueR complexes confirm this coordination environment and demonstrate that the copper is unequivocally in the +1 oxidation state [53]. The metal binding domain in Ag^+ and Au^+ is identical to that found in Cu^+-CueR, except for longer metal-sulfur distances (Ag-S: 2.35 Å, Au-S: 2.32 and 2.39 Å). All sulfur-to-metal distances and bond angles (176º and 177º) are consistent with the average values for linear, two-coordinate Cu^+, Ag^+, and Au^+ complexes found in the small-molecule Cambridge Structural Database.

Mutation of Cys112 or Cys120 in CueR abrogates all response to Cu^+, Ag^+, and Au^+ in vivo [54] and in vitro [53], and the structure confirms that the surrounding environment does not allow for other types of interactions with the metal. The closest residue with metal-binding

potential is Ser77 from the second monomer, which has a Cu-Oγ distance (4.4 Å) that is too long for primary or secondary covalent-bonding interactions. Instead, Ser77 is within hydrogen-bonding distance to Asp115 and several main-chain atoms of the metal-binding loop. These interactions may be important in maintaining the conformation of the metal-binding loop and stabilizing the quaternary interactions in the dimer interface.

In ZntR, both metal-binding domains are well ordered with two Zn^{2+} ions bound at each site through Cys114, Cys115, His119, Cys124, and Cys79, which is supported by biochemical data showing a diminished response to Zn^{2+} upon mutation of any of these residues [54]. The dinuclear Zn-binding site shows each Zn^{2+} ion in a tetrahedral coordination environment with a Zn-Zn distance of 3.6 Å. Cys114 and Cys124 (equivalent to Cys112 and Cys120 in CueR) serve as ligands to Zn1, whereas Cys115 and His119 coordinate Zn2. Cys79 (equivalent to Ser77 in CueR) from the other monomer acts as a bridging ligand to the two Zn^{2+} ions, thus linking the metal-binding domain of one monomer and the dimerization domain of the other monomer. An oxygen atom of a bridging phosphate or sulfate group acts as a fourth ligand to each Zn^{2+} ion. In contrast to the CueR metal-binding site, the coordination environment of the metal-binding loop in ZntR is optimal for binding of divalent metals.

ZntRC79S does not bind Zn(II) in vitro

The structural data shows the bridging ligand Cys79 in ZntR corresponds to Ser77 in CueR. To evaluate the importance of amino acid in metal selectivity, we mutated Cys79 to Ser79 in ZntR to investigate the metal binding and metal responsiveness.

Both purified ZntR and ZntRC79S contain no Zn and Ag (less than 0.1%). ZntR binds ~0.5 Zn/monomer and ~0.5 Ag/monomer in the buffer of 50 mM Tris, 250 mM NaCl, pH 8 and

1~2 mM DTT. Whereas ZntRC79S does not bind Zn(II), instead it binds ~3 Ag/monomer.

Purified CueR contains no Zn and Ag, CueR does not bind Zn(II), CueR binds ~4

Ag(I)/monomer. Collectively, mutation of Cys79 to Ser79 in ZntR totally abolishes its Zn

binding ability, suggesting Cys79 is crucial for +2 metal ion binding. Additionally, both ZntR

and ZntRC79S binds +1 ion, except that ZntRC79S binds more metals and behaves more like

CueR.

In vitro run-off transcription of ZntR and ZntRC79S

In vitro run-off transcription assay using the P_{zntA} template is shown in Figure 2-14. All

reactions contain 20 μM TPEN. Lane 1 shows the 100 bp marker. RNAP alone (lane 2) produces

basal level of the *zntA* transcripts. Both ZntR (lane 3) and ZntRC79S (lane 7) act as repressor and

inhibit basal level of *zntA* transcription. ZntR activates the transcription in the presence of 70 μM

Zn(II) (lane 6); ZntRC79S shows no response to added Zn(II) and still represses *zntA*

transcription (lane 10). ZntR also activates transcription in the presence of Ag(I) and the

transcription level is similar to that of Zn(II) (lane 4 and 5), presumably Cu(I) would behave in

the similar way as Ag(I); ZntRC79S is responsive to Ag(I), but with a much lower transcription

level compared with wild type ZntR (lane 8 and 9). When ZntR is titrated with 10-200 μM of

Ag(I), transcription levels do not show substantial variations. The *in vitro* assay reveals that

mutation of Cys79 to Ser79 of ZntR has subtle but important effect on the protein functions. The

mutated protein can still activate transcription in response to Ag(I) but has lost its Zn

responsiveness.

Dimerization helix

Metal-binding domain

α5

α6

C

α3

α4

N

α1

β1

β2

α2

DNA-binding domain

(A) CueR

C

C

N

N
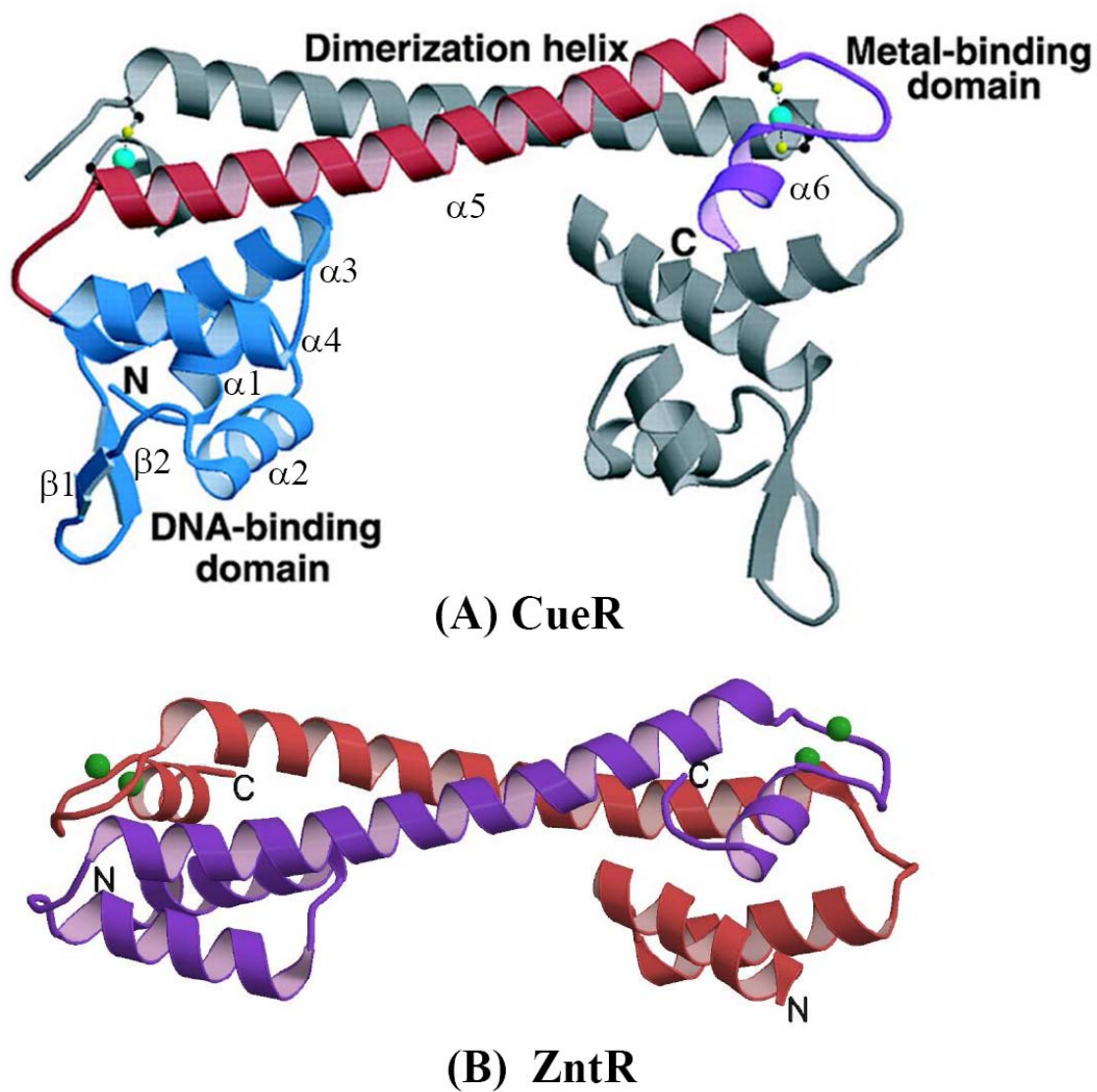
(B) ZntR

Figure 2-8. Overall structures of (A) CueR and (B) ZntR
The metal ions are shown as cyan spheres in CueR and green spheres in ZntR

Figure 2-9. MtaN-DNA recognition pattern

(A) MtaN-DNA binding diagram. The figure is adapted from reference 49. Blue arrows, side chain-DNA hydrogen bonds; green arrows, backbone amide-DNA contacts; yellow arrows, *van der Waals* contacts. Two areas with extensive non-specific interactions are enclosed in red box and labeled as 'T' for terminal and 'C' for center, respectively. R17 and Y38 are involved in the major groove and minor groove specific recognition, respectively.

(B) Structure of MtaN-DNA binding. It shows the positions of the interaction areas and labeled in the similar way as in (A). $\alpha2$ is the major groove recognition helix, it collaborates with $\alpha4$ and loop to interact with C and T regions, respectively.

This protein/DNA binding pattern is similar to that of BmrR/DNA, and mostly likely applies to CueR/DNA as well.

58
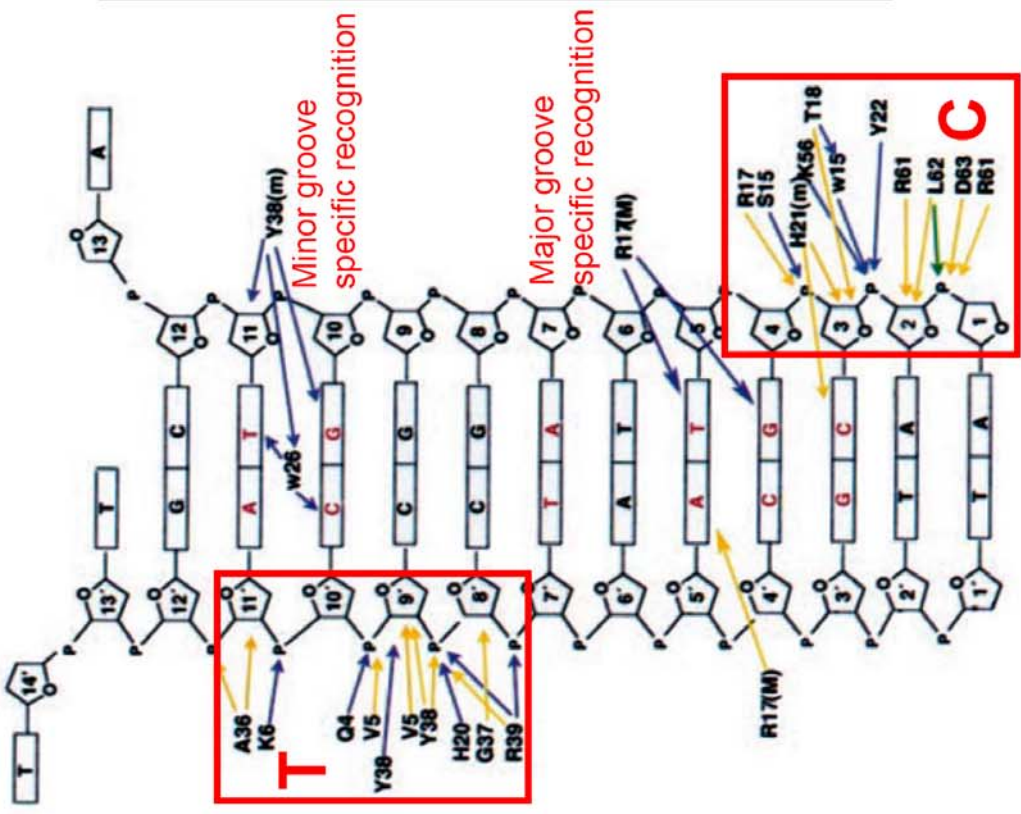
(B) Structure of MtaN-DNA binding
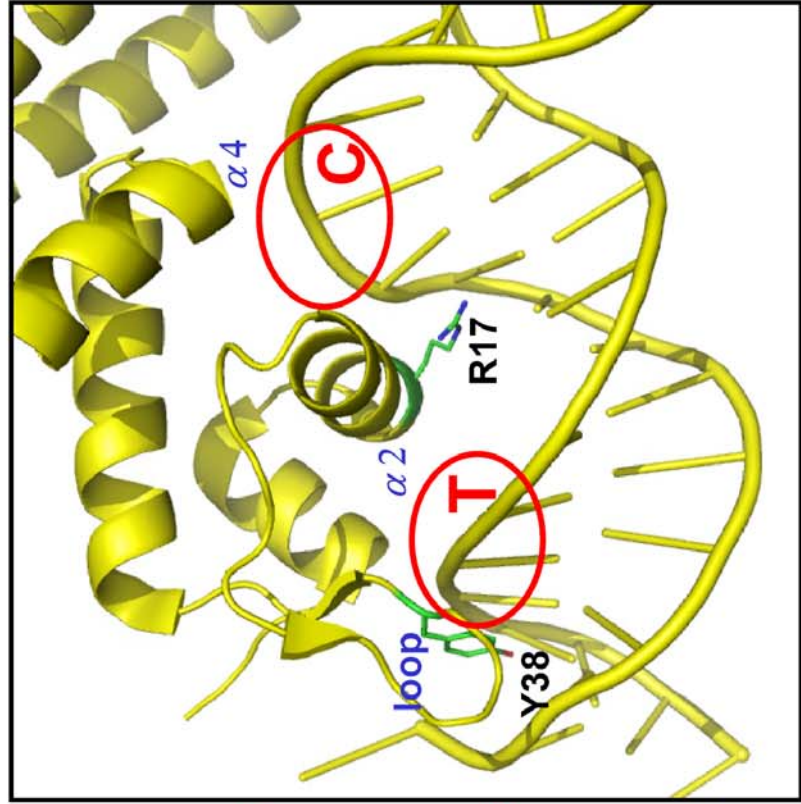
(A) MtaN-DNA binding diagram

Figure 2-10. Several important residues involved in interactions with T region of DNA

(A) DNA-binding domain sequence alignment. The figure is adapted from reference 49. The residues are labeled with red number under it. 1 and 2 are in the loop involved in the minor groove interactions. 3, 4 and 5 are on the major groove recognition helix.

(B) Cartoon diagram MtaN/DNA structure. The residues are highlighted with stick representation. 1 and 2 interacts with the minor groove; 3 and 4 interacts with the major groove; 5 interacts with 2 and thus provide the communication between the major groove and minor groove interactions.
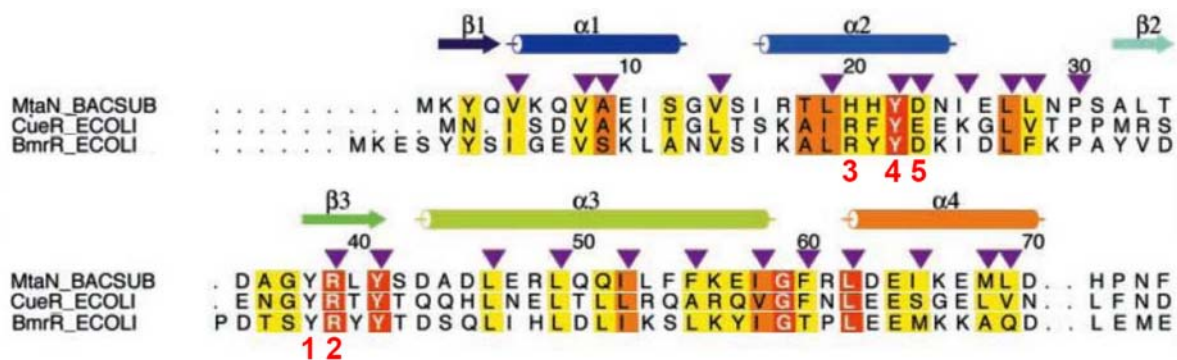
(A) DNA-binding domain sequence alignment
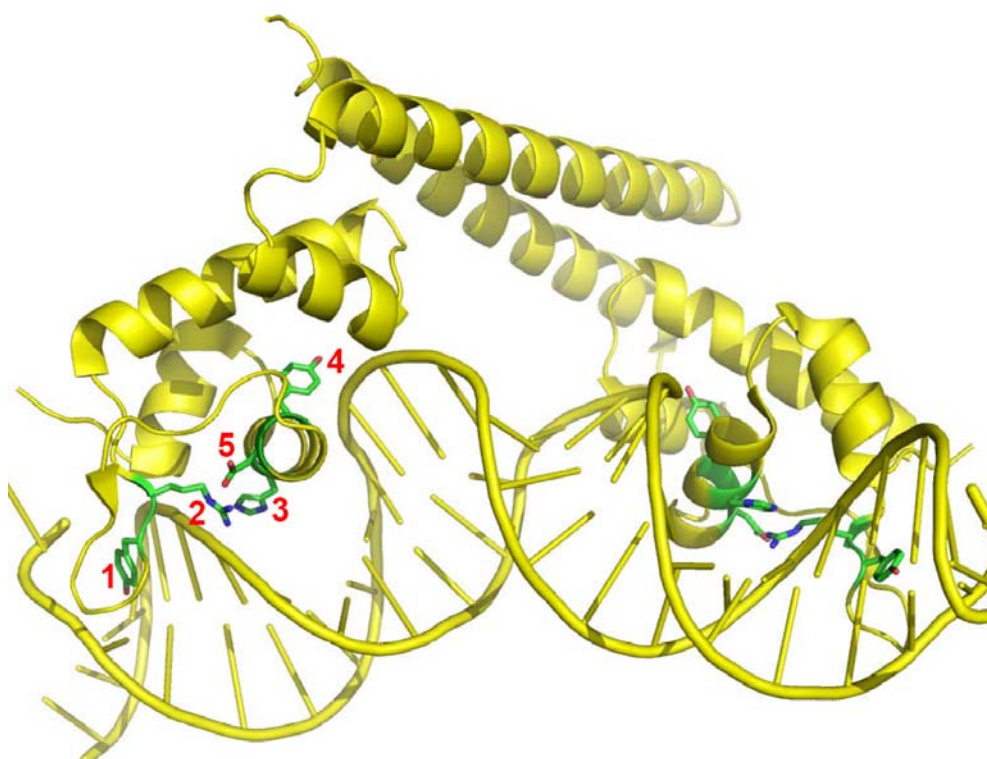
(B) Cartoon diagram MtaN/DNA structure

Figure 2-11. Close-up view of the 2, 3 and 5 network in absence and presence of DNA

(A)   MtaN and MtaN/DNA. In absence of DNA, residues 2,3 and 5 form a H-bond network to stabilize the tertiary structure. 2(Arg) adopts a more extended conformation to contact with 3(His). In presence of DNA, residue 3(His) reorients to contact with DNA backbone, and the interaction between 2(Arg) and 3(His) weakens. 2(Arg) uses NE and NH2 to contact with DNA backbone. The interactions between 2(Arg) and 5(Asp) are similar in both states.

(B)   Cu-CueR and BmrR/DNA/TPP. In absence of DNA, residues 2,3 and 5 form a H-bond network to stabilize the tertiary structure. 2(Arg) exhibits a less extended conformation using NH1 to interact with 3(Arg). In presence of DNA, residue 3 (Arg) reorients to contact with DNA backbone. The H-bond interaction between 2 and 3 loses. 2 (Arg) uses NH1 and NH2 to contact with DNA backbone. The interactions between 2(Arg) and 5(Asp/Glu) are similar in both states.
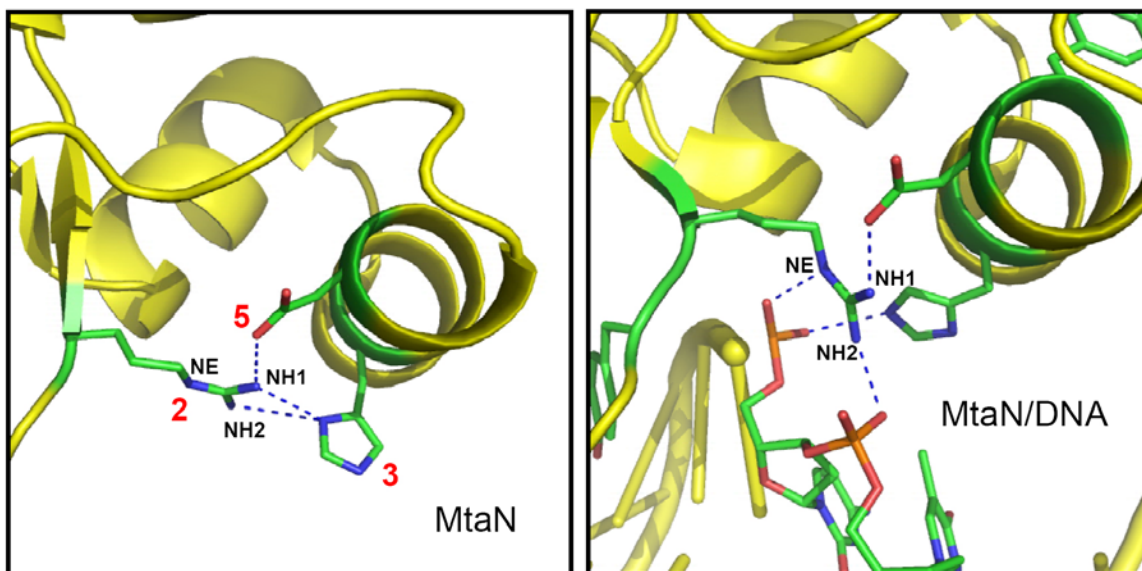
(A) MtaN and MtaN/DNA

(B) Cu-CueR and BmrR/DNA/TPP

L106
Q99
E92
L85

D103
E96
A89
R82

e'

b'

a'

A109
R102
I95
V88
K81

A107
S100
R93
E86

f'

d'

L105
L98
I91
T84

c'

g'

Q104
E97
E90
R83

L108
M101
H94
K87

L85
E92
Q99
L106

K81
V88
I95
R102
A109

e

a

b

R82
A89
E96
D103

T84
I91
L98
L105

d

f

E86
R93
S100
A107

g

c

K87
H94
M101
L108

R83
E90
E97
Q104

(A) CueR

Q108
Q101
E94
Q87

R105
A98
Q91
G84

e'

b'

a'

N111
Q104
I97
L90
K83

R109
S102
A95
E88

f'

d'

L107 V86
L100 V93
V93 L100
V86 L107

c'

g'

S106
E99
E92
I85

L110
M103
R96
R89

Q87
E94
Q101
Q108

K83
L90
I97
Q104
N111

e

a

b

G84
Q91
A98
R105

d

f

E88
A95
S102
R109

g

c

R89
R96
M103
L110

I85
E92
E99
S106

(B) ZntR

Figure 2-12. Schematic diagram of the antiparallel coiled-coil of (A) CueR (B) ZntR. Residues at the *a* and *d* positions form the hydrophobic core of the interacting helices.
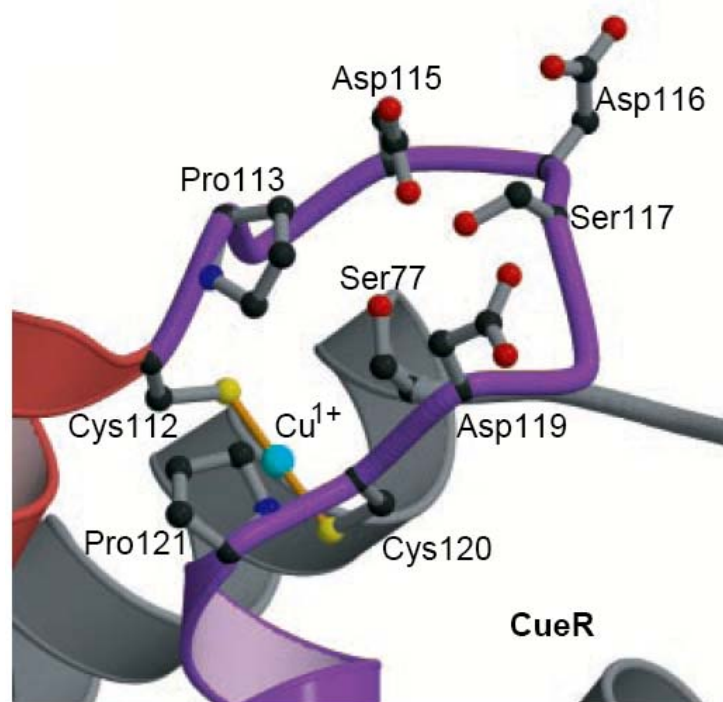
(A) CueR
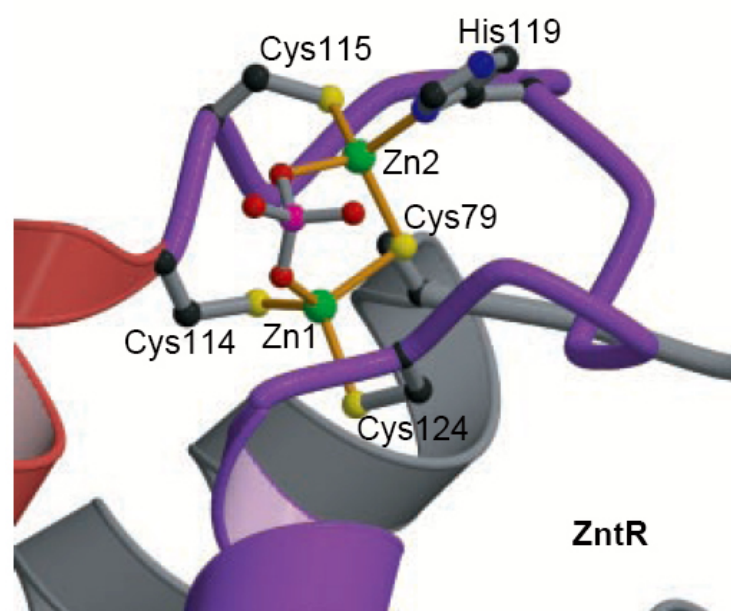
(B) ZntR

Figure 2-13. Metal binding site of (A) CueR and (B) ZntR

Protein	-	WT	WT	WT	WT	WT	M	M	M	M	WT	WT	WT	WT	WT
Zn or Ag	-	-	Ag	Ag	Zn	-	Ag	Ag	Zn	Ag	Ag	Ag	Ag	Ag	Ag
Conc. (μM)			50	100	70		50	100	70	10	20	50	100	200

72 bp *zntA* transcripts

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Figure 2-14. Run-off transcription assay of ZntR and ZntRC79S. WT and M represents wild type ZntR and mutant ZntRC79S, respectively. Protein concentration is 25 nM.

Discussion

The crystal structures of CueR and ZntR reveal several determinants of metal-ion selectivity, which have been discussed in some detail in our published paper. However, here I focus on several additional facts and some new data.

A residue at the N-terminus of the dimerization helix extends across the dimer interface to contact either the metal-binding loop (Ser77 in CueR) or directly to the metal ion (Cys79 in ZntR) and plays a unique role in discriminating between +1 and +2 ions (Figure 2-13). A serine is found at this position in all MerR homologs responsive to +1 ions, whereas a cysteine is present in homologs that are known to respond to +2 ions (Figure 2-16). Mutating the Cys79 to Ser in ZntR causes the lost of Zn responsiveness, which suggests that the tetrahedral metal coordination environment could only be implemented with the conserved Cys79. Meanwhile, both wild-type and mutated ZntR proteins bind Ag(I), however, the coordination environments of Ag(I) in wild-type and mutated proteins are probably different, as reflected in their different behaviors for transcription activation. The fact that Ag-ZntR could activate the transcription at a similar level as Zn-ZntR, while Ag-ZntRC79S exhibits much lower activity comparable to basal level transcription, indicates the Cys79 is a key factor for the gene regulation function of the protein. It will be very interesting to further investigate the binding of +1 ions to ZntR and compare with the physiological activator Zn to fully understand coordination chemistry and derive the important common features crucial for the gene activation.

Structurally, the first level of discrimination is the number of coordinate covalent bond formed because +2 metal ions typically prefer higher coordination numbers than +1 ions of the coinage metals; Second, CueR restricts the bound metal to a low coordination number by

hydrophobic and steric restrictions, which contribute to a shielded coordination environment.

Third, electrostatic features of the metal receptor cavity contribute additionally to the metal affinity and selectivity. The +2 ions that require higher coordinate numbers (i.e., Zn^{2+}, Co^{2+}) thus face significant energy penalties. This is important in discriminating against Hg^{2+}, which, aside from the +2 class, shows many similarities and chemistry of Cu^+ including the propensity to low coordination numbers. In CueR, the buried S-Cu-S center formally has a net negative charge arising from two thiolate anions and one +1 charged metal. The structure reveals a series of weak interactions that compensate this buried charge and provide an electrostatic component to the metal-binding free energy beyond the Cu-S bond energy (Figure 2-15). Charge neutralization of the thiolate of Cys120 likely arises from interactions with the positively charged end of the helix dipole and two backbone H bonds that originate from the short two-turn α helix extending from the metal-binding loop. In CueR, Pro121 constrains the S atom of Cys120 such that it is centered directly over the N-terminus of the short helix (Figure 2-15C). A Cys-Pro motif at the N-terminus of an α helix allows for favorable interactions of the cysteine with the helix dipole [55] (Figure 2-15C), and such a motif is conserved at the end of the metal-binding loop in many of the MerR family members. Mutation of this proline diminishes the Hg^{2+} response of MerR, consistent with a key role for the Cys-Pro motif [56]. Cys112 accepts only one hydrogen bond (Figure 2-15B) and the closest positively charged residue is a conserved lysine, Lys81 (the N-S distance is 5.5 Å). The charge-charge interaction of Lys81 with the Cys112 thiolate, although distant, could contribute additional charge neutralization. These electrostatic- and hydrogen-bonding interactions lead to charge neutrality when a +1 ion, but not a +2 ion, binds. This allows CueR to discriminate against Hg^{2+}-binding in what is otherwise a stable linear dithiolate Hg-

coordination environment. The spacefilling model of CueR (Figure 2-15A) reveals the solvent

inaccessibility of the bound metal, which helps to explain the zeptomolar sensitivity of CueR to

$[Cu(I)]_{free}$ reported by calibration with strong copper binding agent CN^-.

A structure-based sequence alignment of MerR metalloregulatory homologs allows for

predictions of metal selectivity in other family members (Figure 2-14). A serine or cysteine at the

N-terminal end of the dimerization helix provides a clear distinction between putative +1 and +2

metal-responsive subgroups within the MerR family. In the case of PmtR, a role in Zn^{2+}

homeostasis has been proposed [57]; however, the structure-based alignment (Figure 2-16) suggests

a role in Cu^+ regulation. For SoxR, a MerR homolog that uses an Fe-S cluster to sense oxidative

stress [58], the alignment suggests that the 2Fe-2S cluster will bind in a loop resembling the ZntR

binuclear Zn^{2+} site. Furthermore, the ZntR structure reveals how the prototypical

metalloregulatory protein in this family, MerR [59], could bind Hg^{2+} using a trigonal metal-thiolate

coordination environment [60] involving two cysteines from one monomer and a third cysteine, the

equivalent to Cys79 in ZntR, from the other monomer [56, 61, 62]. The alignment shows proteins

responding to +1 ions share a unique signature SerXXVal(Lys/Arg) (where X is any amino acid)

at the N-terminal region of the dimerization helix, which begins at Ser77 in CueR; A

CysXGlyX$_4$AspCys metal-binding loop followed immediately by a Pro, which is the N1 residue

of the next helix. Alternative signatures are found for members responding to +2 ions.

$$CueR: \quad Ser \; X \; X \; Val \; (Lys/Arg) \; \text{———} \; Cys \; X \; Gly \; X_4 \; Asp \; Cys \; Pro$$
$$\qquad\qquad \mathbf{77} \qquad\qquad\qquad\qquad\qquad\qquad \mathbf{120}$$

The Hg^{2+}-sensing proteins are predicted to have a Cys at the N-terminus of the dimerization

helix and a CysX$_8$Cys loop also followed by a Pro as the first residue in a helix.

MerR : *Cys* ———————————————— *Cys X$_8$ Cys Pro*
       82   **117**

The Zn^{2+}-sensing proteins have a Cys at the N-terminus of the dimerization helix followed by a

$CysX_9Cys$ motif with no Pro following the last Cys.

ZntR : *Cys* ———————————————— *Cys X$_9$ Cys*
       79   **114**

Finally, SoxR appears to have a shorter dimerization helix followed by a $CysX_{10}Cys$ loop

containing no Pro residues, but an additional CysXCys motif characteristic of Fe-S binding sites

in other proteins.

SoxR : ———————————————— *Cys X$_2$ Cys X Cys X$_5$ Cys*
                            119

The combination of thermodynamic and structural studies reveals how metalloregulatory

proteins can select among +1 and +2 transition metal ions. The extraordinary copper sensitivity

of CueR suggests that the prokaryotic cytoplasm operates under conditions of copper

deprivation. Given the absence of known copper chaperones in *E. coli*, these results raise the

question of how bacterial copper-dependent proteins obtain their cofactor. All copper dependent

enzymes known to date in *E. coli* are found in the cell envelope [28, 63-67]. The answer may be that

available cytosolic copper ions, bound or free, are detected by Cu^+ sensors and rapidly

transported to the cell envelope for incorporation or ejection. In this respect, copper homeostasis

in Gram-negative bacteria shows a striking parallel to eukaryotic copper trafficking pathways.

Both use a family of homologous copper-specific P-type adenosine triphosphatases to clear

rapidly any available copper ions from the cytosol into more specialized compartments: the cell

envelope in prokaryotes and the trans-Golgi secretory pathway in eukaryotes [68].

(B)

(A)

(C)
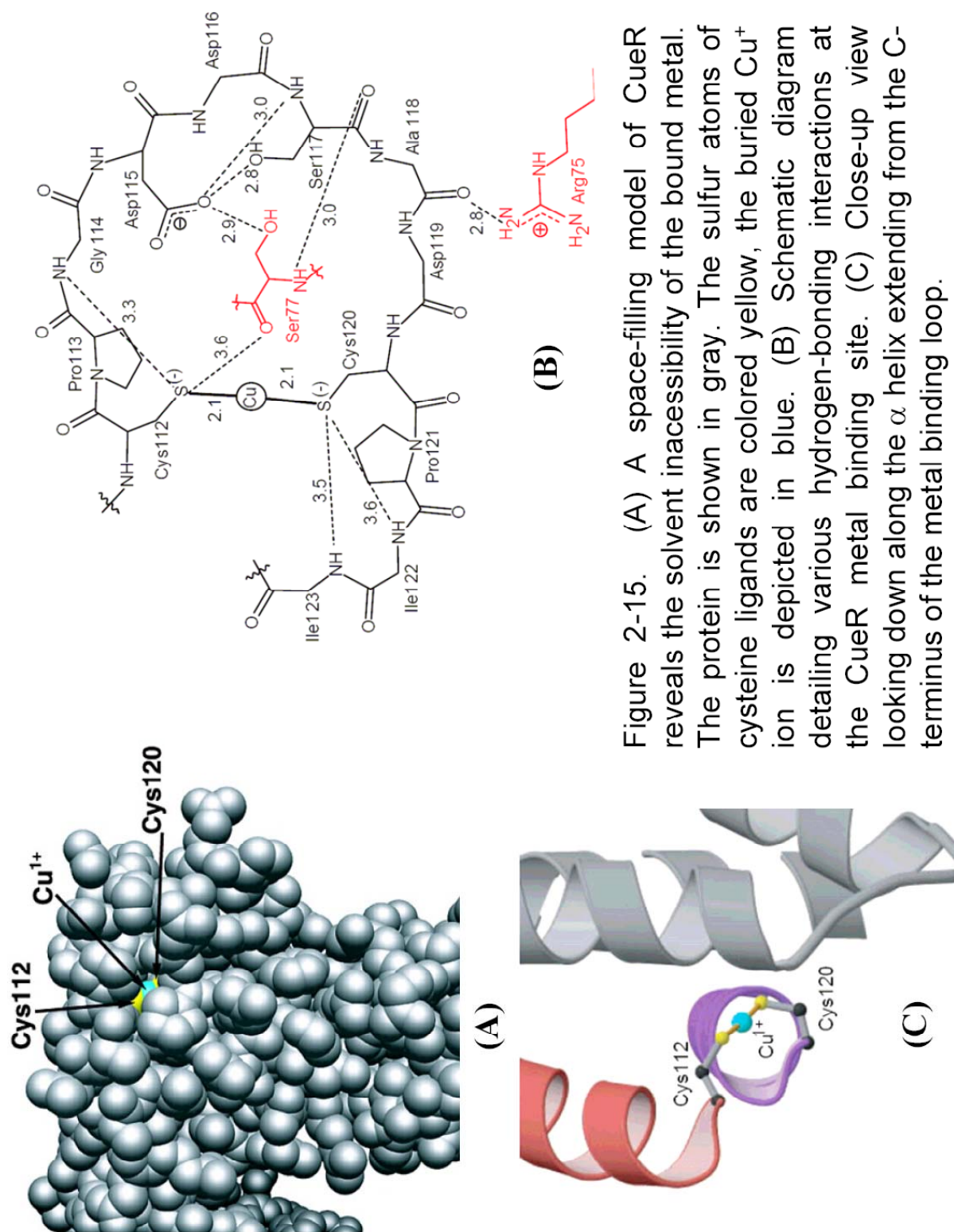
Figure 2-15. (A) A space-filling model of CueR reveals the solvent inaccessibility of the bound metal. The protein is shown in gray. The sulfur atoms of cysteine ligands are colored yellow, the buried Cu$^+$ ion is depicted in blue. (B) Schematic diagram detailing various hydrogen-bonding interactions at the CueR metal binding site. (C) Close-up view looking down along the α helix extending from the C-terminus of the metal binding loop.

Figure 2-16. Structure-based alignment of sequences corresponding to the dimerization and metal binding domains of various metalloregulatory MerR family members.

Residues that bind metal in the CueR and ZntR structures or are predicted to bind metal in other members are highlighted in yellow and marked by black arrowheads above each subgroup. Identical residues within each subgroup are highlighted in pink, and conserved residues are colored in blue.

Dimerization helix   Metal binding loop   2-turn α-helix

```
Cu,Ag,Au(I)
          CueR(E coli)          72 DPQ.RHSADVKRRTLEKVAEIERHIEELQSMRDQLLALANACPG.DDSA..DCPIIENLS.GCCHHRAG.
          CueR(P putida)        72 DRQ.RASADVKALARQHIDELNQKIRELGELRDTLQDLVEHCSG.DHRP..DCPILKELASGCCAQPARA
          CueR(Y pestis)        72 NPA.RHSADVKAATLLKVAEIEQHINDLNQMRMRLLALAAEECPG.DEGA..DCPIINSIA.GCCHSSSSL
          CueR(V cholerae)      72 DPN.RTSAAVRARAQEKWQEISRKLSEITMIKQQLEEWIASCPG.DQGS..DCPIIEQIKGHCCSNNKTK
          CueR(S typhi)         72 DPR.RHSADVKKRTLEKVAEIERHISELQSMRDQLLAMAESCPG.DDSA..DCPIIDNIS.GCCHHKAQK
          HmrR(S meliloti)      72 DRS.RASSDVKAVALEHIAELERKIAAIQDMTRTLKHLASHCHG.DGRP..DCPIIEEMAKGGGAAKTEI
          HmrR(R leguminosarum) 71 DKD.RASADVRDIAQTKLITEIDRKIRELTELRRTLEHIVHACHGND.RP..DCPILEELSDGA.......
          PmtR(P mirabilis)     72 NRE.RTSADVKAIALSHIDELNRKITQLQRMTQTLSHLAQECQG.DNNP..DCPIIAKLVEPQTGTEH..

Zn,Co,Cd,Pb(II)
          ZntR(E coli)          73 DPEHHTCQESKGIVQERLQEVEARIAELQSMQRSLQRLNDACCG.TAHSSVYCSILEALEQGASGVKSGC
          ZntR(S typhimurium)   73 DPEHHTCQESKSIVQERLQEVEARIAELQTMQRSLQRLNDACCG.TAHSSVYCSILEALEQGASGAKSGC
          ZntR(Y pestis)        73 DPDHHTCQESKSIVDSRLISDVEGKIRELERMRDSLKRISEACCG.TSHATTYCSILEALEQGATEELIKV
          ZntR(V cholerae)      74 EATEHSCAEVKAITSAKLAVIDQKIEELTRIRSALKKINDACCGHVEDNASHCSILAALE.........
          ZntR(S oneidensis)    73 DKSNWACADVKGMVDLKLAQVQAKIAELLHFQTSLQSLSNACCG.GPRSAEHCSILEALESSTERVRFEH
          ZccR(B pertussis)     70 DNHEAGCAPINQVFDEHIAHVDARIAELTQLKAQLGELRQRCASARPD.AEDCGILHGLSEMQVEERPER
          CadR(P putida)        71 DSPDDSCGGSVNALIIDEHIEHVQARIDGLVALQEQIVELRRRCN..AQGA.ECAILQQLETNGAVSVPET
          PbrR(R metallidurans) 74 R.PDQDCGEVNMLLDEHIRQVESRIGALLELKHHIVELREACSG..ARPAQSCGILQGLSDCVCDTRGTT

Hg(II)
          MerR(Tn501)           78 DGTH..CEEASSLAEHKLKDVREKMADLARMEAVLSELVCACHA..RRGNVSCPLIASLQGGASLAGSAM
          MerR(Tn21)            78 DGTH..CEEASSLAEHKLKDVREKMADLARMETVLSELVCACHA..RKGNVSCPLIASLQEAGLARSAM
          MerR(S marcescens)    78 DGTH..CEEASSLAEHKLQDVREKMTDLARMETVLSELVFACHA..RQGNVSCPLIASLQGEKEPRGADA
          MerR(S aureus)        73 DQDGERCKDMYAFTVQKTKEIERKVQGLLRIQRLEELKEKCP..DEKAMYTCPIIETIMGGPDK......
          MerR(B sp. RC607)     73 DRDEAKCRDMYDFTILKIEDIQRKIEDLKRIERMDLKERCP..ENKDIYECPIIETIMKK.......

2Fe2S center
          SoxR(E coli)          82 EGHTLSAKEWKQLSSQWREELDRRIHTIVALRDELDG.....CIGCGCLSRSDCPLRNPGDRLGEEGTGAR
          SoxR(S typhi)         82 EGHTLSAKEWKQLSSQWREELDRRIHTIVALRDELDG.....CIGCGCLSRSDCPLRNPGDRLGEHGTGAR
          SoxR(P aeruginosa)    80 AGRSPSAADWARLSAQWKEDLTERIDKLLILRDQLDG.....CIGCGCLSLQACPLRNPGDQLSAEGPGAH
```

CHAPTER 3

Biophysical Studies of Interaction between CueR and Its Responsive Promoter

<u>Abstract</u>

CueR is one of the MerR family metalloregulatory members and controls the transcription of the copper detoxification and export genes in response to copper stress. Structures of metal bound forms of CueR reveal the structural basis of the metal selectivity and sensitivity (Chapter 2). Here, we focus on the DNA binding domain and characterize the protein-DNA interactions with various biophysical methods. Different assays show that the CueR-DNA binding affinity is in μM range. The specific minor groove interaction is extensively addressed with mutagenesis and gel shift assay, the experimental data reveal that the specific minor groove contact is crucial for protein/DNA recognition. Furthermore, the predominant contribution of the residue Tyr36 to the contact is from the aromatic ring instead of the previously assumed hydroxyl group. Finally, one basepair right upstream of the -10 element is shown to be important for the binding and may have some implications related with the open complex formation.

3.1 Introduction

Copper is an essential trace metal element for both eukaryotes and prokaryotes, and is associated with various metal enzymes, which are involved in electron transport, reduction of nitrite and nitrous oxides, and electron carriers. At the same time, the intracellular copper must be maintained below a toxic threshold.

E. coli uses numerous copper homeostasis pathways to control a strict cellular copper quota within a narrow range, about 10^4 atoms per cell (~10 μM) [1, 2]. In the presence of copper stress, CueR, a MerR-family transcription factor, stimulates copper-induced transcription of both *copA* encoding Cu(I)-translocating P-type ATPase pump to export excess copper [3, 4], and *cueO* encoding a periplasmic multicopper oxidase for detoxification [3, 5]. As described in chapter 2, CueR exhibits a zeptomolar (10^{-21} M) sensitivity to free Cu(I) [6], reminiscent of metal sensitivity in other MerR-family proteins such as Hg-MerR [7, 8] and Zn-ZntR [2, 9, 10]. The atomic model of Cu-CueR reveals a linear CuS2 geometry and demonstrate the molecular basis of the metal ion selectivity and ultrasensitivity [6].

The transcription activation mechanism of MerR family members has been studied intensively [8, 11] and proved by the crystal structures of BmrR-DNA-TPP [12] and MtaN-DNA complexes [13] : the center base pair of its pseudopalindromic promoter binding site was dramatically bent, undertwisted and unstacked thus shortening the distance between the -10 and -35 promoter hexamers, as well as repositioning them for recognition by RNA polymerase.

Based on the homologous DNA binding domain (DBD) among MerR family members, we superpose the DBD of CueR (2-72) with that of MtaN/DNA (1-71) to predict the specific recognition between CueR and DNA. Structure superposition (Figure 3-1) reveals a tyrosine

(Y36 for CueR, Y38 for MtaN, and Y42 for BmrR) specifically interacts with the minor groove near the both ends. As discussed in chapter 2, Y36 is adjacent to R37, which is thought to coordinate the protein/DNA interactions to accomplish the allosteric switch from repressor to activator conformations. The specificity of tyrosine has been suggested to be the hydrogen bonds between hydroxyl group of tyrosine and the bases. However, the aromatic ring of the tyrosine could also stack against the bases and form strong interactions. To further investigate the minor groove recognition, we mutate both protein and DNA, and study the effects on the protein-DNA recognition. To access the protein-DNA recognition, we develop a fast and convenient approach to inspect the binding qualitatively and to determine the activity of the DNA qualitatively. The study reveals that the specificity of tyrosine is attributed to the aromatic ring rather than hydroxyl group, and the minor groove recognition is indispensable for the recognition between CueR and its responsive promoter.
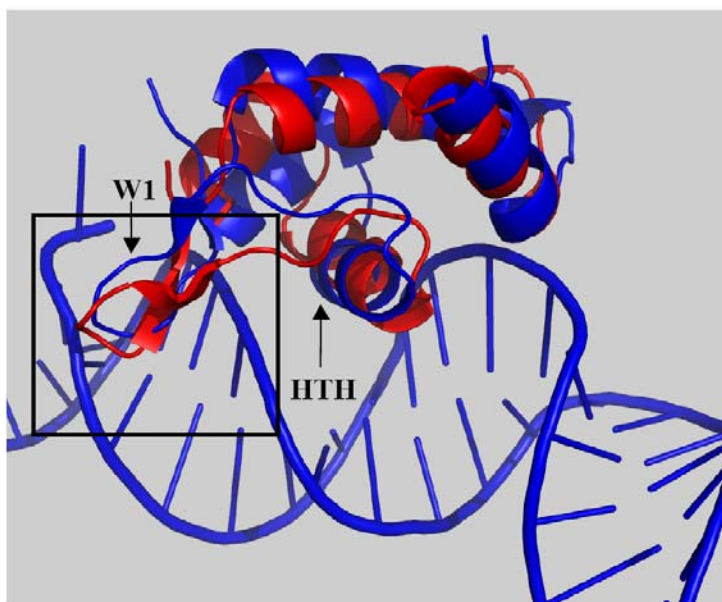
(A) Structure superposition of DNA binding domain of CueR and MtaN-DNA
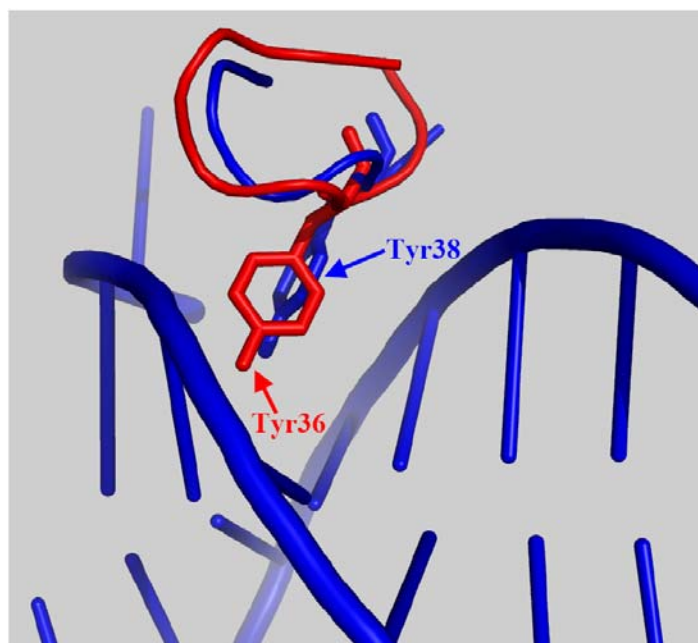
(B) Closer view of rectangular box in (A)

Figure 3-1. Structural superposition of the DNA binding domains of CueR (red) and MtaN-DNA(blue). HTH is the classical helix-turn-helix motif involved in the major groove interaction. W1 is the loop wing involved in the minor groove interaction.

3.2 Experimental procedures

Mutagenesis, Expression, and Protein Purification.

The wide-type CueR protein was prepared as previously described [11]. The CueR mutant genes were prepared by site-directed mutagenesis (QuickChange® Site-Directed Mutagenesis Kit, Stratagene) of the over-expression vector pET24a inserted with the wild-type cueR gene. The sequences of the oligonucleotides used are shown below, with the modified base shown in lower case:

Y36K: 5'- GCAGCGAAAACGGT**aag**CGCACCTACACGCAG -3';

Y36F: 5'- TGCGCAGCGAAAACGGT**ttc**CGCACCTACACGCAGCAG –3'.

The sequence of each mutated gene was confirmed by DNA sequencing analysis. Similar over-expression and purification procedures as those described for the wild-type CueR [11] yielded the mutant proteins with >90% purity indicated by SDS-PAGE analysis. The molecular mass of each of the CueR mutant proteins was determined by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI, PerSpective Biosystems Voyager-DE) and by calculation based on the sequences as follows: CueRY36K, 15203 Da (MALDI), 15200 Da (calc); CueRY36F, 15220 Da (MALDI), 15219 Da (calc). Protein concentrations were measured by Bradford assay using IgG as the calibration standard with the correction factor of 3.71 [14].

Circular Dichroism (CD)

CD spectra of CueR and the mutant proteins were measured in a 0.1-cm path length quartz cuvette with a Jasco J-715 spectropolarimeter at room temperature. The protein samples were exchanged twice into a buffer containing 10 mM K phosphate (pH 7.5) and 100 mM NaCl by dialysis (Slide-A-Lyzer dialysis cassette, 10 K MWCO, Pierce). The bandwidth was 2 nm and

the spectra were measured and averaged over three scans at 100 nm/min with a response time of 0.5 s.

Analytical Ultracentrifugation

AUC experiments were performed by the Center for Analytical Ultracentrifugation of Macromolecular Assemblies (CAUMA) at The University of Texas Health Science Center at San Antonio. The AUC data were analyzed by Dr. Borries Demeler at The University of Texas Health Science Center at San Antonio.

Sedimentation velocity runs were performed on a Beckman Optima XL-A analytical ultracentrifuge at rotor speed of 60,000 rpm at 20°C. CueR protein samples were exchanged twice into the buffer containing 50 mM sodium phosphate (pH 8.0) and 100 mM NaCl (Slide-A-Lyzer dialysis cassette, 10 K MWCO, Pierce). Low- and medium-concentration runs were performed at 230 nm (OD_{230} = 0.8) and 280 nm (OD_{280} = 0.4, 0.6 and 0.8) respectively. The OD values correspond to CueR concentrations of 0.2~1.0 mg/mL. The resulting sedimentation velocity data were analyzed with the van Holde and Weischet method using the computer program UltraScan. Molecular weight and shape distributions were determined with the 2-dimensional spectrum analysis [15] and the genetic algorithm analysis [16]. Confidence intervals were determined with the Monte Carlo analysis [17].

Sedimentation equilibrium experiments were performed at 20°C using a four-hole AN-60 rotor. Eight samples of apo-CueR protein were prepared in the buffer of 50 mM sodium phosphate (pH 8.0) and 100 mM NaCl at the concentrations equivalent to OD_{280} of 0.7, 0.5 and 0.3, OD_{230} of 0.7, 0.5 and 0.3, OD_{220} of 0.7 and 0.5, respectively (86.6 − 3.44 μM). The samples were loaded and sedimented to equilibrium at 22900, 30300, 36200, 41300, and 45800 rpm.

Absorbances scans were measured at the corresponding wavelength. The program PC-NONLIN [18] was used for global fitting of the equilibrium data.

DNA Preparation

The DNA duplex P*copA*-27 covering the full interaction region between CueR and its responsive promoters P*copA* was identified by footprinting experiment [11]. The oligonucleotides of HPLC grade purity were purchased from Integrated DNA Technologies, Inc. Equal amounts of complementary strands were dissolved in annealing buffer (10 mM Tris pH 8.0, 1 mM EDTA), completely denatured at 90°C for 5 minutes, and annealed by slowly cooling to room temperature. The concentrations were determined via UV spectrometry at 260 nm in the same way as described in chapter 2.

Fluorescence Poloarization

FP measurements were collected with a PanVera Beacon Fluorescence Polarization System. Inc. The binding buffer consisted of 10 mM Tris (pH 7.5), 100 mM K Glutamate, 1 mM MgCl$_2$, 5% glycerol, 0.1 mM EDTA, and 5 mM fresh DTT. CueR was added in increasing amounts to a constant amount (1 nM) of the fluoresceinated DNA [FAM]P*copA*-27 (Table 1: II) and incubated for 30 min. All experiments were carried out at 25°C. The excitation wavelength was 490 nm, and units of fluorescence polarization (milipolarization) were read at 530 nm. The binding curve was fit to Equation 1 as described previously [13] using Origin 6.0 (MicroCal, LLC).

$$A = (A_{bound} - A_{free}) \frac{[Protein]_{total}}{K_{diss} + [Protein]_{total}} + A_{free} \quad (Eq.\,\mathbf{1})$$

where A is the polarization measured at a given total protein concentration, A_{free} is the initial polarization of free fluorescein-labeled DNA, and A_{bound} is the maximum polarization of specifically bound DNA. [Protein]$_{free}$ = [Protein]$_{total}$ is assumed because the concentration of fluorescein-labeled DNA was almost 1000-fold below the K_{diss}. Nonlinear least squares analysis was used to determine A_{free}, A_{bound}, and K_{diss}.

Quantitative Electrophoretic Mobility Shift Assays

Binding reactions were performed in the same binding buffer as that used in the fluorescence polarization experiment. EMSA probe [γ-^{32}P] P*copA*-27 was prepared by 5'-end labeling P*copA*-27 with [γ-^{32}P] ATP (MP Biomedicals) using T4 polynucleotide kinase. Different amounts of CueR protein (0.1-10 μM) were incubated with 0.1 μM radioactively labeled DNA for 30 minutes at room temperature and then applied to 20% nondenaturing polyacrylamide gel (made from the 40% ExplorER gel solution, J.T.Baker) prepared in 0.5× TBE buffer (44.5 mM Tris base, 44.5 mM boric acid, 1 mM EDTA, pH 8.0, 25°C), whereas 0.5× TBE buffer was also used as the electrophoresis running buffer. Samples were loaded and subjected to electrophoresis at 300-400V for 5 hours. The amounts of complexed and free DNA were quantified on a STORM 840 PhosphorImager (Molecular Dynamics). The data were plotted and exhibited strong cooperativity. The generated curve was fit with the Hill equation (Equation 2) with Origin 6.0 (MicroCal, LLC).

$$\theta = \frac{[Protein]_{free}{}^{n}}{K_{diss}{}^{n} + [Protein]_{free}{}^{n}} \qquad (Eq.\ \mathbf{2})$$

where θ is the protein bound fraction, n is the Hill coefficient, [Protein]$_{free}$ is calculated by deducting the amount of the bound protein from the total protein amount, since the concentration

of the DNA is less than 10 times lower than K_{diss}. Nonlinear least squares analysis was used to determine n and K_{diss}.

Quantitative Protein Gel Shift Assays

Binding buffer, gel preparation and gel running conditions are the same as described in EMSA experiment. Increasing amounts of PcopA-27 were incubated with 39.8 μM CueR in a total volume of 10 μl for 30 minutes at room temperature before applied to 20% nondenaturing gel. The gel was stained with Coomassie R dye and the stained protein bands were quantified by Kodak DC-290 digital camera and Kodak 1D image analysis software. The stoichiometric titration could determine the percentage of DNA binding activity which can hardly reach 100% due to the unavoidable random and systematic errors during experimental processes.

Qualitative Protein Gel Shift Assays

Instead of the 20% nondenaturing polyacrylamide gel, the 4-15% pre-cast Bio-Rad Ready Gels® were used to resolve the samples. The gel was run at 100V for 1 hour in Tris/Tricine buffer (0.1 M Tris, 0.1 M Tricine, pH 8.25, 25°C). The method is fast and convenient but only qualitative. The binding affinity revealed is significantly lower compared with the quantitative assay. The reasons are uncertain, possibly because of some unknown gradients in the commercially available ready gels.

3.3 Results and discussion

Mutagenesis, Expression, and Protein Purification

Typically, a 4L culture yielded ~ 87 mg CueRY36K, ~ 4 mg CueRY36F, or ~ 15 mg wild type CueR. The purified proteins were of >90% homogeneity verified by SDS-PAGE

(Figure 3-2). The experimentally determined mass of the proteins by MALDI were 15203 Da

(Calc.: 15200 Da) and 15220 Da (Calc.:15219 Da) for CueRY36K and CueRY36F, respectively

(Figure 3-3 and 3-4). Protein concentrations are measured by Bradford assay using the correction

factor of 3.71 derived from the amino acid analysis of CueR using IgG as the standard.

Circular Dichroism (CD)

The CD studies were performed to check if the mutation have an effect on the secondary

structure. The measured CD spectra (Figure 3-5) of the proteins revealed the same characteristic

feature of α–helix, which agreed well with the X-ray structure of CueR [6]. Furthermore, the

spectra of CueRY36F and CueR are remarkably similar, suggesting the mutation has little effects

on the secondary structure. Whereas the CD spectrum of CueRY36K exhibits features similar

features to that of CueR, with the exception of a negative offset, which could be due to the

inaccurate CueRY36K concentration estimated from Bradford assay. In general, the mutants

retain the secondary structural features as the wild type protein.

Determination of CueR Oligomerization State in solution

Analytical ultracentrifugation experiments were preformed to study the solution state

properties of apo-CueR. Sedimentation velocity data (Figure 3-6) showed that at lower

concentrations of 12.1 μM (OD_{230}=0.8) and 33.3 μM (OD_{280}=0.4), WT CueR sedimented as a

2.2-2.3 S species, which was indicated by the vertical G(s) plots (▲ and △), As the concentration

of CueR increased to 50.0 μM (OD_{280}=0.6) and 66.7 μM (OD_{280}=0.8), the sedimentation

coefficient increased to 2.36 S and 2.45 S, respectively, furthermore the G(s) plots (♦ and ■)

were slightly shifted, both of which suggested trace amount of higher molecular weight

aggregates started to appear. The sedimentation velocity data indicates CueR exists as a self-

association system [19] in solution. Further analysis of the velocity data by genetic algorithm/Monte Carlo analysis suggests that the major species of the oligomers is dimer at the experimental concentrations.

To determine the details of the self-association equilibrium, the apo-CueR protein was analyzed by sedimentation equilibrium. 40 data sets were obtained by sedimenting eight different loading concentrations to equilibrium at five different speeds in the buffer of 50 mM sodium phosphate (pH 8.0) and 100 mM NaCl. The data were fit globally using the program PC-NONLIN [18], and it followed that the variance for the single ideal species was the same as that for the monomer-dimer model, suggesting the dimerization was so tight that no appreciable amount of monomer could be detected. The exact dissociation constant cannot be accurately determined with the current data, but at least it is below micromolar range. The molecular weight fitted based on the one component model is 26 kDa, which is suggestive of as dimer and agrees with the sedimentation velocity data. The molecular weight is a little bit low compared with the expected 30 kDa, possibly due to the inaccurate estimation of the protein density.

Taken together, AUC analysis reveals that, in solution CueR exists as a self-association system with the predominant species as the dimer at the experimental concentration range (3 ~ 80 μM).

It needs to point out that because DTT could affect the absorption measurement, the experiments were performed in the absence of the reducing agent. Since CueR has the four cysteines, one might argue that the dimerization could be caused by the formation of the intermolecular disulfide bond. To clarify the ambiguities, additional experiments are required. One approach is to detect the intermolecular disulfide bond formation at the experimental

conditions by AMS modification followed by SDS-PAGE gel. The other approach is to repeat

the AUC experiment with TCEP as the reducing agent.

Determination of CueR-DNA equilibrium dissociation constant

Fluoresence polarization and EMSA were utilized to determine the protein/DNA

dissociation constant. The simplified 1: 1 binding model was employed to avoid the complication

of the association/dissociation of CueR oligomers.

$$DNA + CueR_{dimer} \rightleftharpoons DNA - CueR_{dimer} \qquad\qquad K_{diss} = \frac{[DNA] \bullet [CueR_{dimer}]}{[DNA - CueR_{dimer}]}$$

As CueR dimer predominated at μM (3~ 80 μM) concentration range revealed by AUC data,

which included the major experimental region containing the most information of the binding.

The simplified assumption was quite realistic.

The fluorescence polarization data were fit with the simple 1:1 binding equilibration

(Figure 3-7). The K_{diss} obtained was 0.80 ± 0.25 µM. The EMSA data exhibited strong binding

cooperativity, and was fit with Hill equation (Figure 3-8). The apparent K_{diss} and the Hill

coefficient were 0.58 ± 0.01 μM and 5.33 ± 0.41, respectively. Both assays confirmed the

binding affinity of CueR and the responsive promoter was in μM range, much lower than the nM

affinity of MtaN/DNA and BmrR/DNA [13].

The two assays showed different behaviors, which could be caused by the different

experimental conditions. Fluorescence polarization measurements were performed in solution

without any competitors, thus non-specific recognition could interfere with the specific-

recognition and obscured the binding cooperativity. Whereas in EMSA, the gel matrix may have

played a role in selecting the specific recognition protein/DNA complex during the shifting

process, and the experimental data demonstrated the strong cooperativity for the formation of the specific complex. The speculation could be easily verified by repeating the two assays in the presence of competitors (heparin or poly (dI-dC)).

The protein/DNA binding cooperativity presents a new way to facilitate the genetic switching function of the metalloregulatory proteins in addition to the metal induced ultrasensitivity [6, 7, 9]. The exact reasons for the strong cooperativity are not clear, one of which could be related with association/dissocation of CueR protein.

Protein gel shift assay

We developed a fast and unprecedented method for studying the protein-DNA interactions. The principle is the same as the classical EMSA, the only difference is that protein instead of DNA being used as the probe. The detection and quantification of the probe are through Coomassie brilliant blue R staining and Kodak gel imaging system. Qualitatively, the assay could access the protein/DNA binding properties without any special treatment of the macromolecules. Qualitatively, the stoichiometric titration could be used to measure the DNA binding activity.

The assay was applied to the investigation of CueR/DNA interaction and the shift gel clearly showed two distinctive bands representing free protein and bound protein, respectively (Figure 3-10). The DNA bound protein moved faster toward the positive electrode than the free protein due to the effect of negative charge of DNA exceeding the effect of large molecular weight. The qualitative analysis revealed that 73.2 % of the DNA were in the active binding form.

Attempts have also been made to determine the equilibrium binding constant. Due to the relatively low sensitivity of the Coomassie dye, the titration was performed mostly well above equilibrium dissociation constant of the specific protein/DNA recognition. Since for the stoichiometric titration, the $[DNA]_{free}$ could have large uncertainties or errors, which would have significantly impacts on the determination of the binding constants. Hence, the new assay is not fitful for the quantitative study of the binding equilibrium because of its intrinsic low sensitivity disadvantage.

Investigation of Tyr36 involved in the Specific Minor Groove interaction between CueR and PcopA-27

Superposition of the DNA binding domain of CueR against MtaN/DNA (Figure 3-1) suggested Tyr36 in CueR could be involved in the specific minor groove interaction. While Tyr contained both aromatic ring and hydroxyl group, enabling it to interact with the DNA via both aromatic hydrophobic interaction and hydrogen-bonding. To investigate which one was the major contribution, we mutated the Tyr to Phe with only aromatic ring, and Lys with only H-bonding group, respectively. The effects of the mutation on the protein-DNA binding were accessed with the qualitative protein gel shift assay. Results (Figure 3-10) showed that CueRY36F binds P*copA*-27 in a similar pattern as that of CueR, indicated by two distinctive bands with densities varying with different protein/DNA ratio, whereas CueRY36K exhibited smeared gel shift bands, and the bands became more and more spread as DNA concentration increased, suggesting the interaction between protein and DNA was mostly non-specific. Ag^+ ion did not have substantial effects on the complex formation and protein/DNA affinity in this assay, however, slightly tighten bands were observed, suggesting the conformations of the protein and

the complex being more rigid and uniform in the presence of Ag^+ ion. The result demonstrated that while CueRY36F retained the binding specificity with PcopA-27, CueRY36K has much lower (if any) specific DNA binding activity. This similarity between WT and Y36F demonstrates that the major contribution of Tyr36 to the specific minor-groove interaction is from the aromatic ring rather than hydroxyl group. Furthermore, it also indicated that the minor groove interaction is important for the recognition between CueR and its responsive promoters.

DNA Basesteps Involved in Specific Minor Groove Interaction between CueR and PcopA-27

Both of the published structures of MtaN/DNA(26mer) and BmrR/DNA (23mer)/TPP revealed that the basestep of 'AC/GT' and its palindromic sequence 'GT/AC' on the 5' sides of the operator are involved in the specific minor groove interaction [12, 13]. Sequence alignment of P*copA*-27 and the oligos in the structures of MtaN/DNA and BmrR/DNA/TPP showed the conserved basestep 'AC/GT' and palindromic 'GT/AC' (Figure 3-11A). Considering all the responsive promoters of BmrR, MtaN and CueR have 19-bp spacer between the −10 and −35 elements, and the activated proteins distort and untwist the 19-bp spacer to the consensus 17-bp spacer in a similar fashion, furthermore the distances between the tyrosines involved in the minor groove binding extracted from the crystal structures were 57.60 Å, 59.51 Å and 58.77 Å for CueR, MtaN and BmrR, respectively, thus it is very likely that Tyr36 of CueR also specifically interacts with the 'AC/GT' and 'GT/AC' close to the two ends of P*copA*-27 for the minor-groove recognition. However, is the seemingly conserved sequence a coincidence or does it really have some binding advantages over other random sequences? To test it, we mutated the 'AC/GT' (including the palindromic 'GT/AC' sequence) to 'AG/CT', 'AT/AT', 'TC/GA', and 'GC/GC', respectively, and inspected the binding properties with both CueR and CueRY36F.

The protein gel shift assay (Figure 3-11B) showed that mutation of the conserved 'AC' in

P*copA*-27 changes the protein/DNA affinity in a sequence of TC/GA>AC/GT~AG/CT~

AT/AT>GC/GC for CueR and TC/GA~AG/CT>AT/AT>AC/GT>GC/GC for CueRY36F. In

both cases, mutation of 'TC/GA' showed the strongest binding affinity, mutation of 'GC/GC'

was the weakest one and the wild type 'AC/GT' was in the middle. The result demonstrates that

mutation the conserved 'AC/GT 'does not harm the protein-DNA recognition, but does change

the binding affinity, and the changes could come from both the direct interaction and indirect

recognition of the local features of the DNA. Meanwhile, the difference in DNA binding affinity

sequence for CueR and CueRY36F indicates that tyrosine and phenylalanine have different

recognition preference, which suggests that even though not as the major contribution, hydrogen

bonding might still mediate the protein-DNA interaction in some delicate way.

Minimum DNA Length Required for Recognition between CueR and PcopA-27

Prvious footprinting data [11] (Figure 3-12A) showed that the region of the promoter

protected by CueR expanded 27 bp. Considering the real binding length could be shorter, we

performed a coarse screen started with P*copA*-27 and 2 bps subtracted from both termini each

time. It turned out (Figure 3-12B) that only P*copA*-27 could shift the CueR, which indicated that

the minimum binding length is between 24 and 27 bp. However, based on the analysis based on

the MtaN/DNA structure superposition and DNA sequence alignment, the P*copA*-23 was

anticipated to cover the full binding sequences. Comparing the P*copA*-23 and the 23 mer in

BmrR/DNA/TPP (Figure 3-11A), one prominent difference is the oligonucleotide on the 3'

terminus, which is adjacent to the conserved 'AC/GT' base step in the minor groove interaction.

At this position, a 'T' is found in P*copA*-23, while a 'C' is present in the 23mer of the

BmrR/DNA/TPP complex. We mutated the 'T:A' to 'G:C' or C:G', and the one terminus mutation surprisingly restored the protein-DNA recognition as well as PcopA-27 (Figure 3-12C). The hypothesis is that terminus 'T:A' in P*copA*-23 is not base-paired, which is not uncommon in the DNA structures with T:A terminus, and it hurts the minor groove interaction severely, whereas mutation of 'T:A' to 'G:C' or 'C:G' increases the interaction with one more H-bond, thus 'locking' the terminus, tightening the minor groove interaction, and restoring the recognition. The result is another indirect proof of the indispensability of the minor groove recognition.
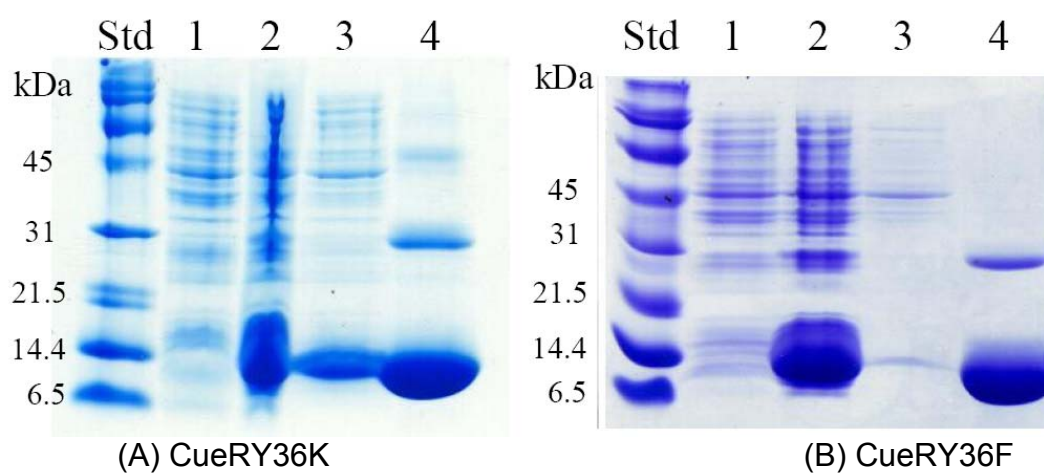
Figure 3-2. SDS-PAGE gels of (A) CueRY36K and (B) CueRY36F.
1: non-induced cells; 2: induced cells; 3: crude protein; 4: purified protein
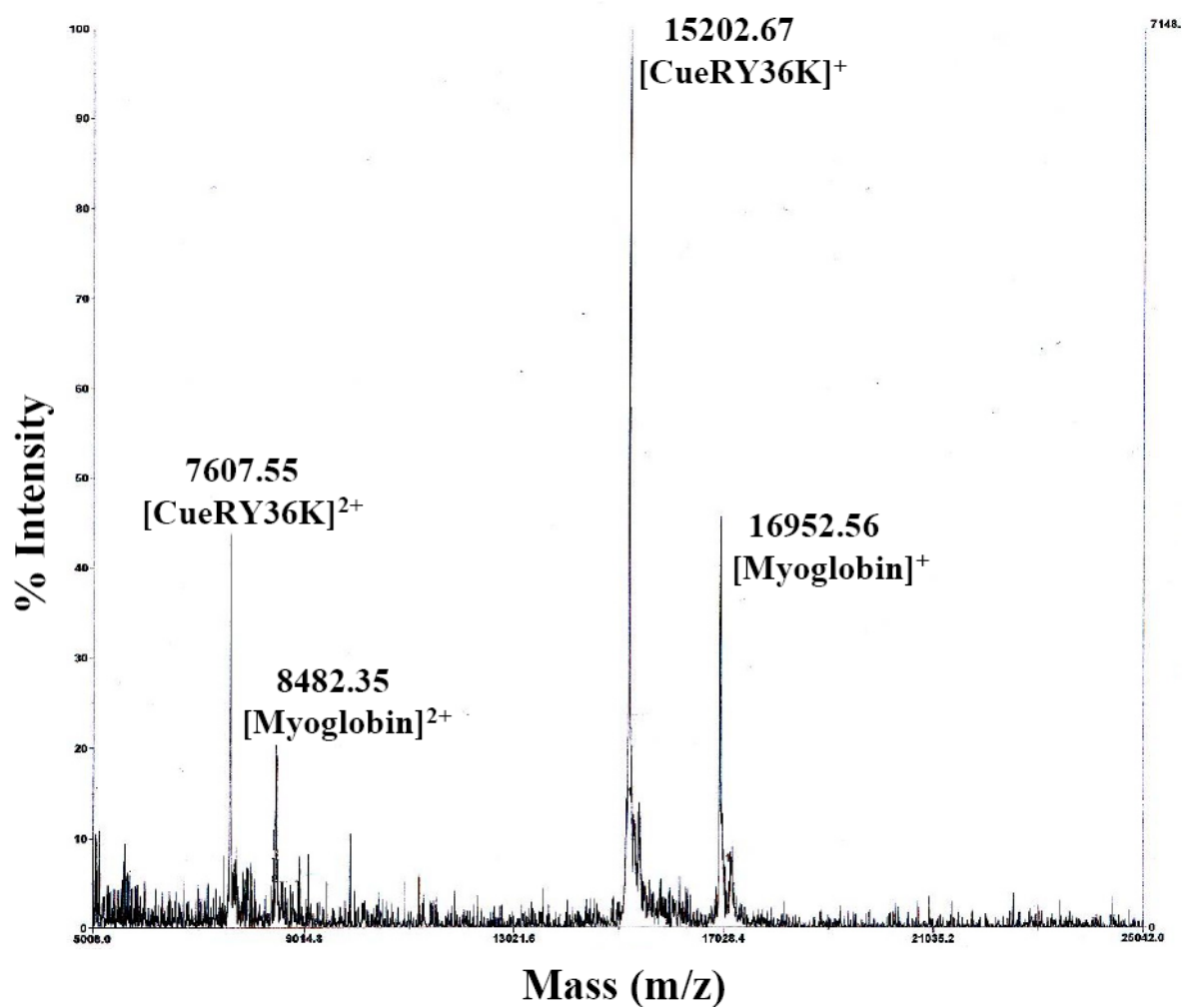
Figure 3-3. CueRY36K MALDI-TOF spectrum
1~2 μg protein sample and ~2 μg myoglobin (internal standard) were used for the measurement.
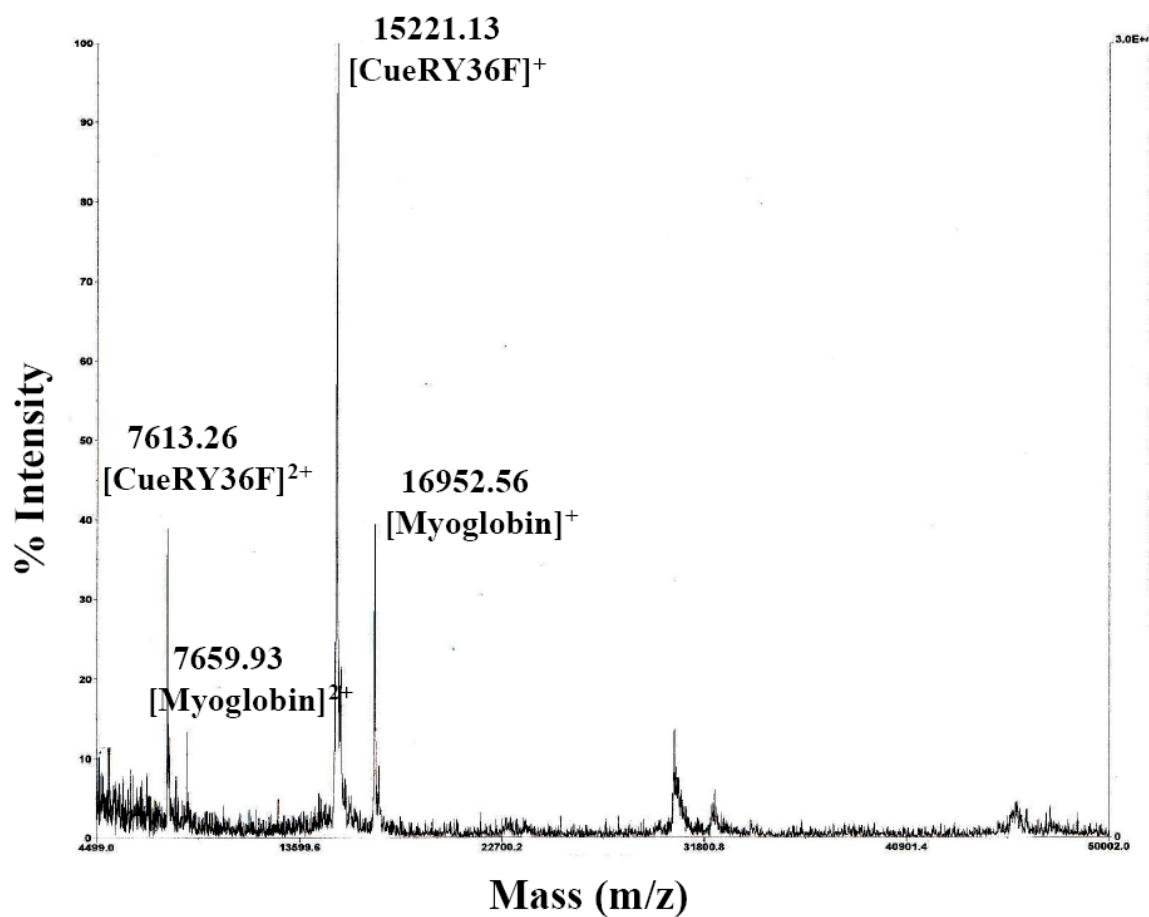
Figure 3-4. CueRY36F MALDI-TOF spectrum
1~2 μg protein sample and ~2 μg myoglobin (internal standard) were used for the measurement.
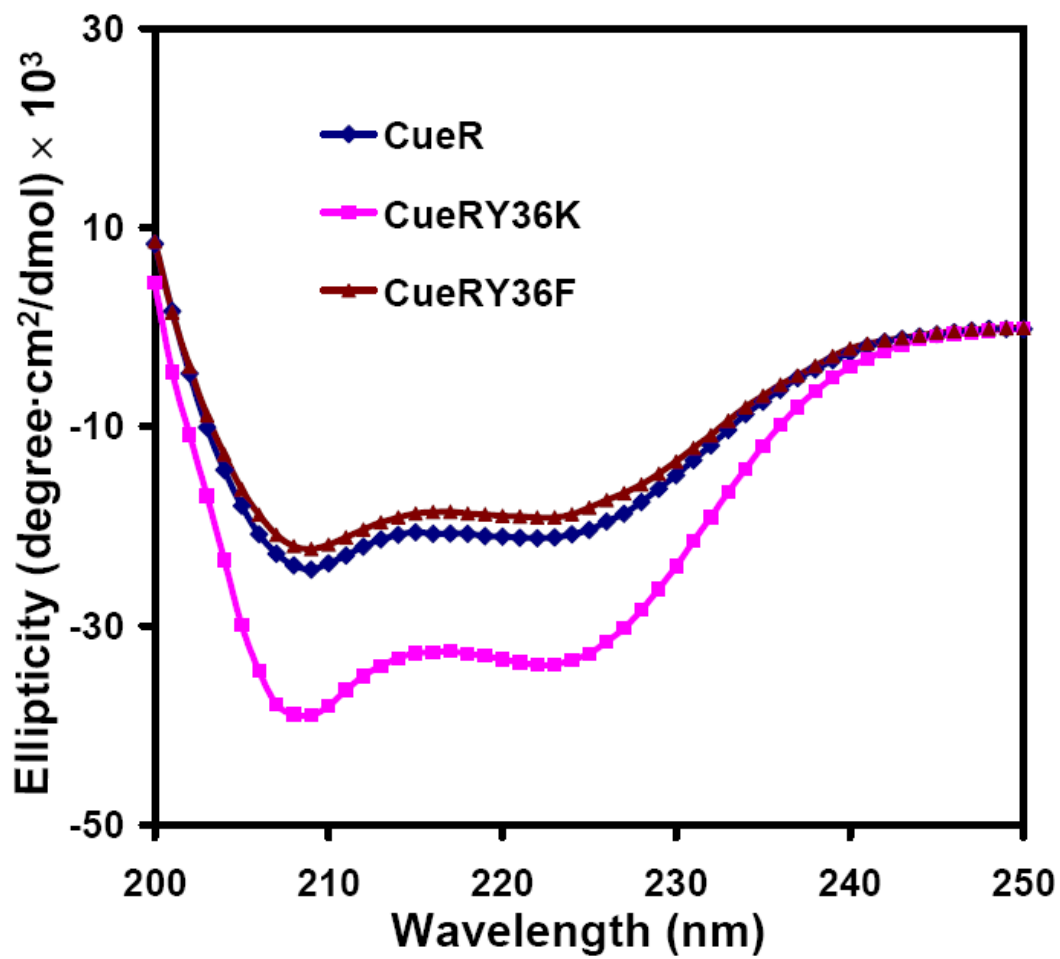
Figure 3-5. CD spectra of CueR and mutants in the buffer of 10 mM K phosphate (pH 7.5) and 100 mM NaCl. — CueR 15.5 μM; ■ — CueRY36F 12.8 μM; ▲ — CueRY36K: 15.6 μM
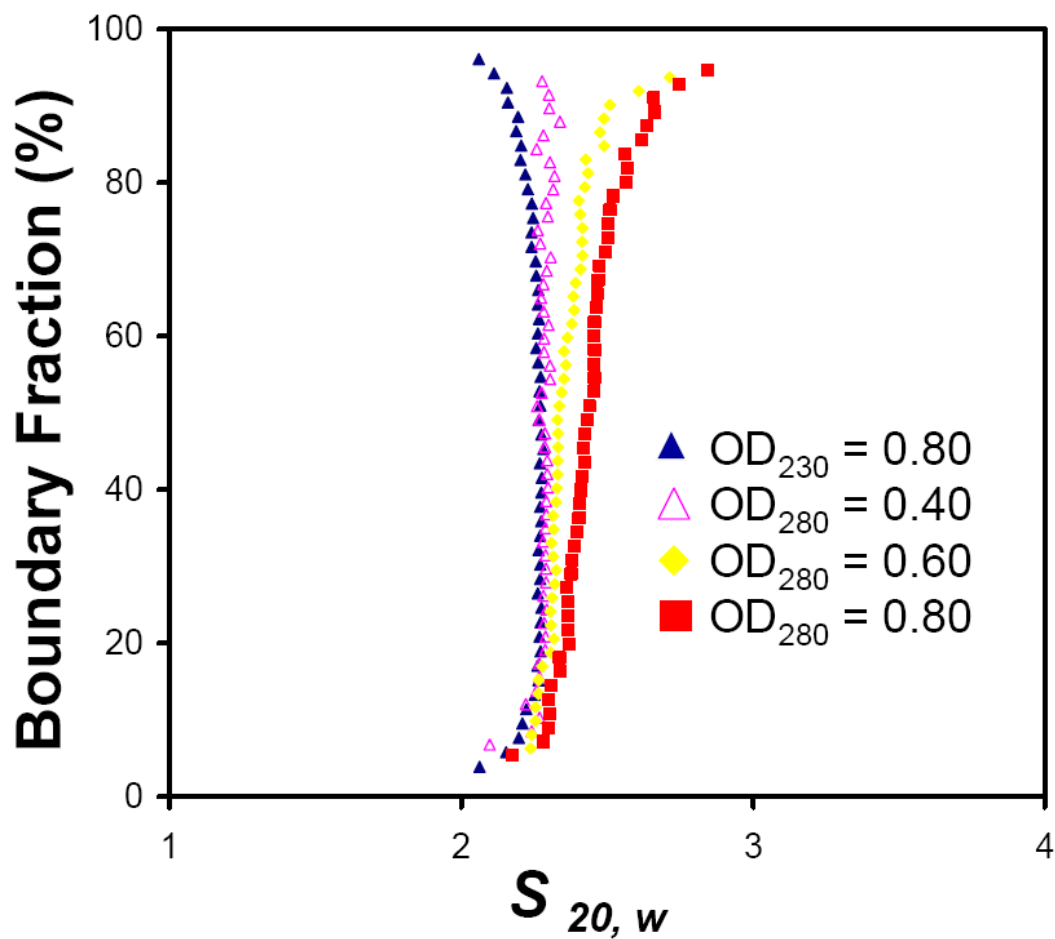
Figure 3-6. *G*(*s*) plots of apo-CueR obtained after analysis of the boundaries by the method of van Holde and Weischet. OD_{230} of 0.8, OD_{280} of 0.4, 0.6 and 0.8 correspond to the concentrations of 12.1, 33.3, 50.0, and 66.7 μM, respectively.
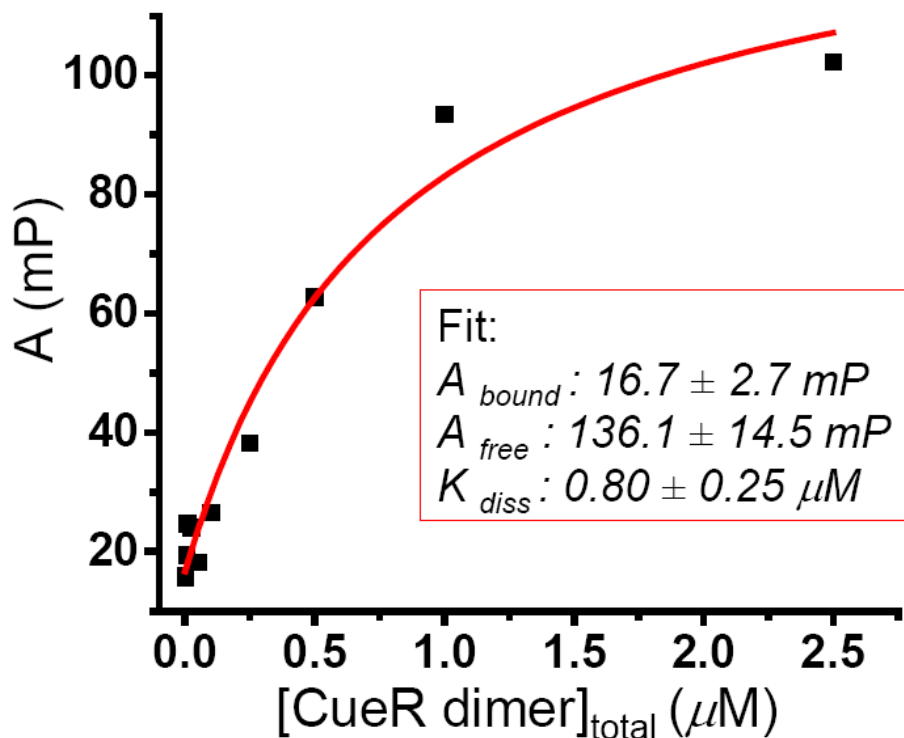
Figure 3-7. Determination of K_{diss} of CueR bound to P*copA-27* with fluorescence polarization.

Different amounts of CueR protein (0.10 nM, 1.0 nM, 10.0 nM, 20.0 nM, 50.0 nM, 0.1 μM, 0.2 μM, 0.5 μM, 1.0 μM, 2.0 μM, and 5.0 μM) were titrated into the solution containing 1 nM fluorescein-labeled DNA and incubated for 10 minutes before the measurement was taken. *mP* represents millipolarization. $[Protein]_{free}$ = $[Protein]_{total}$ is assumed because the concentration of fluorescein-labeled DNA was almost 1000-fold below the K_{diss}. The fitting equation is described in experimental section.

Figure 3-8. Determination of K_{diss} of CueR bound to P*copA-27* with EMSA

(A) EMSA gel. The affinity of CueR and P*copA*-27 were measured by EMSA. Typical reaction contained a fixed amount of labeled oligonucleotides (0.1 μM) and increasing concentrations of CueR protein plus equal amount of Ag(I). Bound and free DNA was visualized by autoradiography.

(B) Graph of fraction of protein bound DNA as a function of the concentration of free CueR dimer. Since the concentration of the DNA is less than 10 times lower than K_{diss}, the free CueR concentration could not be simply approximated to the total CueR concentration. The exact free CueR amount was calculated by deducting the bound fraction from the total amount.
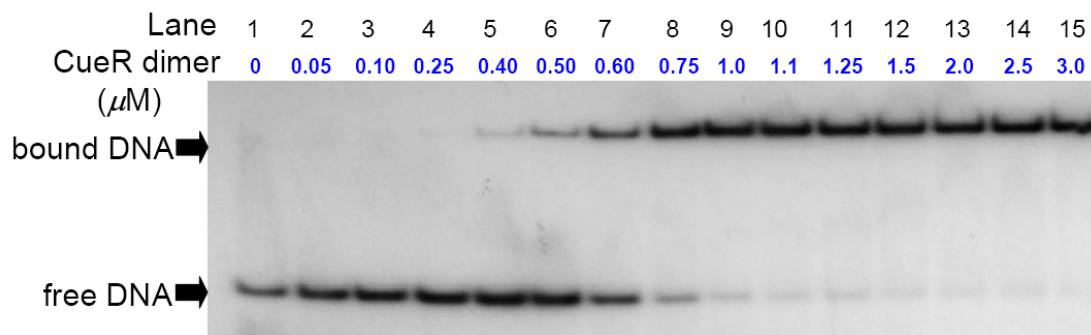
(A) EMSA gel
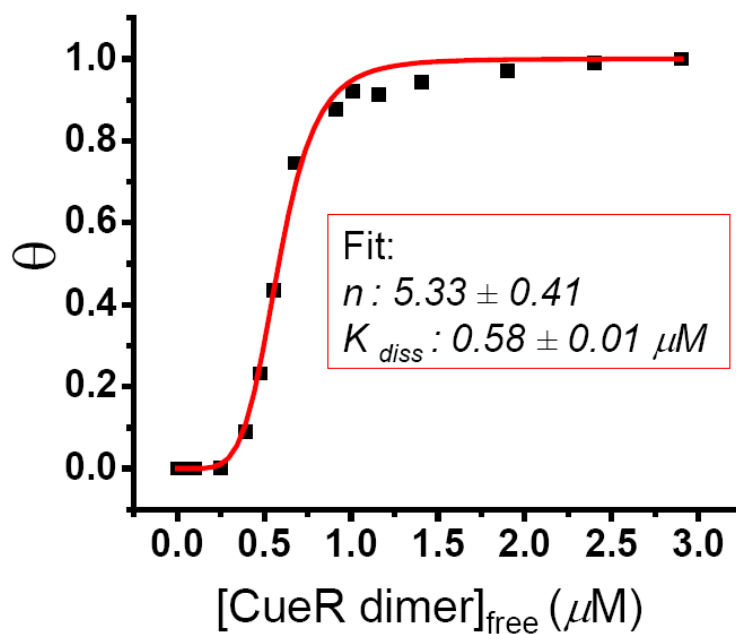
Fit:

$n : 5.33 \pm 0.41$

$K_{diss} : 0.58 \pm 0.01 \ \mu M$

(B) Graph of fraction of protein bound DNA as a function of the concentration of free CueR dimer

Figure 3-9. Determination of the percentage of DNA binding activity by quantitative protein gel shift assay.
(A) The experimental gel. Typical reaction contained a fixed amount of CueR protein, equal amount of Ag, and increasing concentrations of DNA P*copA*-27. Bound protein shifted faster than free protein due to the negatively charged DNA. (B) Graph of protein bound fraction versus the total DNA concentration. The linear fit of the data revealed that 27.2 μM DNA bound 19.9 μM CueR protein, suggesting 73.2 % of the DNA in active binding form.
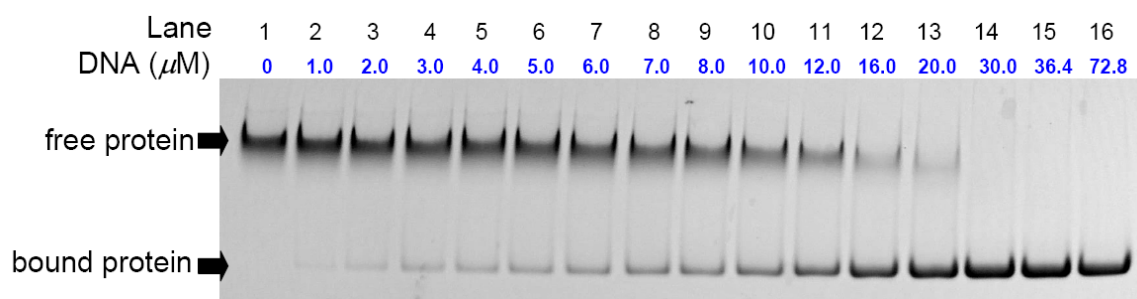
Lane	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
DNA (μM)	0	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	10.0	12.0	16.0	20.0	30.0	36.4	72.8

free protein ➤

bound protein ➤

(A) The experimental gel

27.2 μM
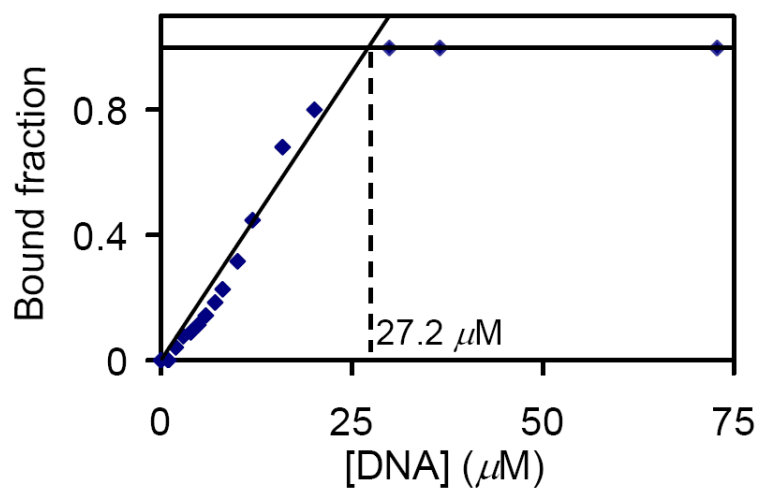
(B) Graph of protein bound fraction versus the total DNA concentration
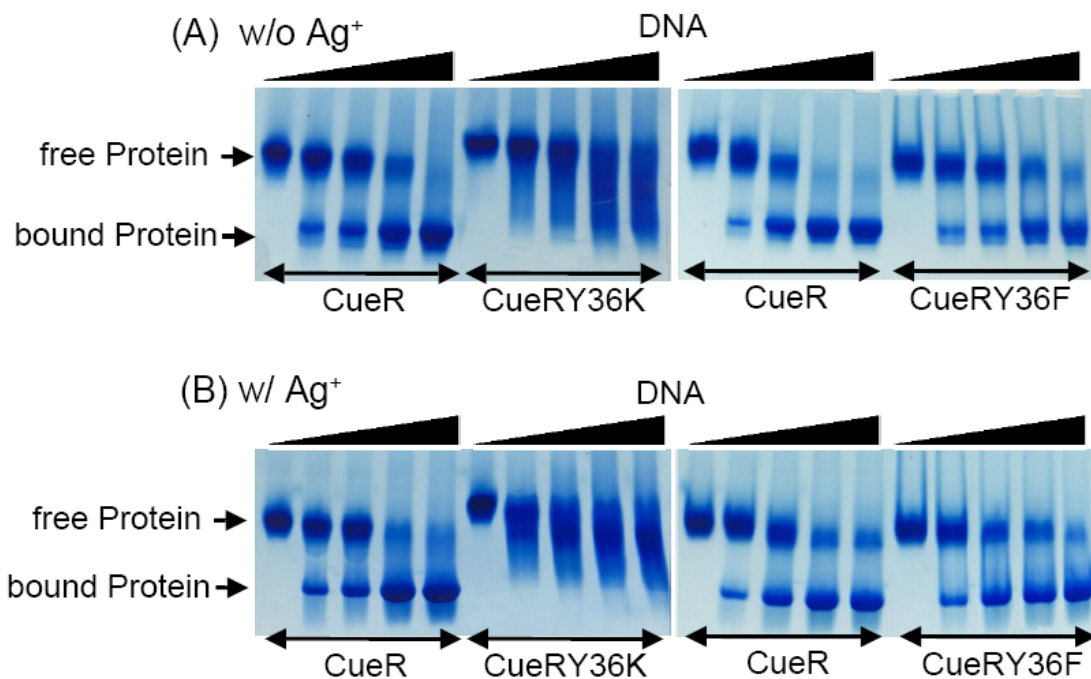
Figure 3-10. Qualitative protein gel shift assays of CueR and its mutants with P*copA*-27 (A) in the absence and (B) in the presence of Ag ion. 75 μM of protein (dimer) was titrated with 0, 0.4, 0.8, 1.6, 2 molar ratio of DNA in a total volume of 15 μL. Ag concentration was 300 μM. CueRY36F was shifted by P*copA*-27 in a similar pattern as that of CueR, whereas CueRY36K exhibited smeared gel shift bands, suggesting non-specific interactions with DNA. Ag ion did not have substantial effects on the complex formation and protein/DNA affinity, however, it slightly tightened the bands.
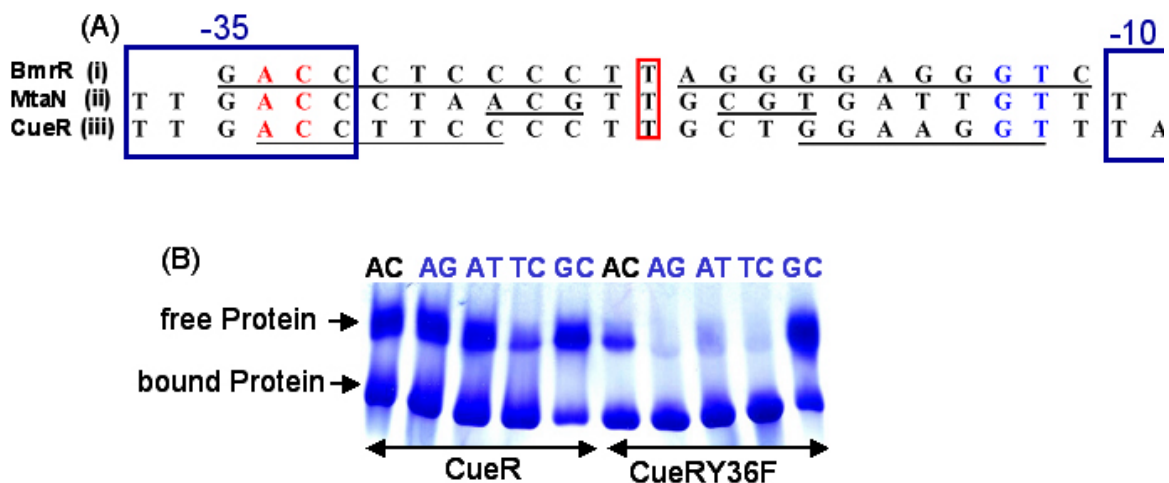
Figure 3-11. (A) Sequence alignment of (i) 23mer oligos used in crystallization of BmrR/DNA/TPP complex, (ii) 26 mer oligos used in crystallization of MtaN/DNA complex and (iii) P*copA*-27. Red 'AC' and its palindromic sequence blue 'GT' were conserved among all three DNA oligos. The palindromic sequences were underlined. The centers were indicated by red box. (B) The qualitative protein gel shift assay of P*copA*-27 and its mutants with CueR and CueRY36F in the presence of Ag ion. 75 μM of protein (dimer) was incubated with 1.2 molar ratio of DNA and 4 molar ratio Ag in a total volume of 15 μL. Mutation of the conserved 'AC' in DNA changes the protein/DNA affinity in a sequence of TC>AC~ AG~AT>GC for CueR and TC~AG>AT>AC>GC for CueRY36F.
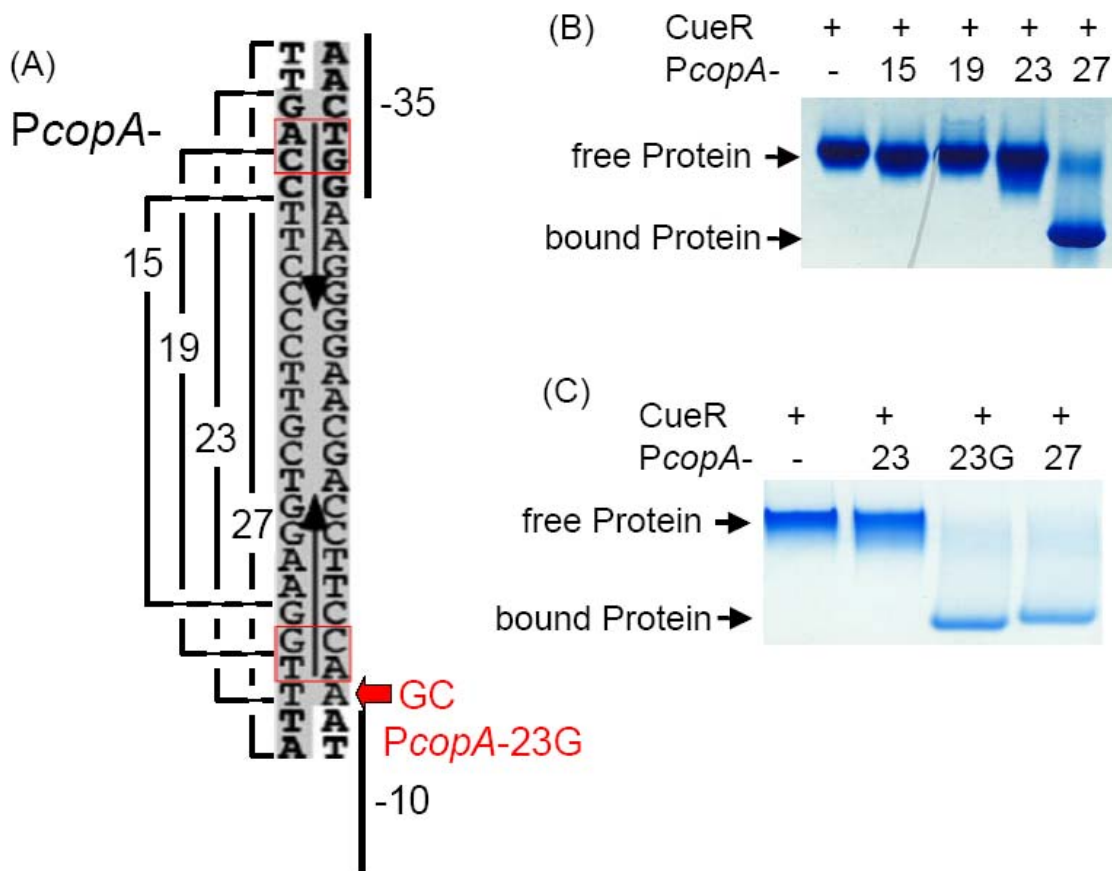
Figure 3-12. Native protein gel shift assay of DNA coarse screen for minimum binding length. (A) P*copA* templates. The figure is adapted from reference 11. The sequence in shaded box is the area of protection from Dnase I in footprinting assay. The vertical arrows indicate the palindromic sequence. The base step 'AC/GT' in red box is the proposed minor groove interaction site. P*copA*-27 is the full-length DNA template, P*copA*-23, P*copA*-19 and P*copA*-15 are shown as PcopA-27 subtracting two basepairs on both termini one ,two and three times, respectively. P*copA*-23G is based on mutating the 'T:A' basepair indicated by the red block arrow to 'G:C' basepair in P*copA*-23.(B) Qualitative protein gel shift assay of CueR with P*copA*-27,23,19 and 15 in the presence of Ag. It shows that only PcopA-27 shifted CueR protein. (C) Qualitative protein gel shift assay of CueR with P*copA*-23, 23G and 27. P*copA*-23 did not shift protein,while P*copA*-23G shifted protein as well as P*copA*-27.

Discussion

Here we identify Tyr36 of CueR is involved in the minor groove recognition and demonstrate the recognition of the tyrosine is mainly attributed to the aromatic ring. In BmrR and MtaN the tyrosine is not close to the base steps and does not cause any unusual local twist or sharp kink of the DNA, which can occur in the minor groove DNA-protein interactions. A common theme involves a protein inserting one or more hydrophobic side chains into the minor groove to unstack two contiguous base pairs and produce substantial kinks of the DNA duplex [20-29]. Compared with the kink of 45º by phenylalanine of TBP [20], the local roll angles by tyrosines are ~11º for MtaN [13] and 13º for BmrR [12], thus, the interaction of tyrosine with the minor groove might not be base-stacking intercalation, or at least, the extent of the intercalation is much more gentle. Additional contacts may exist between the aromatic ring of tyrosine and deoxyribose of the DNA backbone with the distances of 3~4 Å, which has been also observed in the structures of Sac7d mutant/DNA [30], LEF1/DNA [28] and TBP/TATA box [31]. Theoretical calculation suggests that the stabilizing interaction energy of the fucose-benzene could amount to 3.0 kcal/mol [32]. In summary, the minor groove recognition need not involves the classical intercalation, it can be a sum of aromatic stacking, hydrophobic, hydrogen-bonding and non-canonical CH-π interactions [33]. Sequence alignment of the DNA binding domain of MerR homologues (Figure 3-13) reveals that except MerR and SoxR, other members all have a π-system residue (Tyr, Phe or His) at the predicted minor groove interaction position, suggesting the minor groove recognition mode involving an aromatic ring is widespread among MerR family proteins. Meanwhile, the loss of recognition of CueRY36K to the DNA indicates that the minor groove interaction is critical for CueR-DNA recognition.

Contrary to the specific amino acid, the 5' DNA sequence 'AC/GT' step of the copA promoter does not appear to be essential for the minor groove recognitionin this case. Mutation (Figure 3-11) leads to only changes the binding affinity, and the wild type sequence 'AC/GT' does not exhibit the highest binding affinity. Additionally CueR and CueRY36F respond differently to the mutations, suggesting hydrogen bonding can model the minor groove interaction in some delicate way. The detailed mechanisms underlying the difference in the binding affinity to the mutants are not clear and open to further investigation.

In the process of searching for the minimum DNA binding length, we find that, 23mer, which in the case of BmrR includes all the interaction sequences, fails to bind CueR. Intriguingly one mutation of the 3' terminus A:T to G:C or C:G could restore the recognition as well as 27mer (Figure 3-12). Based on the crystal structures of MtaN/DNA (terminus A:T not basepaired) and BmrR/DNA/TPP (terminus G:C basepaired), the hypothesized explanation is that in solution, the A:T terminus is not base-paired, whereas G:C terminus is base-paired, and the base-paired terminus is critical for the minor groove interaction right next to it, hence, P*copA*-23 fails to bind CueR and P*copA*-23G recognizes CueR. Moreover, the basepair A:T is located at the −14 position, right upstream of the −10 element which is melted by σ factor upon open complex formation [34] and within the extended −10 element region [35] (Figure 3-14). How does the basepair mediate the −10 elements recognition with RNAP is open to further investigations.

Additionally, we develop a convenient protein gel shift assay to access the protein-DNA recognition. Compared with the classical EMSA and fluorescence polarization, no special treatment of the macromolecule probe is required. When performing the assay, depending on the

net charge of the species, reverse polarity might be applied, and the migration rate of the complex could be either faster or slower than that of the protein probe. Qualitatively, the assay is universal. Quantitatively, the stoichiometric titration can be used to measure the DNA binding activity. Due to the low sensitivity of the Commassie blue dye, the equilibrium constant (usually nM to μM for transcription factor bound to the responsive promoter) cannot be determined accurately with this method.

Taken together, we identify the amino acid of CueR involved in the minor groove recognition and demonstrate the specificity is attributed to the aromatic ring rather than hydroxyl. Furthermore, the minor groove recognition is indispensable for the recognition between CueR and its responsive promoter. Meanwhile, an unprecedented and convenient approach is developed for accessing the protein-DNA binding.
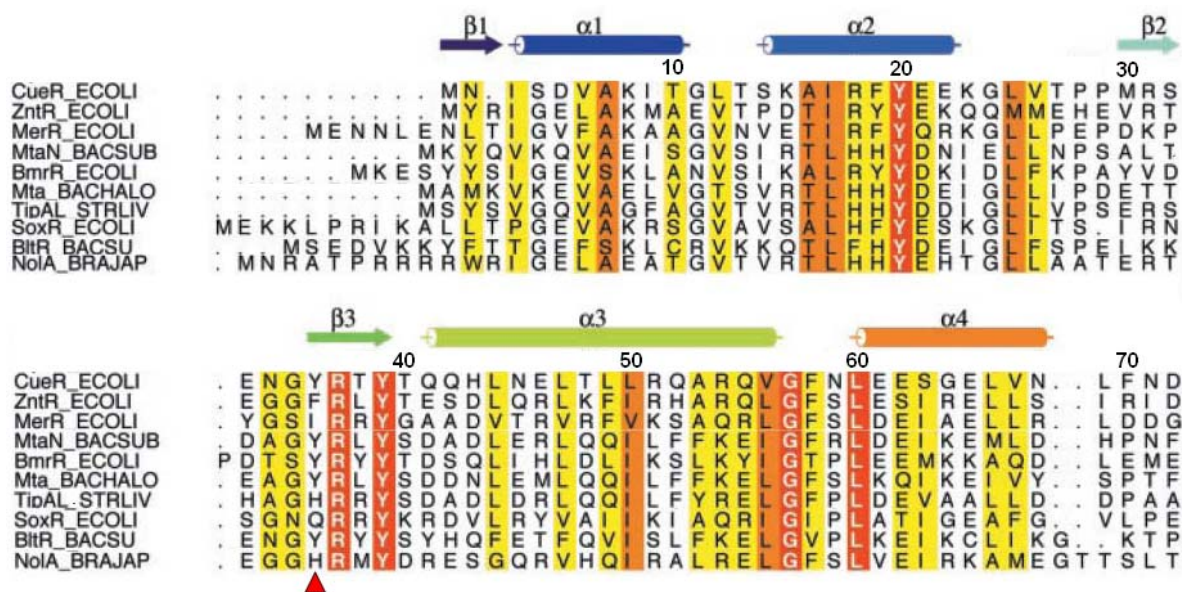
Figure 3-13. Sequence alignment of DNA binding domain of MerR family proteins. Conserved residues are color coded, with decreasing conservation from red shading (absolute) to white (more variable). Red arrow head denotes the residue contacting with the minor groove. Figure is adapted from reference 13.

copA promoter

-14 -13 -12 -11 -10 -9 -8

23 mer

T T G A C C T T C C C C T G C T G G A A G G T T A A C C T
A A C T G G A A G G G G A C G A C C T T C C A A T T G G A

27 mer

-35

Extended –10
5' – TG – 3'

mutation

-10
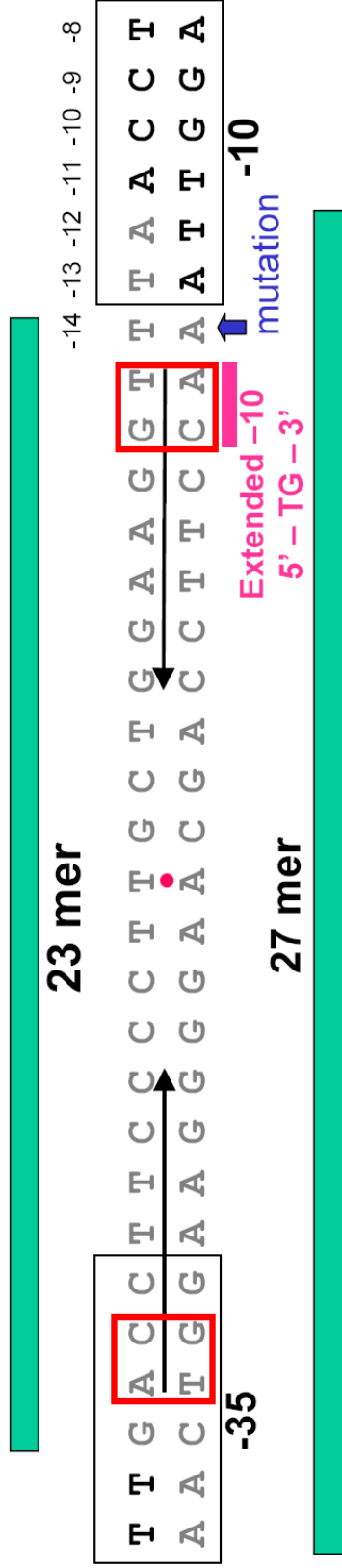
Figure 3-14. Schematic diagram of the *copA* prmoter studied Red box encloses the bases involved in specific major groove interation. Line arrows indicates the palindromic sequences. Sequence in gray is the protected by CueR from DNaseI digestion. Pink block is the position of extended –10 element, where the general consensus sequence is 5'- GT –3' [35]. For the conserved –10 element, positions at –8, -12, and –13 are the most conserved with consensus percentage of 90%, 87% and 79%, respectively, among 554 *E. coli.* Promoters [36]. The rest three positions only have the consensus percentage of ~50. Block blue arrow is the mutation site which could restore the protein/DNA recognition for 23mer, which sit right between the classical –10 element and the extended –10 element.

CHAPTER 4

Crystallographic Studies of Metal bound forms of CusF Protein

Abstract

Methionine-rich motifs have an important role in copper trafficking factors, including the CusF protein. Here we described the crystallization and structure determination of the Cu^I and Ag^I bound forms of CusF. The structures show that CusF uses a new metal recognition site wherein Cu^I/Ag^I is tetragonally displaced from a Met2His ligand plane toward a conserved tryptophan. This novel active site chemistry affords mechanisms for control of adventitious metal redox and substitution chemistry.

4.1 Introduction

Metal ion homeostasis plays an essential role in the health and growth of microbes, and microorganisms often prove to be adept at adjusting to environmental metal stress. Metal ion nutrient acquisition, for example, appears to play an important role in the infectivity and growth of pathogenic organisms.

In *E. coli*, copper tolerance is effected by at least three systems: *Cop*, *Cue*, and *Cus* pathways, with each including regulatory as well as copper-buffering or efflux elements [1]. The present study focuses on one element of the *Cus* system, CusF.

The CusCFBA operon encodes a group of copper efflux proteins with CusA, B and C interacting to span the inner and outer *E. coli* cell membranes [2, 3]. Deletion of the CusCFBA operon leads to silver sensitivity and to copper sensitivity under anaerobic conditions [4]. Unlike the CusA,B and C proteins, CusF is a small (10 kD) soluble periplasmic protein. It is not anchored to a cell membrane, but has been found to interact with part of the ABC complex [3]. Characterization of the copper binding properties of CusF has been somewhat enigmatic, with different studies and different methods providing varied results. An initial EPR study characterized a Cu(II) site comprised of 2 nitrogen and 1-2 oxygen donor atoms but was not able to identify which amino acids were involved in the metal coordination [5]. A later x-ray crystallographic study of the apo protein combined with NMR Cu(I) titration work, proposed a Met, Met, His Cu(I) site [6].

To elucidate the properties of the metal binding sites, we performed X-ray crystallographic studies on the metal bound forms of CusF protein. Both Cu-CusF and Ag-CusF were crystallized and the structures were solved by molecular replacement. The atomic models

reveal a trigonal planar metal coordination for both Cu and Ag, which is formed by two methionine sulfurs, one histidine nitrogen. The Met,Met,His metal binding site represents a recently discovered feature of copper homeostasis proteins and biological copper coordination [7, 8]. Furthermore, one tryptophan is positioned at a remarkablely close distance to the metal center (~ 3.3 Å), suggesting strong cation-π interactions. The CusF active site chemistry affords a distinct class of copper receptor proteins with a means to both control metal exchange and prevent adventitious redox reactions.

4.2 Experimental procedures

Materials

The pure CusF proteins were kindly provided by research technician Mr. Benjamin M. Staehlin.

Crystallization and data collection

Crystallization of CusF was carried out by the hanging drop method mixing 1 μl of protein solution and 1 μl of precipitating solution. For the copper bound CusF, the protein solution consisted of ~30 mg/ml CusF and 2 molar excess $CuSO_4$, and the precipitating solution composed of 2.5 M $(NH_4)_2SO_4$ and 5% isopropanol. For the silver bound CusF, the protein solution consisted of ~30 mg/ml CusF and 2 molar excess $AgNO_3$, and the precipitating solution contained 0.1 M Na Citrate (pH 5.6), 2.5 M $(NH_4)_2SO_4$ and 5% isopropanol. Crystals appeared in 1-2 days, and reached the maximum dimensions of 0.08 mm × 0.08 mm× 0.40 mm for CuCusF and 0.05 mm × 0.05 mm× 0.30 mm for AgCusF in one week (Figure 4-1). Before data collection, crystals were dipped into the cryosolution of 2 M Li_2SO_4 for 1~3 seconds and flash-cooled in liquid nitrogen. All data sets were collected at 100 K using synchrotron radiation. Data of

CuCusF were processed using Denzo and Scalepack [9], and data of AgCusF were processed with Mosflm and SCALA [10]. The crystals belong to the space group $P2_12_12_1$ with one monomer per asymmetric unit. Data statistics are listed in Table 6-1.

Structure determination

The CuCusF and AgCusF structures were solved by molecular replacement with *Phaser* [11] using the apo-CusF structure (PDB code: 1ZEQ) [6] as the search model. A clear solution was found and the rigid body refinement with highest resolution cutoff at 3.5 Å resulted in R and R_{free} factor of 38.9% and 40.3%, 41.0% and 35.4%, 37.8% and 47.2% for 35%, 70% copper occupancy of CuCusF and AgCusF, respectively. Subsequent restrained refinement of the coordinates and the isotropic B factors were performed using *REFMAC5* [10] with intermittent cycles of manual model correction using *Coot* [12]. Final TLS refinement was carried out with the monomer as one TLS group, and R and R_{free} decreased about 1~2%. The final model of 35% copper occupancy CuCusF contains protein residues 14-87, whereas the models of 70% copper occupancy CuCusF and AgCusF contain protein residues 13-87. For CuCusF, the copper atom has been modeled with an appropriate occupancy with a B factor similar to that of the ligating atoms. The assignment of higher occupancies gave a model with high B-factor for Cu and negative density around the metal center in the difference Fourier electron density map. Refinement and model statistics are shown in Table 4-1. Structure evaluation with MOLPROBITY [13] indicated that the final models of Cu bound CusF have good stereochemistry with all residues in favorable regions of the Ramachandran plot, while the model of AgCusF has one outlier, Gln14, which is probably due to the ambiguous electron density at the N-terminus (Figure 4-2).
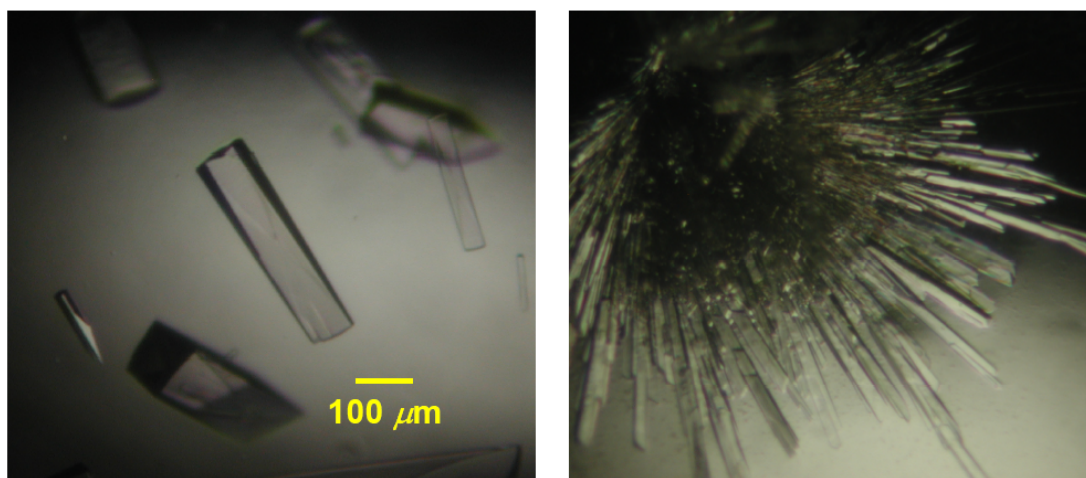
(A) CuCusF   (B) AgCusF

Figure 4-1. Crystals of (A) CuCusF: crystals are either colorless o
light pink. (B) AgCusF: stick-shaped crystals are easily clustered
together.

Table 4-1. Crystallographic statistics.

Values in parentheses are for the highest resolution shell.

Data Statisitics	CuCusF (35%)[1][5]	CuCusF (70%)[5]	AgCusF
Crystal			
Wavelength (Å)	0.9787[1]	1.0[2]	1.0[2]
Space group	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$
Unit-cell parameters (Å °)	a = 38.87, b = 41.65, c = 44.00, a = 90.00, b = 90.00, g = 90.00	a = 38.47, b = 41.26, c = 44.57, a = 90.00, b = 90.00, g = 90.00	a = 39.66, b = 44.23, c = 44.24, a = 90.00, b = 90.00, g = 90.00
Resolution limits (Å)	50.0 - 1.9	50.0 - 1.7	44.2 - 2.3
No. of observed/unique reflections	43308/5471	49919/8250	24011/3611
Completeness (%)	99.2 (99.8)	96.7 (80.5)	99.9 (100)
Data redundancy	6.3 (6.5)	3.5 (1.9)	6.6 (7.0)
R_{sym} (%) [2]	9.4 (28.6)	5.0 (22.4)	6.3 (21.2)
I/$<\sigma>$	15.4 (3.6)	27.2 (4.0)	26.1 (7.0)
Refinement Statistics			
R_{work}/R_{free} (%) [3]	20.8 / 23.6	23.8 / 26.3	24.1 / 26.4
R.m.s. bond lengths (Å)	0.009	0.009	0.011
R.m.s. bond angles (°)	1.336	1.268	1.569
No. of protein residues	74	75	75
No. of water molecules	40	44	19
Validation Statisitics			
Residues in Ramachandran plot regions (%)			
Most favoured	100.0	100.0	95.9
Additional allowed	100.0	100.0	98.6

[1] Dataset was collected at 100 K at the beamline SBC-CAT at the Advance Photon Source.

[2] Both datasets were collected at 100 K at the beamline IMCA-CAT at the Advance Photon Source.

[3] $R_{sym} = \sum |I_{obs} - I_{avg}| / \sum I_{obs}$

[4] $R_{values} = \sum |F_{obs} - F_{cald}| / \sum F_{obs}$, 5% of the reflections were reserved for the calculation of R_{free}

[5] Percentages in parentheses represent the copper occupancy.
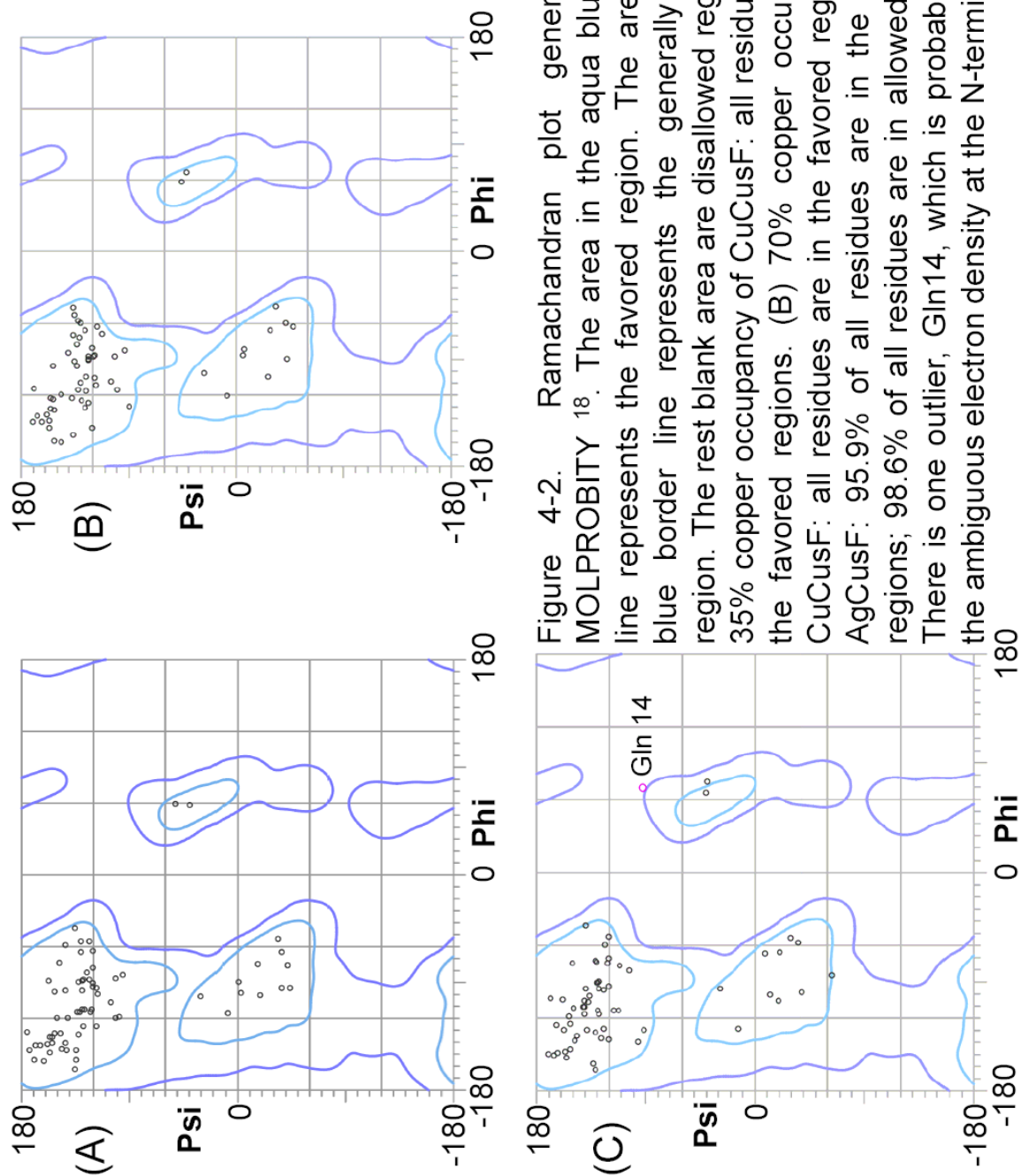
Figure 4-2. Ramachandran plot generated by MOLPROBITY [18]. The area in the aqua blue border line represents the favored region. The area in the blue border line represents the generally allowed region. The rest blank area are disallowed regions. (A) 35% copper occupancy of CuCusF: all residues are in the favored regions. (B) 70% copper occupancy of CuCusF: all residues are in the favored regions. (C) AgCusF: 95.9% of all residues are in the favored regions; 98.6% of all residues are in allowed regions. There is one outlier, Gln14, which is probably due to the ambiguous electron density at the N-terminus

4.3 Results and Discussions

Crystallization of metal bound forms of CusF and metal occupancies

After adding Cu(II) ions to CusF, the solution turns into pink immediately, whereas addition of Ag(I) ion does not change the color of CusF solution. Some of CuCusF crystals also present light pink color, while AgCusF crystals are colorless.

The addition of Cu(II) changes the colorless CusF solution into pink, the color change associated with Cu(II) coordination is very similar to the classical amino terminal Cu(II)- and Ni(II)-binding (ATCUN) motif [14], which includes the N-terminal amine, two deprotonated amides and a His imidazole of the third residue forming the square planar coordination geometry, inspecting the N-terminus of CusF, three histidines are positioned from 3-5. EPR investigation of the Cu(II) binding site in CusF also indicates it as a type II copper center comprised of a His imidazole nitrogen and a deprotonated peptide nitrogen [5], thus the Cu(II) binding is probably due to the ATCUN motif in CusF. Furthermore, deletion of the first five residues of CusF abolishes the characteristic color change upon Cu(II) addition (unpublished data done by Anna and Ben), which corroborates the N-terminus Cu(II) binding hypothesis. Meanwhile, some of the crystallographic data are collected from light pink color crystals, and in the final model, the N-terminus containing residues 1-12 are disordered and no Cu(II) metal binding sites are observed in the atomic model, which suggests the Cu(II) binding happens at the N-terminal region. More detailed and quantitative spectroscopic studies by Anna (unpublished data) further confirm the color change associated with Cu(II) are due to the ATCUN motif in CusF.

The crystals of CuCusF appear in one week with the occupancy of copper only about 35%. After sitting for another 1 to 2 months, the Cu occupancy increases to 70%. The crystals of AgCusF grown in one week contain full occupancy of Ag atom. The 2Fo-Fc electron density maps of the metal binding sites (Figure 4-3) clearly reveal the different metal occupancies. In addition, in CuCusF structure with 35% copper occupancy, the difference Fourier map strongly indicates the existences of dual conformations of both Met47 and Met49, one for the coordination state, and one for the apo state. The copper binding site is determined to be a Cu(I) metal center based on the coordination geometry and the X-ray absorption spectroscopic studies (unpublished data by Anna). It is not clear how Cu(II) is reduced to Cu(I) under the environment with no cysteine residues in CusF and no externally added reducing agent in the crystallization experiment.

Overall Structure

The global fold of metal bound forms of CusF (Figure 4-4) is that of a small barrel formed by two orthogonally packed β-sheets, where β-sheet 1 is composed of β1, β2, and β3, and β-sheet 2 comprises β5, β4, and β1. A bulge in the middle of β1, accommodated by a glycine residue (Gly 20, Figure 4-5), allows this strand to be part of both β-sheets. Glycines and prolines are frequently found on either sides of the strands where they break the regular secondary structure and facilitate the conformation of the interstrand loops (Figure 4-5).

The topology of metal bound forms of CusF is similar to a canonical OB-fold [15], except that the β-barrel is frequently capped by an α-helix between β3 and β4 in the OB fold. Furthermore, the loop between β2 and β3 is more extended in the metal bound forms of CusF compared with the OB fold. The OB fold is a compact structural motif for a variety functions,

but is particularly well represented for oligonucleotide-binding proteins. At present, there are

10 OB-fold superfamilies in the SCOP (Structural Classification of Proteins) database [16], none of

these include proteins with a copper The β-strand framework of the β-barrel leads to a generally

alternating pattern of amino acid side-chains pointing into the barrel interior (Figure 4-6),

forming the hydrophobic core. The interior residues pack in three layers consistent with those

observed in many OB-folds [15, 17].

The first twelve N-terminal amino acids and the final C-terminal residue were invisible in

the electron density map, MALDI-TOF mass spectra of the crystals indicated the existence of the

full length CusF, indicating residues 1-12 and residue 88 were disordered in the crystals.

Superimposition of CuCusF and AgCusF on apo-CusF$_{6-88}$ gave average rms differences of 0.55

Å and 0.54 Å for the mainchain atoms, respectively, suggesting the almost identical overall fold.

Metal Binding Site

The metal ions sit on the loop region between β2 and β3 (Figure 4-4) and are coordinated

to His36, Met47 and Met49 (Figure 4-6). In CuCusF, Cu coordinates to N$^{\varepsilon}$ of His36, S$^{\delta}$ of M47

and S$^{\delta}$ of M49 at distances of 1.98 Å, 2.25 Å and 2.25 Å, respectively, which agrees well with

the EXAFS results (one Cu-N: 1.94 Å, two Cu-S: 2.24 Å).   In AgCusF, Ag is coordinated to

His36, Met47 and Met49 with longer distances of 2.33 Å, 2.49 Å, and 2.67 Å, respectively. The

coordination geometry is almost trigonal planar with the Cu and Ag displaced by 0.56 Å and

0.54 Å, respectively, from the plane defined by the coordinating atoms of His36, Met47 and

Met49 toward Trp44. A search of the Metalloprotein Database (http://metallo.scripps.edu)

indicates that CuCusF and AgCusF are the first crystallographically characterized Met, Met, His

metal binding centers. The trigonal planar coordination geometry is prevalent for Cu(I) metal

center and the bond distances agrees well with the established distances from small molecules analogues in Cambridge Structure Database [*].

In both cases Trp44 is surprisingly close to the metal center, suggesting strong cation-π interactions between the indole ring and the metal cations [18], which is also evidenced by UV-resonance raman studies and strong Trp fluorescence quenching (unpublished data by Anna). Of the cataloged published structures in the cation-π database [18], the structures with the closest Cu-Trp distances exhibit Cu-centroid distances of over 4 Å to be compared to the 3.3 Å distance for CusF. In particular, the $C^{\varepsilon 3}$ and $C^{\zeta 3}$ of the aromatic ring are at the distances of 2.67 Å and 2.86 Å from the copper, and 2.99 Å and 3.29 Å from the silver. Search of small molecules of Cu-arene and Ag-arene complexes in Cambridge Structure Database established that Cu-C and Ag-C distances in a η^2 fashion of 2.08-2.75 Å and 2.39-2.93 Å, respectively. The distances observed in the CuCusF and AgCusF structures are almost at the cutoff ranges, suggesting the possible existence of weak metal-arene interaction. Moreover, copper is exclusively in +1 oxidation state for all the Cu-arene compounds, consistent with Cu being Cu(I) in CuCusF. However, so far, no direct evidences have been observed to support the existence of the metal-arene interactions.

Taken together, the metal coordination is trigonal planar, with two methionine sulfurs and a histidine nitrogen. A nearby tryptophan shows strong cation-π interactions with the metal center. The coordination geometry parameters are listed in Table 4-2.

Superposition of the metal binding region (residue 36-49) of the structures of apo-CusF and CuCusF reveals that upon metal binding (Figure 4-8A), Met47 and Met49 rotate and reorient

---

[*] Cambridge Structural Database (CSD), associated software systems (including CONQUEST), and documentation are proprietary products in which the rights are held by the Cambridge Crystallographic Data Centre, Cambridge, U.K.

the sidechains to make favorable contacts with the metal centers, whereas His36 and Trp44 only shift slightly. Comparison between the metal binding region of AgCusF and CuCusF suggests that the sidechains of Met47 and Met49 also rotate differently to facilitate coordination with metals of different radius (Figure 4-8B). In addition, the Ag atom displaces from the Cu atom by 0.5 Å, which, combined with the longer ionic radius, leads to the tilting of the aromatic ring of Trp44 by ~20º away from the metal center. In both cases, the mainchain displacements are accumulated on Met49. The structural information indicates the major conformational changes involved are the orientations of the ligated methionine sidechains.

Spacefilling models of the metal binding regions of apo-CusF, CuCusF and AgCusF reveal interesting changes of the solvent accessibility of ligated atoms (Figure 4-9). In all three models, the H36 is buried inside and the W44 is at the surface. The S^δ atoms of M47 and M49 are both exposed to solvent in apo-CusF. When metals are bound to CusF, the S^δ atoms of M47 and M49 both rotates toward the metal center, in CuCusF, the S^δ of M49 is totally buried, and the S^δ of M47 is still kept exposed, whereas in AgCusF, the S^δ of M49 is at the surface, and the S^δ of M47 is partially buried. Collectively, the S atoms of methionines are similar to two arms at the surface, and upon binding the metal, they grab it back to the metal site, and protects them from the bulky solvent. Depending on metal identity, the S atoms exhibit different orientation and solvent accessibility.

Table 4-2. Metal-site geometry

Bond Distances (Å)	CuCusF	AgCusF
M-N$^\varepsilon$(H36)	1.98	2.33
M-S$^\delta$(M47)	2.25	2.49
M-S$^\delta$(M49)	2.25	2.67
Bond Angles (°)		
N$^\varepsilon$(H36)-M-S$^\delta$(M47)	116	107
N$^\varepsilon$(H36)-M-S$^\delta$(M49)	108	134
S$^\delta$(M47)-M-S$^\delta$(M49)	115	106
Other Distances (Å)		
M⋯NS$_2$ plane [1]	0.56	0.54
M-C$^{\varepsilon 3}$(W44)	2.67	2.99
M-C$^{\zeta 3}$(W44)	2.86	3.29
M⋯W44 centroid [2]	3.3	3.4

[1] Plane defined by N$^\varepsilon$(H36), S$^\delta$(M47), S$^\delta$(M49).
[2] Centroid of W44 is defined as the centroid of six-member ring of indole ring
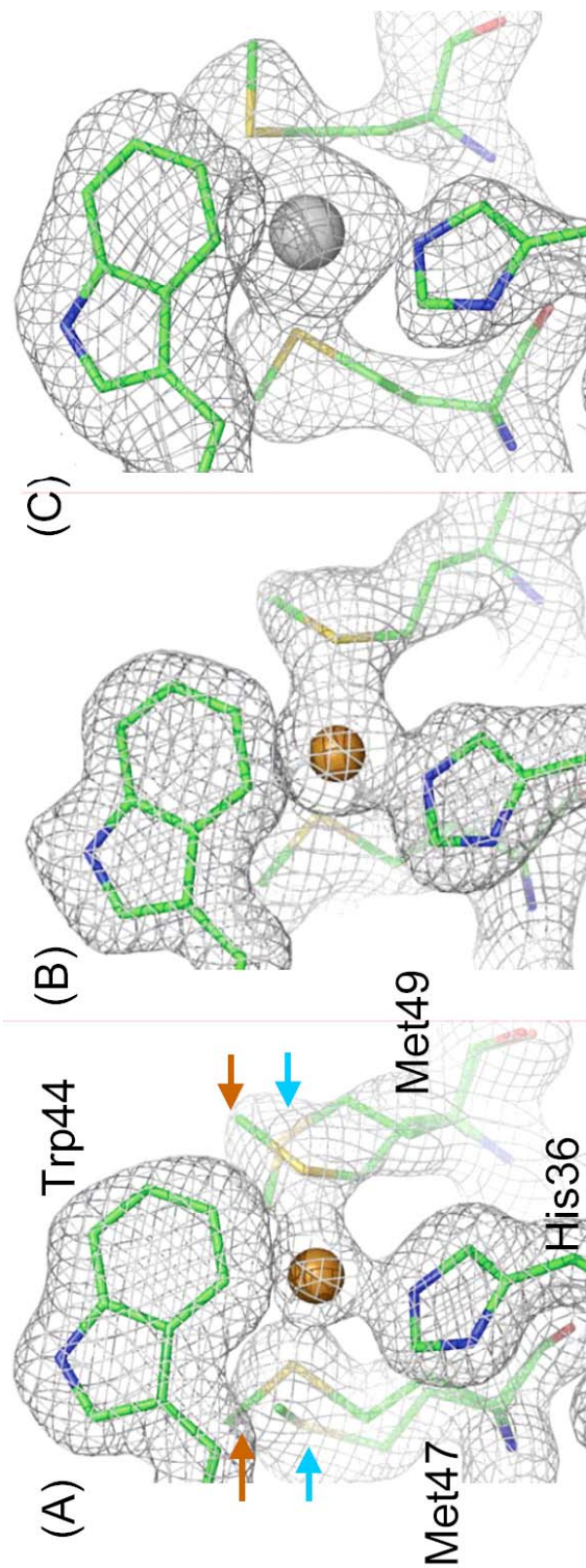
Figure 4-3. 2Fo-Fc maps contoured at 1s level of the metal coordination sites of (A) 35% copper occupancy CuCusF (B) 70% copper occupancy CuCusF (C) 100% silver occupancy AgCusF. The Cu ion and Ag ion are depicted as a bronze sphere and a white sphere, respectively. In (A), Met47 and Met49 have dual conformations due to low Cu occupancy, cyan arrows denote the apo state conformation, bronze arrows denote the coordination state conformation. (A) and (B) clearly illustrate the increasing density levels around the Cu ion, suggesting the elevated copper occupancy. In (C), much higher density around Ag ion is due to the more electrons in silver than in copper.
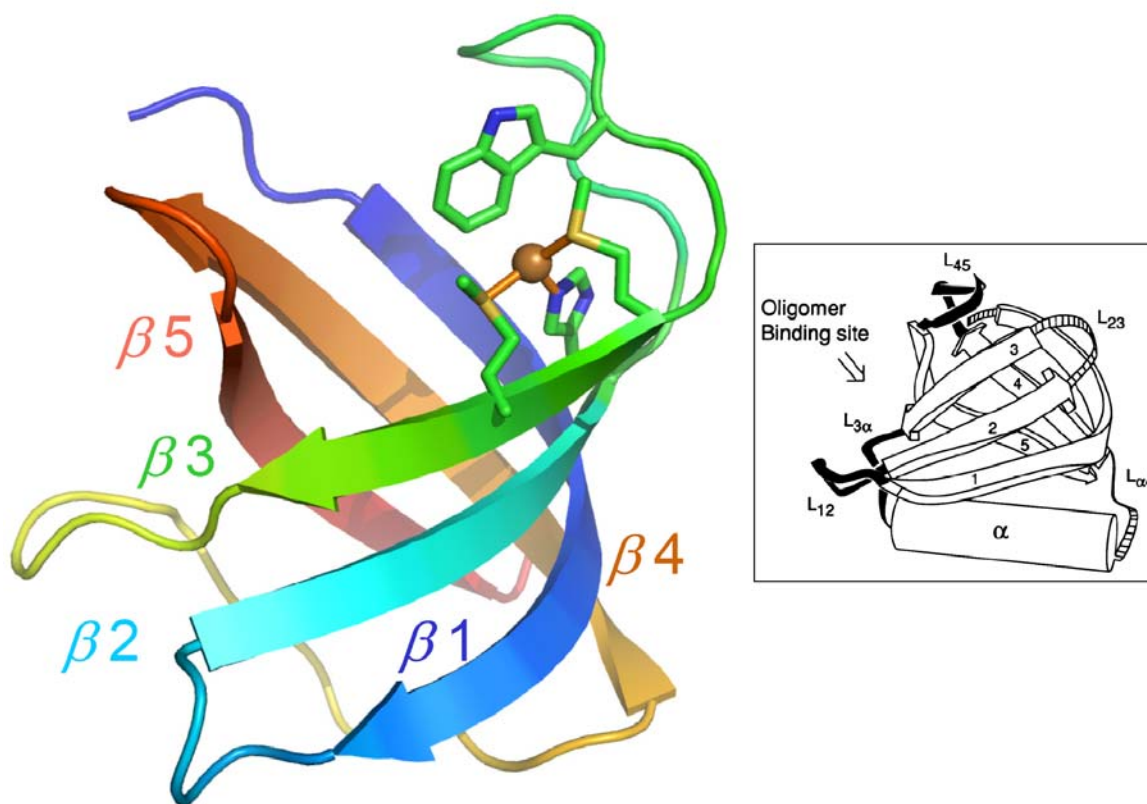
Figure 4-4. Overall structure of CuCusF. The ribbon diagram is prepared by PyMOL. The copper ion is shown as a bronze sphere. The β strands are color coded and labeled. The global fold is that of a small barrel formed by two three-stranded β sheets, one consists of β1, β2 and β3, the other consists of β1, β4 and β5. The overall structure of AgCusF is almost identical to that of CuCusF except copper ion is replaced by silver ion. The small inset picture is a representative description of the OB fold, adapted from reference 19. The closed β - barrel is capped by an α-helix between β3 and β4.
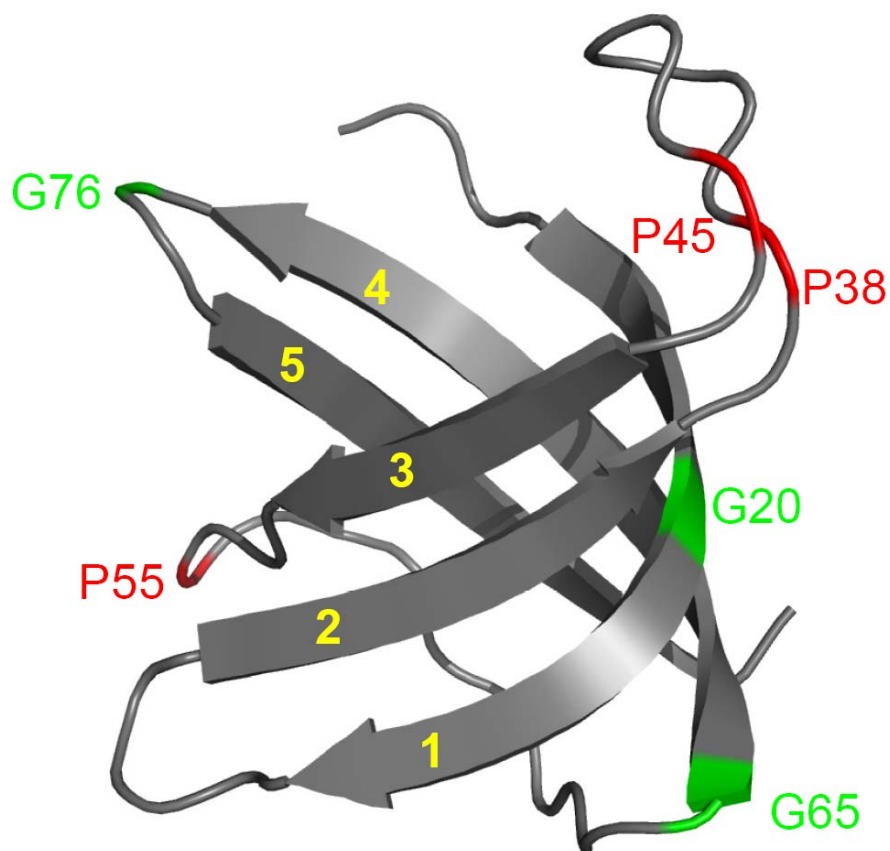
Figure 4-5. Glycines and prolines in metal bound forms of CusF. Glycines are shown in green, prolines are shown in red. The b strands are numbered in yellow. A bulge in the middle of b1, accommodated by a glycine residue (Gly 20), allows this strand to be part of both b-sheets. Glycines and prolines are frequently found on either sides of the strands where they break the regular secondary structure and facilitate the conformation of the interstrand loops.
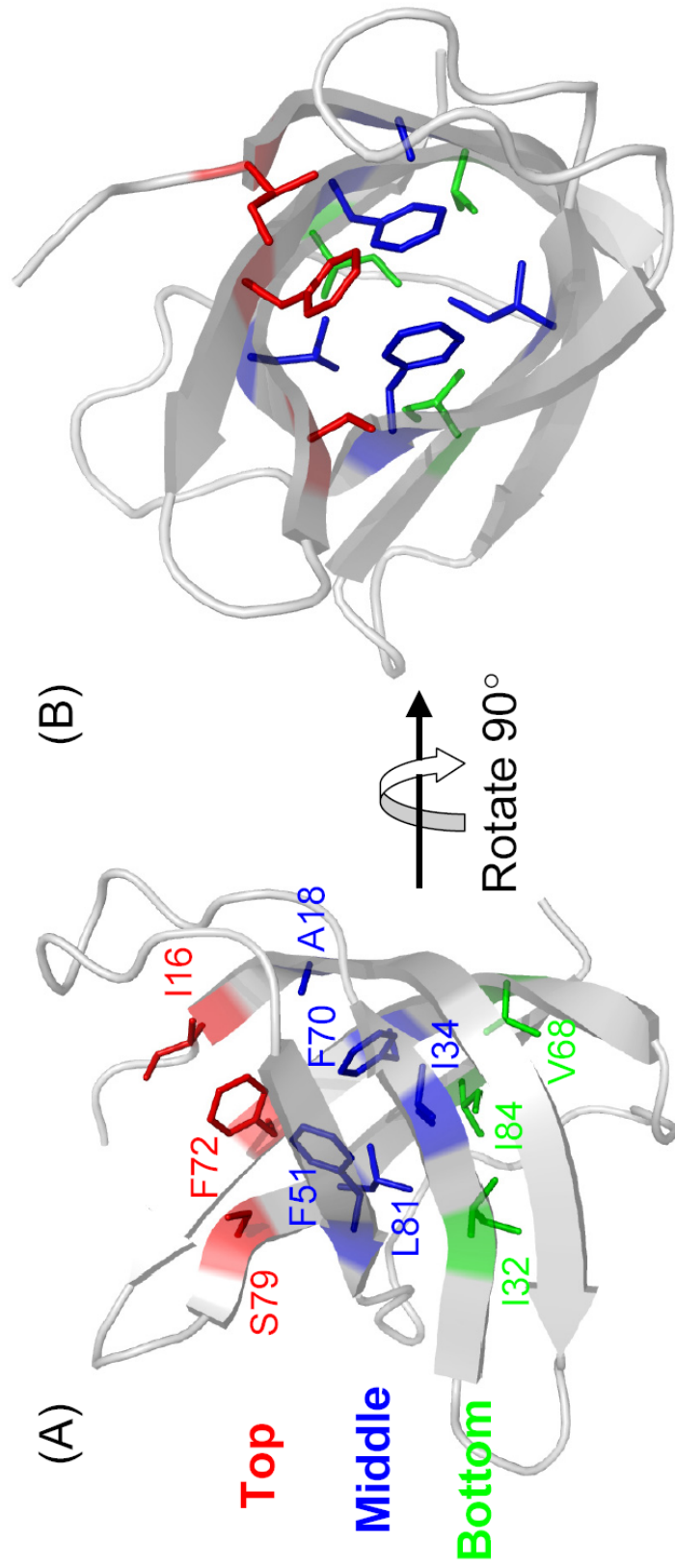
Figure 4-6. The hydrophobic core of the CusF β-barrel. (A) View from the side (B) View from the top. The protein backbone is shown in gray cartoon. Internal residues are color-coded by layer, with the top, middle, and bottom layer (as originally described by Murzin[19]) in red, blue, and green, respectively.
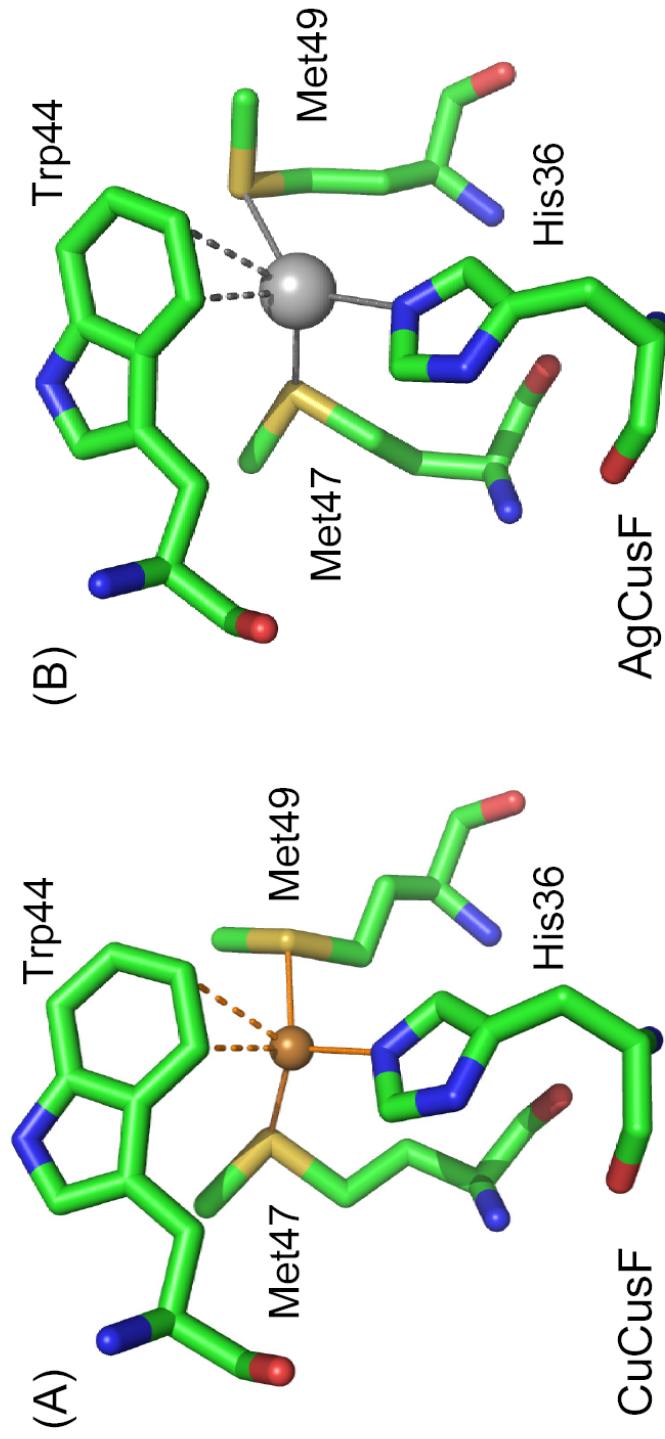
Figure 4-7. Close-up views of metal coordination of (A) CuCusF and (B) AgCusF. In both structures, the metal ion coordinate to N atom of His36 and S atoms of Met47 and Met49 with almost trigonal planar geometry. The proximity of the aromatic ring of Trp44 to the metal center suggests the existences of strong cation-π interactions.
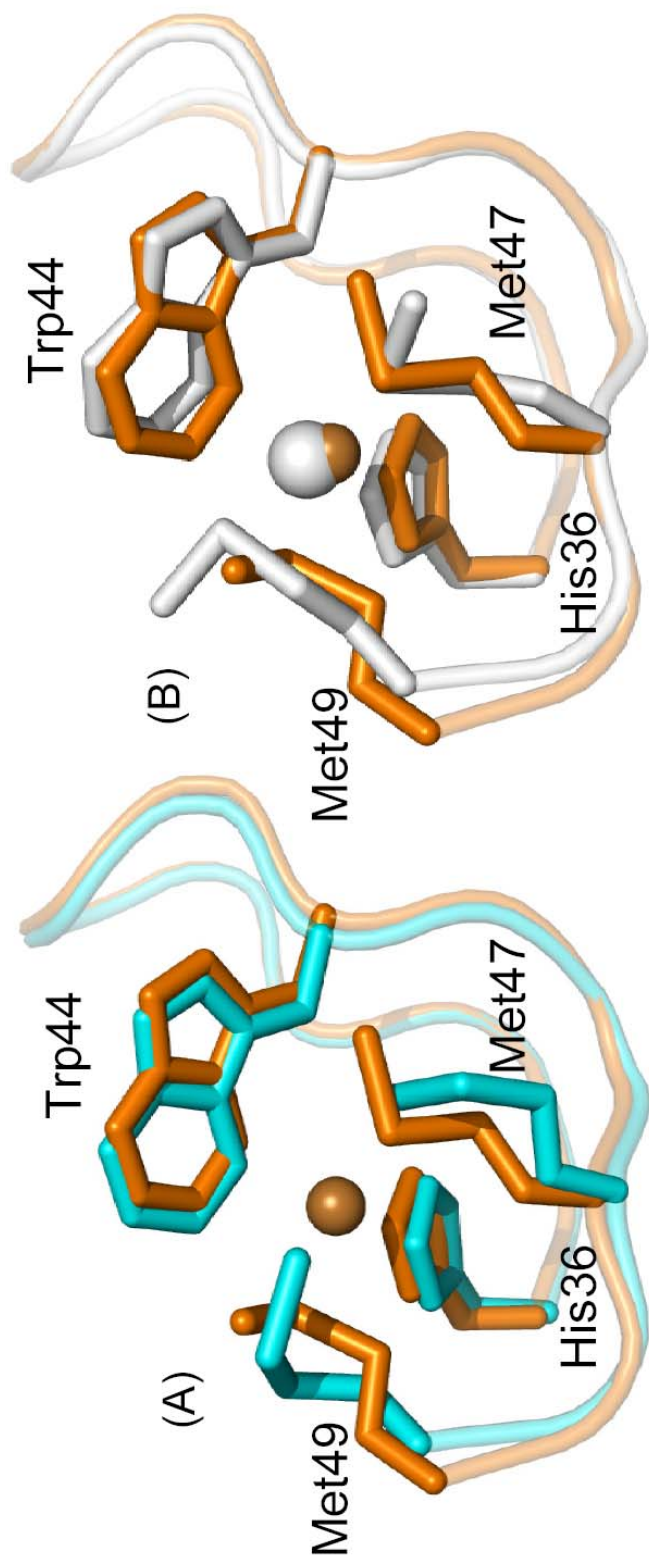
Figure 4-8. Superposition of metal binding region (resi 36-49) of (A) apo-CusF and CuCusF (B) AgCusF and CuCusF. Apo-CusF is cyan, CuCusF is bronze, and AgCusF is white. (A) reveals that upon Cu binding, Met47 and Met49 go through large conformational changes, while His36 and Trp44 only shift slightly. (B) reveals that in order to accommodate Ag ion with larger radius, Met47 and Met49 also undergo large conformational readjustments. In addition, tilting of the aromatic rings of both His36 (~10°) and Trp44 (~20°), together with displaced Ag ion (~0.5 Å from Cu ion), contribute to more favorable coordination contacts.
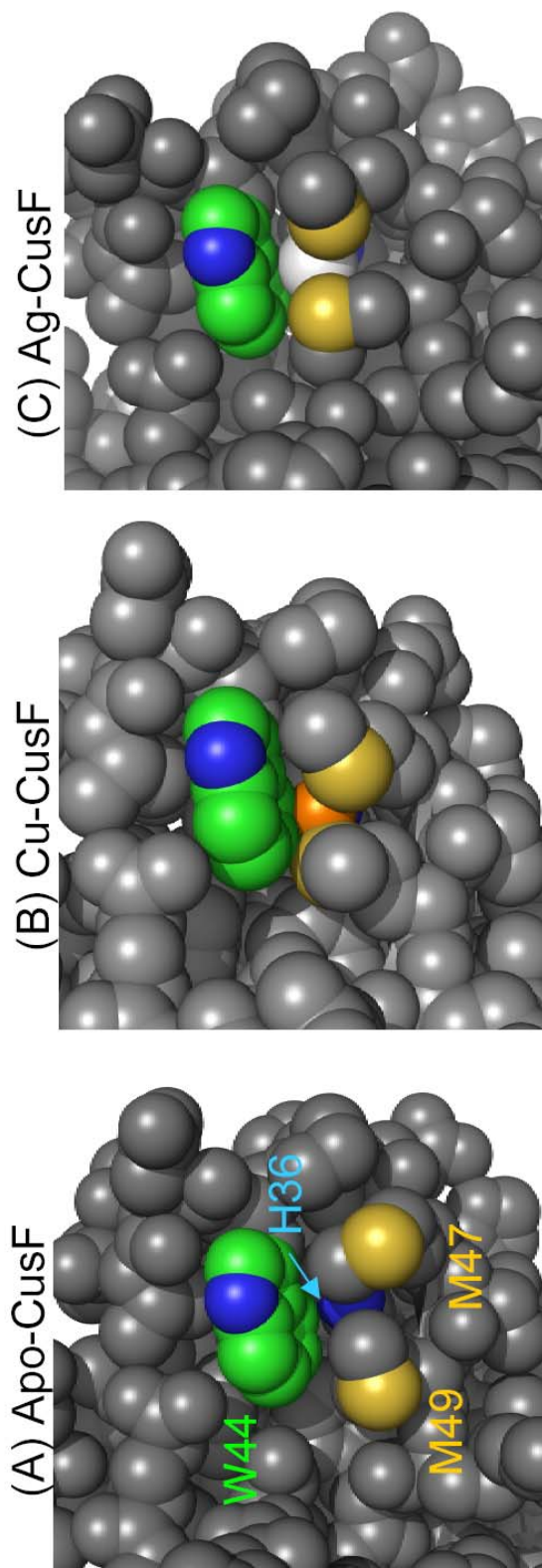
Figure 4-9. Spacefilling models of metal binding sites. The indole ring of W44, S^δ atoms of M47 and M49, N^ϵ of His36 are color coded with conventional color scheme (C: green, N: blue, S: yellow). Cu and Ag are shown in yellow and white, respectively. The rest atoms in the structures are all colored with gray. In all three models, H36 is buried inside, W44 is at surface. For apo-CusF, the S^δ atoms of both M47 and M49 are exposed, whereas in CuCusF, S^δ of M49 rotates and is totally buried, while S^δ of M47 rotates toward the metal and still keeps exposed. For AgCusF, both S^δ atoms also rotate toward the metal center. Contrary to CuCusF, S^δ of M47 is partially buried while S^δ of M49 exposes to solvent.

Discussions

Methionine-rich sequences involved in copper binding are emerging as more copper transport and homeostasis systems are being uncovered. Numerous other proteins which are involved in copper transport utilize this motif, such as the proteins from the prokaryotic *cop*, *cue*, and *pco* systems [4], which are involved in copper resistance, the bacterial copper chaperone DR1885 [19], and the eukaryotic Ctr permease copper transport proteins [20]. In the reducing cytoplasmic environment, copper chaperones, such as yeast Atx1 or Cox17, utilize cysteines as copper ligands [21]. In contrast, periplasmic proteins involved in copper binding and trafficking often utilize methionines and histidines to coordinate copper. To further understanding the Methionine-rich motifs in periplamic Cu proteins, we performed crystallographic studies of Cu and Ag bound forms of CusF protein.

The crystal structures of the metallated forms of the periplasmic Cu- and Ag-binding protein CusF reveal an unprecedented soft-metal binding coordination site in biology and underscore the copper ligating function of methionine residues in biological systems. Cu- and Ag-CusF have been crystallographically characterized as discrete mononuclear 1:1 protein-metal complex complexes. While the previously reported structures of Cu^{I},Cu^{II}-CopC demonstrated Met-rich Cu^{I} coordination sites, Cu^{I} coordination occurred at a dimer interface, so that copper ligation served to bridge two monomers [7]. Perhaps owing to the unique protection of the metal center by Trp44, CusF exhibits discrete monomeric protein-Cu^{I} interactions, suggesting that its Cu^{I} chemistry and function may be significantly different from the similar CopC and PcoC proteins [7,8].

The close proximity of Trp44 is itself an unprecedented feature in metallo-protein chemistry. With two carbons at 2.67 Å and 2.86 Å in the CuCusF structure, the structure hints at an organometallic designation. These distances are certainly at the longer edge of Cu-C bonds, and with out clear spectroscopic evidence we determined that designating Trp44 as a η^2-coordinated arene was not supported at this time. Examination of the cation-π database [18], demonstrates that the Cu-arene interaction in CusF is exceptional in its Cu-C distances. Of the cataloged published structures in the database, the structures with the closest Cu-Trp distances exhibit Cu-centroid distances of over 4 Å and Cu-C distances of at least 3.4Å to be compared to the 3.3 Å Cu-centroid and 2.6 Å Cu-C distance observed for CusF [22, 23].

Recognition and chemical description of methionine-based biological Cu(I) coordination has just begun to be explored. Yet, from yeast to mammals, the eukaryotic membrane protein Ctr1 functions as the primary cellular copper importer, relying on extracellular Met-rich domains to handle the metal ion [24, 25]. Valuable insight into the Cu(I) coordination chemistry properties of Met-rich sequences comes from the prokaryotic *pco*, *cop* and *cus* systems which seem to be able to protect redox active Cu(I) ions from aberrant chemistry in oxidizing environments. It remains to be seen the extent to which Cu^I-arene interactions also play a role in handling the soft metal ion.

CHAPTER 5

Crystallographic Studies of Atx1-Cu-Tetrathiomolybdate(TM) Complex

Abstract

Tetrathiomolybdate (TM) has been used to treat Wilson's disease, a genetic disease that causes toxic accumulation of copper. The therapeutic effect of TM can be traced back to the observed biological Mo-Cu antagonism through the Mo-S-Cu interactions. It is generally believed that TM, copper and proteins form non-bioabsorbable complex to lower the cellular copper level. To elucidate the drug binding mechanism, we choose the yeast copper chaperone as the model system and solved the structure of Atx1-Cu-TM complex. It is the first structural demonstration of TM-Cu-protein complex. The model reveals that TM is linked to the Atx1 trimer via coordination to the Atx1 bound copper. The novel Mo-Cu cluster composes of a mixture of tri- and tetra- coordinated copper atoms and one TM group. The metal cluster geometry agrees well with the previous EXAFS studies of the TM treated animal model of Wilson's disease. Thus, the atomic structure will provide great insights in understanding the molecular mechanisms of the therapeutic effects of TM, and will serve as a promising model to design more efficient and targeted TM-based drugs.

5.1 Introduction

The biological antagonism between copper and molybdenum, first observed when cattle developed copper deficiencies after being pastured on soils rich in molybdenum, has been amply documented [2-5]. However, high Mo levels themselves do not necessarily induce Cu deficiency, it also requires a high dietary S intake, furthermore, it was found that thiomolybdate forms in rumen cultures [6]. It is widely believed that the Cu-Mo antagonism involves the Mo-S-Cu interactions.

Tetrathiomolybdate (TM), a copper chelator, has been used as a therapeutic agent for treatment of Wilson's disease, which is caused by the accumulation of copper in the liver due to the dysfunction of Cu-transporting ATPase (Atp7b) [7, 8]. TM has the appropriate qualities of fast action, low toxicity and oral delivery. Recent phase III clinical data show TM is more effective than other treatments at reducing the disease's effects [9].

The dependence of tumor growth on angiogenesis [10-12] and that copper being critical to angiogenesis [13] resulted in successful trials of anticopper drugs, primarily TM, as antiangiogenic and anticancer agents [14-19]. The activity of TM suppresses the nuclear factor-κB signaling cascade [18] and reduces expression of angiogenic mediators, including vascular endothelial growth factor-1, fibroblast growth factor-1, IL-1a, IL-6, and IL-8 [18]. In the most recently published clinical trial, TM suppressed the proangiogenic factors: basic fibroblast growth factor 1, IL-6, IL-8, and vascular endothelial growth factor-1 in patients with advanced metastatic kidney cancer [20]. More clinical trials with TM in combination with other agents are under way.

The mechanism of TM lowering copper levels has been suggested to be the formation of a non-bioabsorbable ternary complex between TM, Cu and proteins through Mo-S-Cu

interactions [9, 10]. It is still unclear how does TM binds Cu and proteins. Recently, George et al. studied the isolated liver lysosomes of LEC rat, an animal model of Wilson's disease, with X-ray absorption spectroscopy [21]. Data indicated that each molybdenum has approximately three copper neighbors, while every copper has slightly less than one molybdenum neighbor. Molybdenum is four coordinated, whereas copper is more like a mixture of three-coordinate and four-coordinate species.

In order to understand the precise mechanism by which TM binds copper in the body and with the idea that the copper homeostasis always involving the transfer between the chaperone and the target [22], we started with Cu chaperone of yeast, Atx1 [23], which binds copper in the 1+ oxidation state and delivers it to the P-type ATPase Ccc2 for translocation across intracellular membranes and loading into the multicopper oxidase Fet3. We made the Atx1-Cu-TM complex and performed crystallographic studies on the tripartite complex. The structure is solved by a combination of molecular replacement and Cu-MAD. It is the first structural demonstration of the biologically related Mo-S-Cu interactions. The model reveals a very novel Mo-Cu cluster, which includes both tri- and tetra- coordinated Cu atoms. Each Mo is surrounded by three tetra-coordinated Cu atoms, the arrangement of metals is nest-shaped. There is one more Cu atom sitting just opposite to the Mo atom, which is coordinated purely by Atx1 proteins. This model can very well explain the EXAFS data of the isolated LEC rat liver lysosomes [21]. The atomic structure of Atx1-Cu-TM complex will provide insights on understanding the therapeutic effects of TM as a promising anticancer drug.

5.2 Experimental procedures

Crystallization and data collection

The Atx1-Cu-TM sample was prepared by Mr. Hamsell Alvarez with analytical gel filtration column. Mr. Chandler Robinson, undergraduate student, identified the successful crystallization condition through extensive screening process. The crystals were prepared by mixing 1 μl of sample solution and 1 μl of precipitating solution using hanging drop method. The ratio of [Atx1]:[Cu]:[TM] in the sample was analyzed to be 1.00 :0.97 :0.35. The concentration of the sample was about 3.0 mM in the buffer of 20 mM MES, pH 6.0. The precipitating solution consisted of 0.15 M DL-malic acid, pH 7.0 and 20% PEG3,350. Deep purple crystals (Figure 5-1) appeared in one week and reached the maximum dimension of 0.15 mm × 0.06 mm× 0.04 mm. Crystals were soaked in the cryosolution of precipitation solution plus 25% glycerol for three minutes and flash-cooled in liquid nitrogen before data collection.

Native data was collected to 2.30 Å resolution at the APS SBC-CAT beamline. Multiple-wavelength anomalous dispersion (MAD) data were collected at the copper edge at the APS IMCA-CAT beamline on the same crystal. Three data sets were collected at peak, inflection and low remote wavelengths. All data were processed with Denzo and Scalepack [24]. A summary of the data-collection statistics is given in Table 7-1. The crystal belong to spacegroup $P2_12_12$ with possible 10~16 monomers per asymmetric unit based on solvent content analysis (assuming the solvent content ranges from 40% to 65%).

Structure determination

The molecular replacement (MR) was performed with program *Phaser* [25] using the high resolution (1.02 Å) Hg-Atx1 structure (PDB code: 1CC8) [26] as the search model. All water,

hetero-ligands and Hg were removed from the search model. The initial MR strategy was to find the top 10 solutions in each round's search, and feed them to start the next round's search. Due to the enormous calculation amount, and the high copy numbers (10~16) of the monomers per asymmetric unit decreasing the solution signals, all trials with different settings of search parameters stuck at the 4th monomer and could no longer to proceed to find the 5th one after running for 2~3 days. The partial model did not make much structural sense, and rigid body refinement showed the high R factors comparable to those of a random structure, which suggested the MR failed. We also employed Cu-HAH1 [27] (PDB code: 1FEE) as the search model, which was an Atx1 homologous protein in human and formed dimer with copper bound. However, the MR search was also unsuccessful.

The final successful MR strategy was straightforward and simple: each round, find the top solution; fix the solution(s), find the next one. Using the Atx1 monomer model, it found all 12 copies in six hours. Rigid body refinement including the data of 30~3Å showed R and R_{free} of 34% and 35%, clearly indicating the right solution identified. The 12 Atx1 molecules were arranged as four trimers. Each trimer enclosed one large electron density blob in the Fo-Fc map, which revealed the positions of the TM-Cu clusters (Figure 5-2).

Since the native data only diffracted to 2.30 Å, the resolution was not high enough to build the density blobs without the prior knowledge of the coordinates of copper atoms. With the Cu-MAD data, it should be able to provide the positional information of the copper atoms.

Initial evaluation of Cu-MAD data with program mmtbx.xtriage [28] and shelxc [29] suggested the anomalous signals were only good to around 5.7Å. Heavy atom search using peak data found four major sites (occupancy > 80%) right in the four density blobs with program

phenix.hyss [30]. It was probably due to the resolution of the anomalous signals was only good to 5.7Å, which was too low to resolve the copper atoms with distances of 3~4 Å, thus only giving the overall positions of the Mo-Cu clusters.

On the other hand, since the phases of the model had already being determined using MR, it was feasible to generate the anomalous difference map using the phases from MR model and the anomalous signals in the peak data. All data (50~2.30 Å) were included for the calculation. The anomalous difference map clearly showed total 16 peaks with a pattern of four peaks clustered together. In each 4-peak cluster (Figure 5-3), three of them were close the metal binding loop regions of the surrounding Atx1 molecules, one of them sit at the top of the other three peaks. Based on the fact that Mo:Cu ratio of crystallization sample being 1:3, we assumed in each cluster, three bottom peaks were copper atoms, one top peak was molybdenum atom.

We imported the 12 positions of assumed copper atoms and the three MAD data sets into program sharp/autosharp [31, 32] for heavy atom refinement, phasing and density modification. Two more sites corresponding the assumed Mo atoms were found. After protein phasing and solvent flattening, the electron density map showed clear secondary structural features.

Model building started with the experimental phases using the MR model in COOT [33]. The cysteines involved in copper binding were initially mutated to alanines. Obvious wrong rotamers and wrong placement of carbonyl of main chain were fixed. Restrained refinement of the coordinates and the isotropic B factors was performed using REFMAC5 [34] in CCP4 suite. The metal clusters were built in three steps: first, the coordinates of the refined copper atoms from sharp/autosharp was directly imported into the model; second, the mutated alanines were corrected back to cysteines and built carefully according to the difference density; third, after

finishing previous two steps, a clear tetrahedral shaped difference density was perfectly fit by

a MoS_4^{2-} moiety (Figure 5-4). The metal clusters were refined with the bond distances restraints

derived from the EXAFS studies (Cu-S: 2.29 Å; Mo-S: 2.25 Å; Cu-Mo: 2.77 Å) (Appendix 3). In

the process of model building, refinement started with tight NCS (non-crystallographic

symmetry) restraints on the 12 Atx1 molecules. With the improvement of the model, the

restraints were released to medium, and four NCS groups were defined with each consisted of

the Atx1 trimer surrounding the Mo-Cu cluster. Finally, TLS refinement was carried out with

four TLS groups (each group being one trimer). Compared with restrained refinement, both R

and R_{free} decreased to nearly 3%. Refinement and model statistics are listed in Table 5-1.

The structure was validated with MOLPROBITY [35]. Ramachandran plot (Figure 5-5)

showed that all residues were in allowed regions and 99.0% were in favored regions.

All protein diagrams were generated with program Pymol [36].

Table 5-1. Crystallographic Statistics

Values in parentheses are for the highest resolution shell.

Data collection and Statistics		Cu-MAD		
Data set	Native	Peak	Inflection	Low remote
Beamline[1]	19-BM	17-ID-B	17-ID-B	17-ID-B
Wavelength (Å)	0.9787	1.3799	1.3805	1.3850
Resolution limits (Å)	50.0 - 2.30	50.0 - 2.30	50.0 -2.50	50.0 - 2.72
Completeness (%)	97.0 (84.0)	94.3 (64.0)	99.1 (92.4)	99.7 (98.5)
Data redundancy	5.0 (2.9)	3.3 (1.4)	3.6 (2.0)	3.7 (3.0)
R_{sym} (%) [2]	8.1 (47.0)	8.3 (34.2)	6.9 (51.6)	6.8 (54.0)
$I/<\sigma>$	16.7 (2.0)	11.9 (1.7)	15.5 (1.6)	17.1 (2.1)
Space group	$P2_12_12$			
Unit-cell parameters (Å °)	109.9 182.2 52.7	90.0 90.0 90.0		
Phasing Statistics	Peak	Inflection	Low remote	
Phasing Power (iso) [3]	0.44	1.4	N/A	
Phasing Power (ano) [3]	1.5	0.86	0.17	
Cu sites found	14			
FOM after SHARP	0.19			
FOM after SOLOMON	0.64			
FOM after DM	0.86			
Refinement Statistics				
R_{work}/R_{free} (%) [4]	21.3 / 26.5			
R.m.s. bond lengths (Å)	0.014			
R.m.s. bond angles (°)	1.393			
Protein atoms	6746			
No. of water molecules	372			
Residues in Ramachandran plot regions (%) [5]				
Most favoured	99.0			
Additional allowed	100.0			

[1] All data sets were collected at 100 K at the Advance Photon Source.

[2] $R_{sym} = \Sigma |I_{obs} - I_{avg}| / \Sigma I_{obs}$

[3] The phasing power is defined as the ratio of the rms value of heavy atom structure factor amplitudes to the rms value of the lack-of-closure error.

[4] $R_{values} = \Sigma |F_{obs} - F_{calc}| / \Sigma F_{obs}$, 5% of the reflections were reserved for the calculation of R_{free}

[5] The validation was performed by program MOLPROBITY [33]
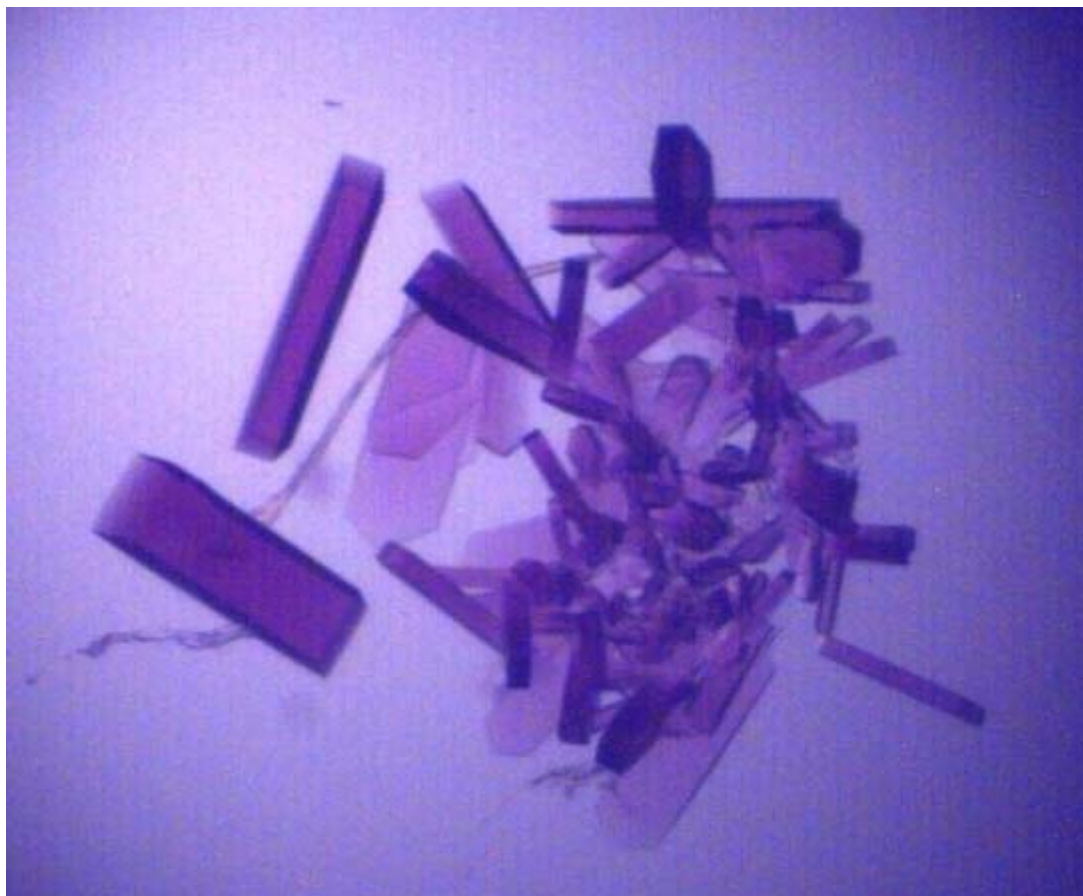
Figure 5-1. Deep purple crystals of Atx1-Cu-TM grown in hanging drops
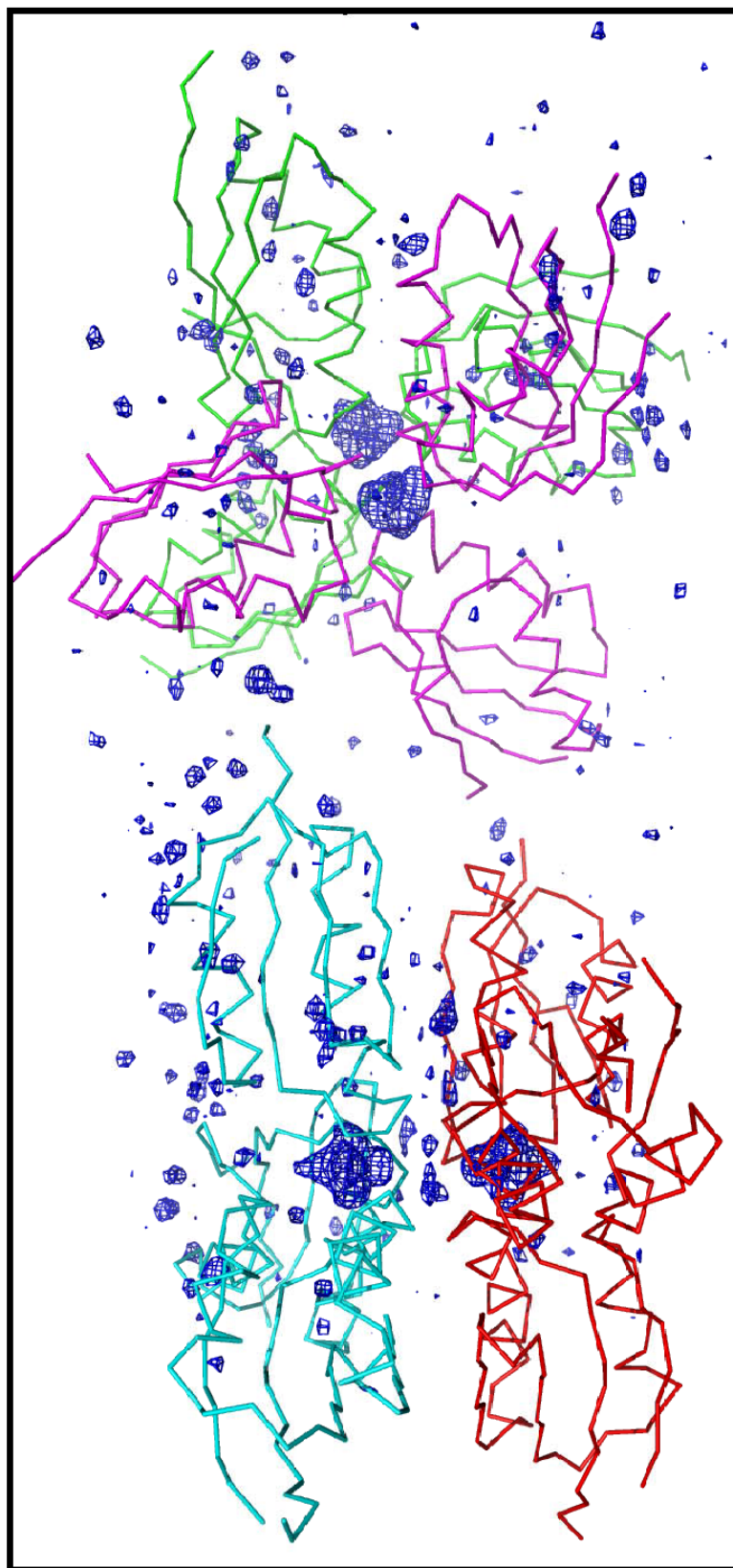
Figure 5-2. Fo-Fc map showing the positions of Mo-Cu clusters. 12 Atx1 molecules are color-coded in groups of trimer. Each trimer encloses one large electron density blob, which indicates the position of the Mo-Cu cluster. The map is contoured at 3σ level.
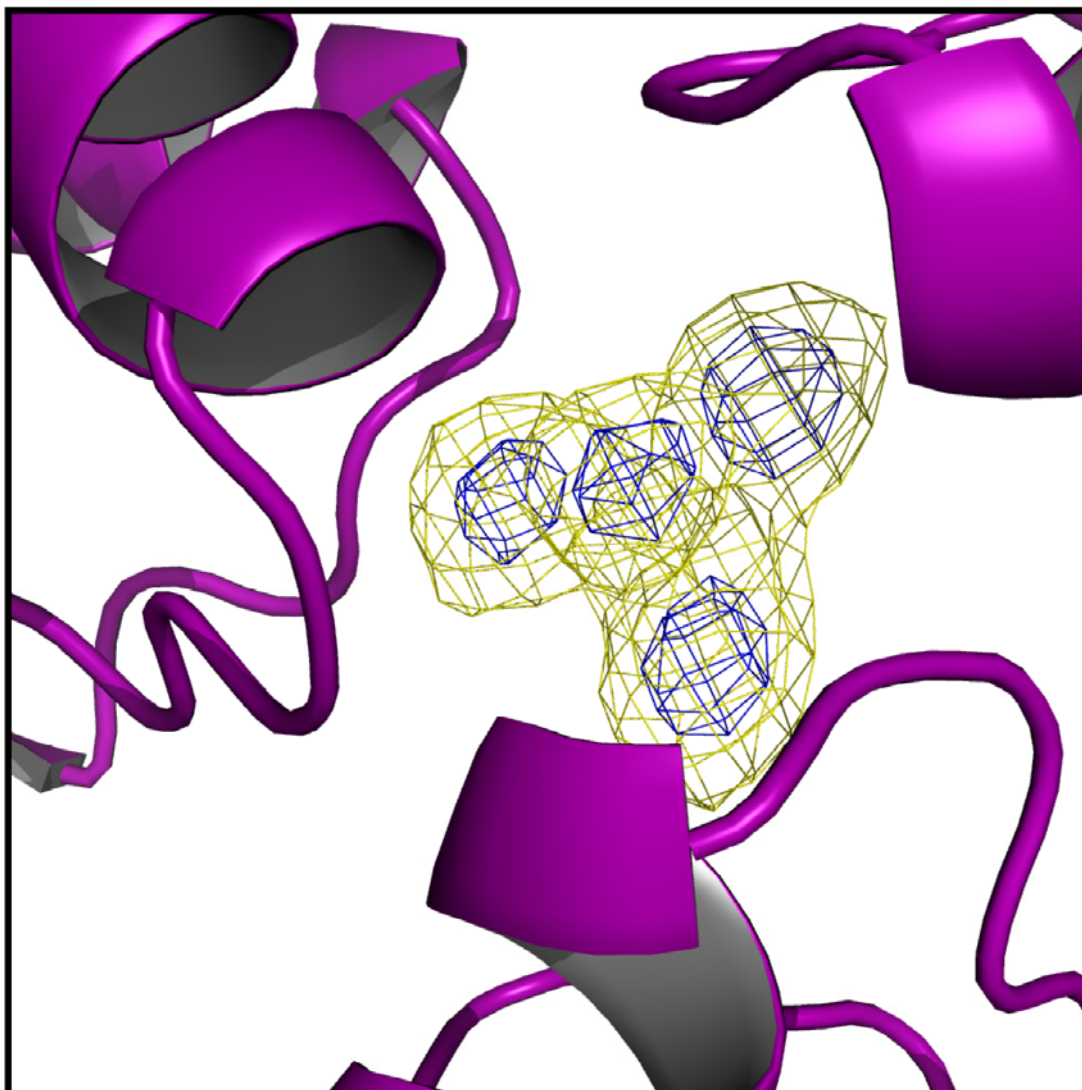
Figure 5-3. Anomalous map generated by combining the phase of the MR model and the anomalous signals in peak data set. There are total 16 peaks with each four peaks clustered together in the identical way. The figure shows one of the clusters. The yellow mesh and blue mesh are contoured at 4σ and 10σ levels, respectively. The peaks are suggestive of the copper sites. Three out of the four peaks are close the metal binding loops of the Atx1 trimer, indicating the copper atoms bind to the proteins, the rest one sits upon the other three, the coordination is unclear at this point.
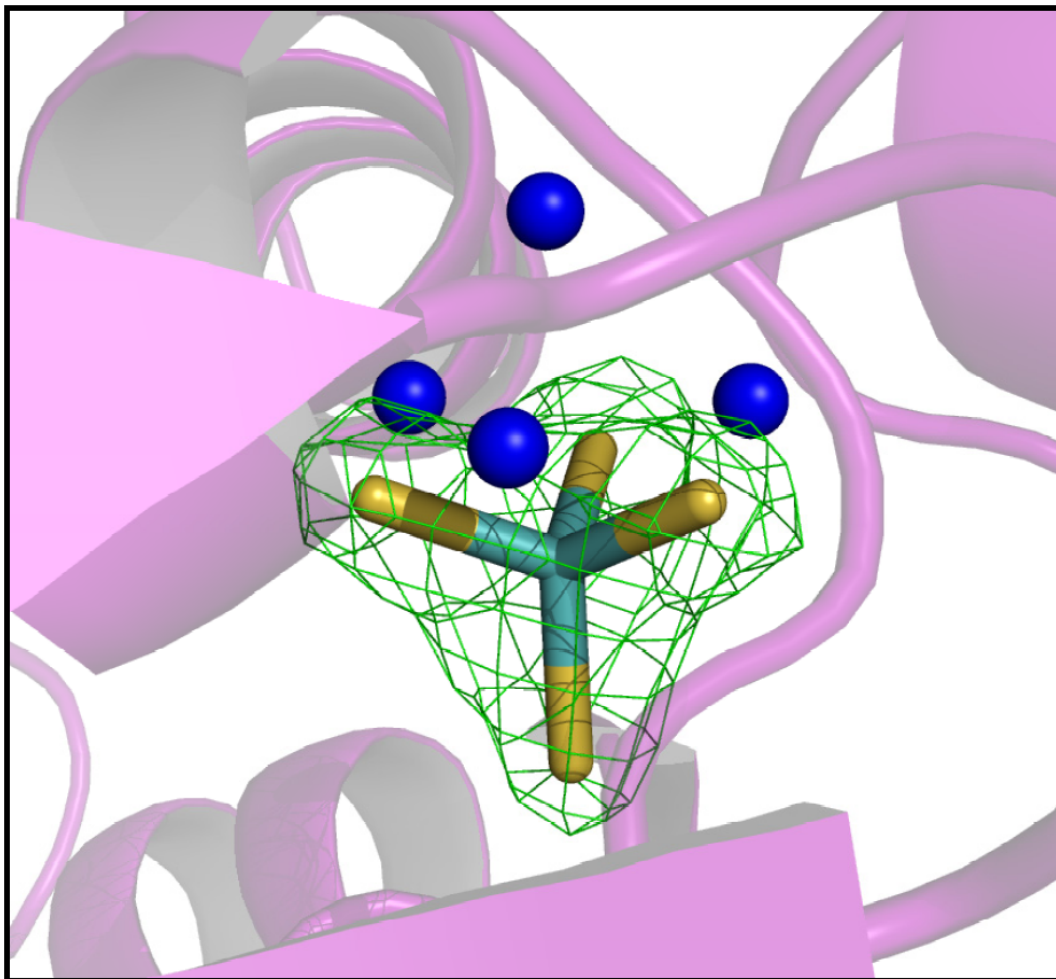
Figure 5-4. Fo-Fc map contoured at 4σ level after building copper atoms and ligated cysteines. The tetrahedrally shaped green density unambiguously reveals the position of TM (MoS_4^{2-}) and fits perfectly of the drug. Proteins are shown in purple ribbon diagram. Copper atoms are shown as blue spheres. TM is represented by stick diagram with sulfurs in yellow and molybdenum in cyan.
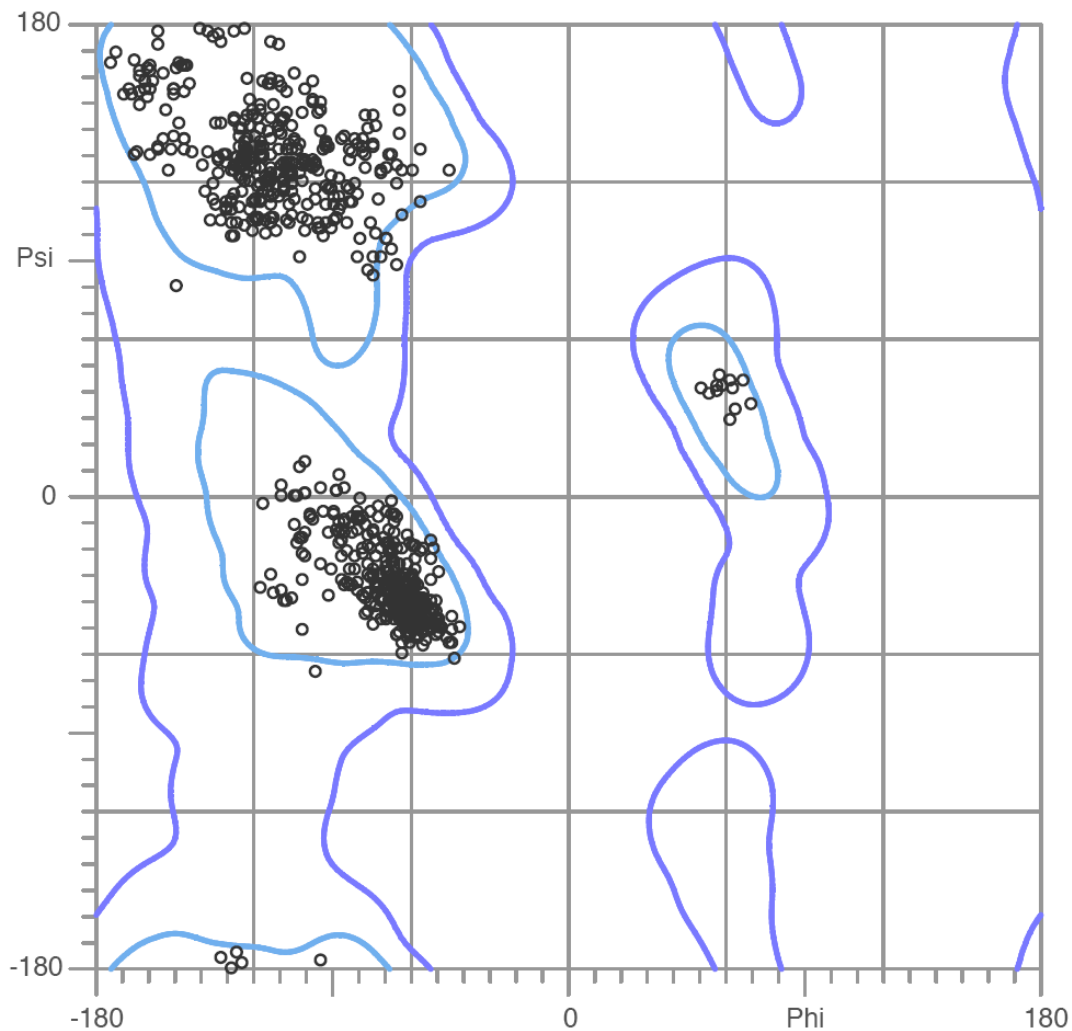
Figure 5-5.   Ramachandran plot of Atx1-Cu-TM model generated by MOLPROBITY [33].
The regions in the blue line are the generally allowed regions, whereas the regions in the aqua blue line are the favored regions. 100% residues are in the generally allowed regions and 99.0% residues are in the favored regions.

5.3 Results and Discussions

Overall Structure of Atx1-Cu-TM complex

As described previously [26], the overall structure of Atx1 comprises a 'ferredoxin-like' [37] $\beta\alpha\beta\beta\alpha\beta$ fold (Figure 5-6). The hydrophobic core of the protein is composed of residues Phe9, Val11, Met13, Val22, Leu26, Leu29, Val33, Ile36, Ile38, Val45, Val47, Ile56, Ile60 and Val67. The two cysteines involved in metal binding are located at the junction of the first β strand and the first α helix and the metal binding site is exposed at the protein surface.

There are total 12 Atx1 molecules per asymmetric unit. They cluster in two hexamer groups. Each hexamer is composed of two parallel layers of trimers related by a 3-fold non-crystallographic symmetry (NCS) axis (Figure 5-7). The two 3-fold NCS axes orient against each other at a degree of approximate 66°. The lower layer trimer is twisted by ca. 33° clockwisely against the upper layer trimer, and the relative orientation of α helices and β strands in Atx1 molecule in the two layers of trimers are reversed (Figure 5-8). On the other hand, observing from the side, the hexamer group is also related by three 2-fold NCS axes which are perpendicular to the 3-fold NCS axis (Figure 5-8).

In the 12 Atx1 molecules (chains), chain L has disordered C-terminus including residues 70-73, chains A, I and H miss one or two N-terminal residues.

Analysis of the model suggests there are few contacts between the chains. The inter-chain H-bonds in the ABC trimer exist only between A/Arg68/NH2 and C/Lys62/O and between B/Arg68/NH2 and A/Lys62/O (chain/resi number/atom) (Figure 5-9). No indirect H-bonds via water are observed. The van der Waals contacts are present mainly in the metal binding loop and the loop region connecting the α helix 2 and the β strand 4. Between the two layers of trimers,

use chain B as the example for inter-chain analysis with chains from the other trimer, the H-bond contacts include: the sidechain of Glu41 from one chain (B or F) with the sidechains of Glu41, backbone amide of Leu40 and backbone carbonyl of Ile38 from the other Chain (F or B); the sidechain of Ser16 from chain B with the backbone amide of Ser16 from chain F (Figure 5-10). The van der Waals contacts exist mainly in the metal binding loop and β strand 2 between chain B and chain F. For the two hexamers, interactions exist only between B/Ala2 and H/Gln72 (Figure 5-11). In summary, there are no substantial protein-protein interactions in the structure. The H-bonds area and the van der Waals contacts region virtually superpose.

Copper sites and anomalous peaks

The positions of copper atoms could not be directly derived from Cu-MAD data due to the low resolution of the anomalous signals (~5.7 Å). A smart alternative way is to combine the phases from the MR model and the anomalous signal from the peak data (all data included) to generate the anomalous map, which shows clearly 16 outstanding peaks (Figure 5-3). Utilizing the anomalous signals from edge, low remote and even native data sets also show the similar results, which indicates the peaks are universally present and real. The result reveals a very interesting fact: although the anomalous signals beyond the cutoff resolution, as defined by the data evaluation programs, cannot be detected due to noise, they do exist.

Initially, we assumed that there were three copper atoms and one molybdenum atom in the structure based on the X-ray absorption spectroscopic data and the ratio measurement of the sample (unpublished data by Mr. Hamsell Alvarez). The 16 peaks arrange in four clusters with four peaks each cluster. Therefore, we thought the three peaks near the metal binding loops of Atx1 molecules should be copper atoms, and the peak above the three copper peaks was

molybdenum atom. However, with the model building and refinement, it turns out the 'Mo' atom is trigonal-planarly coordinated by the cysteines from Atx1; furthermore, the height of the anomalous peak of the 'Mo' atom is quite comparable to those of Cu atoms, at the wavelength of the MAD data collection, the anomalous signal of Mo should be only half of that of Cu. Thus, we confirm that the 'Mo' atom is actually another 'Cu' atom which we did not expect initially. After extensive model building and refinement, the final positions of copper atoms show little shifts compared with the initial anomalous peaks (Figure 5-12), suggesting the good consistency between the model and the data.

Cu(I) bound Atx1

Atx1 and its homologues contain a conserved N-terminal GMXCXXC (single-letter amino acid code where X is any amino acid) sequence motif that binds metal ions with the two cysteines. The structures of Hg(II)-Atx1 was solved at atomic resolution, whereas the crystals of Cu(I)-Atx1 did not contain an ordered, fully occupied copper site, which could be attributed to the crystallization conditions [26]. Thus, the exact nature of the copper-binding site in Atx1 remains unresolved.

Here, with the binding of the drug TM, we successfully obtain the Cu(I) bound form of Atx1. The oxidation state of the copper atoms is confirmed X-ray absorption spectroscopy studies (unpublished data by Mr. Hamsell Alvarez). Similar to Hg(II), Cu(I) is also bound by the two conserved cysteines (Cys15 and Cys18) in the metal binding loop. Hg(II) adopts a nearly linear coordination geometry with S-Hg-S bond angle of 167 °. While in Atx1-Cu-TM complex, Cu(I) is tetrahedrally coordinated with two cysteine S atoms from Atx1 and two S atoms from TM. Compared with Hg, Cu is more exposed to the surface (Figure 5-13A). The Cys15/S-Cu-

Cys18/S bond angle of ranges from 118~ 124 °. The nearly 120 ° bond angle suggests that Cu(I) may adopt trigonal coordination in Cu-Atx1, as revealed by the EXAFS data of Cu-Atx1: Cu(I) binds to two sulfur ligands of 2.25 Å and a third ligand, postulated to be an exogenous thiol, at a distance of 2.40 Å [23] . Superposition of the coordinates of Hg-Atx1 and Cu-Atx1-TM (chain B) shows an rms deviation of 0.56 Å for Cα positions, Cys15 and Cys18 shift by 0.47 Å and 0.59 Å, respectively (Figure 5-13). The similar metal binding loop region suggests: (1) the Hg-Atx1 does serve as a good structural model for CuAtx1; (2) the binding of Cu by Atx1 in Atx1-Cu-TM is not disturbed by TM and probably reflects the real Cu-Atx1 coordination. Likewisely, the metal binding loop of Cu-Atx1-TM (chain B) also superposes well with the corresponding region of Cu-Hah1 (Figure 5-13B), the Atx1 homologoue in human, however, the metal coordination is different. In Cu-Atx1, EXAFS indicates trigonal Cu^I coordination geometry, while the Cu^I-$(Hah1)_2$ is tetrahedrally coordinated with one ligand at longer distance across the homodimer interface. Collectively, the three structures of Hg-Atx1, Cu-Atx1 and Cu-Hah1 exhibit very similar metal binding loop conformations but accomodate quite different coordination geometries. These different coordination conformations have been suggested to represent the different stages in the metal transfer between the copper chaperone and its target (Figure 5-14) [38]. The observation of the different metal coordination geometries requiring little rearrangement of the metal binding loop, conforms to the fact that shallow thermodynamic gradient involved in the copper transfer between Atx1 and Ccc2 [39], and facilitates their major function as metallochaperones.
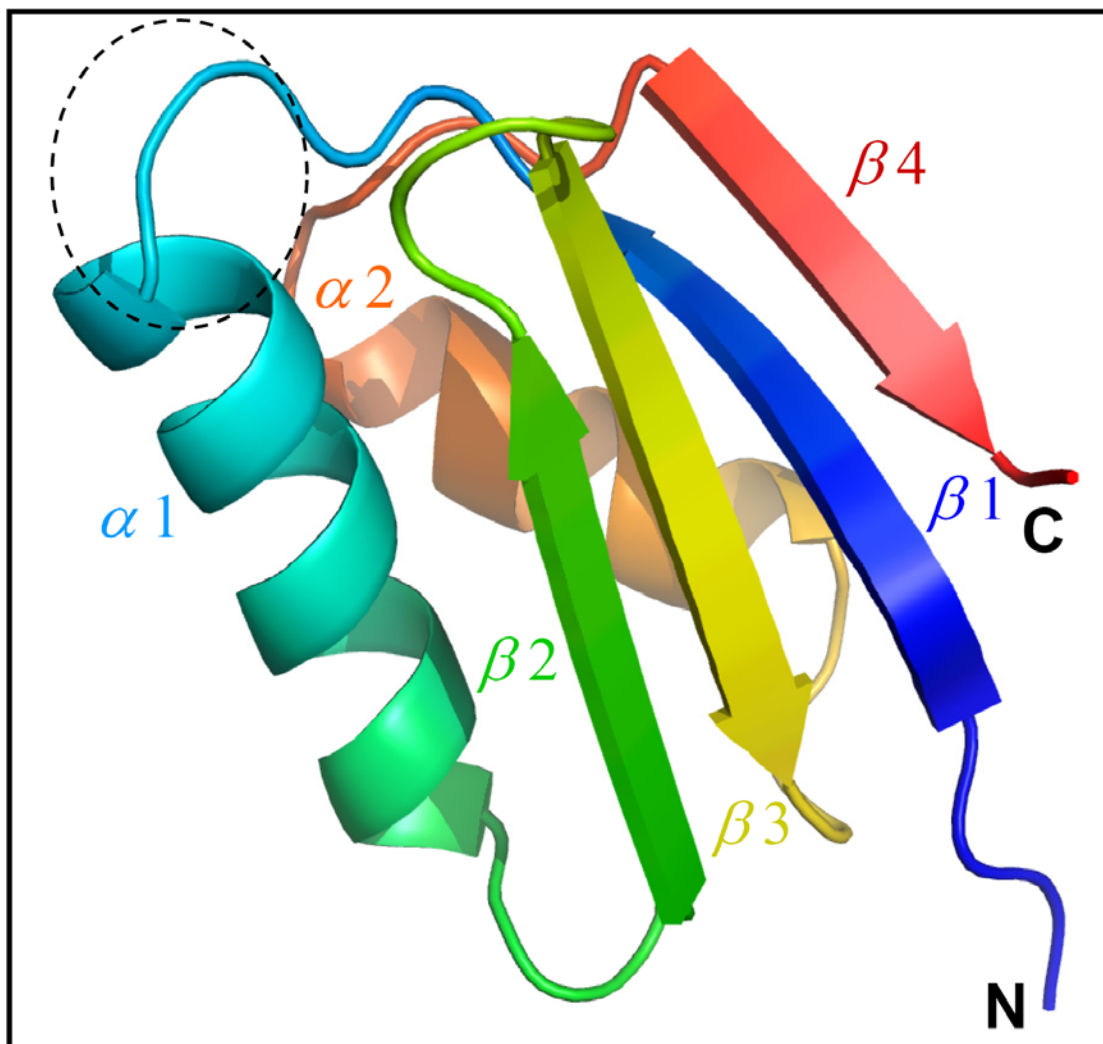
Figure 5-6. Overall structure of Atx1. It comprises a 'ferredoxin-like' $\beta\alpha\beta\beta\alpha\beta$ fold. The loop region in the dotted round circle is where metal binds.

148
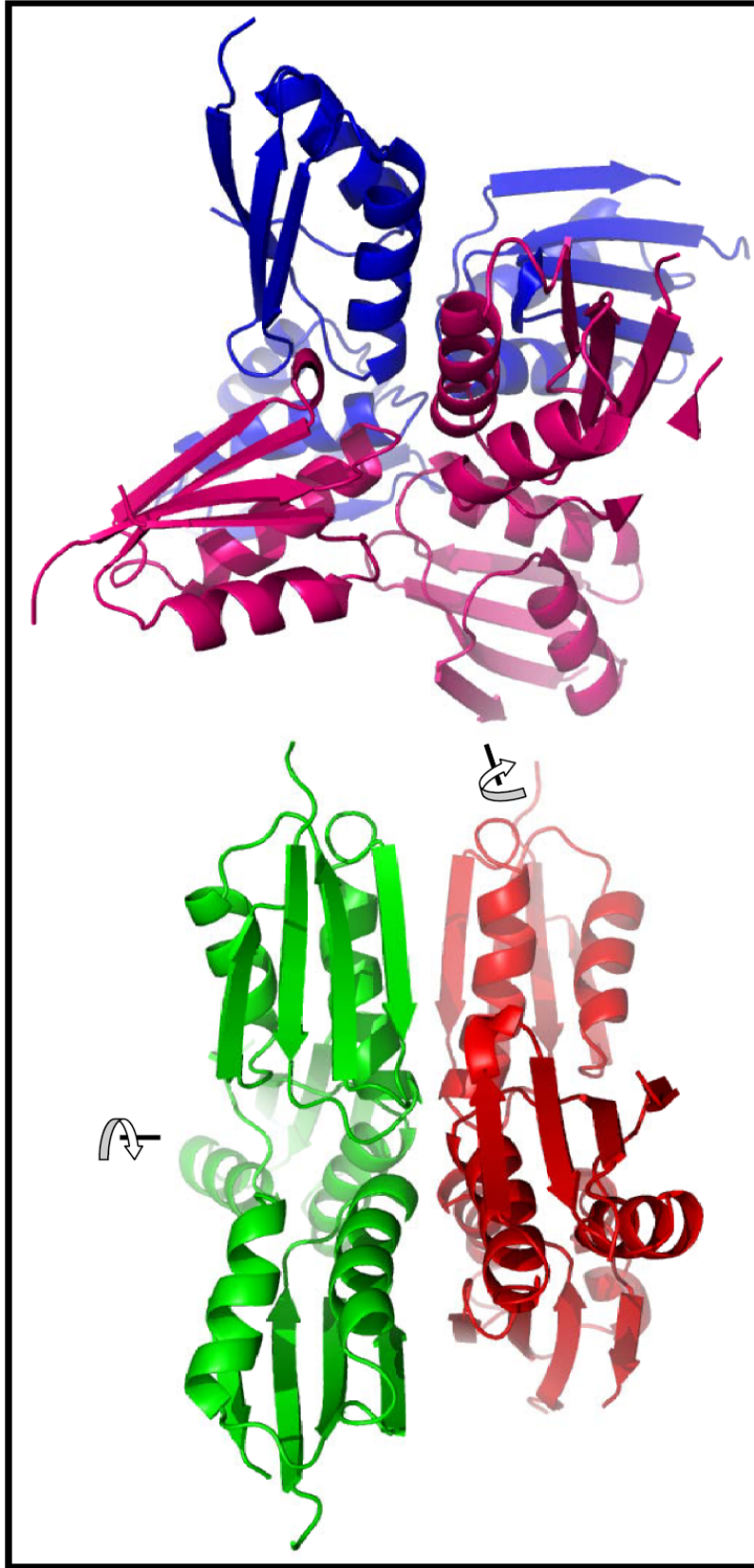
Figure 5-7. Overall structure of Atx1-Cu-TM. The 12 Atx1 molecules cluster in two hexamer groups. Each hexamer is composed of two parallel layers of trimers related by a 3-fold non-crystallographic symmetry (NCS) axis. The Atx1 molecules are color-coded into four trimers. The two short lines and arrows indicate the 3-fold axis. Copper and TM are not shown in the diagram.
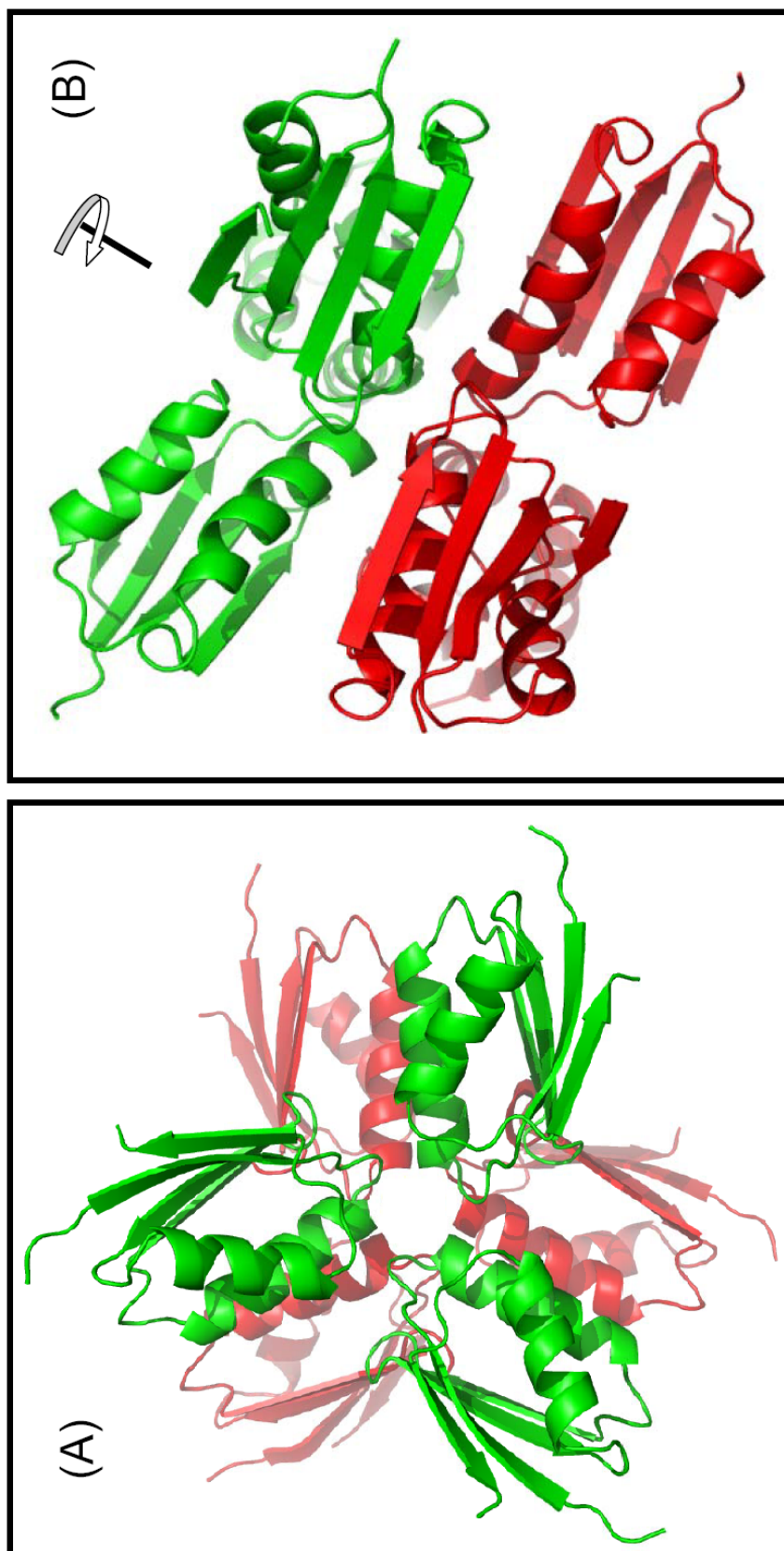
Figure 5-8. The Non-crystallographic rotational symmetries of one hexmaer group. (A) viewed from 3-fold NCS axis. (B) viewed from one 2-fold NCS axis. The 3-fold NCS axis is indiated with arrow and line. There are total three 2-fold NCS axes, only one is shown here.
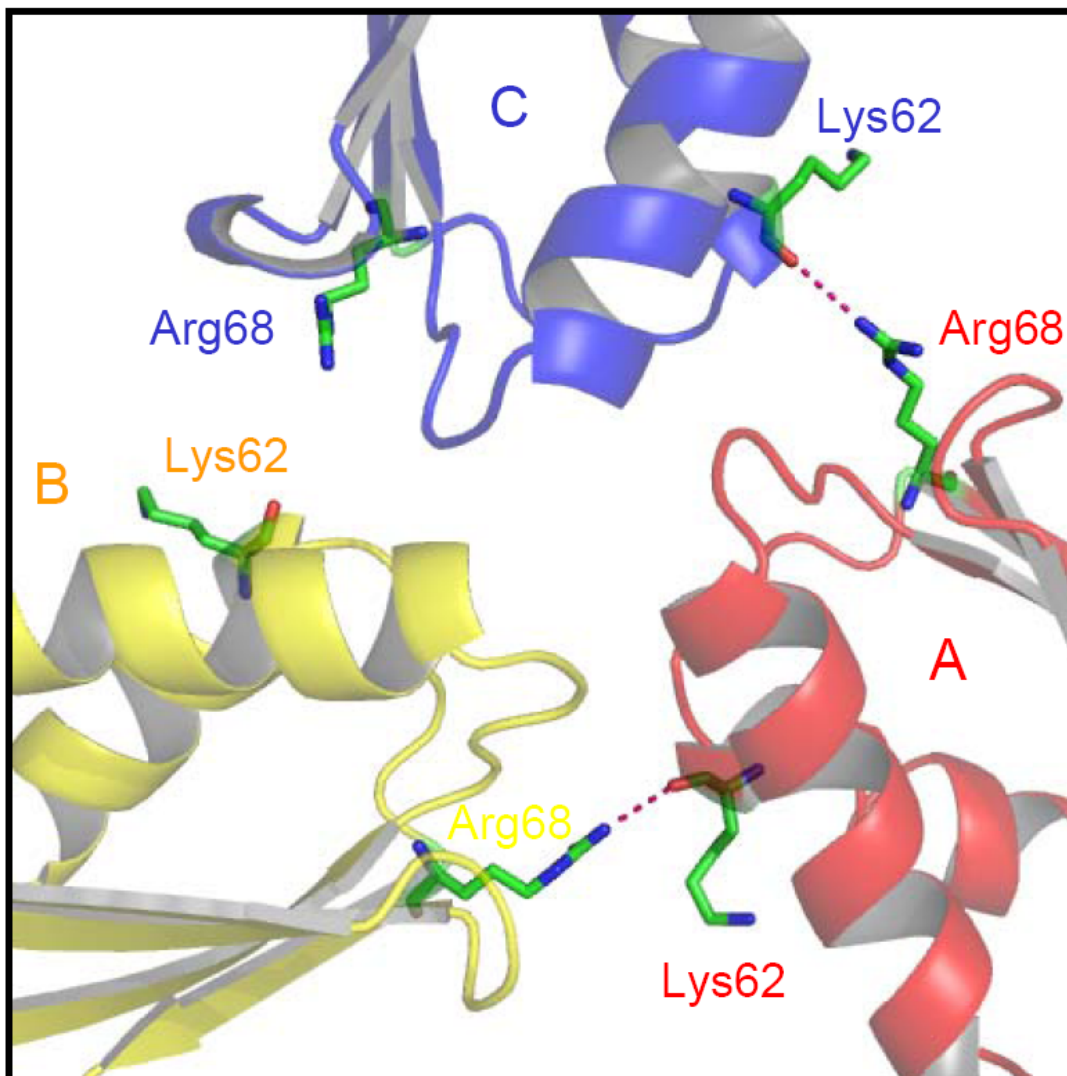
Figure 5-9. The inter-chains H-bonds in the trimer (ABC). Proteins are color-coded in ribbon diagram. The H-bonds are represented by purple dotted line. The H-bonds involving Arg68 and Lys62 are present only at two interfaces (B/A and A/C), which applies to all four trimers in the structure.
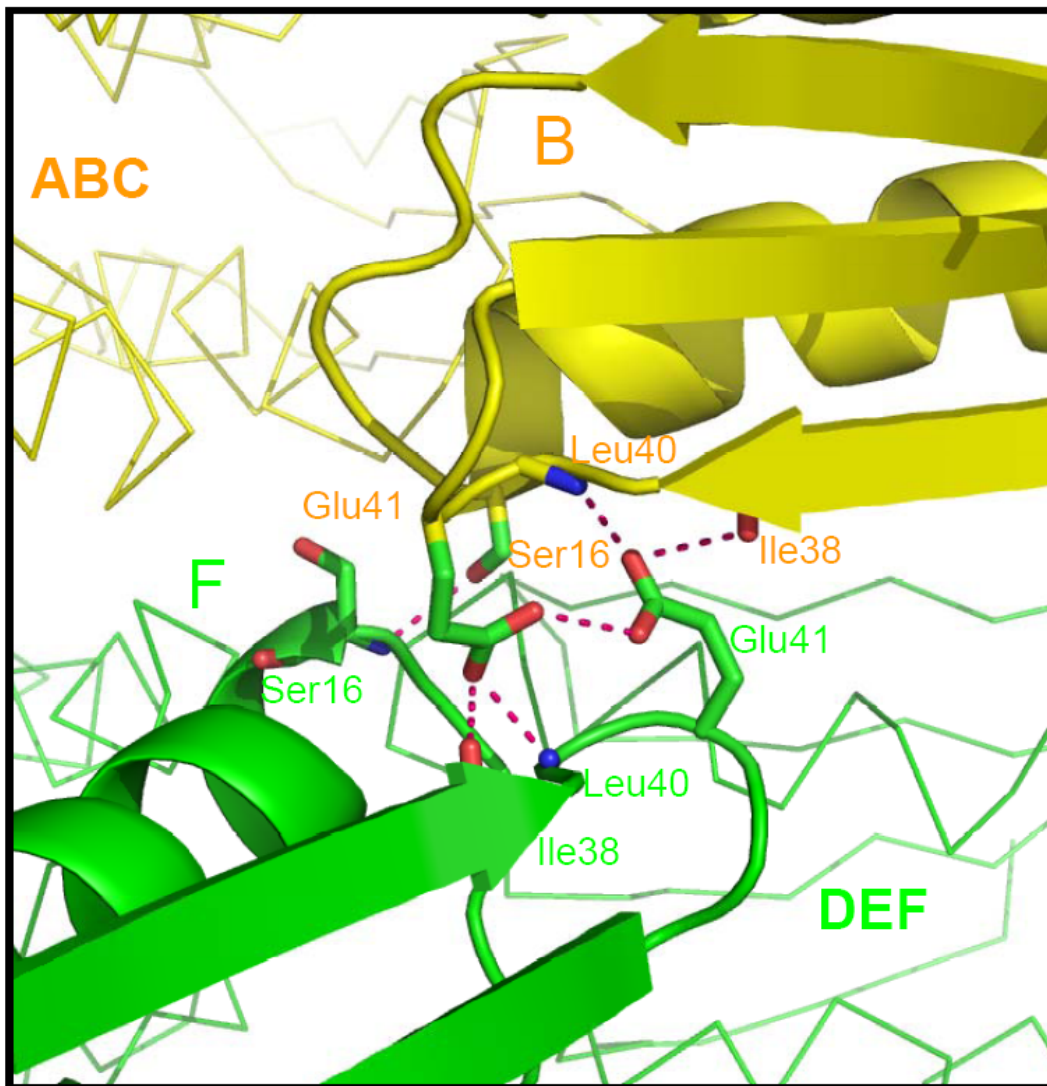
Figure 5-10.   The H-bonds between two layers of trimers. Each monomer from one trimer has H-bonds to only one monomer from the other trimer. For clarity, the chains B and F are shown as the demonstration example, and represented by cartoon ribbons, the rest two monomers in the trimers (ABC and DEF) are shown in Cα trace. The H-bonds are represented by purple dotted line. The sidechains of Leu40 and Ile38 are not shown.
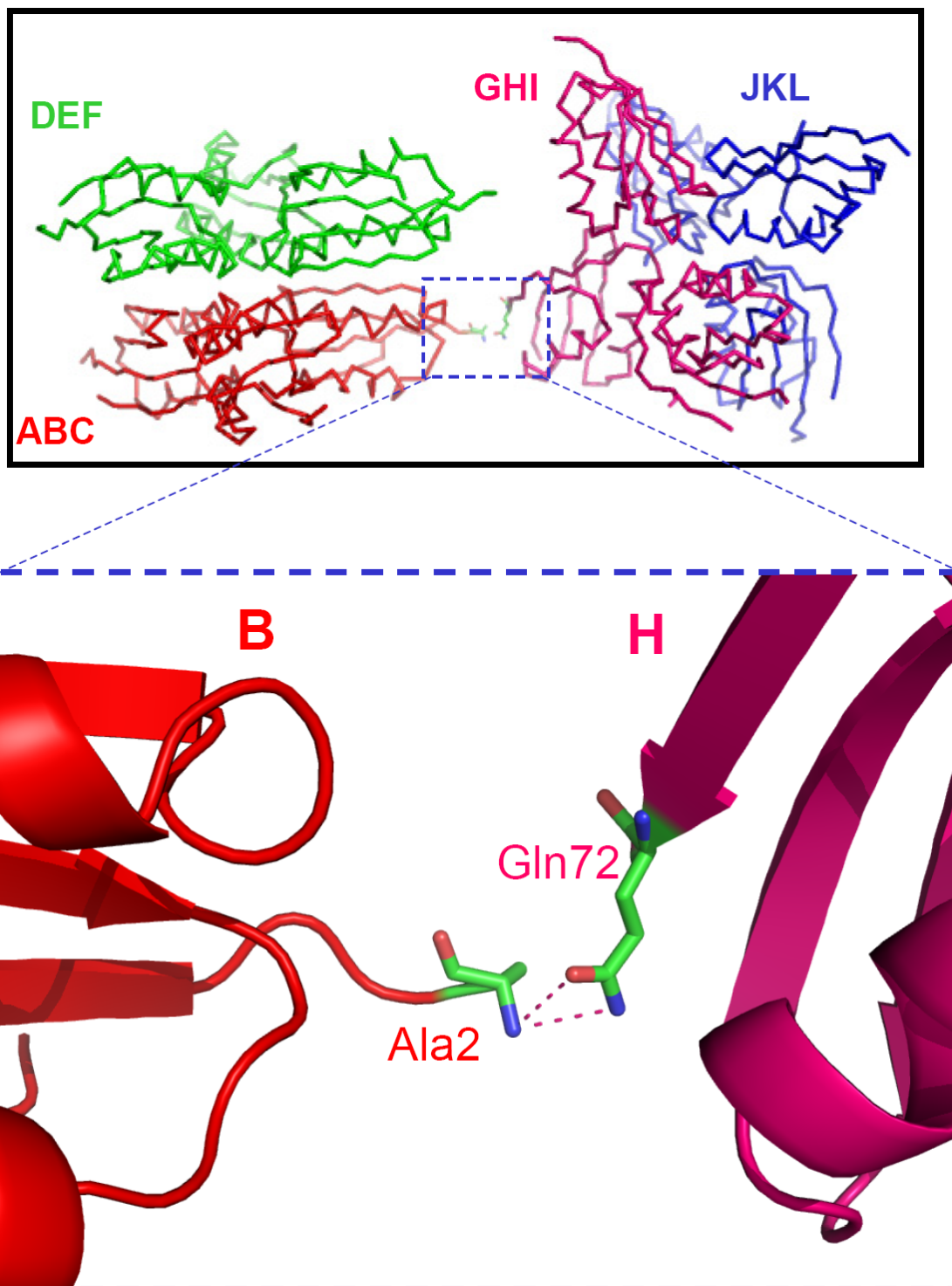
Figure 5-11. The H-bonds between the two hexamers. The H-bonds are represented by dotted line. The interactions between two hexamers only present between B/Ala2 and H/Gln72.
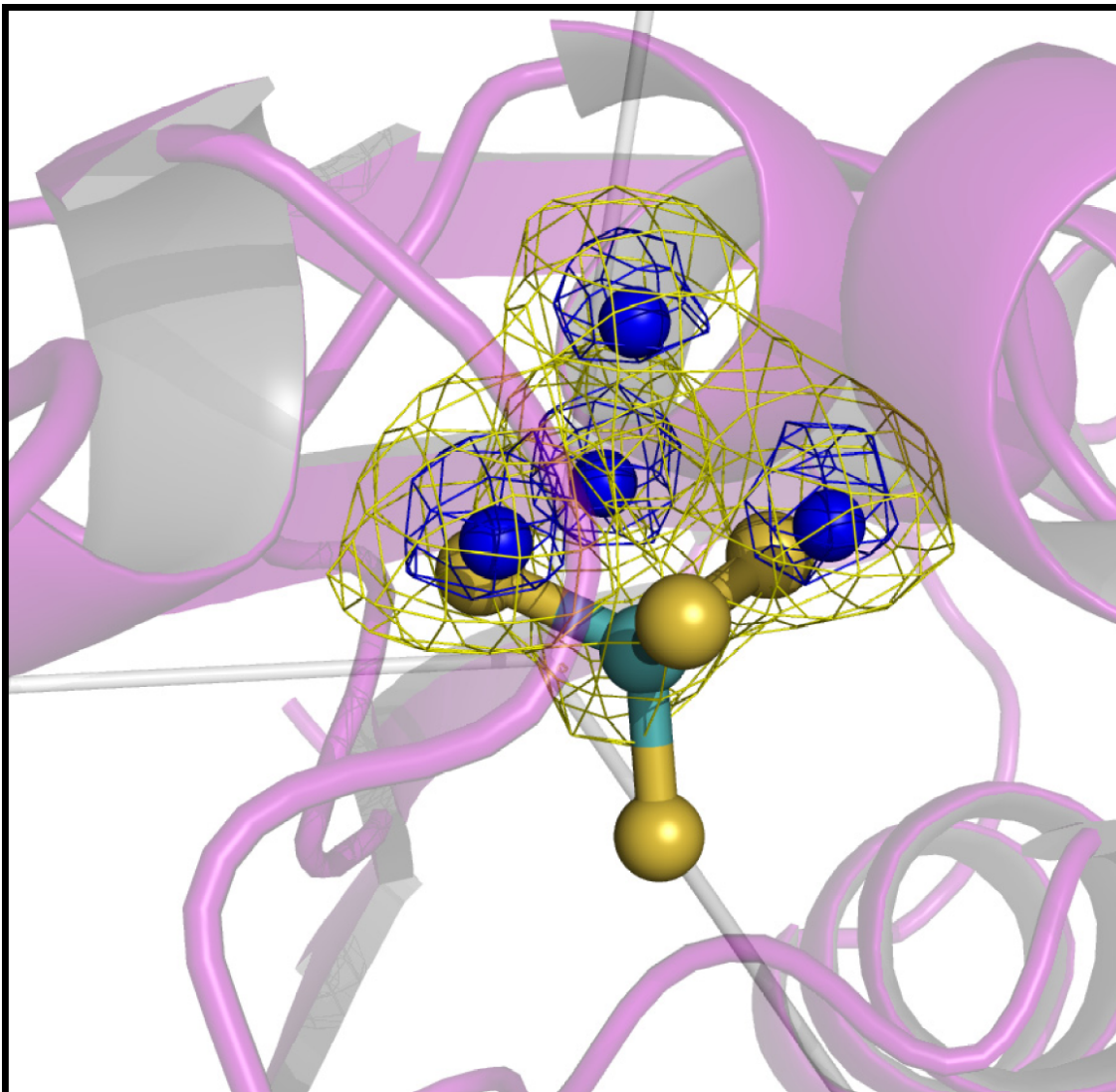
Figure 5-12. The initial anomalous peaks and the copper atoms in the final model. The copper atoms are shown in blue spheres, TM are shown in spheres and sticks, with sulfurs in yellow color and molybdenum in cyan color. The yellow mesh and the blue mesh of the anomalous peaks are contoured at 4.0 σ and 10.0 σ levels, respectively. The copper atoms are lcoated right in the peaks, suggesting the initial anomalous map is very accurate.
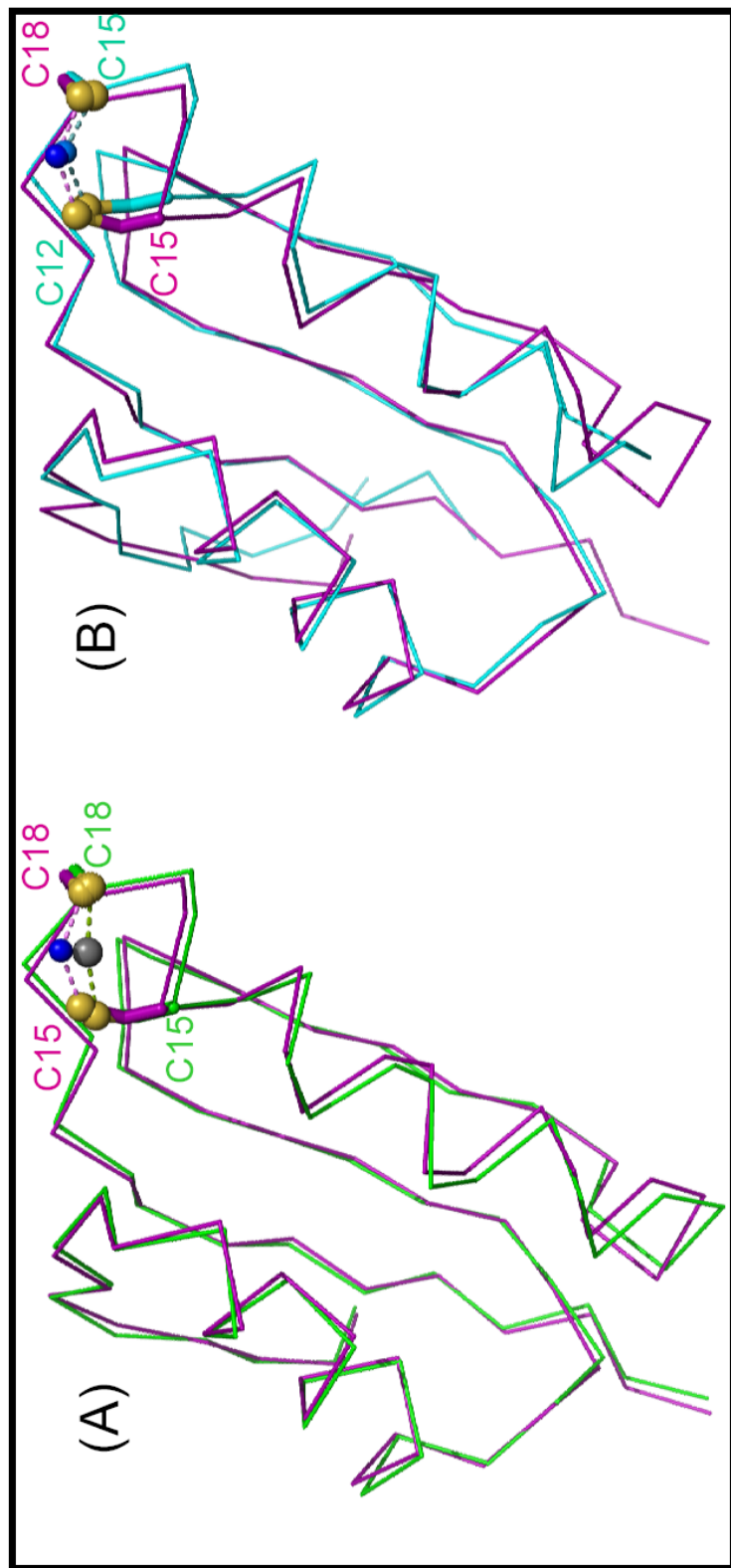
Figure 5-13. Comparison of chain B (Atx1-Cu-TM) with Hg-Atx1 and Cu-Hah1. (A) Superposition of chain B (purple) and Hg-Atx1 (green). Cu and Hg are denoted in blue and gray spheres, respectively. Cu is more exposed compared with Hg. (B) Superposition of chain B (purple) and chain A of Cu-Hah1 (cyan). Cu in chain B is shown with blue sphere, Cu in Cu-Hah1 is shown in light blue sphere. For clarity, the extra ligands for Cu atoms are not shown. The high similarity of the metal binding loop region in the three structures indicates that the loop could easily provide different coordination environments without significant conformation changes, which facilitate their major function as metallchaperones.
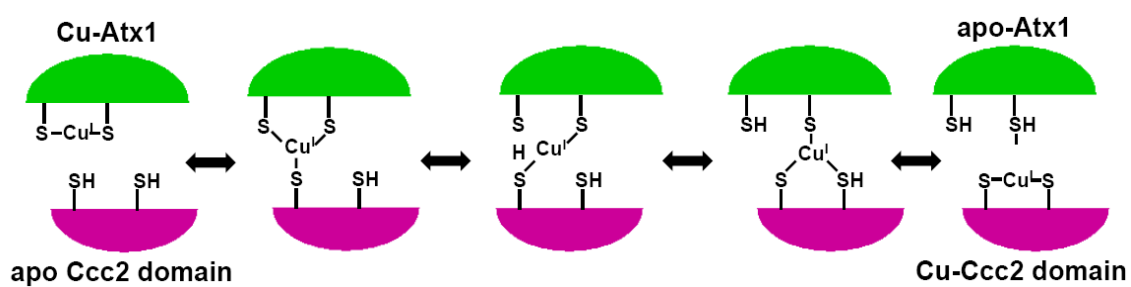
Figure 5-14.   Proposed mechanism of copper transfer between the metallochaperone protein (Atx1) and its target (Ccc2).   Picture is adapted from reference 38.

Mo-Cu Cluster

Each Atx1 trimer houses one Mo-Cu cluster (Figure 5-15) at the center. The Mo-Cu 'nest-shaped' cluster is built up by an unprecedented multinuclear metal cluster that is composed of a MoS_4^{2-} group (TM) and four Cu atoms. The cluster is linked to the Atx1 proteins through the coordination of copper atoms to the conserved metal binding cysteines (Cys15 and Cys18).

The structure of one Mo-Cu cluster is shown in Figure 5-16 and the summary of geometry parameters with standard deviations are listed in Table 5-2. In the following discussion, the values appended in parentheses represent the geometry parameter ranges and the average values.

The Mo atom is tetrahedrally coordinated by four S atoms with the typical Mo-S distances of 2.21~2.27 Å (2.24 Å). Three neighboring Cu atoms exhibits weak interactions with the Mo atom (2.76 ~ 2.82 Å, 2.79 Å). The three Cu atoms adopt distorted tetrahedral coordination geometry by two S atoms from TM (2.28~2.45 Å, 2.31 Å) and two cysteines S atoms (2.27~2.48 Å, 2.31 Å). The rest Cu atom sit above the three copper atoms and is on the opposite side of TM, it exhibits trigonal planar coordination geometry and is trapped by the S atoms of Cys15 (2.24~2.29 Å, 2.27 Å) from the three Atx1 molecules. In particular, three out of the four S atoms in the MoS_4^{2-} group are bridge-ligated to the Mo atom and two tetrahedrally coordinated Cu atoms (denoted with μ_3-S), whereas the S atoms of Cys15 serve as the bridging ligands to two Cu atom adopting tetrahedral geometry and trigonal planar geometry, respectively.

For the 12 Cu atoms adopting distorted tetrahedral geometry, the bond angles of S(Cys15)-Cu-S(Cys18) are substantially larger than the ideal tetrahedral angle of 109.5°, with the

range of 118~124º (123º). The bond angles of (μ_3-S)(TM)-Cu-(μ_3-S)(TM) are in the range of

101~103º (102º), the small deviation is possibly due to the rigidity of TM group. The mixed

bond angels of S(Cys)-Cu-(μ_3-S)(TM) are mainly between the previous two sets of angles and

exhibit large deviation ranging from 98º to 116º (107º). The Mo atoms display less distorted

tetrahedral geometry relative to the Cu atoms, and the bond angles of (μ_3-S)-Mo-(μ_3-S)

(104~111º, 106º) are generally smaller than those of (μ_3-S)-Mo-S (109~115º,112º). The

unexpected Cu atom with trigonal coordination exhibits the S-Cu-S bond angles of 114~126º

(119º), the three bond angles add up to nearly 360°, consistent with the perfect planarity.

The overall Mo-Cu cluster has a net charge of negative four: 1 Mo/+6; 4 Cu/+1; 6 Cys-

SG/-1; 4 S/-2: $6 + (4 \times 1) - (6 \times 1) - (4 \times 2) = -4$. The structure reveals a series of weak

interactions that could compensate the negative charge (Figure 5-17): Lys65, at the

crystallization condition (pH 7.0), is positively charged. For each cluster, three positively

charged lysines balance the majority of the negative charge through forming H-bonds with the

sulfides from TM and the thiolates of Cys18. Relatively, the positively charged lysines show

stronger interactions with terminal ligands (Cys18-SG: 3.2 Å, TM-S: 3.6 Å) than bridging

ligands (μ_3-S: 3.8 Å). Additionally, more H-bonds exist between the thiolates of the metal

binding cysteines and some nearby backbone amides. Furthermore, the positively charged helix

dipole of α helix 2 may also help to neutralize the negative charge via H bonding interaction of

Gly17 backbone amide with the thiolate S of Cys15.

The H-bond network of the thiolates of the metal binding cysteines present some

interesting features among Hg-Atx1, Cu-Hah1 and Atx1-Cu-TM. The H-bonds between the

backbone amides (Thr14, Gly17 in Atx1 corresponding to Thr11 and Gly14 in Hah1) and

thiolates (Cys15 and Cys18 in Atx1 corresponding to Cys12 and Cys15 in Hah1) are the same in all of them, consistent with the observation of the highly similar metal binding region. Relative to the significant H-bond between Lys65 and Cys18 in Atx1-Cu-TM at the distance of 3.2 Å, the corresponding contacts in Hg-Atx1 and Cu-Hah1 are at much longer distances of 3.9 Å and 3.8/3.6 Å (two monomers in Hah1), respectively, suggesting much weaker interactions. While, in Hg-Atx1, a water near the metal binding site displays a significant H-bond with Cys18 (3.3 Å); In Cu-Hah1, due to the dimerization, a strong H-bond appears between Cys12/SG (Cys15 in Atx1) from one monomer and Thr11/OG (Thr14 in Atx1) from the other monomer at a distance of 3.0 Å.

Search of Cambridge Structural Database for similar 'nest-shaped' Mo-Cu cluster reveals several small molecule analogues (Figure 5-18) [40-42]. The small molecule analogue **3** [41] exhibits the closest geometry. All of the S ligands are from TM group, and there is no trigonally-coordinated Cu atom. It has the Mo-Cu, Mo-S, Cu-S distances of 2.69~2.75 Å, 2.06~2.25 Å and 2.29~2.36 Å. Even in the absence of trigonal-coordination Cu atom, the corresponding three S ligands already adopt the appropriate geometry and distances (S-S: 3.85~3.88 Å as compared to 3.87~3.99 Å in the Atx1-Cu-TM) to accept one Cu atom.

The Mo-Cu cluster in Atx1-Cu-TM is, in its composition and geometry, markedly different from any previously observed biological Mo-Cu clusters. The dincuclear [CuSMo(=O)OH] cluster in a CO dehydrogenase [43] is the only structurally characterized Mo-Cu cluster found in Protein Database Bank (Figure 5-19A). In the dinuclear active site, the metals bare bridged by a μ-sulfido ligand. The Cu is linear coordinated with the other S atom for Cys388. The Mo-oxo/hydroxo group are coordinated to the S7' and S8' of the molybdopterin

cytosine cofactor. Another $[S_2MoS_2CuS_2MoS_2]^{3-}$ mixed metal-suflide cluster of the orange protein from *D. gigas* [1] is more related to our Mo-Cu cluster in terms of coordination geometry of metals (Figure 5-19B). However, the clarification of the exact cluster architecture requires structural characterization in addition to the EXAFS studies.

Heterogenous ligand in the Atx1-Cu-TM structure

In the Fo-Fc map, between the two Mo-Cu clusters, there is one extra piece of flat density, which seems to be an aromatic ring (Figure 5-20). Investigation of the crystallization solution and the protein solutions could not identify some chemicals with aromatic rings. We expand the search target to six-membered rings and find the buffering agent in the crystallization solution, malic acid, may be the candidate. Malic acid itself is not a ring, it could adopt a ring-like conformation with H-bonds. With the building of the heterogenous ligand, it does not reveal any direct interactions with the Mo-Cu clusters. However, extensive H-bonds with Ser16 from Atx1 molecules help to anchor the malic acid molecule (Figure 5-21).

Biological implications of the Atx1-Cu-TM model

Wilson's disease is an autosomal recessive hereditary disease, with an incidence of about 1 in 30,000 in most parts of the world. Its main feature is accumulation of copper in tissues, mainly liver and brain, which manifests itself with neurological symptoms and liver diseases. It has served as the main stimulus for the identification of clinically useful anticopper drugs. Dr. John Walshe firstly described the use of penicillamine in Wilson's disease [44]. Penicillamine was soon found to have a large number of toxicities, and Walshe developed trientine as a less toxic substitute for use in penicillamine intolerant patients [45]. Both drugs act as chelators and increases the urinary excretion of copper. In contrast, the zinc therapy serves to block the

intestines' absorption of copper and is ideal for the maintenance therapy of Wilson's disease [46-48]. However, it is too slow-acting to be ideal for the initial treatment of acutely ill patients.

TM was developed to fill the therapeutic void of fast acting and low tocicity drugs. It acts as anticopper agent by forming stable tripartite complexes with copper and proteins, thus blocking the copper absorption or rendering the copper unavailable for cellular uptake. It has been used to treat Wilson's disease since 1982 and proved to be potent, fast acting and low toxicity.

However, little is known about the interaction mechanism of TM with copper and proteins from the atomic level. Our crystallographic model of Atx1-Cu-TM provides the first, and intriguing breakthrough in this blank area. The Mo-Cu cluster not only presents an unprecedented novel architecture, but also reveals a totally unexpected trigonally coordinated copper atom, which could contribute to the therapeutic potency of TM. Furthermore, It is quite remarkable for TM to assemble the Cu-Atx1 in such a compact, efficient and highly symmetrical mode. We speculate that the formation of the trimer blocks the copper transfer from the metal chaperone Atx1 to its target Ccc2, which is under further inspection.

Recently, George et. al investigated the molecular mechanism of tetrathiomolybdate as a treatment of Wilson's disease using X-ray absorption spectroscopic studies of the isolated LEC rat (animal model of Wilson's disease) liver lysosomes, where copper accumulates, after treating with TM [21]. The Mo K-edges data reveals 4 Mo-S at a distance of 2.24 Å, 3 Mo-Cu interactions at 2.70 Å; while the Cu EXAFS indicated 3-3.5 Cu-S at 2.28 Å, with slightly less than one Cu-Mo at 2.70 Å. The results, especially the non-integer Cu data, suggest the copper atoms exist in different coordination environments, which is reflected in our Mo-Cu cluster model. Overall, the

Atx1-Cu-TM structure agrees well with the EXAFS data of the TM treated animal model of

Wilson's disease, suggesting that the model probably reflects the molecular mechanism of the

therapeutic effect of TM in treating Wilson's disease.
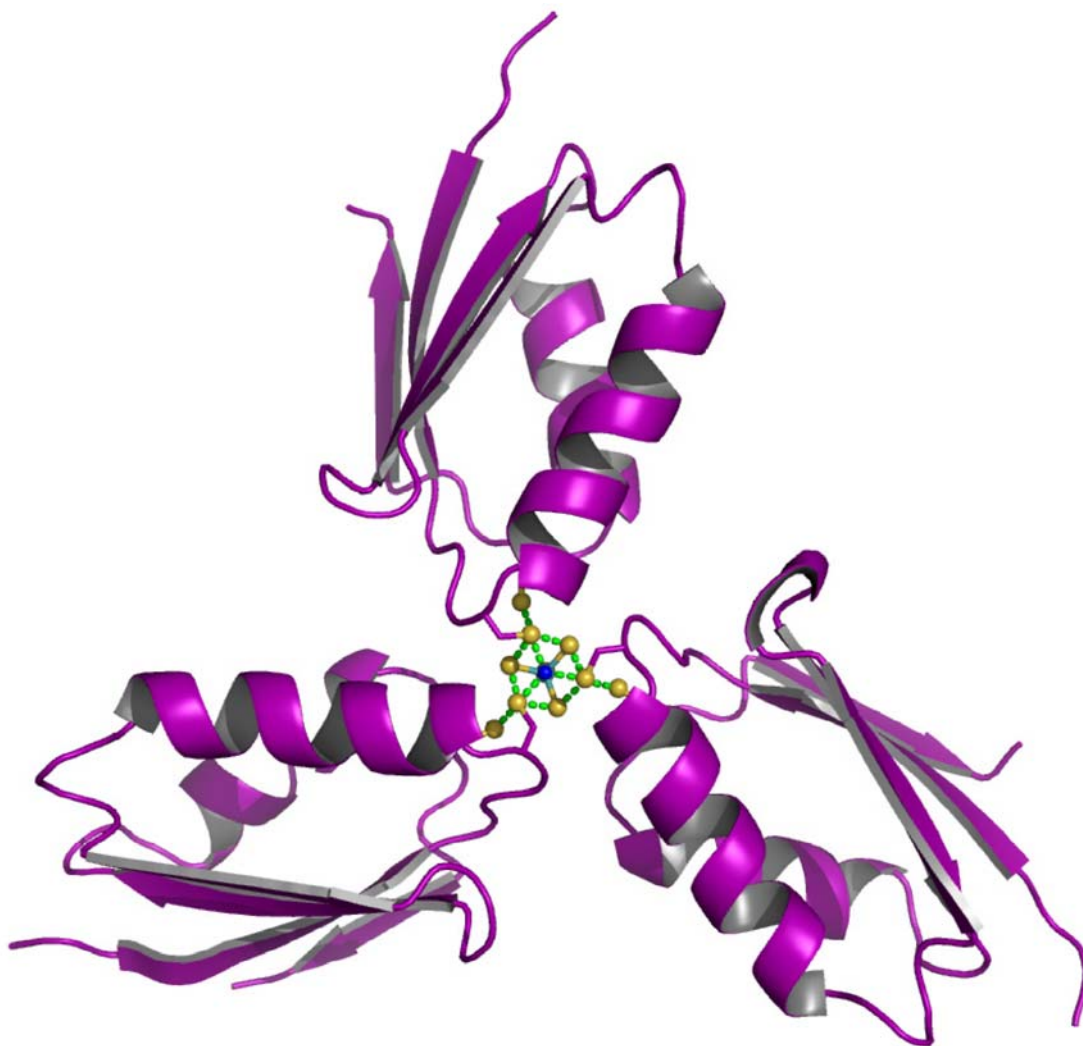
Figure 5-15. (A) The top view of the trimer-cluster (ABC trimer). For clarity, the upper layer of trimer is removed.
The proteins are shown in purple cartoon ribbon diagram. The Cu atoms are shown in blue spheres. The TM and metal binding cysteines are represented with stick and ball model, with S in yellow and Mo in cyan. The coordination bonds are denoted with green dashed lines. The scheme applies to both (B) and (C).
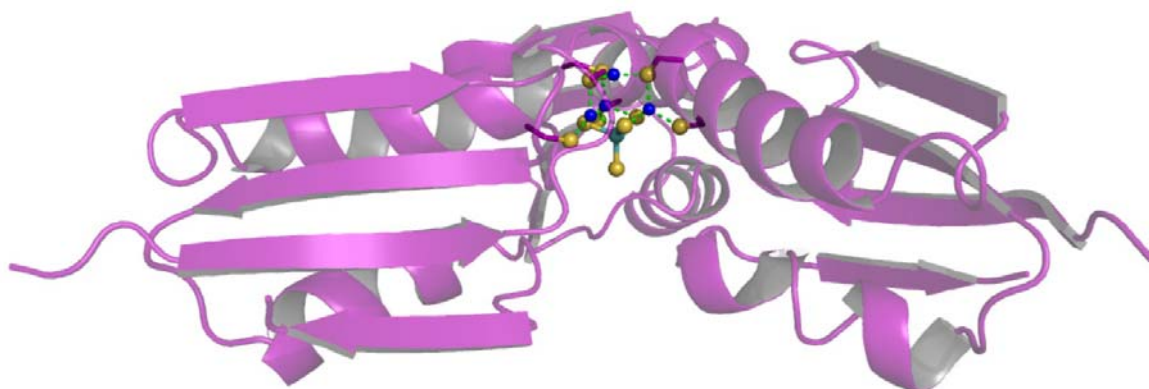
Figure 5-15. (B) The side view of the trimer-cluster (ABC trimer).
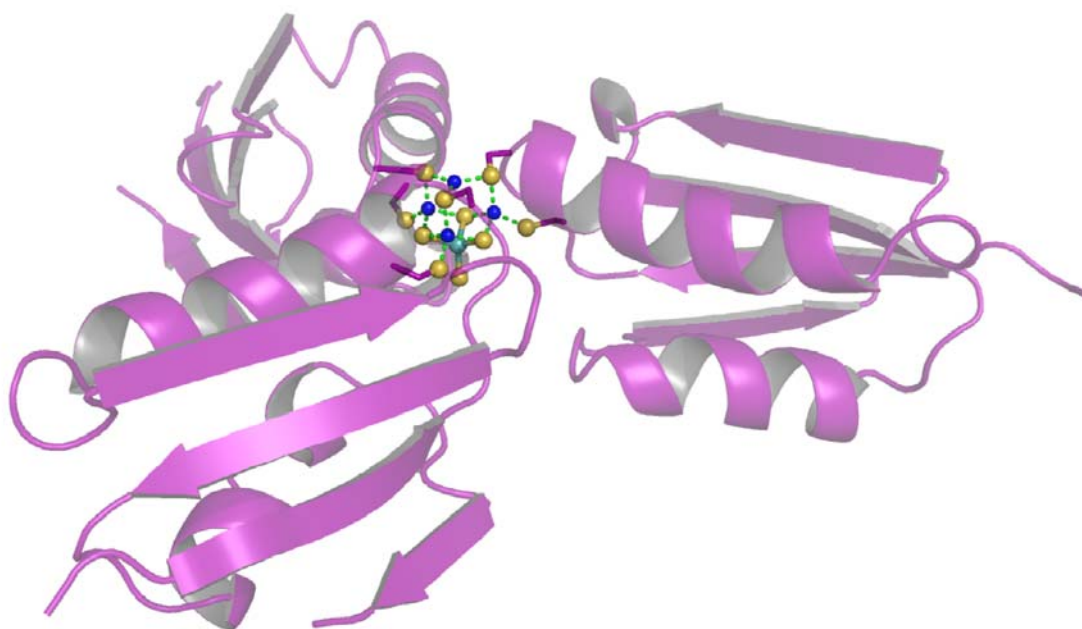
Figure 5-15. (C) The tilted view of the trimer-cluster (ABC trimer).

163
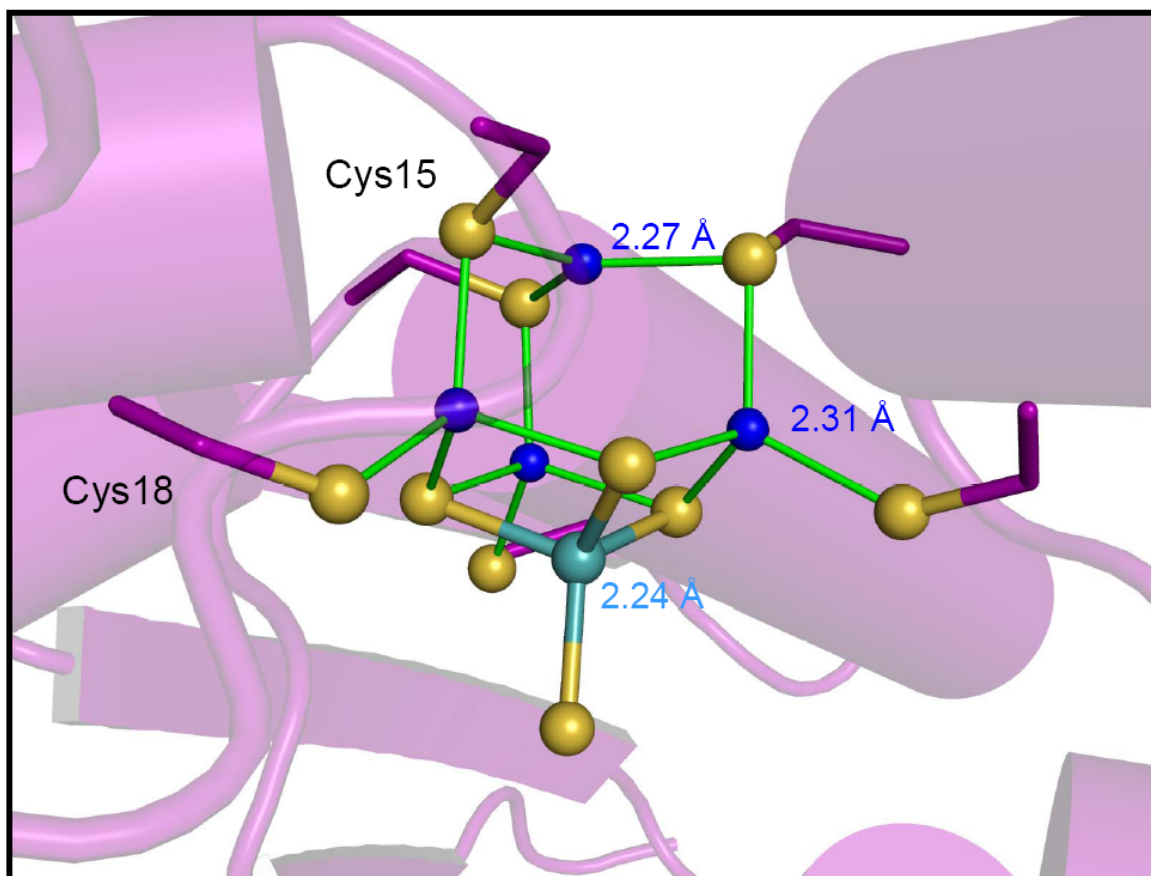
Figure 5-16. The structure of Mo-Cu cluster. Copper atoms are denoted with blue spheres, TM is represented with ball and stick model. The Mo atom is tetrahdrally coordinate by four S atoms with average Mo-S bond distances of 2.24 Å, three neighboring Cu atoms also adopt same coordination geometry with ligands from both TM and Atx1. The top Cu atom displays trigonal-planar geometry coordinated by thiolates from Cys15. The average Cu-S bond distance for tetra- and tri- coordinated copper atoms are 2.31 Å and 2.27 Å, respectively.

Table 5-2. Summary of Cluster Geometry

Tetrahedral Mo	Values	Mean	No. of Observations	S.D.
Mo-S	2.21~2.27 Å	2.24 Å	16	0.02 Å
S-Mo-S	104~115°	109°	24	4°
Tetrahedral Cu				
Cu-S	2.27~2.48 Å	2.31 Å	48	0.04 Å
S-Cu-S	98~124°	109°	72	7°
Trigonal Cu				
Cu-S	2.24~2.29 Å	2.27 Å	12	0.02 Å
S-Cu-S	114~126°	120°	12	4°
Metal-metal distances				
Mo-Cu/4 [1]	2.76~2.82 Å	2.79 Å	12	0.02 Å
Mo-Cu/3 [2]	3.87~4.02 Å	3.97 Å	4	0.07 Å
Cu/4-Cu/4	3.86~3.95 Å	3.91 Å	12	0.03 Å
Cu/4-Cu/3	3.11~3.29 Å	3.24 Å	12	0.05 Å

[1] Cu/4 represents tetrahedrally-coordinated Cu

[2] Cu/3 represents trigonally-coordinated Cu
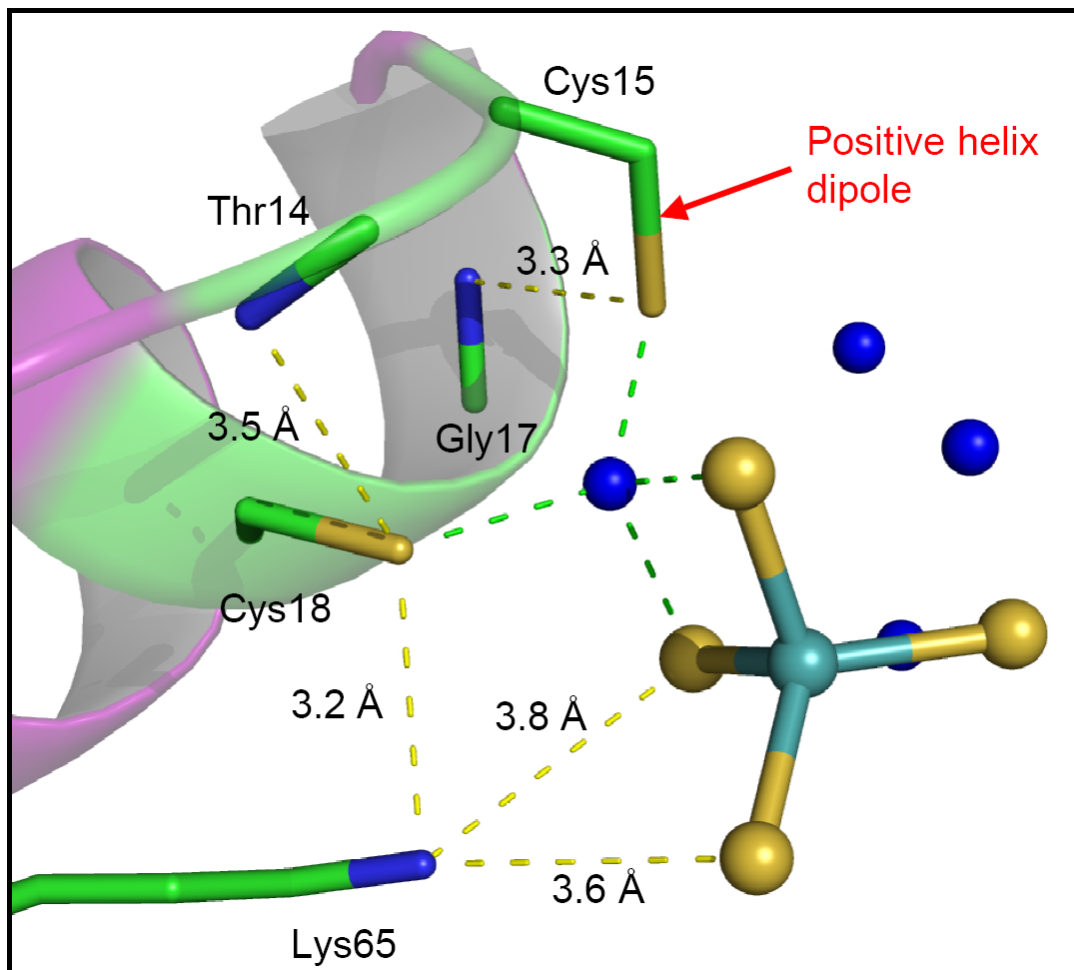
Figure 5-17. The H-bonds network (cutoff distance: 3.8 Å) and positive helix dipole contributing to neutralize the negatively charged metal cluster . The H-bonds are depicted with yellow dotted lines. The Cu-S bonds are shown with yellow dotted lines. Only backbone amides of Thr14 and Gly17 are shown for clarity. The neutralization of the negative charge of the cluster mainly arises from the positively charged Lys65. In addition, more H-bonds with backbone amides and the positive helix dipole also contribute to the charge balancing via interactions with thiolates of metal binding cysteines.
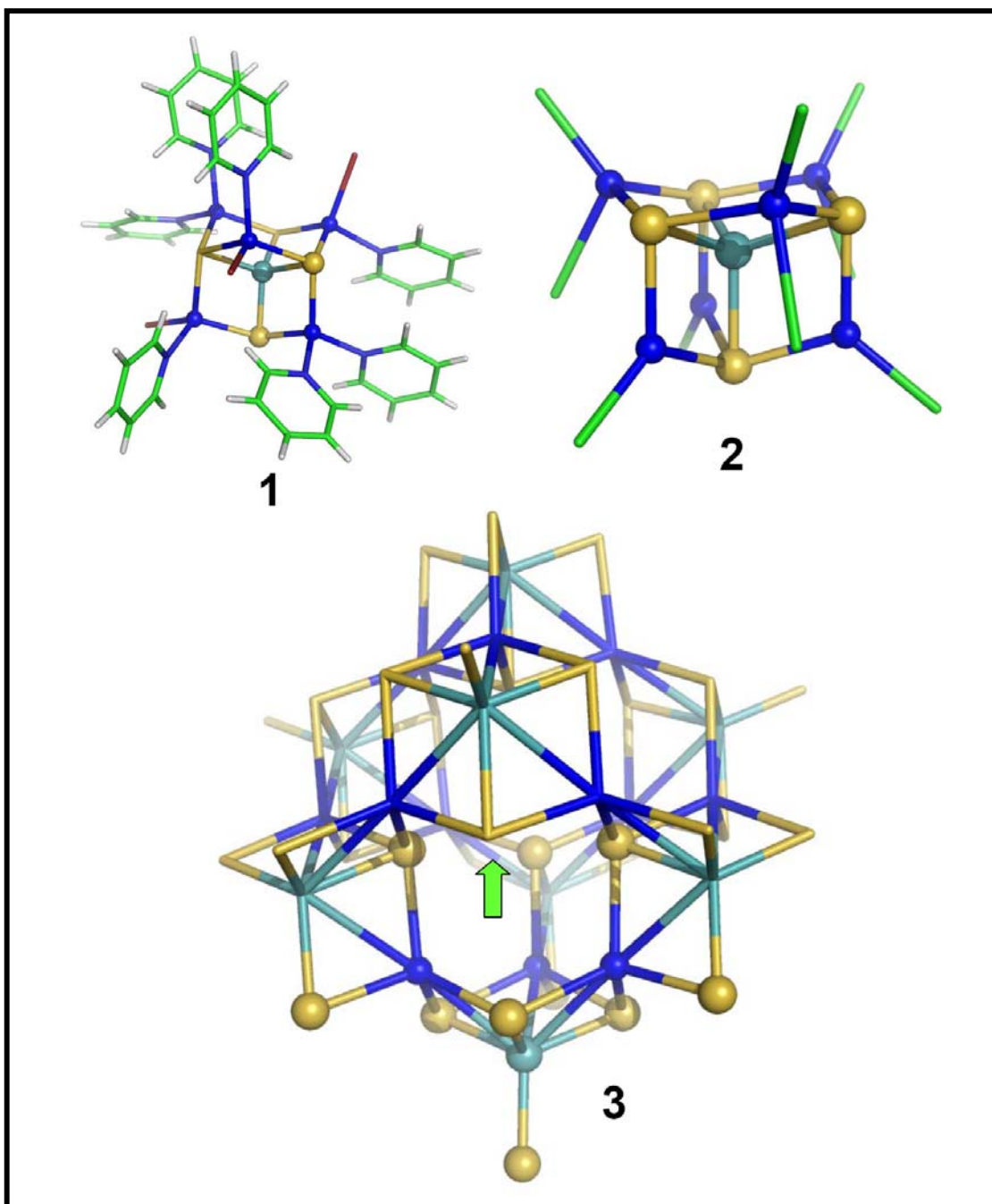
Figure 5-18. The small molecule analogues [40-42] of the Mo-Cu cluster in Atx1-Cu-TM. The cluster is highlighted with sphere and stick representation. **3** [41] exhibits virtually identical architecture except the absence of trigonally coordinated Cu atom. However, the site indicated with green arrow already poses the right geometry to fit the Cu atom.

(A)
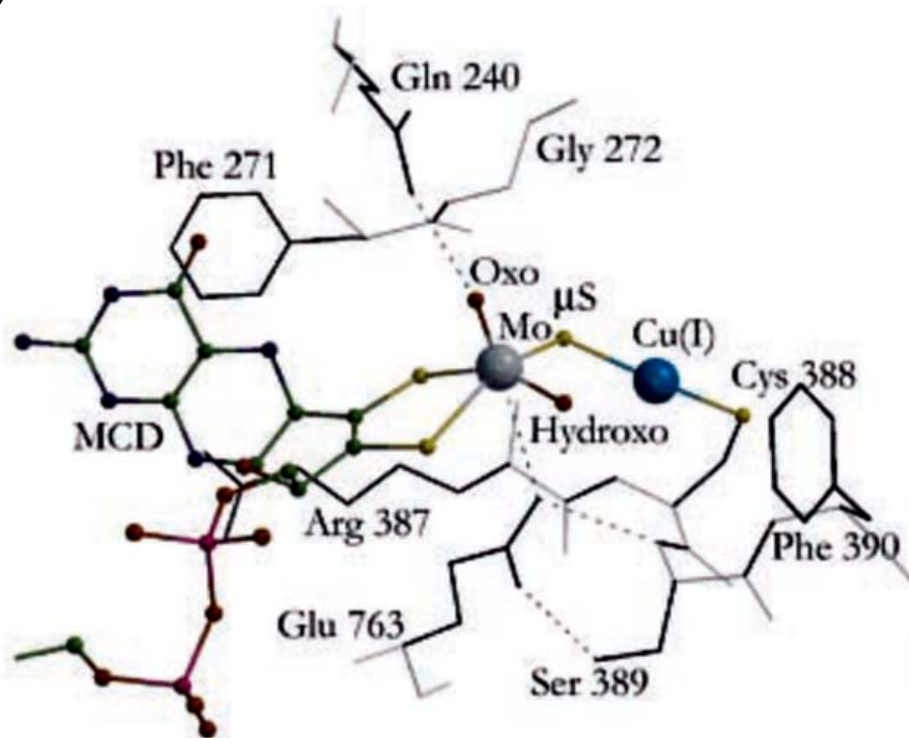
(B)
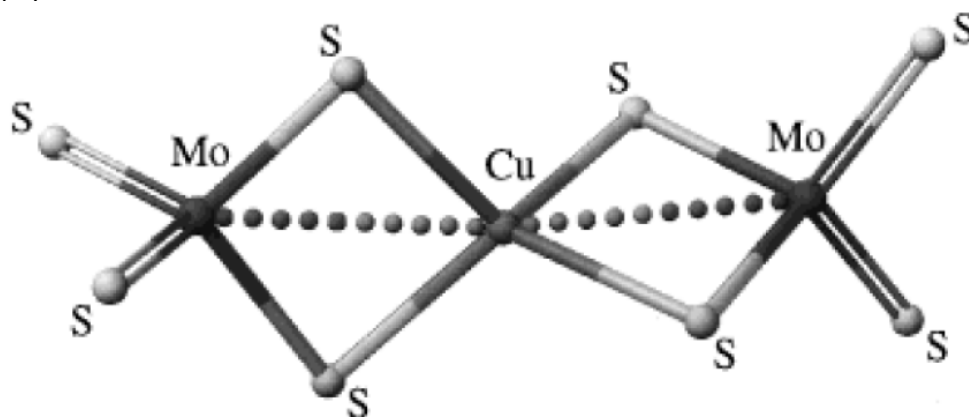
Figure 5-19. Two identified Mo-Cu clusters in biological systems. (A) The dincuclear [CuSMo(=O)OH] cluster in a CO dehydrogenase [42]. The model is solved by X-ray crystallography. (B) The [S₂MoS₂CuS₂MoS₂]³⁻ metal-suflide cluster of the orange protein from *D. gigas* [1]. The model is predicted based on EXAFS studies.
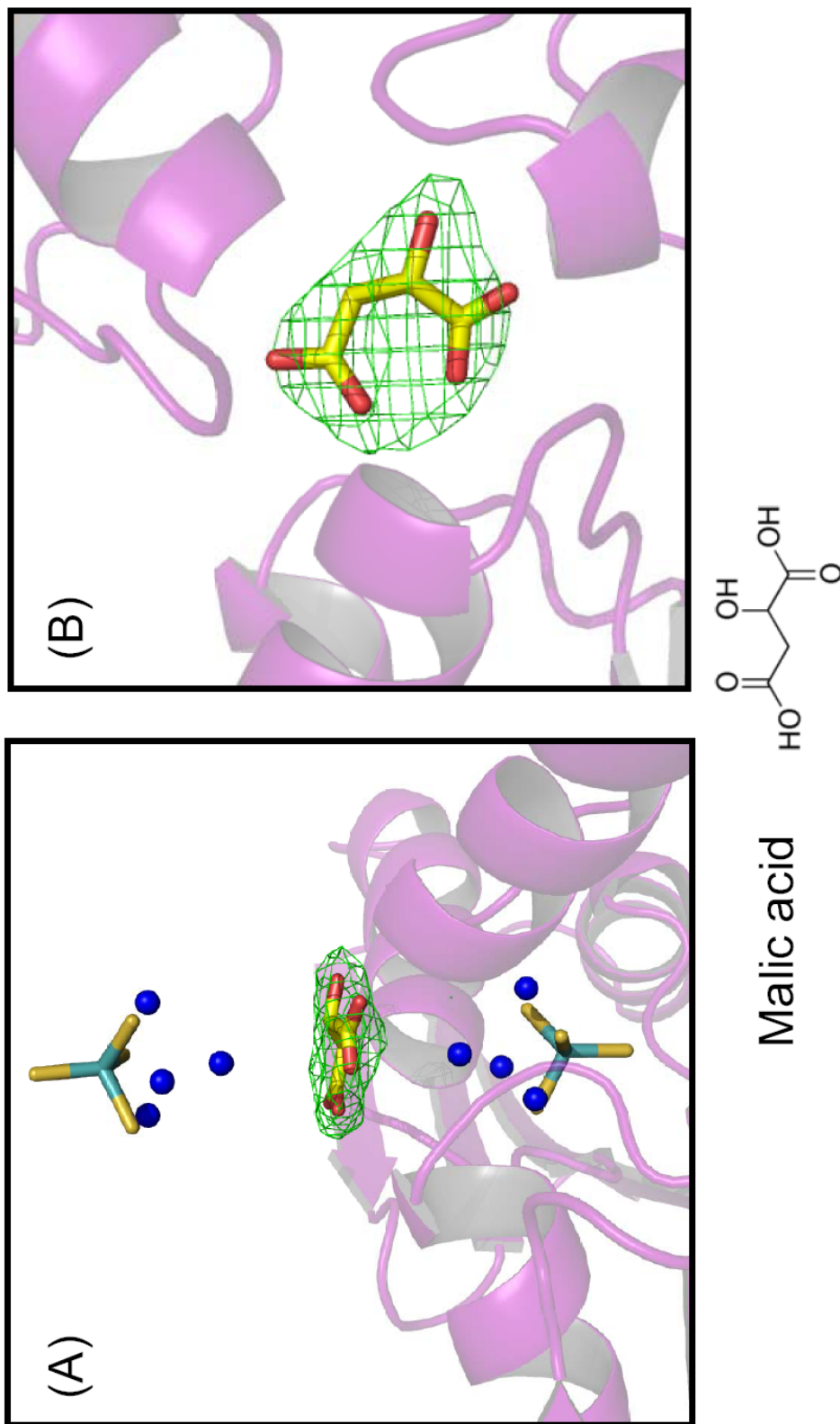
Malic acid

Figure 5-20. Building of heterogenous ligand malic acid. (A) The side view of the flat density and the model (B) The top view of the flat density and the model. For clarity, in (A), the top trimer is omitted; in (B), the top trimer and the Mo-Cu clusters are omitted. The formula of malic acid is shown in (A).
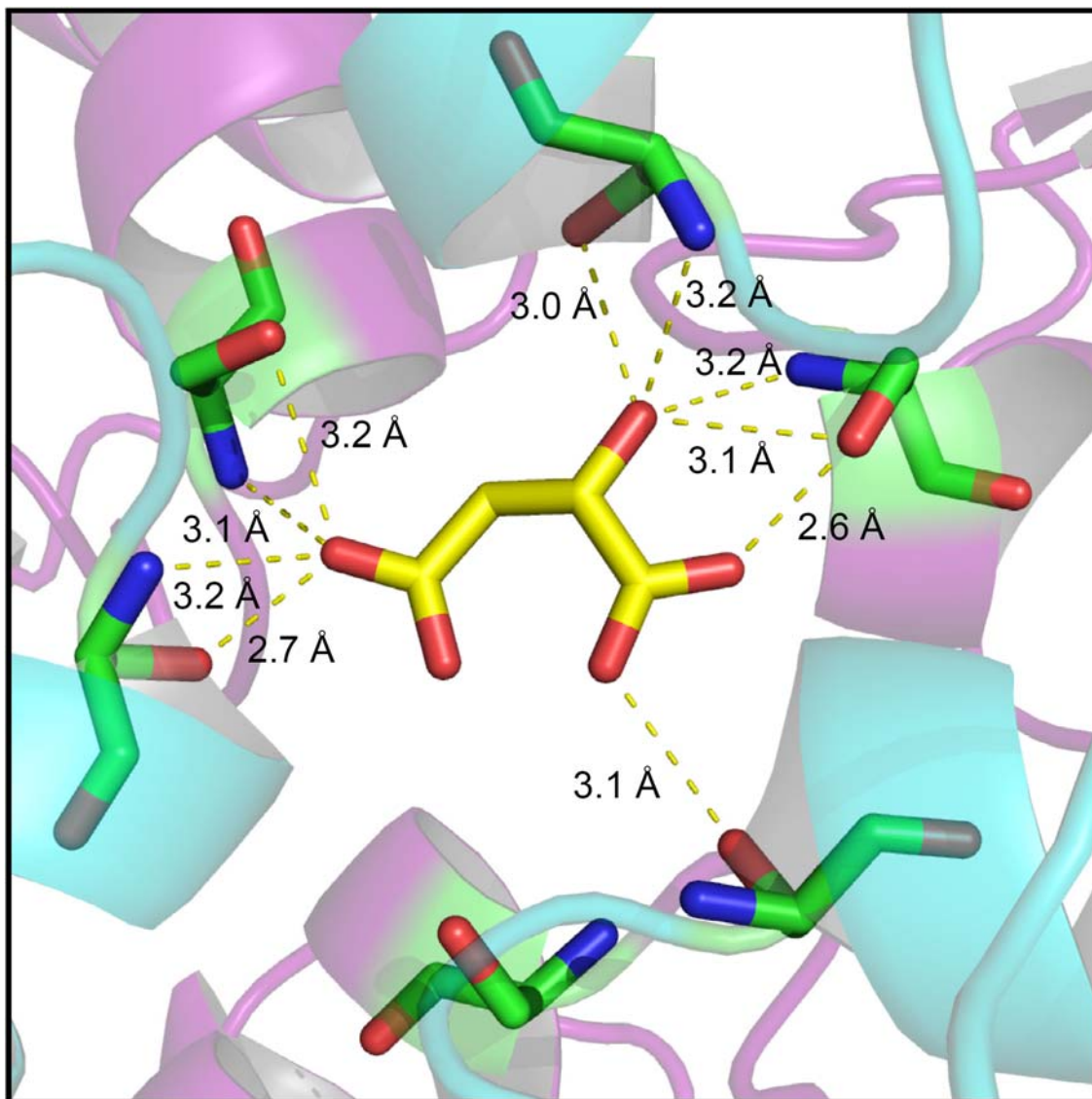
Figure 5-21. Malic acid is anchored between the trimers by forming extensive H-bonds with Ser16. The H-bonds are denoted with yellow dotted lines. The central malic acid and surrounding Ser16 are displayed with stick model. The upper layer trimer and lower layer trimer are shown in cyan and purple cartoon ribbons, respectively. For clarity, the Mo-Cu clusters are omitted.

CHAPTER 6

Crystallization of CueR/DNA/Ag(I) Complex

Abstract

The structures of metal bound forms of CueR and ZntR have revealed the molecular basis for the metal selectivity and ultrasensitivity (Chapter 2). The missing link for the gene regulation function is the allosteric change of the responsive promoter from 'repressed' to 'activated' conformation in response to the metal activator. Here, we described the extensive crystallization trials to obtain the CueR/DNA/Ag(I) crystals, the major focus was on the DNA screening. Several forms of crystals were obtained and diffracted to medium resolution (3~3.2 Å). They all grew in I422 spacegroup with different unit cell parameters. The efforts to solve the structure either failed, or revealed little DNA electron density due to disorder. Finally, the best diffracted (2.95 Å) crystal revealed a different spacegroup of P622. The phase was solved with MR and Se-SAD. The initial electron density map demonstrated the successful incorporation of DNA. Model building and refinement are in process.

6.1 Introduction

With the structures of metal bound forms CueR [2], the puzzles for its ultrasensitivity and selectivity for metals are wonderfully solved. It is tempting to build the full image of the transcription activation with the inclusion of DNA. Even though the DNA distortion mechanisms of MerR family has been well established, and corroborated by the published structures of two drug responsive homologues of the MerR family bound to responsive promoters: BmrR/DNA/TPP [3] and MtaN/DNA [4], it is still of great importance to expand the proofs to the metal responsive members. The complex structure will elucidate the recognition mechanism between CueR and its responsive promoter, additionally, by comparison with the structure of the metal bound protein, it will also reveal the DNA-induced conformation change of the protein. The different levels of models will contribute to the understandings of how the different components work together to function as the genetic switch for copper homeostasis in *E coli.*

One important aspect in crystallization protein/DNA complex is the manipulation of the DNA. It has been observed that DNA length is a key factor that dominates whether a protein/DNA crystallize [5]. Another variable is the sequence of the 'extra' DNA that flanks the central essential region. It is often observed that sticky DNA ends in crystals are packed against other DNA ends or protein molecules. Two types of sticky ends are found in protein-DNA co-crystal structures. In the first type the two ends of a single DNA are complementary. For example, one end has a 5′ protruding "G" and the other end has a 3′ overhanging "C". Such DNA can polymerize in a head–to-tail fashion to form a repetitive linear array that potentially facilitates crystal growth [6]. The second type of sticky end is similar to the first but less regular. It is known that G:C and G or A:T and A can form triplexes, therefore a DNA end with a protruding G or A adjacent to C:G or T:A base pair, respectively, may stack with one another by

forming a two-tier G-C-G or A-T-A triplex at the junction, which is more stable than blunt end stacking. Interestingly, such sticky ends in the CAP-DNA complex crystal was observed to form the DNA stacking by triplex as designed [7], but the overhanging G in the γδ resolvase-DNA complex [8] and MutS-DNA complex [9] were found to interact with protein side chains.

To modify the DNA lengths and overhangs without synthesizing the oligos in their entirety each time, a practical strategy is to divide DNA to two or three segments with adhesive ends, but one prerequisite is the DNA sequence is strictly palindromic. Schultz et al. successfully applied the 'modular' DNA strategy in the crystallization of CAP/DNA complex [10]. As the MerR family promoters are mostly palindromic or pseudo-palindromic, the strategy could serve an efficient way for the initial DNA screen.

In this chapter, we tried to crystallize CueR/DNA/Ag(I) complex by varying the DNA lengths and the terminal overhangs. The 'modular' DNA strategy failed due to the 'annealed' DNA lost the affinity with the protein. The blunt-ended DNAs exhibited better crystallization properties than sticky-ended DNAs. Initial successful crystal was obtained with the DNA of 27mer, however, it only diffracted to 6~7 Å. A fine screen of DNA length from 21mer to 31mer were carried out, finally, three DNAs with odd number length (21mer, 23mer, and 25mer) gave nice crystals diffracted to ~3.2 Å or above. However, the DNA in the 23mer crystal is mostly disordered, and we still have not solved the phases of the 21mer and 25mer crystals. Just recently, a new form of 23mer crystal was grown in a different space group P622 The phase has been solved successfully with MR and Se-SAD. The initial electron density map clearly reveals the successful incorporation of DNA and the DNA is distorted in a similar fashion as the ones in the published structures of BmrR/DNA/TPP [3] and MtaN/DNA [4].

6.2 Experimental procedures

CueR cloning and purification

See chapter 2 experimental procedures: *CueR cloning and purification*. SeMet-CueR protein was prepared by Ms. Monica Canalizo.

Modular DNA design, purification and annealing

See chapter 2 experimental procedures: *DNA design, purification and annealing*

DNA preparation

HPLC grade dry oligos were directly ordered from IDT Corp. and dissolved in TE buffer (10 mM Tris, pH 8, 1 mM EDTA) to 1 mM. Complementary oligos were mixed in equal volume in eppendorf vials sealed well with parafilm, and heated at 95ºC for 5 minutes. After turning off the heat block, the annealed DNA was left in it to cool down slowly to room temperature. The final concentration of the annealed DNA is estimated to be around 0.5 mM.

Protein gel shift assays

Protein gel shift assays were performed to test the binding properties of the DNAs before starting crystallization trials. DNA (100 μM), CueR (141.0 μM) and Ag(I) (240 μM) were incubated together with 5 mM fresh DTT in a total volume of 10 μl for 30 minutes at room temperature before applied to 20% nondenaturing gel. The preparation of the gel is the same as the one described in chapter 3 experimental procedures of *Quantitative Electrophoretic Mobility Shift Assays*.

The gel was stained with Coomassie R dye for about 2~3 hours, then destained, and dried.

Crystallization of CueR/DNA/Ag(I) complex

The crystals were grown with the classical vapor-diffusion hanging drop method by mixing 1 μL of protein solution and 1 μL of crystallization solution. Different lengths of DNAs were tested. Depending on the specific DNA, the successful crystallization conditions may vary, and the morphologies of the crystals also vary.

Data collection and initial structural characterization

The crystals generally appeared in two days. The cryo-preserving solutions consisted of the crystallization conditions and 25 or 30% of the cryoprotectants including glycerol, ethylene glycol, PEG400 and sucrose (25% for glycerol, ethylene glycol and PEG400, 30% for sucrose). The strategies of cryo-preserving included direct transfer and step transfer. Direct transfer is transferring the crystals to the cryo-preserving solution directly, and soaking for about 3 minutes. Step transfer is transferring the crystals in 1-minute steps with each step being 5% increasement of cryoprotectant content. The crystals were then flash-cooled in liquid nitrogen. Paratone oil is also tried as the cryoprotectant. The cryopreseving process is as follows: fish out the crystal, drag into the oil to remove excessive water on the surface, mount the crystals in the loop and flash cool in liquid nitrogen. All data were collected at 100K using synchrotron radiation at APS. Data were processed with Denzo and Scalepack [11]. Molecular replacement was performed with Phaser [12] using Cu-CueR structure (PDB code: 1Q05) as the searching model.

6.3 Results and Discussions

'Modular' DNAs do not bind CueR (ZntR) protein

Initial protein-DNA crystallization started with the 'modular' DNA strategy for both ZntR (zinc-responsive homologue) and CueR (Figure 6-1), and the expected 'protein-DNA' crystals turned out to contain only the protein, which could be caused by two reasons: the

crystallization condition being highly ionic to dissociate the complex, or/and the protein and DNA not forming tightly bound complex. Inspection of the crystallization conditions reveal that, for CueR, the crystallization solution contains 45% Na Citrate, which could easily dissociate the complex, whereas for ZntR, there is no substantial level of salts or other reagents causing the dissociation, however ZntR in the final crystal is a truncated form lacking of the DNA binding domain. Thus from the crystallization experiments only, it was hard to judge whether the 'modular' DNA binds the protein or not. Before starting another round of crystallization screen, it was crucial to determine the binding properties between the DNAs and the protein. We then performed direct protein gel shift assay. Compared with the ZntR modular DNA, the CueR-responsive promoter was not strictly palindromic, mutations in the central area were applied, and one mismatch (central T:T) existed. The assays reveal that, unexpectedly, the modular DNAs lose the specific protein binding abilities as compared with the full-lengths DNAs. The mechanism is not known, but perhaps related with the DNA distortion associated with the protein binding. The existences of nicks in the central area of the modular DNAs could harm the structural tension crucial to the protein-DNA recognition. Collectively, the modular DNA strategy does not apply to the crystallization of DNA complexes of ZntR and CueR, even expandable to the whole MerR family proteins.

Crystallization of CueR/DNA/Ag(I) complexes

Different lengths of DNAs ranging from 21mer to 31 mer were tried (Figure 6-3). Before crystallization trials, protein gel shift assays were carried out to assure their protein binding abilities (Figure 6-4). As revealed in chapter 3, the minor groove interaction is crucial for protein-DNA recognition, in order for shorter lengths (21mer, 23mer) to bind CueR, some mutations should be made on one terminus (A:T to G:C). These mutations restore the protein-

DNA recognition. Due to the inaccurate estimation of the DNA concentrations, the relative binding affinities exhibited in the shift are also rough approximation. It is reasonable that the DNAs of 21mer are of lower affinity for CueR compared with 27mer. The shortest length of DNAs binding CueR is 24. Furthermore, the incorporation of DNA in the crystals are also confirmed by DNA silver staining.

The successful crystallization conditions are listed in Table 6-1. The universal crystallization condition is composed of 8~10% PEG 8K, 0.1 M imidazole, pH 8.0, and 0.2 M $Ca(OAc)_2$. The 0.2 M of $Ca(OAc)_2$ is indispensable, lower concentration decreases the size of the crystals or even results in crystals failing to grow, whereas higher concentration induces precipitation and formation of shower crystals. Substitution with other salts (NaOAc, $Mg(OAc)_2$, KOAc, NH_4OAc and $Zn(OAc)_2$) does not reveal any improvements. Considering the relatively lower binding affinity of CueR/DNA (μM range), excessive (1.6 ~1.7 times) DNA are used. Crystals generally show up in two days, and grow to maximum size in one week. There are two major morphologies (Figure 6-5): stick shape and bipyramidal shape. The length of 27 is roughly the borderline for the two morphologies. With the same crystallization conditions, above 27, most crystals exhibit stick shape; while below 27, bipyramidal shape dominates. The crystals of CueR/31/Ag(I) are mal-shaped and absent of well-defined faces, suggesting 31 might be near the maximum DNA length good for crystallization of the protein-DNA complex. Additionally, the DNAs with overhangs were also tested, and the crystallization properties were worse than those of blunt-ended DNAs, as a result, further experiments were not carried out.

Different cryoprotectants including glycerol, PEG400, ethylene glycol and sucrose were tested. Two general transferring protocols were performed: direct transfer and step transfer.

Comparing the two protocols, the step transfer is more gentle, and imposes gradual osmotic shock to the crystals. Paratone oil is further tested for the crystals showing good diffraction.

Finally, the majority of the crystals diffract to worse than 4 Å. Only the crystals of CueR/25/Ag(I), CueR/23C/Ag(I), and CueR/21CC/Ag(I) diffract to medium resolution of 3.0~3.2 Å (Figure 6-6), and full datasets are collected. Overall, the crystals of odd number length DNAs diffract better than those of even number length DNAs. The Ethylene glycol and glycerol are best for cryopreserving crystals, step transfer and direct transfer do not make significant difference. Paratone oil is not a good cryoprotectant.

CueR/23C/Ag(I)

The crystal diffracts to ~3.2 Å and the data collection statistics is listed in Table 6-2. The crystal belongs to I422 space group with unit cell dimensions of 126.73 Å × 126.73 Å × 66.75 Å. Calculation of solvent content without inclusion of DNA indicates 72% of solvent with one CueR molecule per asymmetric unit (ASU) and 44% of solvent with two CueR molecules (one dimer) per asymmetric unit. Taking DNA into account, one CueR per ASU is more appropriate. Using the CuCueR structure as the search model, one solution is found, after rigid body refinement, the R and R_{free} are 47% and 44%, respectively. Restrained refinement at this stage could decrease the R and R_{free} to 35% and 45%, which suggests the solution is probably right. In the refined solution, the CueR dimer is related by the crystallographic 2-fold axis (symmetry operator: -Y, -X, -1-Z). Inspection of the Fo-Fc and 2Fo-Fc density maps reveals good density around the protein, while little density is observed for the DNA (Figure 6-7). However, the gel shift assays of the crystals with DNA silver staining unambiguously confirm the incorporation of the DNA, and the scattered few densities in the predicted DNA position are still indicative the existence of DNA. Furthermore, there are enough spaces to accommodate the DNA in the crystal

lattice, and the solvent content of 59% with DNA included is in the reasonable range for macromolecular crystals. Hence, the bound DNA in the crystal is mostly disordered. It is not known yet what causes the disorder: the protein/DNA complex itself does not have specific binding? The crystallization condition has some suspicious conditions to weaken some important specific recognition for the complex? The DNA disorder could be due to the multiple conformations formed by the protein sliding along the DNA without specificity.

CueR/25/Ag(I) and CueR/21CC/Ag(I)

Both crystals diffract to ~3 Å. The crystals both belongs to the space group I422, which is the same as that of CueR/23C/Ag(I). The lengths of a and b axes of the unit cell of three crystals are quite close (~126 Å for both CueR/23C/Ag(I) and CueR/25/Ag(I) crystals, ~119 Å for CueR/21CC/Ag(I) crystal), while the length of c axis of the unit cell differs significantly (~67 Å for CueR/23C/Ag(I) crystal, ~ 360 Å for CueR/25/Ag(I) crystal, ~294 Å for CueR/21CC/Ag(I) crystal). Solvent content analysis indicates six CueR molecules per ASU for CueR/25/Ag(I) and four CueR molecules per ASU for CueR/21CC/Ag(I). Molecular replacement using Phaser [12] fails to find the right solutions with the structure of CuCueR as the search model. Even though, solutions with high Z scores are found (higher Z score, solution more outstanding), the rigid body refinement reveals abnormally high R factors of 70-80%, which unambiguously indicate the solutions are wrong. Moreover, such high R factors are normally observed when the space group is wrong, however, in our cases, no other more appropriate space groups could be identified. The approach of direct molecular replacement is currently stuck. Since there are multiple copies of CueR molecules in the asymmetric unit, self-rotation function [13] and native patterson function [13] are calculated to evaluate the existence of non-crystallographic rotational and translational symmetry to gain some insights in the packing of the CueR molecules. The self-

rotation function calculated showed no significant peaks besides those corresponding to crystallographic rotational axes. On the other hand, significant peaks are observed (Table 6-3) in the native Patterson function indicating the presence of non-crystallographic translational symmetry relating the CueR molecules of the asymmetric unit. The largest peaks are 81% and 75% of the origin peak for CueR/25/Ag(I) and CueR/21CC/Ag(I), respectively. The packing of CueR molecules are along c axis, and the distances are 66.2 Å and 64.1 Å for CueR/25/Ag(I) and CueR/21CC/Ag(I), respectively. Based on the knowledge, we perform the rotational search, and choose the solution with the highest Z score, and build a model based on the non-crystallography translational symmetry, then, we use the model and start a new molecular replacement. Unfortunately, the strategy is also unsuccessful, the R factors after rigid body refinement are still nearly 70%. Recently, the effort to make Pt- derivatives by soaking K_2PtCl_4 into the crystals overnight turns out to be a failure: the crystals are completely dead.

SeMet-CueR/23C/Ag(I)

The SeMet-CueR protein used for crystallization was kindly provided by Ms. Monica Canalizo. The crystal diffracts to ~2.95 Å and surprisingly it has a different spacegroup (P622) from that of the CueR/23C/Ag(I) crystals (I422). According to the solvent content calculation, there are one CueR monomer and one DNA strand or one DNA duplex per asymmetric unit (ASU). The molecular replacement has successfully located the CueR monomer. The physiologically relevant CueR dimer is related by the crystallographic symmetry operators. The phases calculated from MR reveal the existence of DNA in the electron density map. However, the phases of the partial model are of poor quality and the derived DNA electron density is not continuous and clear. Thus, we employ the same strategy used in structure determination of Atx1-Cu-TM in chapter 4. The anomalous Se-SAD signals are combined with the calculated MR

phases to generate the Se positions. The positions of the three Se atoms are imported to sharp/autosharp [14, 15] for heavy atom refinement and subsequent protein phasing, solvent flattening procedures. The experimental electron density map reveals remarkably improved DNA density. The preliminary data statistics are listed in Table 6-4. The initial electron density map clearly indicate the DNA is distorted in a similar pattern as the ones observed in the structures of BmrR/DNA/TPP and MtaN/DNA (Figure 6-8), confirming the DNA distortion mechanism also applies to the MerR family metallregulatory member CueR. Additionally, the metal binding loop of CueR undergoes some substantial conformational changes, suggesting the mutual allosteric adaption of the protein and DNA to accomplish the genetic switch function. Further model building and refinement are in process and more structural details are waiting to be revealed.

Discussion

Extensive efforts have been invested into the crystallization of CueR/DNA complexes. Substantial progress has been made: the binding properties of CueR/DNA have been explored and rich experiences have been accumulated in the crystallization studies. The complex crystals with medium resolution (~ 3 Å) are obtained. The final success has been obtained just very recently.

It needs to point out that the success is on a mutated DNA 23mer. The terminus of 23mer is mutated from A/T pair to C/G pair. How much the modification affecting the protein/DNA recognition and the physiological function will be further evaluated. At the same time, the model of the CueR/23C/Ag(I) will serve as a better model for determination of the CueR/25/Ag(I) structure with molecular replacement, which will provide a good structural comparison to reveal the effect of DNA terminus mutation.

So far, all the DNAs for crystallization trials are derived from the promoter of copA. Since CueR also regulates the transcription of CueO, it is straightforward to start the test of the promoter of CueO (Figure 6-9). Comparison of the two promoters suggests that the major differences are mainly in the central area and on one terminus. The palindromic sequences are identical except those of P_{cueO} are one-nucleotide shorter. It is generally accepted that the palindromic sequences are the protein-binding box. In the specific cases, the palindromic sequences cover both major groove and minor groove interaction area. However, the central area and the terminus might also contribute the interaction in some unrecognized ways, thus, the promoter of cueO serves another interesting research target for X-ray crystallography.

In summary, the crystals of CueR/DNA/Ag(I) have been obtained and the structure determination is under way. Combining the CueR/Ag(I) and CueR/DNA/Ag(I) structures, we will get deeper understanding of the mutual adaptiveness of the protein and DNA, as well as the molecular mechanism of the metal-regulated gene activation. The established structures will serve the new starting point for the more ambitious and exciting exploration to include the RNAP into the gene regulation scenario.

zntA promoter: 5' – TTGACT CTGGAGTCGACTCCAGAGTG TATCCT – 3'
                       AACTGA GACCTCAGCTGAGGTCTCAC ATAGGA
                       -35   -10

18 mer: 5' – GATACACTCTGGAGTCGA – 3'
13 mer: 5' – CTCCAGAGTGTAT – 3'

P$_{zntA}$-18/13 mer: 5' – GATACACTCTGGAGTCGACTCCAGAGTGTAT – 3'
                            TATGTGAGACCTCAGCTGAGGTCTCACATAG

copA promoter: 5' – TTGACC TTCCCCTTGCTGGAAGGTT TAACCT – 3'
                       AACTGG AAGGGGAACGACCTTCCAA ATTGGA
                       -35   -10

18 mer: 5' – GTTAAACCTTCCCCTTAG – 3'
12 mer: 5' – GGGAAGGTTTAA – 3'

P$_{copA}$-18/12 mer: 5' – GTTAAACCTTCCCCTTAGGGGAAGGTTTAA – 3'
                            AATTTGGAAGGGGATTCCCCTTCCAAATTG

Figure 6-1. Design of crystallization DNA. Arrow: palindromic sequence; Dot: plaindromic symmetry center; Triangle: nicks in DNA; double underline: identical sequences in designed DNA as in wild type promoters. Adapted from Chapter 2.
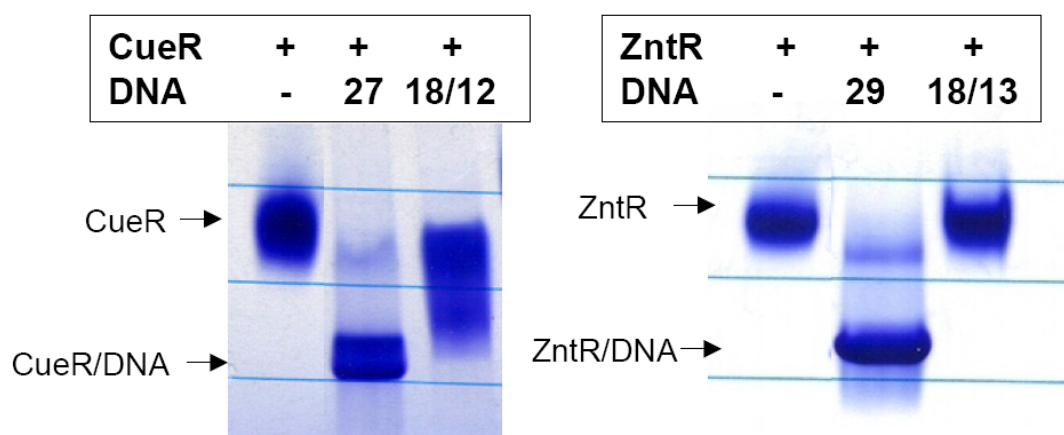
Figure 6-2. Protein gel shift assays of the modular DNAs. The assays reveal that the modular DNAs lose the specific binding to the proteins.

Figure 6-3. The schematic diagram of DNAs used in crystallization trials of CueR/DNA/Ag(I) complex. Number indicates the length of the DNA. Upper case letters suggest the mutations. Lower cases letters a and b stand for two options of the same length DNAs. The red box includes the potential minor groove interaction site (see chapter 3). Mutation of the site (A:T to G:C) can restore the binding affinity of the DNAs with lengths shorter than 24 mer. In 21XX (such as 21GC, etc.), four mutations are tried including 21GC, 21GG, 21CC and 21CG. The red box on the other half site of the 21XX is complementary to the mutated site. In the sequences 27 and 31, the gray sequences are the area protected by CueR in the DNaseI footprinting assay [1]. The inverted arrows suggest the palindromic sequences.

31: T C T T G A C C T T C C C C T T G C T G G A A G G T T T A A C
A G A A C T G G A A G G G G A A C G A C C T T C C A A A T T G

30a:

30b:

29:

28a:

28b:

27: T T G A C C T T C C C C T T G C T G G A A G G T T T A
A A C T G G A A G G G G A A C G A C C T T C C A A A T

26b:

26a:

25:

24:

23C: C
G

23G: G
C

21XX: X X
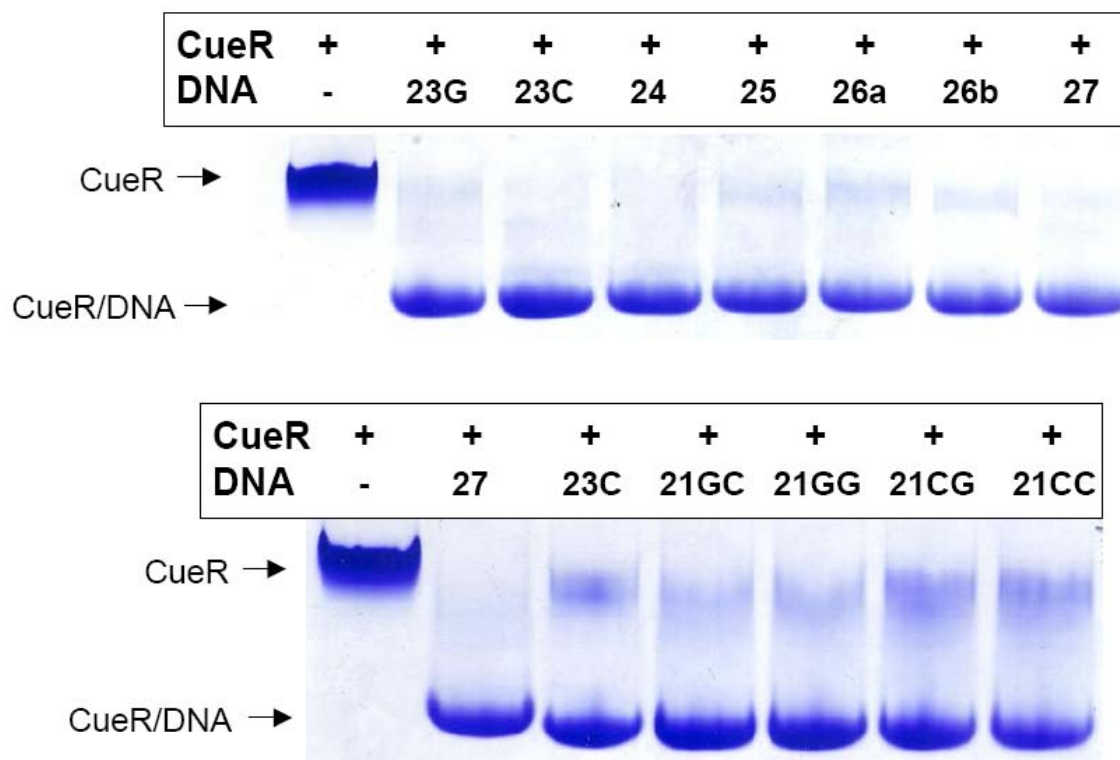
Figure 6-4. Protein gel shift assays of different lengths of DNAs. The binding of the DNAs shifts the protein faster than the unbound protein. The DNAs with lengths ranging from 21 to 27 exhibit similar binding properties. The DNAs with longer lengths of 28~31 all include the sequence of 27, and are expected to own virtually identical binding ability as the DNA of 27.

Table 6-1. Crystallization conditions for different DNAs

No.	DNA	Sample[1] CueR: DNA: AgNO₃	Crystallization condition	Diffraction[2]
1	27	91.1 μM : 82.1 μM : 200 μM	10% PEG 8K , 0.1M MES 6.5/Imidazole 7.0, 0.1M MgCl₂/CaCl₂	~5 Å
2	27	91.1 μM : 82.1 μM : 200 μM	18% PEG 8K, 0.1 M Na Cacodylate pH 6.5, 0.2M Zn(OAc)₂	6~7 Å
3	23G	364.4 μM : 270.0 μM : 320.0 μM	10% PEG 8K, 0.1M Imidazole pH 8.0, 0.1M Ca(OAc)₂	~4 Å
4	23C[3]	250~300 μM: 375~450 μM: 500~600 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	~3.2 Å
5	24	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	~6.5 Å
6	25[3]	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	~3.0 Å
7	26a	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	~5 Å
8	26b	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	~4 Å
9	28a	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	~8 Å
10	28b	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	~7 Å
11	29	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	~5 Å
12	30a	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	N/A
13	30b	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	N/A
14	31	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	N/A
15	21GC	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	N/A
16	21GG	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	N/A
17	21CC[3]	211.5 μM : 350 μM : 400 μM	8~10% PEG 8K, 0.1 M Imidazole pH 8.0, 0.2 M Ca(OAc)₂	~ 2.95 Å

[1] : 5 mM DTT is added freshly.

[2] : The crystals were test shot at beamline 17-ID at APS.

[3] : Full dataset(s) have been collected.

Figure 6-5. The pictures of the CueR/DNA/Ag(I) crystals. The sequence of the crystals corresponds to the sequence in Table 6-1. Generally, the crystals have two major kinds of morphology, one is stick-shaped, represented by (9)~(13), the other one is bipyramidal-shaped as in (6)~(8) and (17). The length of 27 is roughly the borderline for the two morphologies. Above 27, most crystals exhibit stick shape; Below 27, bipyramidal shape dominates. The crystals of CueR/31/Ag(I) are mal-shaped and absent of well-defined faces, suggesting 31 might be near the maximum DNA length good for crystallization of the protein-DNA complex.

~300 μm
(3) CueR/23G/Ag(I)

~300 μm
(6) CueR/25/Ag(I)

150~200 μm
(2) CueR/27/Ag(I)

~400 μm
(5) CueR/24/Ag(I)

~400 μm
(1) CueR/27/Ag(I)

~400 μm
(4) CueR/23G/Ag(I)

~500 x 50 x 50 μm
(9) CueR/28a/Ag(I)

~200 μm
(12) CueR/30a/Ag(I)

~300 μm
(8) CueR/26b/Ag(I)

~400 x 35 x 35 μm
(11) CueR/29/Ag(I)

~300 μm
(7) CueR/26a/Ag(I)

~400 x 25 x 25 μm
(10) CueR/28b/Ag(I)

~ 200 x 40 x 40 μm
(15) CueR/21GC/Ag(I)

< 150 μm
(14) CueR/31/Ag(I)

250~300 μm
(17) CueR/21CC/Ag(I)
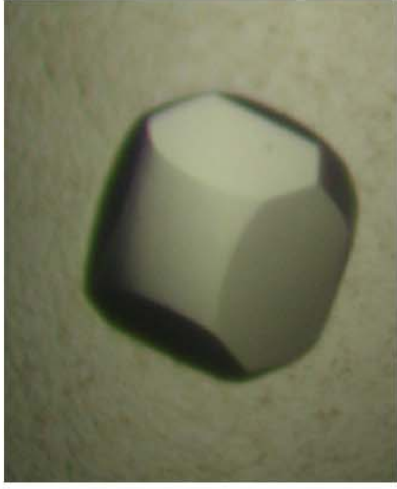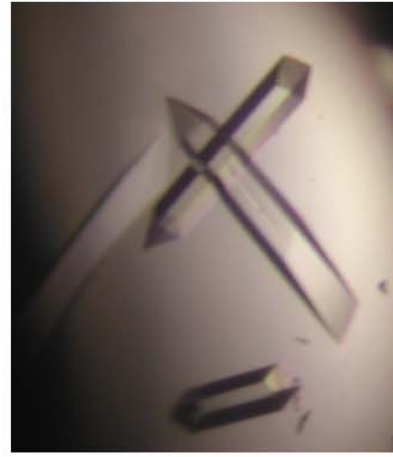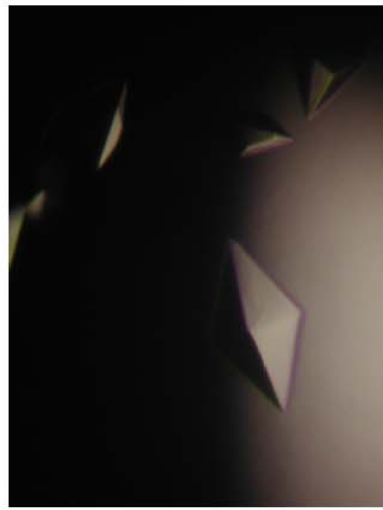
200 ~ 250 μm
(13) CueR/30b/Ag(I)

Small needle crystals
(16) CueR/21GG/Ag(I)
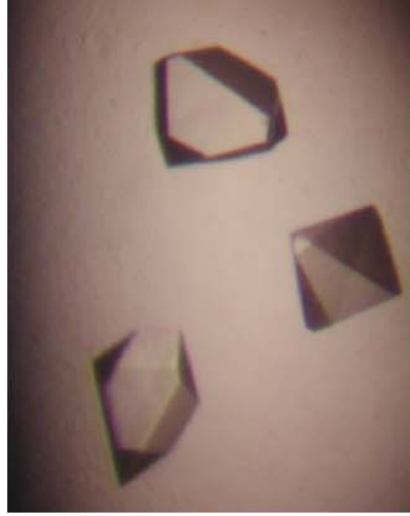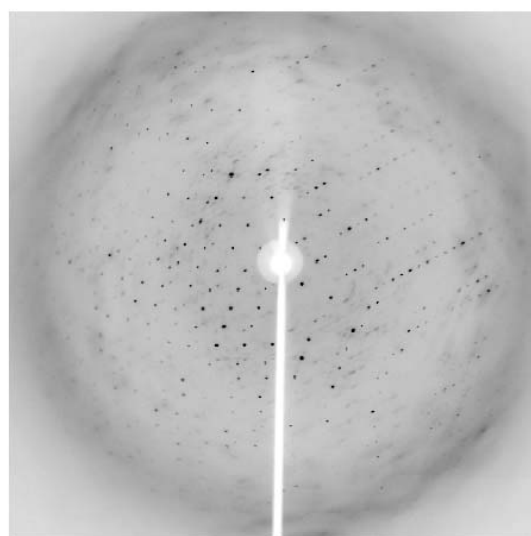
CueR/23C/Ag(I), ~ 3.2 Å
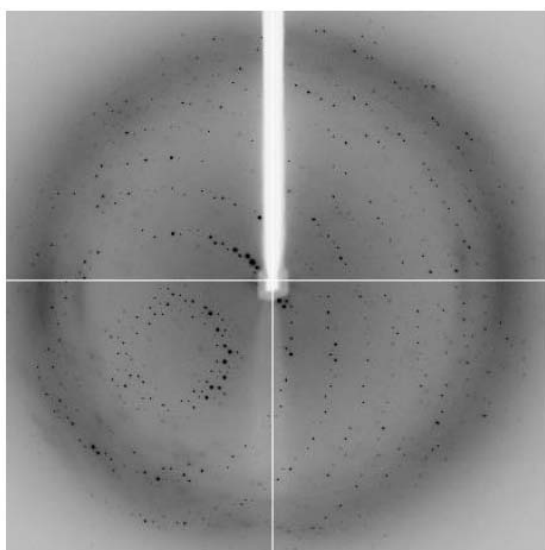
CueR/25/Ag(I), ~ 3.0 Å
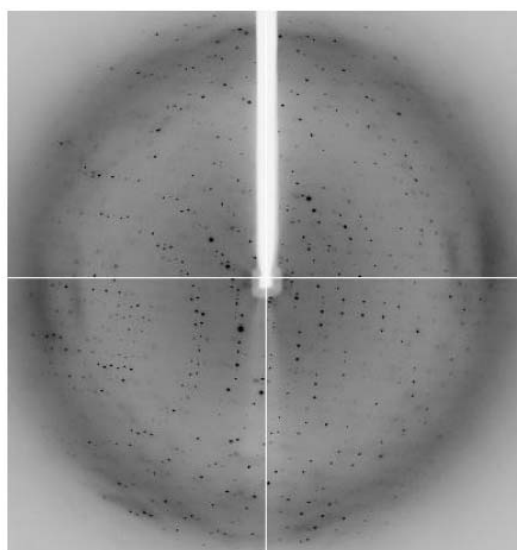
CueR/21CC/Ag(I), ~ 3.0 Å

Figure 6-6. The diffraction images of three CueR/DNA/Ag(I) complexes with medium resolution of 3~3.2 Å. The spots of CueR/23C/Ag(I) are more loosely spaced compared with those of other two crystals, suggesting CueR/23C/Ag(I) has relatively smaller unit cell, as confirmed by data processing program (Table 6-2).

Table 6-2. Data collection Statistics

Values in parentheses are for the highest resolution shell.

Crystal	CueR/23C/Ag(I)[1]	CueR/25/Ag(I)[1]	CueR/21CC/Ag(I)[1]
Wavelength (Å)	1.0	1.0	1.0
Resolution limits (Å)	40.0 - 3.20 (3.37 - 3.20)	50.0 - 3.06 (3.17 - 3.06)	50.0 - 3.0 (3.11 -3.0)
No. of observed/unique reflections	45498 / 4654	397483 / 27819	336193 / 21734
Completeness (%)	98.9 (100)	99.4 (99.5)	99.6 (100)
Data redundancy	9.8 (10.1)	14.3 (14.7)	14.2 (14.7)
R_{sym} (%)[2]	6.5 (40.9)	8.4 (35.3)	8.8 (40.0)
I/$\langle\sigma\rangle$	26.7 (4.2)	44.8 (8.9)	50.8 (8.3)
Space group	I422	I422	I422
Unit-cell parameters (Å °)	$a = b = 126.7$, $c = 66.7$ $\alpha=\beta=\gamma=90.0$	$a = b = 126.5$, $c = 359.8$ $\alpha=\beta=\gamma=90.0$	$a = b = 119.0$, $c = 294.5$ $\alpha=\beta=\gamma=90.0$

[1] Dataset was collected at 100 K at the beamline 17-ID-B at the Advance Photon Source.

[2] $R_{sym} = \sum | I_{obs} - I_{avg}| / \sum I_{obs}$
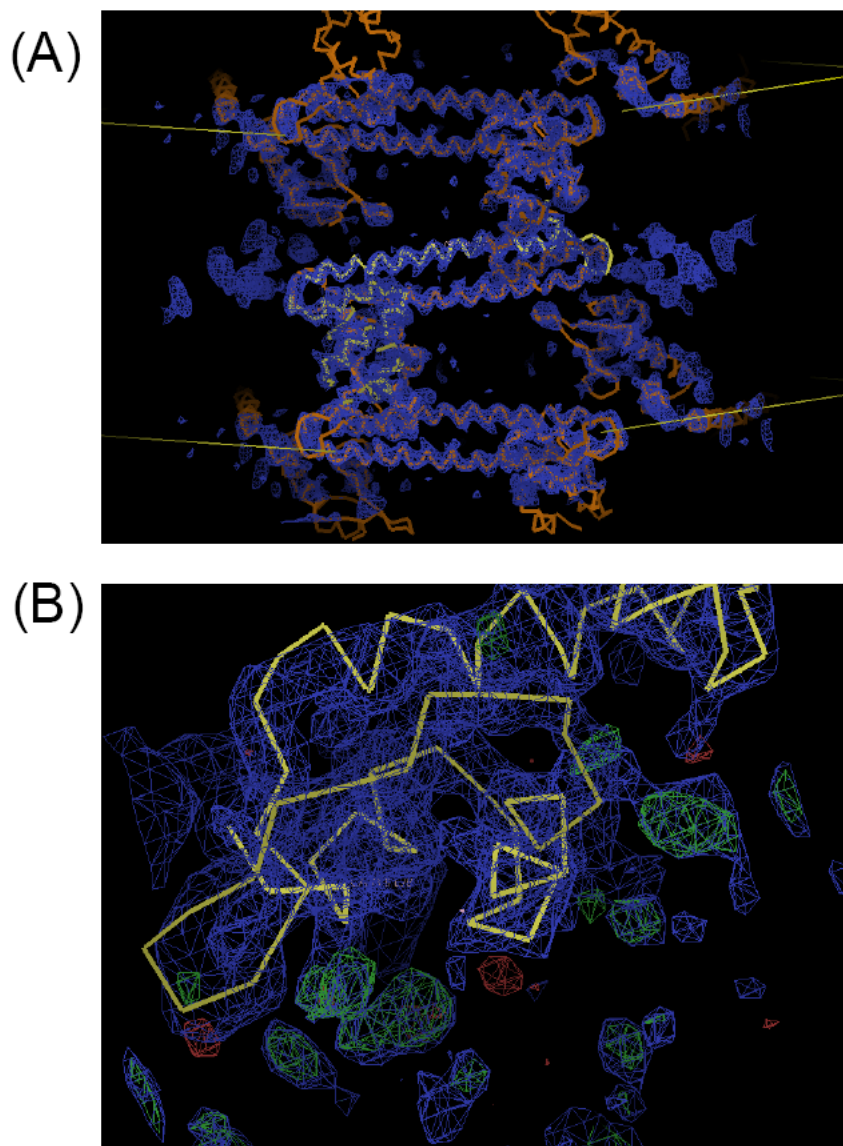
Figure 6-7. 2Fo-Fc and Fo-Fc density maps of CueR/23C/Ag(I) after rigid body refinement and restrained refinement. The blue density represents 2Fo-Fc map contoured at 1σ level, the green and red densities are the Fo-Fc map contoured at 3σ level. (A) The overall view. The monomer in yellow is the original one, all other orange ones are generated by symmetry operations. The yellow straight lines are the unit cell in a and b directions. (B) The close-up view of the DNA binding domain. The scattered but significant densities near the DNA binding domain of CueR indicate the existences of DNA.

Table 6-3. Non-crystallographic translational peaks
(only peaks> 20% origin peak are listed)

CueR/25/Ag(I)

No.	u	v	w	Height
1*	0.00	0.00	0.00	100.0
2	0.00	0.00	0.18	81.2
3	0.50	0.50	0.13	46.5
4	0.50	0.50	0.05	21.2

CueR/21CC/Ag(I)

No.	u	v	w	Height
1*	0.00	0.00	0.00	100.0
2	0.00	0.00	0.22	74.8
3	0.50	0.50	0.07	39.0

*: origin peak

Table 6-4. Data processing and phasing statistics of SeMet-
CueR/DNA/Ag(I) crystal

Values in parentheses are for the highest resolution shell.

Data collection and Statistics	
Data set	Se-SAD
Beamline[1]	LS-CAT
Wavelength (Å)	0.9785
Resolution limits (Å)	2.90
Completeness (%)	99.7 (99.6)
Data redundancy	9.2 (9.6)
R_{sym} (%) [2]	7.2 (34.6)
I/$<\sigma>$	19.6 (5.9)
Space group	P622
Unit-cell parameters (Å °)	162.9 162.9 53.2 90.0 90.0 120.0
Phasing Statistics	
Phasing Power [3]	1.251
Se sites found	3
FOM after SHARP	0.30
FOM after DM	0.86

[1] All data sets were collected at 100 K at the Advance Photon Source.

[2] $R_{sym} = \Sigma |I_{obs} - I_{avg}| / \Sigma I_{obs}$

[3] The phasing power is defined as the ratio of the rms value of heavy atom structure factor amplitudes to the rms value of the lack-of-closure error.
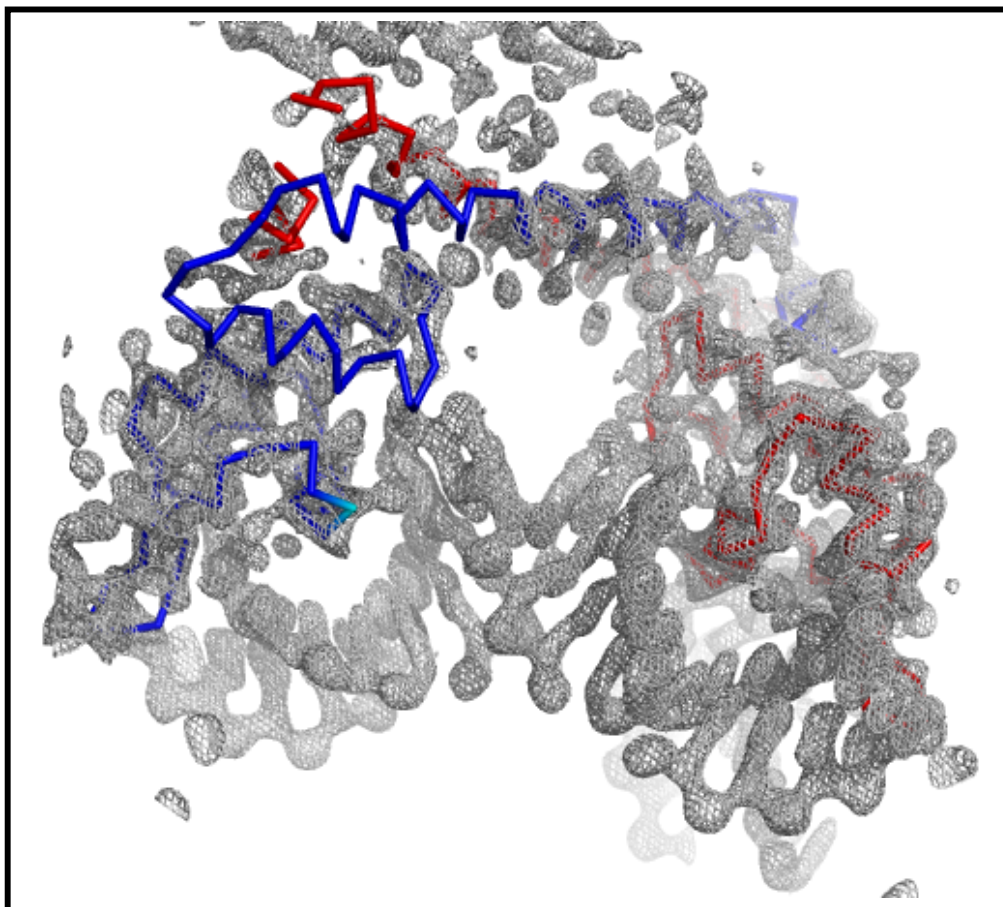
Figure 6-8. The initial Se-SAD electron density map
The CueR dimer is represented with blue and red ribbon. The map clearly revealing the existence of DNA, which is distorted in a similar way as the ones found in BmrR/DNA/TPP and MtaN/DNA.

P_{copA}-27:
```
T T G A C C T T C C C C T T G C T G G A A G G T T T A
A A C T G G A A G G G G A A C G A C C T T C C A A A T
```

P_{cueO}-27:
```
T T G A C C T T C C C G T A A G G G G A A G G A C T A
A A C T G G A A G G G C A T T C C C C T T C C T G A T
```

Figure 6-9. The comparison of the promoters of *copA* and *cueO*. The pink dot indicates the center of the promoter, the inverted lines include the palindromic sequences, the different sequences are shown in red in P_{cueO}-27. The differences are mainly in the central area and on one terminus.

REFERENCES

Chapter 1

1. Rosenzweig, A. C., Metallochaperones: bind and deliver. *Chem Biol* **2002,** 9, (6), 673-7.

2. Outten, C. E.; Outten, F. W.; O'Halloran, T. V., DNA distortion mechanism for transcriptional activation by ZntR, a Zn(II)-responsive MerR homologue in Escherichia coli. *J Biol Chem* **1999,** 274, (53), 37517-24.

3. Finney, L. A.; O'Halloran, T. V., Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* **2003,** 300, (5621), 931-6.

4. Pena, M. M.; Lee, J.; Thiele, D. J., A delicate balance: homeostatic control of copper uptake and distribution. *J Nutr* **1999,** 129, (7), 1251-60.

5. Balamurugan, K.; Schaffner, W., Copper homeostasis in eukaryotes: teetering on a tightrope. *Biochim Biophys Acta* **2006,** 1763, (7), 737-46.

6. Harris, E. D., Celluar copper transport and metabolism. *Annu. Rev. Nutr.* **2000,** 20, 291-310.

7. Puig, S.; Thiele, D. J., Molecular mechanisms of copper uptake and distribution. *Curr Opin Chem Biol* **2002,** 6, (2), 171-80.

8. Bertinato, J.; L'Abbe, M. R., Maintaining copper homeostasis: regulation of copper-trafficking proteins in response to copper deficiency or overload. *J Nutr Biochem* **2004,** 15, (6), 316-22.

9. DiDonato, M.; Sarkar, B., Copper transport and its alterations in Menkes and Wilson diseases. *Biochim Biophys Acta* **1997,** 1360, (1), 3-16.

10. Tanner, M. S., Role of copper in Indian childhood cirrhosis. *Am J Clin Nutr* **1998,** 67, (5 Suppl), 1074S-1081S.

11. Muller, T.; Feichtinger, H.; Berger, H.; Muller, W., Endemic Tyrolean infantile cirrhosis: an ecogenetic disorder. *Lancet* **1996,** 347, (9005), 877-80.

12. Muller, T.; Muller, W.; Feichtinger, H., Idiopathic copper toxicosis. *Am J Clin Nutr* **1998,** 67, (5 Suppl), 1082S-1086S.

13. Bayer, T. A.; Multhaup, G., Involvement of amyloid beta precursor protein (AbetaPP) modulated copper homeostasis in Alzheimer's disease. *J Alzheimers Dis* **2005,** 8, (2), 201-6; discussion 209-15.

14. Bush, A. I., Metals and neuroscience. *Curr Opin Chem Biol* **2000,** 4, (2), 184-91.

15. Brown, D. R., Copper and prion diseases. *Biochem Soc Trans* **2002,** 30, (4), 742-5.

16. Erler, J. T.; Bennewith, K. L.; Nicolau, M.; Dornhofer, N.; Kong, C.; Le, Q. T.; Chi, J. T.; Jeffrey, S. S.; Giaccia, A. J., Lysyl oxidase is essential for hypoxia-induced metastasis. *Nature* **2006,** 440, (7088), 1222-6.

17. Vulpe, C. D.; Kuo, Y. M.; Murphy, T. L.; Cowley, L.; Askwith, C.; Libina, N.; Gitschier, J.; Anderson, G. J., Hephaestin, a ceruloplasmin homologue implicated in intestinal iron transport, is defective in the sla mouse. *Nat Genet* **1999,** 21, (2), 195-9.

18. Grass, G.; Rensing, C., CueO is a multi-copper oxidase that confers copper tolerance in Escherichia coli. *Biochem Biophys Res Commun* **2001,** 286, (5), 902-8.

19. Outten, F. W.; Huffman, D. L.; Hale, J. A.; O'Halloran, T. V., The independent *cue* and *cus* systems confer copper tolerance during aerobic and anaerobic growth in *Escherichia coli*. *J. Biol. Chem.* **2001,** 276, (33), 30670-30677.

20. Rensing, C.; Fan, B.; Sharma, R.; Mitra, B.; Rosen, B. P., CopA: An Escherichia coli Cu(I)-translocating P-type ATPase. *Proc Natl Acad Sci U S A* **2000,** 97, (2), 652-6.

21. Fan, B.; Grass, G.; Rensing, C.; Rosen, B. P., Escherichia coli CopA N-terminal Cys(X)(2)Cys motifs are not required for copper resistance or transport. *Biochem Biophys Res Commun* **2001,** 286, (2), 414-8.

22. Stoyanov, J. V.; Hobman, J. L.; Brown, N. L., CueR (YbbI) of Escherichia coli is a MerR family regulator controlling expression of the copper exporter CopA. *Mol Microbiol* **2001,** 39, (2), 502-11.

23. Franke, S.; Grass, G.; Rensing, C.; Nies, D. H., Molecular analysis of the copper-transporting efflux system CusCFBA of Escherichia coli. *J Bacteriol* **2003,** 185, (13), 3804-12.

24. Lee, S. M.; Grass, G.; Rensing, C.; Barrett, S. R.; Yates, C. J.; Stoyanov, J. V.; Brown, N. L., The Pco proteins are involved in periplasmic copper handling in Escherichia coli. *Biochem Biophys Res Commun* **2002,** 295, (3), 616-20.

25. Wernimont, A. K.; Huffman, D. L.; Finney, L. A.; Demeler, B.; O'Halloran, T. V.; Rosenzweig, A. C., Crystal structure and dimerization equilibria of PcoC, a methionine-rich copper resistance protein from Escherichia coli. *J Biol Inorg Chem* **2003,** 8, (1-2), 185-94.

26. Solioz, M.; Stoyanov, J. V., Copper homeostasis in Enterococcus hirae. *FEMS Microbiol Rev* **2003,** 27, (2-3), 183-95.

27. Odermatt, A.; Krapf, R.; Solioz, M., Induction of the putative copper ATPases, CopA and CopB, of Enterococcus hirae by Ag+ and Cu2+, and Ag+ extrusion by CopB. *Biochem Biophys Res Commun* **1994,** 202, (1), 44-8.

28. Solioz, M.; Odermatt, A., Copper and silver transport by CopB-ATPase in membrane vesicles of Enterococcus hirae. *J Biol Chem* **1995,** 270, (16), 9217-21.

29. Strausak, D.; Solioz, M., CopY is a copper-inducible repressor of the Enterococcus hirae copper ATPases. *J Biol Chem* **1997,** 272, (14), 8932-6.

30. Cobine, P.; Wickramasinghe, W. A.; Harrison, M. D.; Weber, T.; Solioz, M.; Dameron, C. T., The Enterococcus hirae copper chaperone CopZ delivers copper(I) to the CopY repressor. *FEBS Lett* **1999,** 445, (1), 27-30.

31. Wimmer, R.; Herrmann, T.; Solioz, M.; Wuthrich, K., NMR structure and metal interactions of the CopZ copper chaperone. *J Biol Chem* **1999,** 274, (32), 22597-603.

32. Valentine, J. S.; Gralla, E. B., Delivering copper inside yeast and human cells. *Science* **1997,** 278, (5339), 817-8.

33. Zhou, B.; Gitschier, J., hCTR1: a human gene for copper uptake identified by complementation in yeast. *Proc Natl Acad Sci U S A* **1997,** 94, (14), 7481-6.

34. Zhou, H.; Cadigan, K. M.; Thiele, D. J., A copper-regulated transporter required for copper acquisition, pigmentation, and specific stages of development in Drosophila melanogaster. *J Biol Chem* **2003,** 278, (48), 48210-8.

35. Markossian, K. A.; Kurganov, B. I., Copper chaperones, intracellular copper trafficking proteins. Function, structure, and mechanism of action. *Biochemistry (Mosc)* **2003,** 68, (8), 827-37.

36. Arnesano, F.; Banci, L.; Bertini, I.; Huffman, D. L.; O'Halloran, T. V., Solution structure of the Cu(I) and apo forms of the yeast metallochaperone, Atx1. *Biochemistry* **2001,** 40, (6), 1528-39.

37. Walker, J. M.; Tsivkovskii, R.; Lutsenko, S., Metallochaperone Atox1 transfers copper to the NH2-terminal domain of the Wilson's disease protein and regulates its catalytic activity. *J Biol Chem* **2002,** 277, (31), 27953-9.

38. Culotta, V. C.; Klomp, L. W.; Strain, J.; Casareno, R. L.; Krems, B.; Gitlin, J. D., The copper chaperone for superoxide dismutase. *J Biol Chem* **1997,** 272, (38), 23469-72.

39.   Rae, T. D.; Schmidt, P. J.; Pufahl, R. A.; Culotta, V. C.; O'Halloran, T. V., Undetectable intracellular free copper: the requirement of a copper chaperone for superoxide dismutase. *Science* **1999,** 284, (5415), 805-8.

40.   Lamb, A. L.; Wernimont, A. K.; Pufahl, R. A.; Culotta, V. C.; O'Halloran, T. V.; Rosenzweig, A. C., Crystal structure of the copper chaperone for superoxide dismutase. *Nat Struct Biol* **1999,** 6, (8), 724-9.

41.   Lamb, A. L.; Torres, A. S.; O'Halloran, T. V.; Rosenzweig, A. C., Heterodimer formation between superoxide dismutase and its copper chaperone. *Biochemistry* **2000,** 39, (48), 14720-7.

42.   Lamb, A. L.; Torres, A. S.; O'Halloran, T. V.; Rosenzweig, A. C., Heterodimeric structure of superoxide dismutase in complex with its metallochaperone. *Nat Struct Biol* **2001,** 8, (9), 751-5.

43.   Furukawa, Y.; Torres, A. S.; O'Halloran, T. V., Oxygen-induced maturation of SOD1: a key role for disulfide formation by the copper chaperone CCS. *Embo J* **2004,** 23, (14), 2872-81.

44.   Horng, Y. C.; Cobine, P. A.; Maxfield, A. B.; Carr, H. S.; Winge, D. R., Specific copper transfer from the Cox17 metallochaperone to both Sco1 and Cox11 in the assembly of yeast cytochrome C oxidase. *J Biol Chem* **2004,** 279, (34), 35334-40.

45.   Lode, A.; Kuschel, M.; Paret, C.; Rodel, G., Mitochondrial copper metabolism in yeast: interaction between Sco1p and Cox2p. *FEBS Lett* **2000,** 485, (1), 19-24.

46.   Borghouts, C.; Werner, A.; Elthon, T.; Osiewacz, H. D., Copper-modulated gene expression and senescence in the filamentous fungus Podospora anserina. *Mol Cell Biol* **2001,** 21, (2), 390-9.

47.   Voskoboinik, I.; Camakaris, J., Menkes copper-translocating P-type ATPase (ATP7A): biochemical and cell biology properties, and role in Menkes disease. *J Bioenerg Biomembr* **2002,** 34, (5), 363-71.

48.   Voskoboinik, I.; Mar, J.; Strausak, D.; Camakaris, J., The regulation of catalytic activity of the menkes copper-translocating P-type ATPase. Role of high affinity copper-binding sites. *J Biol Chem* **2001,** 276, (30), 28620-7.

49.   Suzuki, M.; Gitlin, J. D., Intracellular localization of the Menkes and Wilson's disease proteins and their role in intracellular copper transport. *Pediatr Int* **1999,** 41, (4), 436-42.

50.   Bull, P. C.; Thomas, G. R.; Rommens, J. M.; Forbes, J. R.; Cox, D. W., The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. *Nat Genet* **1993,** 5, (4), 327-37.

51.   Vanderwerf, S. M.; Cooper, M. J.; Stetsenko, I. V.; Lutsenko, S., Copper specifically regulates intracellular phosphorylation of the Wilson's disease protein, a human copper-transporting ATPase. *J Biol Chem* **2001,** 276, (39), 36289-94.

52.   Palmiter, R. D., The elusive function of metallothioneins. *Proc Natl Acad Sci U S A* **1998,** 95, (15), 8428-30.

53.   Turner, J. S.; Robinson, N. J., Cyanobacterial metallothioneins: biochemistry and molecular genetics. *J Ind Microbiol* **1995,** 14, (2), 119-25.

54.   Ecker, D. J.; Butt, T. R.; Sternberg, E. J.; Neeper, M. P.; Debouck, C.; Gorman, J. A.; Crooke, S. T., Yeast metallothionein function in metal ion detoxification. *J Biol Chem* **1986,** 261, (36), 16895-900.

55.   Bonneton, F.; Theodore, L.; Silar, P.; Maroni, G.; Wegnez, M., Response of Drosophila metallothionein promoters to metallic, heat shock and oxidative stresses. *FEBS Lett* **1996,** 380, (1-2), 33-8.

56.   Thiele, D. J., ACE1 regulates expression of the Saccharomyces cerevisiae metallothionein gene. *Mol Cell Biol* **1988,** 8, (7), 2745-52.

57.   Jensen, L. T.; Posewitz, M. C.; Srinivasan, C.; Winge, D. R., Mapping of the DNA binding domain of the copper-responsive transcription factor Mac1 from Saccharomyces cerevisiae. *J Biol Chem* **1998,** 273, (37), 23805-11.

58.   Keller, G.; Bird, A.; Winge, D. R., Independent metalloregulation of Ace1 and Mac1 in Saccharomyces cerevisiae. *Eukaryot Cell* **2005,** 4, (11), 1863-71.

59.   Heuchel, R.; Radtke, F.; Georgiev, O.; Stark, G.; Aguet, M.; Schaffner, W., The transcription factor MTF-1 is essential for basal and heavy metal-induced metallothionein gene expression. *Embo J* **1994,** 13, (12), 2870-5.

60.   Auf der Maur, A.; Belser, T.; Elgar, G.; Georgiev, O.; Schaffner, W., Characterization of the transcription factor MTF-1 from the Japanese pufferfish (Fugu rubripes) reveals evolutionary conservation of heavy metal stress response. *Biol Chem* **1999,** 380, (2), 175-85.

61.   Zhang, B.; Egli, D.; Georgiev, O.; Schaffner, W., The Drosophila homolog of mammalian zinc finger factor MTF-1 activates transcription in response to heavy metals. *Mol Cell Biol* **2001,** 21, (14), 4505-14.

62.   Zhang, B.; Georgiev, O.; Hagmann, M.; Gunes, C.; Cramer, M.; Faller, P.; Vasak, M.; Schaffner, W., Activity of metal-responsive transcription factor 1 by toxic heavy metals and H2O2 in vitro is modulated by metallothionein. *Mol Cell Biol* **2003,** 23, (23), 8471-85.

63.   Bittel, D.; Dalton, T.; Samson, S. L.; Gedamu, L.; Andrews, G. K., The DNA binding activity of metal response element-binding transcription factor-1 is activated in vivo and in vitro by zinc, but not by other transition metals. *J Biol Chem* **1998,** 273, (12), 7127-33.

64.   Outten, F. W.; Outten, C. E.; Hale, J.; O'Halloran, T. V., Transcriptional activation of an Escherichia coli copper efflux regulon by the chromosomal MerR homologue, cueR. *J Biol Chem* **2000,** 275, (40), 31024-9.

65.   Harley, C. B.; Reynolds, R. P., Analysis of E. coli promoter sequences. *Nucleic Acids Res* **1987,** 15, (5), 2343-61.

66.   Ansari, A. Z.; Bradner, J. E.; O'Halloran, T. V., DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature* **1995,** 374, (6520), 371-5.

67.   Ansari, A. Z.; Chael, M. L.; O'Halloran, T. V., Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR. *Nature* **1992,** 355, (6355), 87-9.

68.   Heldwein, E. E.; Brennan, R. G., Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **2001,** 409, (6818), 378-82.

69.   Newberry, K. J.; Brennan, R. G., The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus. *J Biol Chem* **2004,** 279, (19), 20356-62.

70.   Changela, A.; Chen, K.; Xue, Y.; Holschen, J.; Outten, C. E.; O'Halloran, T. V.; Mondragon, A., Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* **2003,** 301, (5638), 1383-7.

71.   Outten, C. E.; O'Halloran, T. V., Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **2001,** 292, (5526), 2488-92.

72.   Munson, G. P.; Lam, D. L.; Outten, F. W.; O'Halloran, T. V., Identification of a copper-responsive two-component system on the chromosome of Escherichia coli K-12. *J Bacteriol* **2000,** 182, (20), 5864-71.

73.   Loftin, I. R.; Franke, S.; Roberts, S. A.; Weichsel, A.; Heroux, A.; Montfort, W. R.; Rensing, C.; McEvoy, M. M., A novel copper-binding fold for the periplasmic copper resistance protein CusF. *Biochemistry* **2005,** 44, (31), 10533-40.

74. Kittleson, J. T.; Loftin, I. R.; Hausrath, A. C.; Engelhardt, K. P.; Rensing, C.; McEvoy, M. M., Periplasmic metal-resistance protein CusF exhibits high affinity and specificity for both CuI and AgI. *Biochemistry* **2006,** 45, (37), 11096-102.

75. Astashkin, A. V.; Raitsimring, A. M.; Walker, F. A.; Rensing, C.; McEvoy, M. M., Characterization of the copper(II) binding site in the pink copper binding protein CusF by electron paramagnetic resonance spectroscopy. *J Biol Inorg Chem* **2005,** 10, (3), 221-30.

76. Lin, S. J.; Culotta, V. C., The ATX1 gene of Saccharomyces cerevisiae encodes a small metal homeostasis factor that protects cells against reactive oxygen toxicity. *Proc Natl Acad Sci U S A* **1995,** 92, (9), 3784-8.

77. Dancis, A.; Haile, D.; Yuan, D. S.; Klausner, R. D., The Saccharomyces cerevisiae copper transport protein (Ctr1p). Biochemical characterization, regulation by copper, and physiologic role in copper uptake. *J Biol Chem* **1994,** 269, (41), 25660-7.

78. Moller, J. V.; Juul, B.; le Maire, M., Structural organization, ion transport, and energy transduction of P-type ATPases. *Biochim Biophys Acta* **1996,** 1286, (1), 1-51.

79. Pufahl, R. A.; Singer, C. P.; Peariso, K. L.; Lin, S. J.; Schmidt, P. J.; Fahrni, C. J.; Culotta, V. C.; Penner-Hahn, J. E.; O'Halloran, T. V., Metal ion chaperone function of the soluble Cu(I) receptor Atx1. *Science* **1997,** 278, (5339), 853-6.

80. Klomp, L. W.; Lin, S. J.; Yuan, D. S.; Klausner, R. D.; Culotta, V. C.; Gitlin, J. D., Identification and functional expression of HAH1, a novel human gene involved in copper homeostasis. *J Biol Chem* **1997,** 272, (14), 9221-6.

81. Harris, Z. L.; Takahashi, Y.; Miyajima, H.; Serizawa, M.; MacGillivray, R. T.; Gitlin, J. D., Aceruloplasminemia: molecular characterization of this disorder of iron metabolism. *Proc Natl Acad Sci U S A* **1995,** 92, (7), 2539-43.

82. Bull, P. C.; Cox, D. W., Wilson disease and Menkes disease: new handles on heavy-metal transport. *Trends Genet* **1994,** 10, (7), 246-52.

83. Dancis, A.; Yuan, D. S.; Haile, D.; Askwith, C.; Eide, D.; Moehle, C.; Kaplan, J.; Klausner, R. D., Molecular characterization of a copper transport protein in S. cerevisiae: an unexpected role for copper in iron transport. *Cell* **1994,** 76, (2), 393-402.

84. Wernimont, A. K.; Huffman, D. L.; Lamb, A. L.; O'Halloran, T. V.; Rosenzweig, A. C., Structural basis for copper transfer by the metallochaperone for the Menkes/Wilson disease proteins. *Nat Struct Biol* **2000,** 7, (9), 766-71.

85.   Rosenzweig, A. C.; Huffman, D. L.; Hou, M. Y.; Wernimont, A. K.; Pufahl, R. A.; O'Halloran, T. V., Crystal structure of the Atx1 metallochaperone protein at 1.02 A resolution. *Structure* **1999,** 7, (6), 605-17.

86.   Steele, R. A.; Opella, S. J., Structures of the reduced and mercury-bound forms of MerP, the periplasmic protein from the bacterial mercury detoxification system. *Biochemistry* **1997,** 36, (23), 6885-95.

87.   Banci, L.; Bertini, I.; Del Conte, R.; Markey, J.; Ruiz-Duenas, F. J., Copper trafficking: the solution structure of Bacillus subtilis CopZ. *Biochemistry* **2001,** 40, (51), 15660-8.

88.   Yamaguchi, Y.; Heiny, M. E.; Gitlin, J. D., Isolation and characterization of a human liver cDNA as a candidate gene for Wilson disease. *Biochem Biophys Res Commun* **1993,** 197, (1), 271-7.

89.   Brewer, G. J.; Hedera, P.; Kluin, K. J.; Carlson, M.; Askari, F.; Dick, R. B.; Sitterly, J.; Fink, J. K., Treatment of Wilson disease with ammonimu tetrathiomolybdate III: Initial therapy in a total of 55 neurologically affected patients and follow up with zinc therapy. *Arch. Neurol.* **2003,** 60, 378-385.

90.   Pan, Q.; Bao, L. W.; Kleer, C. G.; Brewer, G. J.; Merajver, S. D., Antiangiogenic tetrathiomolybdate enhances the efficacy of doxorubicin against breast carcinoma. *Mol Cancer Ther* **2003,** 2, (7), 617-22.

91.   Pan, Q.; Bao, L. W.; Merajver, S. D., Tetrathiomolybdate inhibits angiogenesis and metastasis through suppression of the NFkappaB signaling cascade. *Mol Cancer Res* **2003,** 1, (10), 701-6.

92.   George, G. N.; Pickering, I. J.; Harris, H. H.; Gailer, J.; Klein, D.; Lichtmannegger, J.; Summer, K. H., Tetrathiomolybdate causes formation of hepatic copper-molybdenum clusters in an animal model of Wilson's disease. *J Am Chem Soc* **2003,** 125, (7), 1704-5.

93.   Summers, A. O., Untwist and shout: a heavy metal-responsive transcriptional regulator. *J Bacteriol* **1992,** 174, (10), 3097-101.

94.   Hobman, J. L.; Brown, N. L., bacterial mercury-resistance genes. *Met Ions Biol Syst* **1997,** 34, 527-68.

95.   Summers, A. O., Organization, expression, and evolution of genes for mercury resistance. *Annu Rev Microbiol* **1986,** 40, 607-34.

96.   Barkay, T.; Miller, S. M.; Summers, A. O., Bacterial mercury resistance from atoms to ecosystems. *FEMS Microbiol Rev* **2003,** 27, (2-3), 355-84.

97. Brown, N. L.; Stoyanov, J. V.; Kidd, S. P.; Hobman, J. L., The MerR family of transcriptional regulators. *FEMS Microbiol Rev* **2003,** 27, (2-3), 145-63.

98. Brocklehurst, K. R.; Hobman, J. L.; Lawley, B.; Blank, L.; Marshall, S. J.; Brown, N. L.; Morby, A. P., ZntR is a Zn(II)-responsive MerR-like transcriptional regulator of zntA in Escherichia coli. *Mol Microbiol* **1999,** 31, (3), 893-902.

99. Outten, C. E.; Tobin, D. A.; Penner-Hahn, J. E.; O'Halloran, T. V., Characterization of the metal receptor sites in Escherichia coli Zur, an ultrasensitive zinc(II) metalloregulatory protein. *Biochemistry* **2001,** 40, (35), 10417-23.

100. Ding, H.; Hidalgo, E.; Demple, B., The redox state of the [2Fe-2S] clusters in SoxR protein regulates its activity as a transcription factor. *J Biol Chem* **1996,** 271, (52), 33173-5.

101. Gaudu, P.; Weiss, B., SoxR, a [2Fe-2S] transcription factor, is active only in its oxidized form. *Proc Natl Acad Sci U S A* **1996,** 93, (19), 10094-8.

102. Tsaneva, I. R.; Weiss, B., soxR, a locus governing a superoxide response regulon in Escherichia coli K-12. *J Bacteriol* **1990,** 172, (8), 4197-205.

103. Rutherford, J. C.; Cavet, J. S.; Robinson, N. J., Cobalt-dependent transcriptional switching by a dual-effector MerR-like protein regulates a cobalt-exporting variant CPx-type ATPase. *J Biol Chem* **1999,** 274, (36), 25827-32.

104. Lee, S. W.; Glickmann, E.; Cooksey, D. A., Chromosomal locus for cadmium resistance in Pseudomonas putida consisting of a cadmium-transporting ATPase and a MerR family response regulator. *Appl Environ Microbiol* **2001,** 67, (4), 1437-44.

105. Borremans, B.; Hobman, J. L.; Provoost, A.; Brown, N. L.; van Der Lelie, D., Cloning and functional analysis of the pbr lead resistance determinant of Ralstonia metallidurans CH34. *J Bacteriol* **2001,** 183, (19), 5651-8.

106. Reeve, W. G.; Tiwari, R. P.; Kale, N. B.; Dilworth, M. J.; Glenn, A. R., ActP controls copper homeostasis in Rhizobium leguminosarum bv. viciae and Sinorhizobium meliloti preventing low pH-induced copper toxicity. *Mol Microbiol* **2002,** 43, (4), 981-91.

107. Noll, M.; Petrukhin, K.; Lutsenko, S., Identification of a novel transcription regulator from Proteus mirabilis, PMTR, revealed a possible role of YJAI protein in balancing zinc in Escherichia coli. *J Biol Chem* **1998,** 273, (33), 21393-401.

108. Chiu, M. L.; Folcher, M.; Katoh, T.; Puglia, A. M.; Vohradsky, J.; Yun, B. S.; Seto, H.; Thompson, C. J., Broad spectrum thiopeptide recognition specificity of the Streptomyces

lividans TipAL protein and its role in regulating gene expression. *J Biol Chem* **1999,** 274, (29), 20578-86.

109.   Baranova, N. N.; Danchin, A.; Neyfakh, A. A., Mta, a global MerR-type regulator of the Bacillus subtilis multidrug-efflux transporters. *Mol Microbiol* **1999,** 31, (5), 1549-59.

110.   Ahmed, M.; Borsch, C. M.; Taylor, S. S.; Vazquez-Laslop, N.; Neyfakh, A. A., A protein that activates expression of a multidrug efflux transporter upon binding the transporter substrates. *J Biol Chem* **1994,** 269, (45), 28506-13.

111.   Ahmed, M.; Lyass, L.; Markham, P. N.; Taylor, S. S.; Vazquez-Laslop, N.; Neyfakh, A. A., Two highly similar multidrug transporters of Bacillus subtilis whose expression is differentially regulated. *J Bacteriol* **1995,** 177, (14), 3904-10.

Chapter 2

1.   Buffoni, F.; Ignesti, G., The copper-containing amine oxidases: biochemical aspects and functional role. *Mol Genet Metab* **2000,** 71, (4), 559-64.

2.   Camakaris, J.; Voskoboinik, I.; Mercer, J. F., Molecular mechanisms of copper homeostasis. *Biochem Biophys Res Commun* **1999,** 261, (2), 225-32.

3.   Cobine, P.; Wickramasinghe, W. A.; Harrison, M. D.; Weber, T.; Solioz, M.; Dameron, C. T., The Enterococcus hirae copper chaperone CopZ delivers copper(I) to the CopY repressor. *FEBS Lett* **1999,** 445, (1), 27-30.

4.   Weissman, Z.; Berdicevsky, I.; Cavari, B. Z.; Kornitzer, D., The high copper tolerance of Candida albicans is mediated by a P-type ATPase. *Proc Natl Acad Sci U S A* **2000,** 97, (7), 3520-5.

5.   Vallee, B. L.; Auld, D. S., Zinc coordination, function, and structure of zinc enzymes and other proteins. *Biochemistry* **1990,** 29, (24), 5647-59.

6.   Coleman, J. E., Zinc proteins: enzymes, storage proteins, transcription factors, and replication proteins. *Annu Rev Biochem* **1992,** 61, 897-946.

7.   Pena, M. M.; Lee, J.; Thiele, D. J., A delicate balance: homeostatic control of copper uptake and distribution. *J Nutr* **1999,** 129, (7), 1251-60.

8.   Petris, M. J.; Mercer, J. F.; Culvenor, J. G.; Lockhart, P.; Gleeson, P. A.; Camakaris, J., Ligand-regulated transport of the Menkes copper P-type ATPase efflux pump from the Golgi apparatus to the plasma membrane: a novel mechanism of regulated trafficking. *Embo J* **1996,** 15, (22), 6084-95.

9.   Petris, M. J., The SLC31 (Ctr) copper transporter family. *Pflugers Arch* **2004,** 447, (5), 752-5.

10.   Petris, M. J.; Smith, K.; Lee, J.; Thiele, D. J., Copper-stimulated endocytosis and degradation of the human copper transporter, hCtr1. *J Biol Chem* **2003,** 278, (11), 9639-46.

11.   Ooi, C. E.; Rabinovich, E.; Dancis, A.; Bonifacino, J. S.; Klausner, R. D., Copper-dependent degradation of the Saccharomyces cerevisiae plasma membrane copper transporter Ctr1p in the apparent absence of endocytosis. *Embo J* **1996,** 15, (14), 3515-23.

12. Valentine, J. S.; Gralla, E. B., Delivering copper inside yeast and human cells. *Science* **1997,** 278, (5339), 817-8.

13. Thorvaldsen, J. L.; Sewell, A. K.; McCowen, C. L.; Winge, D. R., Regulation of metallothionein genes by the ACE1 and AMT1 transcription factors. *J Biol Chem* **1993,** 268, (17), 12512-8.

14. Garcia-Dominguez, M.; Lopez-Maury, L.; Florencio, F. J.; Reyes, J. C., A gene cluster involved in metal homeostasis in the cyanobacterium Synechocystis sp. strain PCC 6803. *J Bacteriol* **2000,** 182, (6), 1507-14.

15. Georgatsou, E.; Mavrogiannis, L. A.; Fragiadakis, G. S.; Alexandraki, D., The yeast Fre1p/Fre2p cupric reductases facilitate copper uptake and are regulated by the copper-modulated Mac1p activator. *J Biol Chem* **1997,** 272, (21), 13786-92.

16. Borghouts, C.; Werner, A.; Elthon, T.; Osiewacz, H. D., Copper-modulated gene expression and senescence in the filamentous fungus Podospora anserina. *Mol Cell Biol* **2001,** 21, (2), 390-9.

17. Beaudoin, J.; Labbe, S., The fission yeast copper-sensing transcription factor Cuf1 regulates the copper transporter gene expression through an Ace1/Amt1-like recognition sequence. *J Biol Chem* **2001,** 276, (18), 15472-80.

18. Himelblau, E.; Mira, H.; Lin, S. J.; Culotta, V. C.; Penarrubia, L.; Amasino, R. M., Identification of a functional homolog of the yeast copper homeostasis gene ATX1 from Arabidopsis. *Plant Physiol* **1998,** 117, (4), 1227-34.

19. Ruzsa, S. M.; Scandalios, J. G., Altered Cu metabolism and differential transcription of Cu/ZnSod genes in a Cu/ZnSOD-deficient mutant of maize: evidence for a Cu-responsive transcription factor. *Biochemistry* **2003,** 42, (6), 1508-16.

20. Zhou, H.; Cadigan, K. M.; Thiele, D. J., A copper-regulated transporter required for copper acquisition, pigmentation, and specific stages of development in Drosophila melanogaster. *J Biol Chem* **2003,** 278, (48), 48210-8.

21. Lin, S. J.; Pufahl, R. A.; Dancis, A.; O'Halloran, T. V.; Culotta, V. C., A role for the Saccharomyces cerevisiae ATX1 gene in copper trafficking and iron transport. *J Biol Chem* **1997,** 272, (14), 9215-20.

22. Grotz, N.; Fox, T.; Connolly, E.; Park, W.; Guerinot, M. L.; Eide, D., Identification of a family of zinc transporter genes from Arabidopsis that respond to zinc deficiency. *Proc Natl Acad Sci U S A* **1998,** 95, (12), 7220-4.

23. Pence, N. S.; Larsen, P. B.; Ebbs, S. D.; Letham, D. L.; Lasat, M. M.; Garvin, D. F.; Eide, D.; Kochian, L. V., The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator Thlaspi caerulescens. *Proc Natl Acad Sci U S A* **2000,** 97, (9), 4956-60.

24. Westin, G.; Schaffner, W., A zinc-responsive factor interacts with a metal-regulated enhancer element (MRE) of the mouse metallothionein-I gene. *Embo J* **1988,** 7, (12), 3763-70.

25. Zhao, H.; Eide, D. J., Zap1p, a metalloregulatory protein involved in zinc-responsive transcriptional regulation in Saccharomyces cerevisiae. *Mol Cell Biol* **1997,** 17, (9), 5044-52.

26. Langmade, S. J.; Ravindra, R.; Daniels, P. J.; Andrews, G. K., The transcription factor MTF-1 mediates metal regulation of the mouse ZnT1 gene. *J Biol Chem* **2000,** 275, (44), 34803-9.

27. Rouch, D. A.; Parkhill, J.; Brown, N. L., Induction of bacterial mercury- and copper-responsive promoter: functional differences between inducible systems and implications for their use in gene-fusion for *in vivo* metal biosensors. *J Ind Microbiol* **1995,** 14, 249-253.

28. Outten, F. W.; Huffman, D. L.; Hale, J. A.; O'Halloran, T. V., The independent *cue* and *cus* systems confer copper tolerance during aerobic and anaerobic growth in *Escherichia coli*. *J. Biol. Chem.* **2001,** 276, (33), 30670-30677.

29. Rensing, C.; Fan, B.; Sharma, R.; Mitra, B.; Rosen, B. P., CopA: An Escherichia coli Cu(I)-translocating P-type ATPase. *Proc Natl Acad Sci U S A* **2000,** 97, (2), 652-6.

30. Outten, F. W.; Outten, C. E.; Hale, J.; O'Halloran, T. V., Transcriptional activation of an Escherichia coli copper efflux regulon by the chromosomal MerR homologue, cueR. *J Biol Chem* **2000,** 275, (40), 31024-9.

31. Patzer, S. I.; Hantke, K., The ZnuABC high-affinity zinc uptake system and its regulator Zur in Escherichia coli. *Mol Microbiol* **1998,** 28, (6), 1199-210.

32. Brocklehurst, K. R.; Hobman, J. L.; Lawley, B.; Blank, L.; Marshall, S. J.; Brown, N. L.; Morby, A. P., ZntR is a Zn(II)-responsive MerR-like transcriptional regulator of zntA in Escherichia coli. *Mol Microbiol* **1999,** 31, (3), 893-902.

33. Frantz, B.; O'Halloran, T. V., DNA distortion accompanies transcriptional activation by the metal-responsive gene-regulatory protein MerR. *Biochemistry* **1990,** 29, (20), 4747-51.

34.	Ansari, A. Z.; Chael, M. L.; O'Halloran, T. V., Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR. *Nature* **1992,** 355, (6355), 87-9.

35.	Outten, C. E.; Outten, F. W.; O'Halloran, T. V., DNA distortion mechanism for transcriptional activation by ZntR, a Zn(II)-responsive MerR homologue in Escherichia coli. *J Biol Chem* **1999,** 274, (53), 37517-24.

36.	Ralston, D. M.; O'Halloran, T. V., Ultrasensitivity and heavy-metal selectivity of the allosterically modulated MerR transcription complex. *Proc Natl Acad Sci U S A* **1990,** 87, (10), 3846-50.

37.	O'Halloran, T. V., Transition metals in control of gene expression. *Science* **1993,** 261, (5122), 715-25.

38.	Outten, C. E.; O'Halloran, T. V., Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **2001,** 292, (5526), 2488-92.

39.	Stoyanov, J. V.; Hobman, J. L.; Brown, N. L., CueR (YbbI) of Escherichia coli is a MerR family regulator controlling expression of the copper exporter CopA. *Mol Microbiol* **2001,** 39, (2), 502-11.

40.	Leslie, A. G. W., Oxford University Press, Oxford: 1990.

41.	Kabsch, W., Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **1993,** 26, 795-800.

42.	The CCP4 suite: programs for protein crystallography. *Acta. Cryst. D* **1994,** 50, 760-763.

43.	de La Fortelle, E.; Bricogne, G., Maximum-likelihood heavy-atom parameter refinement for the multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In *Methods in Enzymology*, Academic Press: New York, 1997; Vol. 276, pp 472-494.

44.	Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **1991,** 47 (Pt 2), 110-9.

45.	Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L., Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **1998,** 54, (Pt 5), 905-21.

213

46.   Gajiwala, K. S.; Burley, S. K., Winged helix proteins. *Curr Opin Struct Biol* **2000,** 10, (1), 110-6.

47.   Huffman, J. L.; Brennan, R. G., Prokaryotic transcription regulators: more than just the helix-turn-helix motif. *Curr Opin Struct Biol* **2002,** 12, (1), 98-106.

48.   Heldwein, E. E.; Brennan, R. G., Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **2001,** 409, (6818), 378-82.

49.   Newberry, K. J.; Brennan, R. G., The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus. *J Biol Chem* **2004,** 279, (19), 20356-62.

50.   Godsey, M. H.; Baranova, N. N.; Neyfakh, A. A.; Brennan, R. G., Crystal structure of MtaN, a global multidrug transporter gene activator. *J Biol Chem* **2001,** 276, (50), 47178-84.

51.   Ansari, A. Z.; Bradner, J. E.; O'Halloran, T. V., DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature* **1995,** 374, (6520), 371-5.

52.   Nishikawa, K.; Scheraga, H. A., Geometrical criteria for formation of coiled-coil structures of polypeptide chains. *Macromolecules* **1976,** 9, (3), 395-407.

53.   Chen, K.; Yuldasheva, S.; Penner-Hahn, J. E.; O'Halloran, T. V., An atypical linear $Cu(I)-S_2$ center constitutes the high-affinity metal-sensing site in the CueR metalloregulatory protein. *J. Am. Chem. Soc.* **2003,** 125, (40), 12088-12089.

54.   Stoyanov, J. V.; Brown, N. L., The Escherichia coli copper-responsive copA promoter is activated by gold. *J Biol Chem* **2003,** 278, (3), 1407-10.

55.   Kortemme, T.; Creighton, T. E., Ionisation of cysteine residues at the termini of model alpha-helical peptides. Relevance to unusual thiol pKa values in proteins of the thioredoxin family. *J Mol Biol* **1995,** 253, (5), 799-812.

56.   Ross, W.; Park, S. J.; Summers, A. O., Genetic analysis of transcriptional activation and repression in the Tn21 mer operon. *J Bacteriol* **1989,** 171, (7), 4009-18.

57.   Noll, M.; Petrukhin, K.; Lutsenko, S., Identification of a novel transcription regulator from Proteus mirabilis, PMTR, revealed a possible role of YJAI protein in balancing zinc in Escherichia coli. *J Biol Chem* **1998,** 273, (33), 21393-401.

58.   Demple, B.; Ding, H.; Jorgensen, M., Escherichia coli SoxR protein: sensor/transducer of oxidative stress and nitric oxide. *Methods Enzymol* **2002,** 348, 355-64.

59. O'Halloran, T.; Walsh, C., Metalloregulatory DNA-binding protein encoded by the merR gene: isolation and characterization. *Science* **1987,** 235, (4785), 211-4.

60. Utschig, L. M.; Bryson, J. W.; O'Halloran, T. V., Mercury-199 NMR of the metal receptor site in MerR and its protein-DNA complex. *Science* **1995,** 268, (5209), 380-5.

61. Shewchuk, L. M.; Helmann, J. D.; Ross, W.; Park, S. J.; Summers, A. O.; Walsh, C. T., Transcriptional switching by the MerR protein: activation and repression mutants implicate distinct DNA and mercury(II) binding domains. *Biochemistry* **1989,** 28, (5), 2340-4.

62. Helmann, J. D.; Ballard, B. T.; Walsh, C. T., The MerR metalloregulatory protein binds mercuric ion as a tricoordinate, metal-bridged dimer. *Science* **1990,** 247, (4945), 946-8.

63. Abramson, J.; Riistama, S.; Larsson, G.; Jasaitis, A.; Svensson-Ek, M.; Laakkonen, L.; Puustinen, A.; Iwata, S.; Wikstrom, M., The structure of the ubiquinol oxidase from Escherichia coli and its ubiquinone binding site. *Nat Struct Biol* **2000,** 7, (10), 910-7.

64. Wilmot, C. M.; Hajdu, J.; McPherson, M. J.; Knowles, P. F.; Phillips, S. E., Visualization of dioxygen bound to copper during enzyme catalysis. *Science* **1999,** 286, (5445), 1724-8.

65. Roh, J. H.; Takenaka, Y.; Suzuki, H.; Yamamoto, K.; Kumagai, H., Escherichia coli K-12 copper-containing monoamine oxidase: investigation of the copper binding ligands by site-directed mutagenesis, elemental analysis and topa quinone formation. *Biochem Biophys Res Commun* **1995,** 212, (3), 1107-14.

66. Pesce, A.; Capasso, C.; Battistoni, A.; Folcarelli, S.; Rotilio, G.; Desideri, A.; Bolognesi, M., Unique structural features of the monomeric Cu,Zn superoxide dismutase from Escherichia coli, revealed by X-ray crystallography. *J Mol Biol* **1997,** 274, (3), 408-20.

67. Roberts, S. A.; Weichsel, A.; Grass, G.; Thakali, K.; Hazzard, J. T.; Tollin, G.; Rensing, C.; Montfort, W. R., Crystal structure and electron transfer kinetics of CueO, a multicopper oxidase required for copper homeostasis in Escherichia coli. *Proc Natl Acad Sci U S A* **2002,** 99, (5), 2766-71.

68. O'Halloran, T. V.; Culotta, V. C., Metallochaperones, an intracellular shuttle service for metal ions. *J Biol Chem* **2000,** 275, (33), 25057-60.

Chapter 3

1.   Finney, L. A.; O'Halloran, T. V., Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* **2003,** 300, (5621), 931-6.

2.   Outten, C. E.; O'Halloran, T. V., Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **2001,** 292, (5526), 2488-92.

3.   Outten, F. W.; Huffman, D. L.; Hale, J. A.; O'Halloran, T. V., The independent cue and cus systems confer copper tolerance during aerobic and anaerobic growth in Escherichia coli. *J Biol Chem* **2001,** 276, (33), 30670-7.

4.   Rensing, C.; Fan, B.; Sharma, R.; Mitra, B.; Rosen, B. P., CopA: An Escherichia coli Cu(I)-translocating P-type ATPase. *Proc Natl Acad Sci U S A* **2000,** 97, (2), 652-6.

5.   Munson, G. P.; Lam, D. L.; Outten, F. W.; O'Halloran, T. V., Identification of a copper-responsive two-component system on the chromosome of Escherichia coli K-12. *J Bacteriol* **2000,** 182, (20), 5864-71.

6.   Changela, A.; Chen, K.; Xue, Y.; Holschen, J.; Outten, C. E.; O'Halloran, T. V.; Mondragon, A., Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* **2003,** 301, (5638), 1383-7.

7.   Ralston, D. M.; O'Halloran, T. V., Ultrasensitivity and heavy-metal selectivity of the allosterically modulated MerR transcription complex. *Proc Natl Acad Sci U S A* **1990,** 87, (10), 3846-50.

8.   Ansari, A. Z.; Bradner, J. E.; O'Halloran, T. V., DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature* **1995,** 374, (6520), 371-5.

9.   Outten, C. E.; Outten, F. W.; O'Halloran, T. V., DNA distortion mechanism for transcriptional activation by ZntR, a Zn(II)-responsive MerR homologue in Escherichia coli. *J Biol Chem* **1999,** 274, (53), 37517-24.

10.   Hitomi, Y.; Outten, C. E.; O'Halloran, T. V., Extreme zinc-binding thermodynamics of the metal sensor/regulator protein, ZntR. *J Am Chem Soc* **2001,** 123, (35), 8614-5.

11.   Outten, F. W.; Outten, C. E.; Hale, J.; O'Halloran, T. V., Transcriptional activation of an Escherichia coli copper efflux regulon by the chromosomal MerR homologue, cueR. *J Biol Chem* **2000,** 275, (40), 31024-9.

12.   Heldwein, E. E.; Brennan, R. G., Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **2001,** 409, (6818), 378-82.

13. Newberry, K. J.; Brennan, R. G., The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus. *J Biol Chem* **2004,** 279, (19), 20356-62.

14. Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **1976,** 72, 248-54.

15. Brookes, E.; Boppana, R. V.; Demeler, B., Computing large sparse multivariate optimization problems with an application in biophysics. SC2006, Tampa, FL, U.S.A. IEEE 0-7695-2700-0/06. **2006**.

16. Brookes, E.; Demeler, B., Genetic algorithm optimization for obtaining accurate molecular weight distributions from sedimentation velocity experiments. Analytical Ultracentrifugation VIII,. *Prog. Colloid Polym. Sci* **2006,** 131, 78-82.

17. Demeler, B.; Brookes, E., Monte Carlo analysis of sedimentation experiments. In *Progress in Colloid and Polymer Science, in press*, 2007.

18. Johnson, M. L.; Correia, J. J.; Yphantis, D. A.; Halvorson, H. R., Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys J* **1981,** 36, (3), 575-88.

19. Demeler, B.; Saber, H.; Hansen, J. C., Identification and interpretation of complexity in sedimentation velocity boundaries. *Biophys J* **1997,** 72, (1), 397-407.

20. Kim, Y.; Geiger, J. H.; Hahn, S.; Sigler, P. B., Crystal structure of a yeast TBP/TATA-box complex. *Nature* **1993,** 365, 512-520.

21. Kim, J. L.; Nikolov, D. B.; Burley, S. K., Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **1993,** 365, 520-527.

22. Nikolov, D. B.; Chen, H.; Halay, E. D.; Hoffmann, A.; Roeder, R. G.; Burley, S. K., Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc. Natl. Acad. Sci. USA* **1996,** 93, 4862-4867.

23. Schumacher, M. A.; Choi, K. Y.; Zalkin, H.; Brennan, R. G., Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. *Science* **1994,** 266, 763-770.

24. Lewis, M.; Chang, G.; Horton, N. C.; Kercher, M. A.; Pace, H. C.; Schumacher, M. A.; Brennan, R. G.; Lu, P., Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **1996,** 271, 1247-1254.

25.   Swinger, K. K.; Lemberg, K. M.; Zhang, Y.; Rice, P. A., Flexible DNA bending in HU-DNA cocrystal structures. *EMBO J.* **2003,** 22, 3749-3760.

26.   Ohndorf, U. M.; Rould, M. A.; He, Q.; Pabo, C. O.; Lippard, S. J., Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature* **1999,** 399, 708-712.

27.   Werner, M. H.; Huth, J. R.; Gronenborn, A. M.; Clore, G. M., Molecular basis of human 46X, Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell* **1995,** 81, 705-714.

28.   Love, J. J.; Li, X.; Case, D. A.; Giese, K.; Grosschedl, R.; Wright, P. E., Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* **1995,** 376, 791-795.

29.   Masse, J. E.; Wong, B.; Yen, Y. M.; Allian, F. H.-T.; Johnson, R. C.; Fegon, J., The *S. cerevisiae* architectural HMGB protein NHP6A complexed with DNA: DNA and protein conformational changes upon binding. *J. Mol. Biol.* **2002,** 323, 263-284.

30.   Chen, C. Y.; Ko, T. P.; Lin, T. W.; Chou, C. C.; Chen, C. J.; Wang, A. H.-J., Probing the DNA kink structure induced by the hyperthermophilic chromosomal protein Sac7d. *Nucleic Acids Res.* **2005,** 33, 430-438.

31.   Umezawa, Y.; Nishio, M., CH/pi interactions in the crystal structure of TATA-box binding protein/DNA complexes. *Bioorg. Med. Chem.* **2000,** 8, 2643-2650.

32.   Fernandez-Alonso, M. d. C.; Canada, F. J.; Jimenez-Barbero, J.; Cuevas, G., Molecular recognition of saccharides by proteins. Insights on the origin of the carbohydrate-aromatic interactions. *J. Am. Chem. Soc.* **2005,** 127, 7379-7386.

33.   Brandl, M.; Weiss, M. S.; Jabs, A.; Suhnel, J.; Hilgenfeld, R., C-H---PI interactions in proteins. *J. Mol. Biol.* **2001,** 307, 357-377.

34.   Gross, C. A.; Lonetto, M.; Losick, R., *Bacterial sigma factors*. Cold Spring Harbor Labortory Press: NY, 1992; p 129-176.

35.   Burr, T.; Mitchell, J.; Kolb, A.; Minchin, S.; Busby, S., DNA sequence elements located immediately upstream of the -10 hexamer in Escherichia coli promoters: a systematic study. *Nucleic Acids Res* **2000,** 28, (9), 1864-70.

36.   Mitchell, J. E.; Zheng, D.; Busby, S. J.; Minchin, S. D., Identification and analysis of 'extended -10' promoters in Escherichia coli. *Nucleic Acids Res* **2003,** 31, (16), 4689-95.

Chapter 4

1.   Rensing, C.; Grass, G., Escherichia coli mechanisms of copper homeostasis in a changing environment. *FEMS Microbiol Rev* **2003,** 27, (2-3), 197-213.

2.   Munson, G. P.; Lam, D. L.; Outten, F. W.; O'Halloran, T. V., Identification of a copper-responsive two-component system on the chromosome of Escherichia coli K-12. *J Bacteriol* **2000,** 182, (20), 5864-71.

3.   Franke, S.; Grass, G.; Rensing, C.; Nies, D. H., Molecular analysis of the copper-transporting efflux system CusCFBA of Escherichia coli. *J Bacteriol* **2003,** 185, (13), 3804-12.

4.   Outten, F. W.; Huffman, D. L.; Hale, J. A.; O'Halloran, T. V., The independent cue and cus systems confer copper tolerance during aerobic and anaerobic growth in Escherichia coli. *J Biol Chem* **2001,** 276, (33), 30670-7.

5.   Astashkin, A. V.; Raitsimring, A. M.; Walker, F. A.; Rensing, C.; McEvoy, M. M., Characterization of the copper(II) binding site in the pink copper binding protein CusF by electron paramagnetic resonance spectroscopy. *J Biol Inorg Chem* **2005,** 10, (3), 221-30.

6.   Loftin, I. R.; Franke, S.; Roberts, S. A.; Weichsel, A.; Heroux, A.; Montfort, W. R.; Rensing, C.; McEvoy, M. M., A novel copper-binding fold for the periplasmic copper resistance protein CusF. *Biochemistry* **2005,** 44, (31), 10533-40.

7.   Zhang, L.; Koay, M.; Maher, M. J.; Xiao, Z.; Wedd, A. G., Intermolecular transfer of copper ions from the CopC protein of Pseudomonas syringae. Crystal structures of fully loaded Cu(I)Cu(II) forms. *J Am Chem Soc* **2006,** 128, (17), 5834-50.

8.   Wernimont, A. K.; Huffman, D. L.; Finney, L. A.; Demeler, B.; O'Halloran, T. V.; Rosenzweig, A. C., Crystal structure and dimerization equilibria of PcoC, a methionine-rich copper resistance protein from Escherichia coli. *J Biol Inorg Chem* **2003,** 8, (1-2), 185-94.

9.   Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode *Methods Enzymol.* **1997,** 276, 307-326.

10.   The CCP4 suite: programs for protein crystallography. *Acta. Cryst. D* **1994,** 50, 760-763.

11.   Storoni, L. C.; McCoy, A. J.; Read, R. J., Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, (Pt 3), 432-8.

12.   Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, (Pt 12 Pt 1), 2126-32.

13. Davis, I. W.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* **2004,** 32, (Web Server issue), W615-9.

14. Harford, C.; Sarkar, B., Amino terminal Cu(II)- and Ni(II)-binding (ATCUN) motif of proteins and peptides: metal binding, DNA cleavage, and other properties. *Acc. Chem. Res.* **1997,** 30, 123-130.

15. Murzin, A. G., OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *Embo J* **1993,** 12, (3), 861-7.

16. Murzin, A. G.; Brenner, S. E.; Hubbard, T.; Chothia, C., SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **1995,** 247, (4), 536-40.

17. Theobald, D. L.; Mitton-Fry, R. M.; Wuttke, D. S., Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **2003,** 32, 115-33.

18. Reddy, A. S.; Sastry, G. M.; Sastry, G. N., Cation-aromatic database. *Proteins* **2007,** 67, (4), 1179-84.

19. Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Katsari, E.; Katsaros, N.; Kubicek, K.; Mangani, S., A copper(I) protein possibly involved in the assembly of CuA center of bacterial cytochrome c oxidase. *Proc Natl Acad Sci U S A* **2005,** 102, (11), 3994-9.

20. Puig, S.; Lee, J.; Lau, M.; Thiele, D. J., Biochemical and genetic analyses of yeast and human high affinity copper transporters suggest a conserved mechanism for copper uptake. *J Biol Chem* **2002,** 277, (29), 26021-30.

21. Arnesano, F.; Banci, L.; Bertini, I.; Huffman, D. L.; O'Halloran, T. V., Solution structure of the Cu(I) and apo forms of the yeast metallochaperone, Atx1. *Biochemistry* **2001,** 40, (6), 1528-39.

22. Ducros, V.; Brzozowski, A. M.; Wilson, K. S.; Ostergaard, P.; Schneider, P.; Svendson, A.; Davies, G. J., Structure of the laccase from Coprinus cinereus at 1.68 A resolution: evidence for different 'type 2 Cu-depleted' isoforms. *Acta Crystallogr D Biol Crystallogr* **2001,** 57, (Pt 2), 333-6.

23. Hakulinen, N.; Kiiskinen, L. L.; Kruus, K.; Saloheimo, M.; Paananen, A.; Koivula, A.; Rouvinen, J., Crystal structure of a laccase from Melanocarpus albomyces with an intact trinuclear copper site. *Nat Struct Biol* **2002,** 9, (8), 601-5.

24. Nose, Y.; Rees, E. M.; Thiele, D. J., Structure of the Ctr1 copper trans'PORE'ter reveals novel architecture. *Trends Biochem Sci* **2006,** 31, (11), 604-7.

25.   Aller, S. G.; Unger, V. M., Projection structure of the human copper transporter CTR1 at 6-A resolution reveals a compact trimer with a novel channel-like architecture. *Proc Natl Acad Sci U S A* **2006,** 103, (10), 3627-32.

Chapter 5

1.   George, G. N.; Pickering, I. J.; Yu, E. Y.; Prince, R. C.; Bursakov, S. A.; Gavel, O. Y.; Moura, I.; Moura, J. G., A novel protein-bound copper-molybdenum cluster. *J Am Chem Soc* **2000,** 122, 8321-8322.

2.   Underwood, E. J., *Trace elements in human and animal nutrition*. Academic Press: New York, 1977.

3.   Chappell, W. R.; Kellogg Petersen, K., *Molydbdenum in the Environment*. Dekker: New York, 1976; Vol. 1.

4.   Mills, C. F., Trace elements in animals. *Philos Trans R Soc Lond B Biol Sci* **1979,** 288, (1026), 51-63.

5.   Clarke, N. J.; Laurie, S. H., The copper-molybdenum antagonism in ruminants. I. The formation of thiomolybdates in animal rumen. *J Inorg Biochem* **1980,** 12, (1), 37-43.

6.   El-Gallad, T. T.; Mills, C. F.; Bremner, I.; Summers, R., Thiomolybdates in rumen contents and rumen cultures. *J Inorg Biochem* **1983,** 18, (4), 323-34.

7.   Bull, P. C.; Thomas, G. R.; Rommens, J. M.; Forbes, J. R.; Cox, D. W., The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. *Nat Genet* **1993,** 5, (4), 327-37.

8.   Yamaguchi, Y.; Heiny, M. E.; Gitlin, J. D., Isolation and characterization of a human liver cDNA as a candidate gene for Wilson disease. *Biochem Biophys Res Commun* **1993,** 197, (1), 271-7.

9.   Brewer, G. J.; Hedera, P.; Kluin, K. J.; Carlson, M.; Askari, F.; Dick, R. B.; Sitterly, J.; Fink, J. K., Treatment of Wilson disease with ammonimu tetrathiomolybdate III: Initial therapy in a total of 55 neurologically affected patients and follow up with zinc therapy. *Arch. Neurol.* **2003,** 60, 378-385.

10.   Folkman, J., Tumor angiogenesis: therapeutic implications. *N Engl J Med* **1971,** 285, (21), 1182-6.

11.   Wiggins, D. L.; Granai, C. O.; Steinhoff, M. M.; Calabresi, P., Tumor angiogenesis as a prognostic factor in cervical carcinoma. *Gynecol Oncol* **1995,** 56, (3), 353-6.

12.   Weidner, N.; Carroll, P. R.; Flax, J.; Blumenfeld, W.; Folkman, J., Tumor angiogenesis correlates with metastasis in invasive prostate carcinoma. *Am J Pathol* **1993,** 143, (2), 401-9.

13. Harris, E. D., A requirement for copper in angiogenesis. *Nutr Rev* **2004,** 62, (2), 60-4.

14. Brewer, G. J., Copper control as an antiangiogenic anticancer therapy: lessons from treating Wilson's disease. *Exp Biol Med (Maywood)* **2001,** 226, (7), 665-73.

15. Brewer, G. J.; Merajver, S. D., Cancer therapy with tetrathiomolybdate: antiangiogenesis by lowering body copper--a review. *Integr Cancer Ther* **2002,** 1, (4), 327-37.

16. Pan, Q.; Kleer, C. G.; van Golen, K. L.; Irani, J.; Bottema, K. M.; Bias, C.; De Carvalho, M.; Mesri, E. A.; Robins, D. M.; Dick, R. D.; Brewer, G. J.; Merajver, S. D., Copper deficiency induced by tetrathiomolybdate suppresses tumor growth and angiogenesis. *Cancer Res* **2002,** 62, (17), 4854-9.

17. van Golen, K. L.; Bao, L.; Brewer, G. J.; Pienta, K. J.; Kamradt, J. M.; Livant, D. L.; Merajver, S. D., Suppression of tumor recurrence and metastasis by a combination of the PHSCN sequence and the antiangiogenic compound tetrathiomolybdate in prostate carcinoma. *Neoplasia* **2002,** 4, (5), 373-9.

18. Pan, Q.; Bao, L. W.; Merajver, S. D., Tetrathiomolybdate inhibits angiogenesis and metastasis through suppression of the NFkappaB signaling cascade. *Mol Cancer Res* **2003,** 1, (10), 701-6.

19. Pan, Q.; Bao, L. W.; Kleer, C. G.; Brewer, G. J.; Merajver, S. D., Antiangiogenic tetrathiomolybdate enhances the efficacy of doxorubicin against breast carcinoma. *Mol Cancer Ther* **2003,** 2, (7), 617-22.

20. Redman, B. G.; Esper, P.; Pan, Q.; Dunn, R. L.; Hussain, H. K.; Chenevert, T.; Brewer, G. J.; Merajver, S. D., Phase II trial of tetrathiomolybdate in patients with advanced kidney cancer. *Clin Cancer Res* **2003,** 9, (5), 1666-72.

21. George, G. N.; Pickering, I. J.; Harris, H. H.; Gailer, J.; Klein, D.; Lichtmannegger, J.; Summer, K. H., Tetrathiomolybdate causes formation of hepatic copper-molybdenum clusters in an animal model of Wilson's disease. *J Am Chem Soc* **2003,** 125, (7), 1704-5.

22. Rosenzweig, A. C., Copper delivery by metallochaperone proteins. *Acc Chem Res* **2001,** 34, (2), 119-28.

23. Pufahl, R. A.; Singer, C. P.; Peariso, K. L.; Lin, S. J.; Schmidt, P. J.; Fahrni, C. J.; Culotta, V. C.; Penner-Hahn, J. E.; O'Halloran, T. V., Metal ion chaperone function of the soluble Cu(I) receptor Atx1. *Science* **1997,** 278, (5339), 853-6.

24. Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode *Methods Enzymol.* **1997,** 276, 307-326.

25.   Storoni, L. C.; McCoy, A. J.; Read, R. J., Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, (Pt 3), 432-8.

26.   Rosenzweig, A. C.; Huffman, D. L.; Hou, M. Y.; Wernimont, A. K.; Pufahl, R. A.; O'Halloran, T. V., Crystal structure of the Atx1 metallochaperone protein at 1.02 A resolution. *Structure* **1999,** 7, (6), 605-17.

27.   Wernimont, A. K.; Huffman, D. L.; Lamb, A. L.; O'Halloran, T. V.; Rosenzweig, A. C., Structural basis for copper transfer by the metallochaperone for the Menkes/Wilson disease proteins. *Nat Struct Biol* **2000,** 7, (9), 766-71.

28.   Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C., PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **2002,** 58, (Pt 11), 1948-54.

29.   Sheldrick, G. M., *SHELXC*. University of Gottingen: Gottingen, Germany, 2003.

30.   Grosse-Kunstleve, R. W.; Adams, P. D., Substructure search procedures for macromolecular structures. *Acta Crystallogr D Biol Crystallogr* **2003,** 59, (Pt 11), 1966-73.

31.   de La Fortelle, E.; Bricogne, G., Maximum-likelihood heavy-atom parameter refinement for the multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In *Methods in Enzymology*, Academic Press: New York, 1997; Vol. 276, pp 472-494.

32.   Vonrhein, C.; Blanc, E.; Roversi, P.; Bricogne, G., Automated Structure Solution With autoSHARP. *Methods Mol Biol* **2006,** 364, 215-30.

33.   Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, (Pt 12 Pt 1), 2126-32.

34.   The CCP4 suite: programs for protein crystallography. *Acta. Cryst. D* **1994,** 50, 760-763.

35.   Davis, I. W.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* **2004,** 32, (Web Server issue), W615-9.

36.   DeLano, W. L. *The PyMOL Molecular Graphics System*, DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org.

37.   Hubbard, T. J.; Murzin, A. G.; Brenner, S. E.; Chothia, C., SCOP: a structural classification of proteins database. *Nucleic Acids Res* **1997,** 25, (1), 236-9.

38.   Finney, L. A.; O'Halloran, T. V., Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* **2003,** 300, (5621), 931-6.

39.   Huffman, D. L.; O'Halloran, T. V., Energetics of copper trafficking between the Atx1 metallochaperone and the intracellular copper transporter, Ccc2. *J Biol Chem* **2000,** 275, (25), 18611-4.

40.   Jeannin, Y.; Sécheresse, F.; Bernès, S.; Robert, F., Molecular architecture of copper(I) thiometallate complexes. Example of a cubane with an extra face, (NPr4)3[MS4Cu4Cl5] M=Mo, W). *Inorg. Chim. Acta.* **1992,** 198, 493-505.

41.   Li, J.; Xin, X.; Zhou, Z.; Yu, K., Synthesis and structural characterization of a large molybdenum-copper-sulphur cluster compound [Bu[n][4]N][4]]Cu[12]Mo[8]S[32]]. *Chem. Commun* **1991**, 249-250.

42.   Chen, J.; Wei, Z.; Xu, Q.; Li, H.; Lang, J., Synthesis and crystal structure of a novel novel twin nido-like hexanuclear cluster [MS4Cu5Br3(Py)7]. *Chinese J. Struct. Chem.* **2004,** 23, 464-469.

43.   Dobbek, H.; Gremer, L.; Kiefersauer, R.; Huber, R.; Meyer, O., Catalysis at a dinuclear [CuSMo(==O)OH] cluster in a CO dehydrogenase resolved at 1.1-A resolution. *Proc Natl Acad Sci U S A* **2002,** 99, (25), 15971-6.

44.   Walshe, J. M., Penicillamine, a new oral therapy for Wilson's disease. *Am J Med* **1956,** 21, (4), 487-95.

45.   Walshe, J. M., Treatment of Wilson's disease with trientine (triethylene tetramine) dichloride. *Lancet* **1982,** 1, 643-647.

46.   Hoogenraad, T. U.; Van Hattum, J.; Van den Hamer, C. J. A., Management of Wilson's disease with zinc sulfate. Experience in a series of 27 patients. *J. Neurol. Sci.* **1987,** 77, 137-146.

47.   Brewer, G. J.; Hill, G. M.; Prasad, A. S.; Cossack, Z. T.; Rabbani, P., Oral Zinc Therapy for Wilson's disease. *Ann. Intern. Med.* **1983,** 99, 314-320.

48.   Brewer, G. J.; Dick, R. D.; Johnson, V.; Brunberg, J. A.; Kluin, K. J.; Fink, J. K., Treatment of Wilson's disease with zinc: XV. Long-term follow-up studies. *J. Lab. Clin. Med.* **1998,** 132, 264-278.

Chapter 6

1.   Outten, F. W.; Outten, C. E.; Hale, J.; O'Halloran, T. V., Transcriptional activation of an Escherichia coli copper efflux regulon by the chromosomal MerR homologue, cueR. *J Biol Chem* **2000,** 275, (40), 31024-9.

2.   Changela, A.; Chen, K.; Xue, Y.; Holschen, J.; Outten, C. E.; O'Halloran, T. V.; Mondragon, A., Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* **2003,** 301, (5638), 1383-7.

3.   Heldwein, E. E.; Brennan, R. G., Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **2001,** 409, (6818), 378-82.

4.   Newberry, K. J.; Brennan, R. G., The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus. *J Biol Chem* **2004,** 279, (19), 20356-62.

5.   Jordan, S. R.; Whitcombe, T. V.; Berg, J. M.; Pabo, C. O., Systematic variation in DNA length yields highly ordered repressor-operator cocrystals. *Science* **1985,** 230, (4732), 1383-5.

6.   Rice, P. A.; Yang, S.; Mizuuchi, K.; Nash, H. A., Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **1996,** 87, (7), 1295-306.

7.   Schultz, S. C.; Shields, G. C.; Steitz, T. A., Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* **1991,** 253, (5023), 1001-7.

8.   Yang, W.; Steitz, T. A., Crystal structure of the site-specific recombinase gamma delta resolvase complexed with a 34 bp cleavage site. *Cell* **1995,** 82, (2), 193-207.

9.   Obmolova, G.; Ban, C.; Hsieh, P.; Yang, W., Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature* **2000,** 407, (6805), 703-10.

10.   Schultz, S. C.; Shields, G. C.; Steitz, T. A., Crystallization of Escherichia coli catabolite gene activator protein with its DNA binding site. The use of modular DNA. *J Mol Biol* **1990,** 213, (1), 159-66.

11.   Otwinowski, Z.; Minor, W., Processing of X-ray diffraction data collected in oscillation mode *Methods Enzymol.* **1997,** 276, 307-326.

12.   Storoni, L. C.; McCoy, A. J.; Read, R. J., Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* **2004,** 60, (Pt 3), 432-8.

13.   The CCP4 suite: programs for protein crystallography. *Acta. Cryst. D* **1994,** 50, 760-763.

14.   de La Fortelle, E.; Bricogne, G., Maximum-likelihood heavy-atom parameter refinement for the multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In *Methods in Enzymology*, Academic Press: New York, 1997; Vol. 276, pp 472-494.

15.   Vonrhein, C.; Blanc, E.; Roversi, P.; Bricogne, G., Automated Structure Solution With autoSHARP. *Methods Mol Biol* **2006,** 364, 215-30.

16.   O'Halloran, T. V.; Lippard, S. J.; Richmond, T. J.; Klug, A., Multiple heavy-atom reagents for macromolecular X-ray structure determination: application to the nucleosome core particles. *J. Mol. Biol.* **1987,** 194, 705-712.

17.   Huang, H.; Chopra, R.; Verdine, G. L.; Harrison, S. C., Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **1998,** 282, (5394), 1669-75.

18.   Huang, H.; Harrison, S. C.; Verdine, G. L., Trapping of a catalytic HIV reverse transcriptase*template:primer complex through a disulfide bond. *Chem Biol* **2000,** 7, (5), 355-64.

APPENDIX 1

Protocol for *in vitro* run-off transcription assay for ZntR and mutant in chapter 2

(1)   Cold DNA template preparation

• Set up PCR reactions to amplify pUC19/Zntfoot (described in Caryn E. Outten's thesis) containing the binding sequences for ZntR (2 tubes)

Table 1. PCR cycling parameters

Segment	Cycles	Temperature (°C)	Time
1	1	95	3 min
2	34	95	1 min
		68	1 min
		72	30 sec
3	1	4	Hold (end)

Table 2. PCR reaction components

Reaction component	Volume	Final concentration
10 ng/μL pUC19Zntfoot	10 μL	1 ng/μL
5 mM dNTP mix	4 μL	200 μM dNTPs
20 μM –47 primer	5 μL	1 μM
20 μM –48 primer	5 μL	1 μM
10 × ThermolPol buffer	10 μL	1×
Deep Vent DNAP (NEB)	1 μL	
MQ H$_2$O	65 μL	

• Transfer PCR reactions from 2 tubes into two1.5 mL eppendorf. To each one, add 200 μL phenol:chloroform:isoamyl alcohol (25:24:1), pH 8.0 and extract by vortexing.

• Spin 3 min at RT to separate phases, pipet off top aqueous phase and transfer to new tube.

• Add 22 μL 3 M NaOAc and 500 μL EtOH to aqueous phase to precipitate DNA. Vortex and put in -80°C freezer for 10 min.

• Spin 10 min at 4°C. Remove SN and add 200 μL 70% EtOH. Incubate 5 min at 4°C, spin 5 min at 4°C.

• Remove SN and speed-vac 2 min. Keep dry DNA fragments in -20°C before use.

• Digest PCR fragment with endonuclease BbvI (New England Biolabs) at 37°C for 4 hours to produce a 172-bp DNA cold template.

Table 3. Enzymatic digestion components

Reaction component	*Volume*
DNA fragements	
100 × BSA	0.5 μL
10 × NEB buffer 2	5 μL
BbvI (NEB, 2000 units/mL)	4 μL
MQ H$_2$O	40.5 μL

• Purify DNA cold template by agarose gel electrophoresis followed by gel extraction using the Qiagen gel extraction kit. Check concentration with ethidium bromide drop test (mix 1~3 μL DNA + 20 μL solution containing ethidium bromide and check intensity on UV light box.

Compare to known standards). Desired concentration is ~20 μg/mL DNA. Store in -20°C freezer.

(2)   Sequencing-gel preparation

• Place 21 g urea in a 125-mL flask, add 5 mL 10 × TBE, add 8.75 mL 40% ExplorER gel solution (J. T. Baker), add MQ H_2O until the liquid level just below 50 mL, place in 37°C shakcer to dissolve urea.

• Cleaning sequencing plates with with soap, MQ H_2O and 95% EtOH in turn.

• Assemble gel using thin spacers (0.25mm) with binder clips along the sides to hold them in space. Wrap ~2/3 of bottom of gel with yellow sequencing tape.

• Remove gel solution from shaker incubator and pour into 50-mL graduated cylinder, add MQ H_2O to bring volume up to 50 mL. Pour gel solution into 100-mL beaker, add 350 μL 10% ammonium persulfate and 35 μL TEMED (Sigma) and swirl to mix. Fill a 60-mL syringe with gel solution.

• Place assembled gel plates on bench with top end raised. Slowly squirt gel solution between gel plates from top of gel, let liquid move between plates by capillary force and gravity. Tap plates constantly to prevent bubble formation. Place notched comb into top of the gel and let dry flat on bench for 30 minutes.

• Remove comb carefully and rinse top and bottom of gel with deionized water. Place gel in electrophoresis apparatus and fill top and bottom apparatus wells with 800 mL 0.5 × TBE.

• Pre-run gel at 40-50 W before loading samples.

(3)   Run-off transcription assay

• Prepare buffer solutions: buffers were prepared with DEPC-treated water (Ambion). Rnase-free microcentrifuge tubes were used in experiments.

$5 \times$ Transcription assay buffer: 50 mM   Tris-Cl, pH 8

10 mM $MgCl_2$

5 mM $CaCl_2$

500 mM   KGlu

5 mM   DTT (fresh)

500 μg/mL   BSA

25%   Glycerol

100 μM   TPEN

Heparin-NTP mix solution: 0.15 mM   ATP

0.15 mM   GTP

0.15 mM   CTP

0.0015 mM   UTP

200 μg/mL   Heparin

Stop Solution: 9 M NH_4OAc

200 μg/mL   Yeast RNA

40 mM   EDTA

— Prepare markers

• The Century Marker Template Set directly ordered from Ambion Corp., it contains mixtures of linearized plasmids. When transcribed with T7 RNA polymerase, it produces 5 RNA transcripts with the lengths of 100, 200, 300, 400, and 500 bases.

• Set up transcriptional reaction (Table 3-4), incubate at 37°C for one hour.

Table 4. Reaction components of markers

Reaction component	*Volume (μL)*
RNA Century™ Marker (0.5 mg/mL)	1
5 × Transcription buffer	4
RNase Inhibitor (20U/μL)	1
ATP (10 mM)	1
CTP (10 mM)	1
GTP (10 mM)	1
UTP (10 mM)	1
[α-^{32}P]-UTP (800C$_i$/mmol)	3
T7 RNA polymerase (10U/μL)	2
DEPC-treated H$_2$O	5

• Add 20 μL stop buffer and 100 μL EtOH, mix well, put in –80°C freezer for 20 minutes.

• Spin 5-10 minutes at 4°C, remove supernatant.

• Wash the pellet with 100 μL 70% cold EtOH, spin 5 minutes at 4 °C, remove supernatant, speed-vac for 5 minutes.

• Resuspend dried pellet in 50 μL formamide gel loading buffer (Ambion).

• Load 0.5~1 μL of standard solution onto gel, store the rest in -20°C freezer for later usage.

— Prepare samples

• Dilute ZntR and ZntRC79S stock to 5 μM with DEPC-treated water. Prepare 2 mM AgNO$_3$ and 1.4 mM ZnSO$_4$ with DEPC-treated water.

• Mix cold template, RNAP (USB), RNase Inhibitor (Ambion), protein and metal ions per tube in 20 μL total volume (Table 3-4), and incubate 20 minutes at 37°C.

Table 5. Premix components for each sample

Component	*Volume*	*Final concentration*
Cold template (40 μg/mL)	0.5 μL	~1 μg/mL
RNAP (~1.2 μM)	0.5 μL	30 nM
RNase Inhibitor(20U/μL)	1 μL	1U/μL
5 × transcription assay buffer	4 μL	1×
ZntR or ZntRC79S (5 μM)	0 or 1 μL	0 or 25 nM
ZnSO$_4$ (1.4 mM) AgNO$_3$ (0.1-2 mM)	0 or 1 μL	0 or 70 μM 0 or 10-200 μM
DEPC-treated H$_2$O	Add up to 20 μL	

- Combine Heparin-NTP mix solution with [α-^{32}P]-UTP (MP Biomedicals, 3000 C$_i$/mmol) in ratio of 9:1, incubate 10 minutes at 37°C.

- To each reaction tube, add 10 μL Heparin-NTP-[α-^{32}P]-UTP mix solution, incubate 20 minutes at 37°C.

- Add 30 μL stop solution to each tube, then add 100 μL EtOH, mix well, put in –80°C freezer for 20 minutes.

- Spin 5-10 minutes at 4°C, remove supernatant.

- Wash the pellet with 50 μL 70% cold EtOH, spin 5 minutes at 4 °C, remove supernatant, speed-vac for 5 minutes.

- Resuspend dried pellet in 10 μL formamide gel loading buffer (Ambion).

- Heat samples 10 minutes in 95°C heat block and quench on ice.

- Load all 10 μL of sample on 7% sequencing gel.

- Run gel 2 hours at 40-50W.

- Transfer gel to chromatography paper and dry for 20-30 minutes in gel dryer.

APPENDIX 2

Dictionary file for refinement of Atx1-Cu-TM with Refmac5 in Chapter 5

The dictionary file contains the geometric description of TM group, malic acid and the bond distances restrained applied to the metal clusters.

File name: Atx1-Cu-TM.cif

```
global_
_lib_name          mon_lib
_lib_version       4.11
_lib_update        15/04/05
# -----------------------------------------------
#
# ---    LIST OF MONOMERS ---
#
data_comp_list
loop_
_chem_comp.id
_chem_comp.three_letter_code
_chem_comp.name
_chem_comp.group
_chem_comp.number_atoms_all
_chem_comp.number_atoms_nh
_chem_comp.desc_level
TTM      TTM '.                                  ' non-polymer      9    5
.
MLT      MLT 'MALIC ACID                         ' non-polymer     13    9
.
#
# ---    LIST OF LINKS ---
#
data_link_list
loop_
_chem_link.id
_chem_link.comp_id_1
_chem_link.mod_id_1
_chem_link.group_comp_1
_chem_link.comp_id_2
_chem_link.mod_id_2
_chem_link.group_comp_2
_chem_link.name
CU-CYS    CU        .        .        CYS        .        .
 bond_CU_=_CYS-SG
CU-MO     CU        .        .        TTM        .        .
 bond_CU_=_TTM-MO1
CU-S1     CU        .        .        TTM        .        .
 bond_CU_=_TTM-S1
CU-S2     CU        .        .        TTM        .        .
 bond_CU_=_TTM-S2
CU-S3     CU        .        .        TTM        .        .
```

```
 bond_CU_=_TTM-S3
# -------------------------------------------------------
#
# --- DESCRIPTION OF LINKS ---
#
data_link_CU-CYS
#
loop_
_chem_link_bond.link_id
_chem_link_bond.atom_1_comp_id
_chem_link_bond.atom_id_1
_chem_link_bond.atom_2_comp_id
_chem_link_bond.atom_id_2
_chem_link_bond.type
_chem_link_bond.value_dist
_chem_link_bond.value_dist_esd
 CU-CYS    1 CU       2 SG       metal       2.290    0.020
#
data_link_CU-MO
#
loop_
_chem_link_bond.link_id
_chem_link_bond.atom_1_comp_id
_chem_link_bond.atom_id_1
_chem_link_bond.atom_2_comp_id
_chem_link_bond.atom_id_2
_chem_link_bond.type
_chem_link_bond.value_dist
_chem_link_bond.value_dist_esd
 CU-MO    1 CU       2 MO1       metal       2.770    0.020
#
data_link_CU-S1
#
loop_
_chem_link_bond.link_id
_chem_link_bond.atom_1_comp_id
_chem_link_bond.atom_id_1
_chem_link_bond.atom_2_comp_id
_chem_link_bond.atom_id_2
_chem_link_bond.type
_chem_link_bond.value_dist
_chem_link_bond.value_dist_esd
 CU-S1    1 CU       2 S1       metal       2.290    0.020
#
data_link_CU-S2
#
loop_
_chem_link_bond.link_id
_chem_link_bond.atom_1_comp_id
_chem_link_bond.atom_id_1
_chem_link_bond.atom_2_comp_id
_chem_link_bond.atom_id_2
_chem_link_bond.type
_chem_link_bond.value_dist
```

```
_chem_link_bond.value_dist_esd
 CU-S2     1 CU      2 S2         metal        2.290    0.020
#
data_link_CU-S3
#
loop_
_chem_link_bond.link_id
_chem_link_bond.atom_1_comp_id
_chem_link_bond.atom_id_1
_chem_link_bond.atom_2_comp_id
_chem_link_bond.atom_id_2
_chem_link_bond.type
_chem_link_bond.value_dist
_chem_link_bond.value_dist_esd
 CU-S3     1 CU      2 S3         metal        2.290    0.020
#
# -----------------------------------------------------
#
# --- DESCRIPTION OF MONOMERS ---
#
data_comp_TTM
#
loop_
_chem_comp_atom.comp_id
_chem_comp_atom.atom_id
_chem_comp_atom.type_symbol
_chem_comp_atom.type_energy
_chem_comp_atom.partial_charge
_chem_comp_atom.x
_chem_comp_atom.y
_chem_comp_atom.z
 TTM           S4      S     SH1        0.000      1.438    7.008    7.851
 TTM           HS4     H     H          0.000      1.995    8.102    7.339
 TTM           MO1     MO    MO         0.000      2.376    5.242    6.989
 TTM           S3      S     SH1        0.000      2.058    5.242    4.842
 TTM           HS3     H     H          0.000      2.629    4.165    4.309
 TTM           S2      S     SH1        0.000      4.512    5.242    7.440
 TTM           HS2     H     H          0.000      5.078    4.161    6.911
 TTM           S1      S     SH1        0.000      1.438    3.477    7.851
 TTM           HS1     H     H          0.000      1.995    2.383    7.339
loop_
_chem_comp_tree.comp_id
_chem_comp_tree.atom_id
_chem_comp_tree.atom_back
_chem_comp_tree.atom_forward
_chem_comp_tree.connect_type
 TTM       S4      MO1     HS4     .
 TTM       HS4     S4      .       END
 TTM       MO1     S1      S4      .
 TTM       S3      MO1     HS3     .
 TTM       HS3     S3      .       .
 TTM       S2      MO1     HS2     .
 TTM       HS2     S2      .       .
 TTM       S1      n/a     MO1     START
```

```
 TTM        HS1     S1      .       .
loop_
_chem_comp_bond.comp_id
_chem_comp_bond.atom_id_1
_chem_comp_bond.atom_id_2
_chem_comp_bond.type
_chem_comp_bond.value_dist
_chem_comp_bond.value_dist_esd
 TTM        S1      MO1     metal       2.250       0.020
 TTM        S2      MO1     metal       2.250       0.020
 TTM        S3      MO1     metal       2.250       0.020
 TTM        MO1     S4      metal       2.250       0.020
 TTM        HS1     S1      single      1.330       0.020
 TTM        HS2     S2      single      1.330       0.020
 TTM        HS3     S3      single      1.330       0.020
 TTM        HS4     S4      single      1.330       0.020
loop_
_chem_comp_angle.comp_id
_chem_comp_angle.atom_id_1
_chem_comp_angle.atom_id_2
_chem_comp_angle.atom_id_3
_chem_comp_angle.value_angle
_chem_comp_angle.value_angle_esd
 TTM        HS4     S4      MO1     109.500     3.000
 TTM        S4      MO1     S3      109.500     3.000
 TTM        S4      MO1     S2      109.500     3.000
 TTM        S4      MO1     S1      109.500     3.000
 TTM        S3      MO1     S2      109.500     3.000
 TTM        S3      MO1     S1      109.500     3.000
 TTM        S2      MO1     S1      109.500     3.000
 TTM        MO1     S3      HS3     109.500     3.000
 TTM        MO1     S2      HS2     109.500     3.000
 TTM        MO1     S1      HS1     109.500     3.000
loop_
_chem_comp_tor.comp_id
_chem_comp_tor.id
_chem_comp_tor.atom_id_1
_chem_comp_tor.atom_id_2
_chem_comp_tor.atom_id_3
_chem_comp_tor.atom_id_4
_chem_comp_tor.value_angle
_chem_comp_tor.value_angle_esd
_chem_comp_tor.period
 TTM        var_1   HS4     S4      MO1     S1      180.000     20.000  1
 TTM        var_2   S4      MO1     S3      HS3     180.000     20.000  1
 TTM        var_3   S4      MO1     S2      HS2     180.000     20.000  1
 TTM        var_4   S4      MO1     S1      HS1     180.000     20.000  1
loop_
_chem_comp_chir.comp_id
_chem_comp_chir.id
_chem_comp_chir.atom_id_centre
_chem_comp_chir.atom_id_1
_chem_comp_chir.atom_id_2
_chem_comp_chir.atom_id_3
```

```
_chem_comp_chir.volume_sign
 TTM        chir_01  MO1    S4      S3      S2          both
#
data_comp_MLT
#
loop_
_chem_comp_atom.comp_id
_chem_comp_atom.atom_id
_chem_comp_atom.type_symbol
_chem_comp_atom.type_energy
_chem_comp_atom.partial_charge
_chem_comp_atom.x
_chem_comp_atom.y
_chem_comp_atom.z
 MLT         O5      O    OC      -0.500     0.000    0.000    0.000
 MLT         C4      C    C        0.000     0.982    0.778    0.013
 MLT         O4      O    OC      -0.500     2.163    0.376    0.066
 MLT         C3      C    CH2      0.000     0.630    2.246   -0.025
 MLT         H31     H    H        0.000    -0.456    2.343    0.035
 MLT         H32     H    H        0.000     0.979    2.660   -0.973
 MLT         C2      C    CH1      0.000     1.269    2.996    1.120
 MLT         H2      H    H        0.000     2.309    2.659    1.232
 MLT         O3      O    OH1      0.000     0.546    2.688    2.319
 MLT         HO3     H    H        0.000     0.988    3.097    3.076
 MLT         C1      C    C        0.000     1.261    4.466    0.846
 MLT         O1      O    OC      -0.500     0.516    4.969   -0.023
 MLT         O2      O    OC      -0.500     2.098    5.128    1.502
loop_
_chem_comp_tree.comp_id
_chem_comp_tree.atom_id
_chem_comp_tree.atom_back
_chem_comp_tree.atom_forward
_chem_comp_tree.connect_type
 MLT       O5      C4       .       .
 MLT       C4      C3       O4      .
 MLT       O4      C4       .       END
 MLT       C3      C2       C4      .
 MLT       H31     C3       .       .
 MLT       H32     C3       .       .
 MLT       C2      C1       C3      .
 MLT       H2      C2       .       .
 MLT       O3      C2       HO3     .
 MLT       HO3     O3       .       .
 MLT       C1      O2       C2      .
 MLT       O1      C1       .       .
 MLT       O2      n/a      C1      START
loop_
_chem_comp_bond.comp_id
_chem_comp_bond.atom_id_1
_chem_comp_bond.atom_id_2
_chem_comp_bond.type
_chem_comp_bond.value_dist
_chem_comp_bond.value_dist_esd
 MLT       O1      C1       deloc    1.250    0.020
```

```
MLT        O2        C1        deloc         1.250     0.020
MLT        C1        C2        single        1.500     0.020
MLT        O3        C2        single        1.432     0.020
MLT        C2        C3        single        1.524     0.020
MLT        C3        C4        single        1.510     0.020
MLT        O4        C4        deloc         1.250     0.020
MLT        C4        O5        deloc         1.250     0.020
MLT        H2        C2        single        1.099     0.020
MLT        HO3       O3        single        0.967     0.020
MLT        H31       C3        single        1.092     0.020
MLT        H32       C3        single        1.092     0.020
loop_
_chem_comp_angle.comp_id
_chem_comp_angle.atom_id_1
_chem_comp_angle.atom_id_2
_chem_comp_angle.atom_id_3
_chem_comp_angle.value_angle
_chem_comp_angle.value_angle_esd
 MLT       O5        C4        O4        123.000     3.000
 MLT       O5        C4        C3        118.500     3.000
 MLT       O4        C4        C3        118.500     3.000
 MLT       C4        C3        H31       109.470     3.000
 MLT       C4        C3        H32       109.470     3.000
 MLT       C4        C3        C2        109.470     3.000
 MLT       H31       C3        H32       107.900     3.000
 MLT       H31       C3        C2        109.470     3.000
 MLT       H32       C3        C2        109.470     3.000
 MLT       C3        C2        H2        108.340     3.000
 MLT       C3        C2        O3        109.470     3.000
 MLT       C3        C2        C1        109.470     3.000
 MLT       H2        C2        O3        109.470     3.000
 MLT       H2        C2        C1        108.810     3.000
 MLT       O3        C2        C1        109.470     3.000
 MLT       C2        O3        HO3       109.470     3.000
 MLT       C2        C1        O1        118.500     3.000
 MLT       C2        C1        O2        118.500     3.000
 MLT       O1        C1        O2        123.000     3.000
loop_
_chem_comp_tor.comp_id
_chem_comp_tor.id
_chem_comp_tor.atom_id_1
_chem_comp_tor.atom_id_2
_chem_comp_tor.atom_id_3
_chem_comp_tor.atom_id_4
_chem_comp_tor.value_angle
_chem_comp_tor.value_angle_esd
_chem_comp_tor.period
 MLT       var_1     O5        C4        C3        C2        175.000   20.000   3
 MLT       var_2     C4        C3        C2        C1        175.000   20.000   3
 MLT       var_3     C3        C2        O3        HO3       175.000   20.000   1
 MLT       var_4     C3        C2        C1        O2        175.000   20.000   3
loop_
_chem_comp_plane_atom.comp_id
_chem_comp_plane_atom.plane_id
```

```
_chem_comp_plane_atom.atom_id
_chem_comp_plane_atom.dist_esd
 MLT      plan-1    C1        0.020
 MLT      plan-1    O1        0.020
 MLT      plan-1    O2        0.020
 MLT      plan-1    C2        0.020
 MLT      plan-2    C4        0.020
 MLT      plan-2    C3        0.020
 MLT      plan-2    O4        0.020
 MLT      plan-2    O5        0.020
 # --------------------------------------------------
```

CURRICULUM VITAE

Yi Xue

BIRTH   Hefei, Anhui, P.R.China, November 25, 1976

EDUCATION

- Ph.D.   Inorganic chemistry, Northwestern University, December 2007
- M.S.   Inorganic chemistry, University of Science and Technology of China, P.R.China, June 2000
- B.S.   Chemistry, Xiamen University, P.R.China, June 1998

PUBLICATIONS

- **Xue, Y**., Canalizo, M., Mondragon, A., O'Halloran, T.V., Structural mechanism of transcription regulation by CueR and the importance of minor groove recognition, manuscript in preparation.

- **Xue, Y**., Alvarez, H., Robinson, C., Focia, P., Mondragon, A., O'Halloran, T.V. , Structure of tetrathiomolybdate (TM) bound to Cu-Atx1 revealing the structural anticopper mechanism of TM as the drug for Wilson's disease, manuscript in preparation.

- **Xue, Y**., Davis A.V., Balakrishnan, G., Stasser, J., Staehlin B.M., Focia, P., Spiro, T. G., Penner-Hahn, J. E., O'Halloran, T. V., Cu(I) recognition via cation-p and methionine interactions in CusF. *Nature Chemical Biology*, in press.

- Changela, A., Chen, K., **Xue, Y**., Holschen, J., Outten, C. E., O'Halloran, T. V., Mondragon, A. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science*, 2003, **301**, 1383-1387.

- Xie, Y. S., Liu, Q. L., Du, C. X., Jiang, H., Xu, X. L., Zhu, Y., **Xue, Y**. 2D hydrogen bonded copper(II) coordination network constructed by the combination of flexible and rigid ligands. *Chinese Journal of Chemistry*, 2001, **19**, 1005-1008.

- Xie, Y. S., **Xue, Y**., Kou, F. P., Lin, R. S., Liu, Q. L. Solution study and crytal structure of a pentacoordinate zinc(II) complex with N, N''-bis-(2-hydroxybenzyl)-dimethylenetriamine. *Journal of Coordination Chemistry*, 2001, **53**, 91-97.

- Xie, Y. S., **Xue, Y**., Lin, R. S., Liu, Q. L. Hydrolysis of p-nitrophenyl acetate promoted by Cu(II) and Zn(II) complexes of an alkoxyl-containing ligand with N3O2 donor set. Wuji Huaxue Xuebao, *2000*, **16**, 745-750.