

Washington University in St. Louis

Washington University Open Scholarship

All Theses and Dissertations (ETDs)

January 2009

Discovery of a Novel Metalloregular Family: Structure/Function Studies of E. Coli Rcnr

Jeffrey Iwig

Washington University in St. Louis

Follow this and additional works at: <https://openscholarship.wustl.edu/etd>

Recommended Citation

Iwig, Jeffrey, "Discovery of a Novel Metalloregular Family: Structure/Function Studies of E. Coli Rcnr" (2009). *All Theses and Dissertations (ETDs)*. 166.

<https://openscholarship.wustl.edu/etd/166>

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY

Division of Biology and Biomedical Sciences

Program in Biochemistry

Dissertation Examination Committee:

Peter Chivers, Chair

Peter Burgers

Tom Ellenberger

Robert Kranz

Timothy Lohman

John Majors

Discovery of a Novel Metalloregulator Family: Structure/Function Studies of
E. coli RcnR

by

Jeffrey Scott Iwig

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2009

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Discovery of a Novel Metalloregulator Family: Structure/Function Studies of

E. coli RcnR

by

Jeffrey Scott Iwig

Doctor of Philosophy in Biochemistry

Washington University in St. Louis, 2009

Professor Peter Chivers, Chairperson

The first row transition metal ions- Mn, Fe, Co, Ni, Cu and Zn are vital cofactors for proteins involved in diverse processes including photosynthesis, oxidative respiration and protein translation. Despite these essential functions, excess metal ions can lead to cell death through the generation of oxidative damage or the occupation of non-native metal sites. Acquiring metal ions from the environment while limiting possible toxic effects requires the coordinated regulation of metal uptake, trafficking and efflux, which in bacteria is often carried out by metal-responsive transcription factors (metalloregulators). Nickel homeostasis in *E. coli* is an ideal model system for understanding the relationships between metal physiology and metalloregulator function, as cellular nickel requirements are primarily limited to Ni-Fe hydrogenases, the levels of which can be easily controlled by different electron acceptors in the medium. While the previously studied Ni-responsive metalloregulator NikR controls Ni

import, this thesis describes the identification of RcnR, a novel regulator of Ni efflux. RcnR is the founding member of the RcnR/CsoR family of metalloregulators that are found throughout the eubacterial kingdom and respond to a range of different environmental stresses including metal ions, aldehydes and oxidative stress. RcnR-dependent gene repression is directly relieved by excess intracellular Ni(II) or Co(II), which bind to the protein at similar high-spin, six-coordinate, distorted octahedral binding sites. The metal site is distinct from the four coordinate high-affinity site of NikR, and these differences lead to the sequential repression of nickel import and activation of nickel efflux with respect to increasing extracellular nickel levels. Unlike other transcription factors, RcnR/CsoR proteins contain a unique DNA binding motif that recognizes DNA sequences containing G-tracts flanked by AT-rich inverted repeats. RcnR binds to a pair of sites in the *rcnA-rcnR* intergenic region leading to DNA wrapping, and repression of both genes. The identification of this new metalloregulator family and subsequent investigation of the functional properties of one member, RcnR, provides a solid foundation for understanding the stress responses mediated by other members of this family and the mechanisms of DNA binding by the unique, all-helical fold.

ACKNOWLEDGEMENTS

I am indebted to many people for both scientific and personal support that has enabled me to successfully navigate my journey through graduate school. First, I would like to thank Peter Chivers for his guidance and mentorship. Peter has given me the freedom to follow my instincts and to learn from my own mistakes. His enthusiasm for scientific discovery has set a solid example for everyone in the lab.

I am also grateful for my thesis committee, who has continually challenged me during my time as a graduate student and driven me to thoroughly investigate the most interesting questions associated with this project. I would especially like to thank Tim Lohman and his lab for sharing equipment and always having time to critique and help improve my work.

The Chivers lab has provided a supportive scientific environment. Although many members have moved on, I will fondly remember the positive atmosphere that made it easier to come in to work every day. I also appreciate the constructive criticism that helped drive me to complete this thesis. I would also like to thank Mike Maroney and his lab for a very fruitful collaboration. I am especially thankful to Sharon Leitch who has been a wonderful colleague.

The personal support I have received from my friends (both the new ones in St. Louis, and the old ones from Kentucky) during my time here has been amazing. They have provided a kind ear to listen to me vent when necessary, and have shared in many great experiences that have provided a welcome escape. I will never forget our fishing, canoeing and homebrewing experiences.

I also owe a tremendous amount of gratitude to Erin who has found a way to put up with me both personally and professionally. She has always helped me keep everything in perspective and been a steady influence in all aspects of my life.

Finally, I would like to thank my family who has supported me through thick and thin. I am sure they never imagined that it would take me over 20 years to finish my education. They have always been anxious to learn about science, while providing a constant reminder of what is important in life.

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	iv
Table of Contents	v
List of Figures	x
List of Tables.....	xvii

Chapter 1 Introduction

Chemical properties of transition metals	2
Cellular transition metal requirements and acquisition	4
Mechanisms of metal toxicity and resistance	7
Regulation of intracellular metal concentrations	10
Nickel is a vital microbial nutrient found in diverse coordination states	16
Nickel homeostasis in <i>E. coli</i>	20
References	25

Chapter 2 Identification of the RcnR metalloregulator and its linkage to NikR Function

Summary	41
Introduction	42
Results	46
Discussion	66
Materials and methods	73
References	79

Chapter 3 Ni(II) and Co(II) sensing by *E. coli* RcnR

Summary	85
Introduction	86
Results	89
Discussion	129
Materials and methods	148
References	161

Chapter 4 DNA recognition and wrapping by RcnR

Summary	171
Introduction	172

Results	174
Discussion	202
Materials and methods	212
References	221

Chapter 5
Perspectives and future directions

Structural studies on RcnR.....	227
A tale of two regulators: NikR and RcnR	230
Untapped potential of the RcnR/CsoR family	232
Relevance and final thoughts	236
References	238

Appendix 1
Mechanistic insights into signal sensing by the RcnR homolog FrmR

Introduction	241
Results and discussion	243
Materials and methods	258
References	264

Appendix 2
Supplementary XAS tables for Chapter 3

Tables	268
--------------	-----

Appendix 3
Metal binding by *Helicobacter pylori* HypA

Results	287
Materials and methods	292
References	294

LIST OF FIGURES

Chapter 1

Figure 1.1 General mechanisms of metal homeostasis.....	5
Figure 1.2 Metalloregulator families	11
Figure 1.3 Nickel-requiring enzymes	18

Chapter 2

Figure 2.1 Metalloregulators of <i>E. coli</i>	43
Figure 2.2 Nickel-dependent regulation of $P_{nikABCDE-lacZ}$ and $P_{rcnA-lacZ}$ in RZ4500	47
Figure 2.3 Regulation of $P_{nikABCDE-lacZ}$ and $P_{rcnA-lacZ}$ in LB media	48
Figure 2.4 Neither nitrate nor NikR affect $P_{rcnA-lacZ}$ expression	50
Figure 2.5 Sequence conservation of the RcnR protein family	51
Figure 2.6 RcnR is necessary for Ni- and Co-dependent P_{rcnA} activation	55
Figure 2.7 NikA does not affect $P_{rcnA-lacZ}$ expression	57
Figure 2.8 RcnR directly interacts with <i>rcnA</i> promoter region	58
Figure 2.9 Cobalt inhibits specific RcnR DNA binding	59
Figure 2.10 The effect of RcnA on P_{nik} expression	62
Figure 2.11 Model for nickel homeostasis in <i>E. coli</i>	70

Chapter 3

Figure 3.1 Divalent metal induction of P_{rcnA} expression.....	90
Figure 3.2 Secondary structure and stability of <i>E. coli</i> RcnR measured by CD spectroscopy	92
Figure 3.3 Oligomeric state of RcnR	93
Figure 3.4 Electronic spectroscopy of Ni or Co-substituted RcnR	96
Figure 3.5 Iodoacetamide modified Ni- and Co-RcnR	98
Figure 3.6 Ni(II)- and Co(II)-binding affinities determined by fluorescence competition assay	100
Figure 3.7 XAS of metal-substituted RcnR proteins	102
Figure 3.8 Effect of exposure time on XAS data for Ni-RcnRC35A	104
Figure 3.9 EXAFS comparison of M(II)-RcnR and M(II)-C35A RcnR	106
Figure 3.10 XAS of metal substituted RcnR proteins in NaCl	111
Figure 3.11 X-band EPR spectrum of Co(II)-RcnR	116
Figure 3.12 Effect of side-chain substitution of conserved residues on RcnR Function in vivo	117
Figure 3.13 Electronic absorption spectra of Ni-substituted RcnR mutant proteins.....	121
Figure 3.14 Electronic absorption spectra of Co-substituted RcnR mutant proteins.....	122

Figure 3.15 RcnR mutant metal titrations monitored by UV-visible spectroscopy	125
Figure 3.16 Fingerprint of the RcnR/CsoR family	136
Figure 3.17 Alignment of predicted Ni-/Co-dependent regulators	141
Figure 3.18 Alignment of RcnR/CsoR family members	145

Chapter 4

Figure 4.1 RcnR binds to an extended region of P_{rcnA}	175
Figure 4.2 RcnR autoregulates its expression	178
Figure 4.3 RcnR protects one face of the DNA duplex	180
Figure 4.4 RcnR DNA affinity increases with length of flanking DNA	182
Figure 4.5 Flanking DNA sequence is not important for RcnR binding ...	184
Figure 4.6 RcnR introduces a single negative supercoil into DNA	186
Figure 4.7 RcnR is tetrameric in Buffer P.....	188
Figure 4.8 One RcnR tetramers binds to single site duplexes by ITC.....	189
Figure 4.9 ITC with dual operator site duplex	192
Figure 4.10 RcnR binds to an extended Site 1 with a 1:1 stoichiometry ...	194
Figure 4.11 RcnR binding to a single site does not lead to DNA distortion ...	195
Figure 4.12 TACT inverted repeats are necessary for high-affinity DNA binding	197
Figure 4.13 P_{rcnA} displays A-DNA properties	199
Figure 4.14 G-tracts impart A-form properties to P_{rcnA} and are required for RcnR binding	200
Figure 4.15 RcnR does not protect G-tracts from DMS.....	205
Figure 4.16 Common sequence features of DNA-binding sites of RcnR/CsoR proteins	209

Appendix 1

Figure A1.1 P_{frm} expression is FrmR dependent.....	244
Figure A1.2 P_{frm} is induced by small aldehydes	245
Figure A1.3 FrmR binds to P_{frm} in the presence of Zn(II)	248
Figure A1.4 Glutathione is not necessary for FrmR formaldehyde sensing ...	250
Figure A1.5 FrmR directly binds Zn(II)	251
Figure A1.6 XAS of Zn-FrmR	253
Figure A1.7 Identification of formaldehyde-insensitive FrmR mutants	255

Appendix 3

Figure A3.1 HypA binds metal ions	288
Figure A3.2 Ni(II) quenches intrinsic HypA fluorescence	290
Figure A3.3 Ni(II) alters secondary structure and stability	291

List of Tables

Chapter 2

Table 2.1 <i>rcnR</i> gene neighbors	53
Table 2.2 Ni accumulation and hydrogenase activity in <i>E. coli</i> strains	64
Table 2.3 <i>E. coli</i> strains, plasmids and oligonucleotides	74

Chapter 3

Table 3.1 EXAFS fits for RcnR complexes	107
Table 3.2 UV-visible spectral features for M(II)-substituted RcnRs	123
Table 3.3 Oligonucleotides	149
Table 3.4 Plasmids and bacterial strains	151

Chapter 4

Table 4.1 Oligonucleotides	213
Table 4.2 Plasmids and bacterial strains	215

Appendix 1

Table A1.1 Oligonucleotides	259
Table A1.2 Plasmids and bacterial strains	260

Appendix 2

Table A2.1 Selected EXAFS fits for Ni(II)-RcnR in NaBr	268
Table A2.2 Selected EXAFS fits for C35A Ni(II)-RcnR in NaBr	272
Table A2.3 Selected EXAFS fits for Ni(II)-RcnR in NaCl	274
Table A2.4 Selected EXAFS fits for Co(II)-RcnR in NaBr	278
Table A2.5 Selected EXAFS fits for C35A Co(II)-RcnR in NaBr	281
Table A2.6 Selected EXAFS fits for Co(II)-RcnR in NaCl	283

Chapter 1
Introduction

The first row transition metals, Mn, Fe, Co, Ni, Cu and Zn, play unique enzymatic and structural roles in all domains of life. While commonly referred to as trace elements, Fe and Zn are actually quite abundant in organisms, present at $\sim 2 \times 10^5$ atoms/cell in *E. coli*¹ and in humans iron is the 12th most abundant element with ~ 2.3 g present in a typical adult². Organisms must concentrate necessary metals from the environment and allocate them to their proper cellular locations while avoiding the toxic consequences of excess metal ions. This thesis is focused on the regulation of nickel levels in *E. coli*, which can be used as a model system to understand general properties of cellular metal homeostasis.

I. Chemical Properties of Transition Metals

While the first row transition metals share many similar features, small differences in ionic radius and number of valence electrons tune their thermodynamic, kinetic and chemical properties, allowing for a staggering number of different functions that are usually specific for a given metal. Protein recognition of metal ions is fundamentally distinct from recognition of small organic molecules, which are commonly distinguished by patterns of hydrogen bond acceptors/donors, hydrophobic interactions and significant size differences. Similar metals have very small discrepancies in ionic radii, for instance Fe(II) and Co(II) only differ by 0.04 \AA^3 . However, there are additional inherent properties of

metal ions that allow them to carry out various specific functions in the cell. The polarizability and number of *d*-electrons help dictate the preferred coordination geometry and types of ligands that can bind, as demonstrated by Mn(II) (d^5) which is generally only found in an octahedral geometry bound to O-containing and, less often, N-containing ligands, while Zn(II) (d^{10}) is often found in tetrahedral, trigonal bipyramidal and octahedral sites bound to both S and N/O donors⁴. Zn(II) is also a stronger Lewis acid than Mn(II), allowing it to polarize bound small molecules, as for carbonic anhydrase, where a tetrahedrally bound Zn(II) activates a H₂O molecule, poising it to attack CO₂⁵.

Affinities of a given small molecule or protein for metal ions generally follows the Irving-Williams series [Mn(II) < Fe(II) < Co(II) < Ni(II) < Cu(II) > Zn(II)], which is reflective of metal ion size and second ionization energy⁴. This indicates that cells require mechanisms other than, or in addition to, binding affinity to ensure proteins receive the correct ion. Ligand exchange rates and mechanisms also differ for different transition metals and coordination geometries, making some sites very stable and others dynamic, capable of rapid chemical reactions or sensory functions. For example, the exchange rate for metal-coordinated water molecules is five orders of magnitude faster for Cu(II) than Ni(II)⁶. Metals also exist in multiple oxidation states, which can be tuned by the bound ligands, allowing metalloproteins to carry out a variety of redox reactions. These include both one- and two-electron transfers such as those associated with photosystem I of oxygen-evolving plants that are necessary for photosynthesis⁷.

II. Cellular Transition Metal Requirements and Acquisition

Historical Perspective

The disinfectant power of transition metals has been known for centuries as ancient Egyptian texts circa 1600 BC describe the practice of sterilizing drinking water and wounds with copper⁸. While metal preparations were used for a variety of therapies in many ancient societies, it was not until the 18th century that the bases for some treatments were discovered. These include the demonstration by Lemery and Geoffroy that iron is a natural constituent of human blood, therefore explaining the effectiveness of iron supplements for anemic patients who exhibited fatigue and paleness⁹. It took still another century until the discovery that hemoglobin binds iron¹⁰. Independently, European chemists were developing the concepts of metal coordination chemistry¹¹. Since the mid-1800's ideas regarding metal biology and chemistry have been integrated, leading to the discovery of thousands of metalloproteins and an estimation that up to one-third of all proteins require metal ions¹. These proteins carry out essential reactions in central metabolic pathways, electron transport and DNA metabolism, and necessitate dedicated metal homeostasis machinery (transporters, chaperones and regulators) to ensure that the correct metal reaches the proper location (Figure 1.1). These components will be explained throughout the following sections on metals in disease and toxicity.

The Role of Metals in Human Disease and Virulence

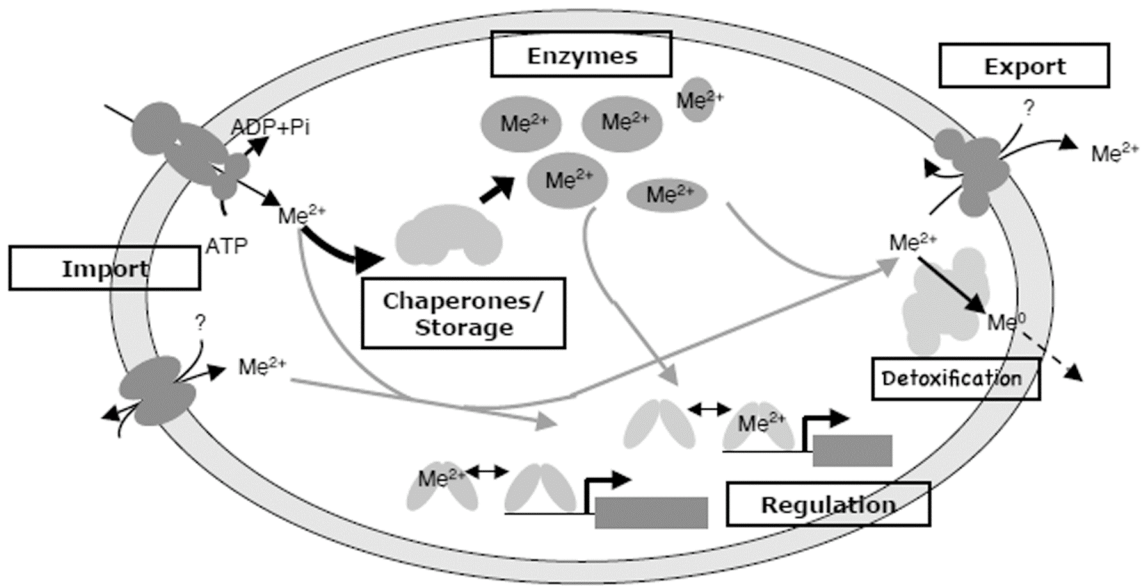


Figure 1.1 General mechanisms of metal homeostasis. Six classes of proteins are important for maintaining proper metal levels in a cell. **Import.** Importers use the energy of ion gradients or ATP to bring metal ions across the membrane, and these are distributed to other pathways depending on the metabolic needs of the cell. **Chaperones/Storage.** Metal chaperones transfer metal ions from transporters into the complex active sites of metalloenzymes, inhibiting promiscuous interactions. Storage proteins such as metallothioneins or ferritin are used to buffer intracellular metal ion concentrations during metal replete conditions. **Enzymes.** Approximately one-third of proteins require metal ions, and these proteins play diverse roles in processes such as photosynthesis, the electron transport chain and DNA metabolism. **Export.** Metal efflux proteins utilize ATP or ion gradients to pump excess metal ions out of the cell when enzyme requirements are fulfilled. **Detoxification.** In rare cases, enzymes can

reduce or oxidize the metal ion to a less toxic state. For example, some bacterial species reduce Hg(II) to Hg(0) which diffuses out of the cell. **Regulation.** The metal homeostasis machinery is regulated to conserve energy and prevent loss of necessary metal ions while ensuring that excess metal does not wreak havoc in the cell. In bacteria, transcriptional regulators that have increased or decreased DNA-binding activity upon metal binding are common.

The cell membrane prevents metal ions from passively entering the cytoplasm, necessitating the use of metal importers to obtain these required nutrients. Transporters are structurally diverse and can utilize ATP or ion gradients to drive import of either free metal ions, or ion complexes such as Fe³⁺-citrate¹². Because of the vital roles of transition metals, deficiencies in humans lead to disease, and pathogenic bacteria with insufficient metal levels are less virulent. These deficiencies are often linked to defects in metal acquisition. Menkes disease is an X-linked recessive disorder linked to the loss of function of a copper transport protein, which results in the inactivation of cuproenzymes, loss of neuronal function and usually death by the age of 10^{13,14}. In pathogens, metalloproteins, metal transporters, and regulatory proteins are known virulence factors. In a *Bordetella bronchiseptica* infection model, bacteria missing key iron acquisition machinery show reduced colonization of the lung and trachea¹⁵. Iron acquisition of pathogens is of utmost importance and hosts have evolved innate immunity mechanisms to sequester iron from bacteria to prevent infection^{16,17}. Similarly, in *Shigella* species, which cause dysentery, deletion of *fur*, encoding an Fe-responsive transcription factor, results in a drastic decrease in the ability of the bacteria to invade a cell monolayer due to the lack of expression of Fur-controlled virulence factors¹⁸.

III. Mechanisms of Metal Toxicity and Resistance

Toxicity

Despite the abundance of unique, beneficial transition metal functions in all living cells, excess metal concentrations are toxic, demanding the evolution of mechanisms of metal detoxification. In addition, metals such as mercury, which have no known biological function, are treated strictly as toxic entities. Just as metal deficiencies can lead to human disease, genetic disorders that cause metal hyperaccumulation are also known^{19,20}. At high levels metals can bind to non-native sites, deplete glutathione levels and generate reactive oxygen species, which damage lipids, proteins and DNA, leading to cell death²¹⁻²⁵. Toxicity can also be caused by direct inactivation of particular enzyme classes. For example, in *E. coli* excess copper has been shown to specifically inactivate dehydratases by displacing Fe from solvent-exposed Fe-S clusters, leading to a defect in branched chain amino acid synthesis²¹. The liberated iron can further damage the cell through generation of hydroxyl radicals by the Fenton reaction²⁵.

Resistance

To combat these deleterious processes, cells utilize metal sequestration, efflux and reduction/oxidation machinery. One common family of metal scavengers is the metallothioneins, which are small, cysteine-rich proteins found in organisms ranging from cyanobacteria to humans that chelate excess zinc, copper or cadmium inside the cell and prevent metal binding to adventitious sites²⁶. Presumably, metal ions stored by scavenger proteins can be recycled to proper metalloproteins as necessary and it is hypothesized that toxicity can also be circumvented by excretion of metal chelators into the environment²⁷.

Metallothioneins are of particular interest due to their ability to inactivate platinum-based cancer drugs (reviewed in ²⁸). Metallochaperones similarly chelate metal ions, but they have the unique ability to specifically transfer cognate metals to metalloenzymes, ensuring that metal fluxes occur in the correct direction and prevent free metal ions from encountering non-native sites²⁹.

A complementary mechanism of metal resistance utilizes export pumps to remove excess metal ions from the cytoplasm. Like importers, these proteins utilize the energy from ATP hydrolysis or ion gradients to scavenge excess metal ions, and are necessary for the virulence of bacteria such as *Streptococcus pneumoniae*³⁰. Efflux proteins generally have relaxed specificity relative to importers, allowing for the decontamination of multiple metals^{31,32}.

Metal ion detoxification is a highly specialized process that reduces or oxidizes metals to a less toxic species and is exemplified by the Mer (mercury resistance) proteins, which are often encoded on plasmids or transposable elements in a wide variety of bacteria. This set of proteins detoxifies Hg(II) by specifically importing the toxic ion into the cell and reducing it to Hg(0), which lacks affinity for any cellular functional group and freely diffuses out of the cell³³⁻³⁵.

For metal ions with specific cellular functions, it is unclear how the native metalloproteins compete with the homeostatic machinery to acquire the proper cofactor. In Chapter 2 of this thesis I have examined the in vivo competition

between a Ni(II)/Co(II)-efflux protein and other Ni(II)-binding proteins to provide insight into this question.

IV. Regulation of Intracellular Metal Concentrations

Because both intracellular metal requirements and environmental metal levels fluctuate, it is vital for cells to not only possess the arsenal of detoxification machinery described above, but also sense these changes and alter protein levels accordingly. This allows cells to obtain metals from the environment when they are metal deficient, and prevent the buildup of toxic levels under metal-replete conditions, while preventing unnecessary energy investments, thereby maintaining metal homeostasis. Metal-responsive regulation has been observed at transcriptional, post-transcriptional³⁶ and post-translational^{37,38} levels. In prokaryotes transcriptional regulation is most common and is mediated by metalloregulators, which are a group of diverse proteins whose DNA-binding properties are modulated by direct interaction with metal ions (reviewed by Giedroc and Arunkumar³⁹). Classic bacterial two-component regulatory systems can also contribute to metal sensing^{40,41}. Metalloregulators must be exquisitely tuned to sense specific metals and utilize the energy of metal binding to allosterically drive conformational changes that affect DNA binding. These mechanisms are functionally similar to the classic examples of the *lac* repressor, in which inducer (allolactose) binding leads to a decrease in DNA-binding affinity and

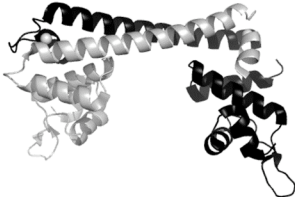
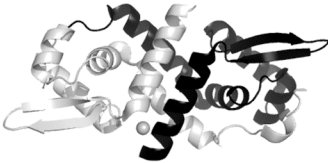

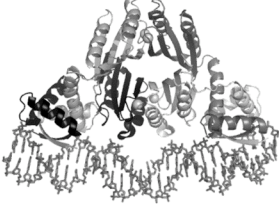

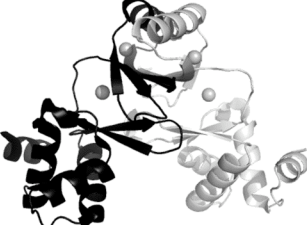
Family	Signals	DNA-binding Motif	DNA Affinity Apo	DNA Affinity Holo	
MerR	Cu, Zn, Ag, Cd, Au, Hg, Pb, redox	Winged H-T-H	High (repressor)	High (activator)	
ArsR/SmtB	Co, Ni, Cu, Zn, Pb, Bi, Ag, Cd, As, Hg, Sb, redox	Winged H-T-H	High (repressor)	Low	
CopY	Cu	Winged H-T-H	High (repressor)	Low	
NikR	Ni	R-H-H	Low	High (repressor/activator)	
DtxR	Mn, Fe	Winged H-T-H	Low	High (repressor)	
Fur	Mn, Fe, Ni, Zn, redox	Winged H-T-H	Low	High (repressor)	

Figure 1.2. Metalloregulator Families. The six characterized metalloregulator structural families are shown, along with the known signals that alter DNA binding activity. Redox signals may include oxidative (eg, hydroxyl radical) or nitrosative (eg, nitric oxide) stresses. All of these regulators have either a winged-helix-turn-helix (H-T-H) or ribbon-helix-helix (R-H-H) DNA-binding motif. Structures are shown for CueR (1Q05⁴²), SmtB (1R23⁴³), Blal (1SD4⁴⁴), NikR-DNA (2HZV⁴⁵), MntR (2F5F⁴⁶) and Fur (2W57⁴⁷).

subsequent derepression of gene expression⁴⁸⁻⁵⁰, and the *trp* repressor, which binds DNA only when bound to the corepressor tryptophan^{51,52}.

Metalloregulator Families Are Structurally Diverse

The known examples of metalloregulators can be divided into six structural families³⁹ (Figure 1.2). Based on their functions, regulators are separated into those for which metal binding leads to increased expression of metal exporters, detoxification or storage systems and those for which metal binding leads to a decrease in importer levels. Several families also contain members that respond to oxidative or nitrosative stress, often utilizing a metal cofactor⁵³⁻⁵⁶. Within each family, except the NikR (Ni) and CopY families, are proteins that respond to a distinct set of signals by utilizing different types or numbers of protein ligands. For example, in the MerR family, MerR, CueR and ZntR all bind their cognate metal ions [Hg(II), Cu(I) and Zn(II), respectively] at an equivalent position in each protein, but specificity is achieved, at least in part, by the presence of different metal coordinating amino acids which lead to trigonal planar mercury coordination by MerR⁵⁷, a two-coordinate linear copper site for CueR⁴² and a tetrahedral bi-metallic zinc cluster in ZntR⁴². In general, functional specificity is not dictated solely by metal affinity, instead only metal ions that can bind with the “correct” coordination geometry will induce the proper allosteric response leading to changes in DNA binding and subsequent alterations in gene expression. This is perhaps best shown for the nickel-responsive regulator NikR from *E. coli*, for which structural, biochemical and in vivo evidence show that of

the physiologically relevant metals, only Ni(II) forms a square planar complex and is capable of acting as a corepressor, whereas other metals form trigonal planar, tetrahedral or octahedral complexes that are incapable of repressing gene expression⁵⁸.

Although each family has a distinct overall structure, the metal and DNA binding functions in each family are separable to distinct protein domains and allosteric communication between these domains is necessary to coordinate the two activities. All known metalloregulators are functional as dimers or tetramers, and the metal-binding domain is often responsible for oligomerization. Metal binding occurs at the interface between monomers, which can lead to alterations in protein quaternary structure (Fur, SmtB/ArsR) and/or oligomeric state (DtxR) and subsequent changes in DNA affinity for some families^{59,43,60}. For other protein families the metal-induced changes are less clear. The DNA-binding domain of each protein in these families is either a winged helix-turn-helix or ribbon-helix-helix, and much of the metalloregulator structural diversity is due to differences in the metal-binding/oligomerization domains.

Sensitivity and Specificity of Metal Recognition

It is unclear how fundamental metal-binding properties control the sensitivity of metalloregulators so they do not alter metal fluxes before cellular metal requirements are fulfilled. In addition to the kinetic and thermodynamic aspects of the metal-protein interaction, the cellular environment is also important. For example, *Mycobacterium* NmtR, which senses Ni(II) and Co(II) in

its native environment, but only Co(II) in a cyanobacterial host⁶¹. This complexity highlights the need for systems in which multiple regulators from the same organism respond to the same metal ion in order to understand metalloregulator function, as the measured biophysical and biochemical parameters for regulators provide more insight when compared to one another under identical conditions.

Unfortunately, very few of these systems have been identified, with an exception being the Zn-responsive ZntR and Zur pair from *E. coli*, where ZntR activates expression of a gene encoding a Zn efflux protein⁶², and Zur represses transcription of a gene encoding a Zn importer⁶³. In vitro transcription experiments demonstrated that the responses of the regulators are staggered such that Zur repression occurs at lower zinc concentrations than ZntR activation⁶⁴. These responses can at least partially be accounted for by the metal-binding affinities of the proteins, as Zur binds Zn ~6-fold more tightly, although the metal coordination sites are both tetrahedral with similar ligand sets^{64,65}. However, this view is oversimplified, as apo-ZntR is preferentially degraded by proteases over Zn-ZntR⁶⁶ in vivo. Yamamoto and coworkers have found that the pair of regulators actually behaves differently in response to chronic Zn stress versus the acute response observed soon after a metal pulse⁶⁷, implying that kinetic components to metal binding are also important. It is not known if other sets of complementary regulators will behave similarly to ZntR and Zur. Metal complexes with different coordination numbers have different ligand

exchange rates and mechanisms (associative versus dissociative) so the coordination numbers themselves will also strongly impact transcriptional responses. In addition, relating the responses of ZntR and Zur to changes in physiology is difficult due to the high Zn-requirement in *E. coli* caused by the large number of zinc-requiring enzymes, many of which are essential⁶⁴.

Overall, much has been discovered about metalloregulators, but there remains a limited understanding of how the biochemical properties of metal binding are evolutionarily tuned for metalloregulator functions in response to changing metabolic states, and this is partly due to the lack of ideal systems to investigate these relationships. In this thesis I have identified a second Ni-responsive metalloregulator in *E. coli* (RcnR) that, when paired with NikR, provides a powerful system capable of providing key insights into many outstanding questions about metal physiology. The RcnR/CsoR family includes members that respond to a variety of metal and redox signals and differs from known metalloregulators with respect to its mechanism of DNA binding.

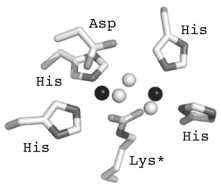
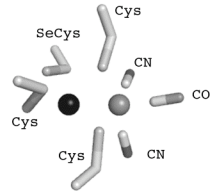
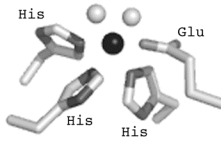
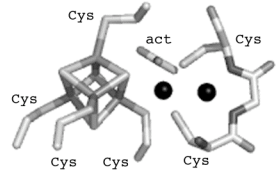
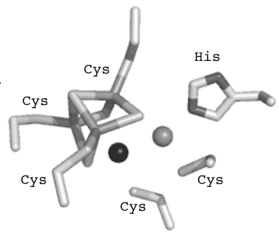
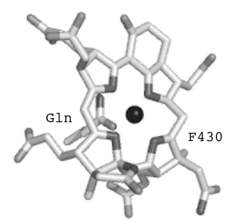
V. Nickel is a Vital Microbial Nutrient Found in Diverse Coordination States

The nickel physiologies of microorganisms provide powerful models to understand the complex interactions of metal transporters, chaperones, storage proteins and metalloregulators, due in part to the small number of Ni enzymes per organism, the levels of which can often be perturbed in a predictable manner⁶⁸. Nickel is a versatile transition metal, with Ni(II) being the most

common oxidation state, although Ni(III) and Ni(IV) can occur^{4,11}. Ni(II) is a d^8 ion capable of forming both diamagnetic and paramagnetic complexes, usually in square planar (four-coordinate), square pyramidal (five) or octahedral (six) geometries, and is a borderline metal in the hard/soft classification meaning that it can bind to both S and N/O ligands^{4,69}.

Functions of Nickel in Biological Systems

The spectrum of inherent features of nickel is nicely illustrated by the diversity of coordination environments seen for the nine known Ni-requiring enzymes, where the nickel ion plays an enzymatic role, often through direct substrate coordination as opposed to a structural function⁶⁹ (Figure 1.3). Many of these enzymes have extremely primitive origins, and were perhaps more widespread when CH₄ and H₂ were prevalent in the atmosphere, but still play key roles in central metabolic processes of many diverse bacterial and archaeal species, such as the utilization of carbon monoxide as a sole carbon source⁷⁰. Of these proteins, only urease, is found in higher organisms such as fungi⁷¹ and plants⁷². There are no known Ni-containing enzymes in humans, though a small amount of Ni (~100 µg/day) is necessary in our diet⁷³, possibly for gut microbiota. Oatmeal, dried beans, nuts and chocolate are known to be rich in nickel⁷⁴. The active sites of many Ni-containing enzymes are complex and contain non-protein ligands such as the tetrapyrrole ring of methyl coenzyme M reductase⁷⁵ and CN and CO ligands of Ni-Fe hydrogenase⁷⁶, and often require metallochaperones for metal insertion. Some Ni-enzymes such as hydrogenase⁷⁷⁻⁷⁹ and superoxide

Enzyme	Reaction	Ni Coordination	Structure
Urease	$\text{H}_2\text{N}-\text{CO}-\text{NH}_2 + 2\text{H}_2\text{O} \rightleftharpoons 2\text{NH}_3 + \text{H}_2\text{CO}_3$	bimetallic octahedral, square pyramidal	
Hydrogenase	$2\text{H}^+ + 2\text{e}^- \rightleftharpoons \text{H}_2$	Ni-Fe cluster tetrahedral	
Acireductone dioxygenase	$1,2 \text{ dihydroxy-3-keto-5-methylthiopentane} + \text{O}_2 \rightleftharpoons \text{HCOOH} + \text{methylthiopropionate} + \text{CO}$	octahedral	
Acetyl-CoA Synthase	$\text{CH}_3-\text{Co}(\text{III})\text{FeSP} + \text{CoASH} + \text{CO} \rightleftharpoons \text{CH}_3-\text{C}(\text{O})-\text{S}-\text{CoA} + \text{Co}(\text{I})\text{FeSP}$	bimetallic square planar, tetrahedral	
CO Dehydrogenase	$2\text{e}^- + 2\text{H}^+ + \text{CO}_2 \rightleftharpoons \text{CO} + \text{H}_2\text{O}$	1Ni-3Fe-S cluster tetrahedral	
MeCoM Reductase	$\text{CH}_3-\text{CoM} + \text{CoB}-\text{SH} \rightleftharpoons \text{CH}_4 + \text{CoM}-\text{S}-\text{S}-\text{CoB}$	F ₄₃₀ cofactor square pyramidal	

Enzyme	Reaction	Ni Coordination	Structure
Superoxide Dismutase	$2\text{H}^+ + 2\text{O}_2^{\cdot -} \rightleftharpoons \text{H}_2\text{O}_2 + \text{O}_2$	square planar	
Glyoxylase I	Methylglyoxal \rightleftharpoons Lactate + H ₂ O	octahedral	
Glycerol-1-Phosphate Dehydrogenase	dihydroxyacetone phosphate \rightleftharpoons glycerol-1-phosphate	unknown	

Figure 1.3. Nickel-requiring enzymes. There are nine known Ni-requiring enzymes that catalyze a diverse set of reactions with four-, five- and six-coordinate Ni sites. Crystal structures are shown for urease (1FWJ⁸⁸), Ni-Fe hydrogenase (1CC1⁸⁹), acireductone dioxygenase (1M4O⁹⁰), acetyl-CoA synthase (1OAO⁹¹), carbon monoxide dehydrogenase (1SU8⁹²), methyl CoM reductase (1MRO⁷⁵), Ni-superoxide dismutase (1T6U⁹³) and glyoxylase I (1F9Z⁹⁴). Nickel ions are shown in black spheres, other metal ions in gray spheres, and water molecules in white spheres.

dismutase (SOD)⁸⁰ have counterparts that perform the same function with a different metal. SODs are found in all domains of life with Ni⁸¹, Fe⁸², Mn⁸³⁻⁸⁵ or Cu and Zn^{86,87} in the active sites and this could reflect the relative availability of each metal to the respective organisms.

Each organism that possesses a Ni-containing enzyme must also have the correct chaperones, dedicated Ni transporters and regulatory proteins. Ni-responsive members of the Fur⁹⁵, NikR^{96,97} and SmtB/ArsR^{61,98} families have been identified. The combination of different transporters, regulators and enzymes result in many unique physiologies⁹⁹ in different bacteria and archaea, but these organisms likely share common mechanisms of maintaining metal homeostasis. A number of genetically tractable microorganisms have only one or two enzymes, making them ideal systems with which to understand how different trafficking processes work in concert to avoid nickel toxicity while ensuring sufficient nickel flux to the appropriate enzymes, which are mechanistic aspects that are difficult to understand in cells with complex metal physiologies.

VI. Ni Homeostasis in *E. coli*

Enzymes and Chaperones

Four Ni-Fe hydrogenase enzymes constitute the major nickel requirement for *E. coli*^{100,101,68}, with glyoxylase I¹⁰² representing a significantly smaller demand. Hydrogenases catalyze the oxidation of molecular hydrogen to electrons and protons or the reverse process, playing key roles in energy

generation and storage of excess reducing equivalents. Each hydrogenase isozyme is expressed under a subset of anaerobic growth conditions. For example hydrogenase-2 cleaves hydrogen gas for energy production and is synthesized in the presence of fumarate⁶⁸, while hydrogenase-3 is part of a large enzyme complex important for the oxidation of formate into carbon dioxide and hydrogen, and is expressed in the presence of its substrate¹⁰³. Therefore, hydrogenase expression and the cellular nickel requirement can be carefully modulated through the addition of these small molecule metabolites.

The complex hydrogenase active site contains a nickel ion bound to four cysteine residues, two of which also coordinate the adjacent iron center, whose coordination sphere is completed by CO and two CN ligands⁷⁶ (Figure 1.3). A dedicated set of chaperone proteins are necessary for sequential generation of the Fe ligands, insertion of the metal ions into the active site, and proteolytic processing of the holoenzyme, with each step necessary to produce an active enzyme. Ni insertion is accomplished by a set of three chaperones^{104,105} that directly bind Ni(II)¹⁰⁶⁻¹⁰⁸ and possess GTPase activity¹⁰⁹, but the details of metal insertion remain elusive. These additional metal-binding sites help ensure that nickel in the cell is tightly bound and not free to associate with non-native sites.

Transport and Regulation

Before these chaperones transfer Ni(II) into the hydrogenase active site, nickel must first be acquired from the environment by the NikABCDE transporter, which is required for hydrogenase activity under physiological metal levels¹¹⁰.

This multicomponent transporter couples ATP hydrolysis to substrate transport, and utilizes a periplasmic binding protein (NikA) to directly bind the substrate, the nature of which is controversial. Recent evidence indicated that NikA interacts with small molecule-Ni complexes^{111,112}, suggesting that bacteria must adapt to recognize the dominant metal complexes in their environment, or synthesize Ni chelators, and that these specific Ni complexes could be present in the cytosol. Lower capacity non-specific transporters such as CorA can also contribute to intracellular nickel levels^{113,114}. *nikABCDE* transcript levels, and therefore, intracellular Ni levels are intimately linked to hydrogenase levels through the action of the global anaerobic activator FNR and NarLX two-component repressor, which control expression of both the hydrogenase genes¹¹⁵ and *nikABCDE*¹¹³ in response to anoxia and nitrate, respectively. NikR provides an additional regulatory input, repressing *nikABCDE* transcription when sufficient nickel is present⁹⁶.

NikR has an N-terminal ribbon-helix-helix DNA binding domain and a C-terminal tetramerization domain^{97,116,117} that is also the site of four high-affinity ($K_d=7$ pM)¹¹⁸, square planar Ni(II)-binding sites per tetramer¹¹⁶. Ni(II) binding occurs across the tetramerization interface, utilizing two histidines and one cysteine from one monomer and a third histidine from the adjacent monomer, and this somehow drives reorientation of the DNA-binding domains to increase DNA-binding affinity, and subsequent transcriptional repression^{116,119}. A second set of Ni(II) binding sites are present in the NikR-DNA complex, these being six-

coordinate octahedral sites with a His₂(N/O)₄ ligand set, but the location and role of these sites is currently unclear⁵⁸.

Efflux

Although NikR can decrease nickel influx by repressing *nikABCDE* expression, at high extracellular concentrations nickel can still enter the cell through other nonspecific pathways. The recently discovered Ni(II)/Co(II) efflux pump, RcnA, plays a key role in reducing the toxicity of these metals in *E. coli*¹²⁰. A *P_{rcnA}-uidA* reporter fusion was induced ~30 fold upon addition of NiCl₂ or CoCl₂ to *E. coli* grown in rich media, indicating that transcription is somehow activated in the presence of these metals, but the mechanism for this was not investigated further¹²⁰.

While much has been learned about nickel homeostasis in *E. coli*, how RcnA levels are controlled is not known and this significantly impedes a complete understanding of the intricacies of this machinery, and using it as a model for more complex systems. The initial goal of these studies was to understand the molecular basis of *rcnA* regulation, and how this is integrated with the other regulatory, transport and trafficking components important for Ni homeostasis in *E. coli*. In Chapter 2, I identify RcnR as a direct transcriptional regulator of *rcnA*, and the founding member of a large, novel family of bacterial transcription factors. Through rigorous investigation of the mechanisms of Ni(II) recognition by RcnR, a comparison with the nickel-binding properties of NikR is possible in Chapter 3, allowing for an understanding of how these two proteins carry out

their complementary regulatory functions in vivo. RcnR has a previously unidentified DNA-binding fold, and the mechanisms of DNA recognition are described in Chapter 4 and compared to the predicted, distinct binding modes of other proteins of this family. Together these studies provide a more complete picture of nickel homeostasis, and new insights into the signal sensing and DNA binding of this new family of bacterial regulators.

References

1. Finney, L.A. & O'Halloran, T.V. Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* **300**, 931-936(2003).
2. Emsley, J. *The Elements*. (Oxford University Press, USA: 1998).
3. Shannon, R. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides . *Acta Crystallogr A* **32**, 751-767(1976).
4. Silva, J.J.R.F.D. & Williams, R.J.P. *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*. (Oxford University Press, USA: 2001).
5. Coleman, J.E. Mechanism of action of carbonic anhydrase. Substrate, sulfonamide, and anion binding. *J. Biol. Chem* **242**, 5212-5219(1967).
6. Lippard, S.J. & Berg, J.M. *Principles of Bioinorganic Chemistry*. (University Science Books: 1994).
7. Amunts, A. & Nelson, N. Plant photosystem I design in the light of evolution. *Structure* **17**, 637-650(2009).
8. Hostýnek, J.J. & Maibach, H.I. *Copper and the Skin*. 305(2006).
9. Boccio, J. et al. Current knowledge of iron metabolism. *Biological Trace Element Research* **92**, 189-211(2003).
10. Steinberg, M.H. et al. *Disorders of hemoglobin*. 1268(Cambridge University Press: 2001).
11. Cotton, F.A., Wilkinson, G. & Gaus, P.L. *Basic Inorganic Chemistry, 3rd Edition*. (Wiley: 1994).

12. Hussein, S., Hantke, K. & Braun, V. Citrate-dependent iron transport system in *Escherichia coli* K-12. *Eur. J. Biochem* **117**, 431-437(1981).
13. Bertini, I. & Rosato, A. Menkes disease. *Cell. Mol. Life Sci* **65**, 89-91(2008).
14. Madsen, E. & Gitlin, J.D. Copper and iron disorders of the brain. *Annu. Rev. Neurosci* **30**, 317-337(2007).
15. Brickman, T.J., Anderson, M.T. & Armstrong, S.K. Bordetella iron transport and virulence. *Biometals* **20**, 303-322(2007).
16. Schaible, U.E. et al. Correction of the iron overload defect in beta-2-microglobulin knockout mice by lactoferrin abolishes their increased susceptibility to tuberculosis. *J. Exp. Med* **196**, 1507-1513(2002).
17. Singh, P.K. et al. A component of innate immunity prevents bacterial biofilm development. *Nature* **417**, 552-555(2002).
18. Payne, S.M. et al. Iron and pathogenesis of *Shigella*: iron acquisition in the intracellular environment. *Biometals* **19**, 173-180(2006).
19. Levi, S., Cozzi, A. & Arosio, P. Neuroferritinopathy: a neurodegenerative disorder associated with L-ferritin mutation. *Best Pract Res Clin Haematol* **18**, 265-276(2005).
20. Logan, J.I. et al. Hereditary caeruloplasmin deficiency, dementia and diabetes mellitus. *QJM* **87**, 663-670(1994).
21. Macomber, L. & Imlay, J.A. The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. *Proc. Natl. Acad. Sci. U.S.A* **106**, 8344-8349(2009).

22. Mahboob, M. et al. Lipid peroxidation and antioxidant enzyme activity in different organs of mice exposed to low level of mercury. *J Environ Sci Health B* **36**, 687-697(2001).
23. Mello Filho, A.C., Hoffmann, M.E. & Meneghini, R. Cell killing and DNA damage by hydrogen peroxide are mediated by intracellular iron. *Biochem. J* **218**, 273-275(1984).
24. Valko, M., Morris, H. & Cronin, M.T.D. Metals, toxicity and oxidative stress. *Curr. Med. Chem* **12**, 1161-1208(2005).
25. Imlay, J.A., Chin, S.M. & Linn, S. Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro. *Science* **240**, 640-642(1988).
26. Robinson, N.J., Whitehall, S.K. & Cavet, J.S. Microbial metallothioneins. *Adv. Microb. Physiol* **44**, 183-213(2001).
27. Clarke, S.E., Stuart, J. & Sanders-Loehr, J. Induction of siderophore activity in *Anabaena* spp. and its moderation of copper toxicity. *Appl. Environ. Microbiol* **53**, 917-922(1987).
28. Knipp, M. Metallothioneins and platinum(II) anti-tumor compounds. *Curr. Med. Chem* **16**, 522-537(2009).
29. Rosenzweig, A.C. Metallochaperones: bind and deliver. *Chem. Biol* **9**, 673-677(2002).

30. Rosch, J.W. et al. Role of the manganese efflux system *mntE* for signalling and pathogenesis in *Streptococcus pneumoniae*. *Mol. Microbiol* **72**, 12-25(2009).
31. Tong, L. et al. A novel histidine-rich CPx-ATPase from the filamentous cyanobacterium *Oscillatoria brevis* related to multiple-heavy-metal cotolerance. *J. Bacteriol* **184**, 5027-5035(2002).
32. Nies, D.H. & Silver, S. Plasmid-determined inducible efflux is responsible for resistance to cadmium, zinc, and cobalt in *Alcaligenes eutrophus*. *J. Bacteriol* **171**, 896-900(1989).
33. Barkay, T., Miller, S.M. & Summers, A.O. Bacterial mercury resistance from atoms to ecosystems. *FEMS Microbiol. Rev* **27**, 355-384(2003).
34. Schottel, J.L. The mercuric and organomercurial detoxifying enzymes from a plasmid-bearing strain of *Escherichia coli*. *J. Biol. Chem* **253**, 4341-4349(1978).
35. Wilson, J.R. et al. MerF is a mercury transport protein: different structures but a common mechanism for mercuric ion transporters? *FEBS Lett* **472**, 78-82(2000).
36. Massé, E. & Gottesman, S. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A* **99**, 4620-4625(2002).

37. Adle, D.J. & Lee, J. Expressional control of a cadmium-transporting P1B-type ATPase by a metal sensing degradation signal. *J. Biol. Chem* **283**, 31460-31468(2008).
38. Petris, M.J. et al. Copper-stimulated endocytosis and degradation of the human copper transporter, hCtr1. *J. Biol. Chem* **278**, 9639-9646(2003).
39. Giedroc, D.P. & Arunkumar, A.I. Metal sensor proteins: nature's metalloregulated allosteric switches. *Dalton Trans* 3107-3120(2007).doi:10.1039/b706769k
40. Anderson, L.A. et al. Characterisation of the molybdenum-responsive ModE regulatory protein and its binding to the promoter region of the modABCD (molybdenum transport) operon of Escherichia coli. *Eur. J. Biochem* **246**, 119-126(1997).
41. Ogawa, T. et al. A two-component signal transduction pathway regulates manganese homeostasis in Synechocystis 6803, a photosynthetic organism. *J. Biol. Chem* **277**, 28981-28986(2002).
42. Changela, A. et al. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* **301**, 1383-1387(2003).
43. Eicken, C. et al. A metal-ligand-mediated intersubunit allosteric switch in related SmtB/ArsR zinc sensor proteins. *J. Mol. Biol* **333**, 683-695(2003).
44. Safo, M.K. et al. Crystal structures of the BlaI repressor from Staphylococcus aureus and its complex with DNA: insights into transcriptional regulation of the bla and mec operons. *J. Bacteriol* **187**, 1833-1844(2005).

45. Schreiter, E.R. et al. NikR-operator complex structure and the mechanism of repressor activation by metal ions. *Proc. Natl. Acad. Sci. U.S.A* **103**, 13676-13681(2006).
46. Kliegman, J.I. et al. Structural basis for the metal-selective activation of the manganese transport regulator of *Bacillus subtilis*. *Biochemistry* **45**, 3493-3505(2006).
47. Sheikh, M.A. & Taylor, G.L. Crystal structure of the *Vibrio cholerae* ferric uptake regulator (Fur) reveals insights into metal co-ordination. *Mol. Microbiol* **72**, 1208-1220(2009).
48. JACOB, F. & MONOD, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol* **3**, 318-356(1961).
49. Jobe, A. & Bourgeois, S. lac Repressor-operator interaction. VI. The natural inducer of the lac operon. *J. Mol. Biol* **69**, 397-408(1972).
50. Riggs, A.D., Newby, R.F. & Bourgeois, S. lac repressor--operator interaction. II. Effect of galactosides and other ligands. *J. Mol. Biol* **51**, 303-314(1970).
51. Arvidson, D.N., Bruce, C. & Gunsalus, R.P. Interaction of the *Escherichia coli* trp aporepressor with its ligand, L-tryptophan. *J. Biol. Chem* **261**, 238-243(1986).
52. Gunsalus, R.P. & Yanofsky, C. Nucleotide sequence and expression of *Escherichia coli* trpR, the structural gene for the trp aporepressor. *Proc. Natl. Acad. Sci. U.S.A* **77**, 7117-7121(1980).

53. Ding, H. & Dimple, B. Direct nitric oxide signal transduction via nitrosylation of iron-sulfur centers in the SoxR transcription activator. *Proc. Natl. Acad. Sci. U.S.A* **97**, 5146-5150(2000).
54. Lee, J. & Helmann, J.D. The PerR transcription factor senses H₂O₂ by metal-catalysed histidine oxidation. *Nature* **440**, 363-367(2006).
55. Bsat, N., Chen, L. & Helmann, J.D. Mutation of the Bacillus subtilis alkyl hydroperoxide reductase (ahpCF) operon reveals compensatory interactions among hydrogen peroxide stress genes. *J. Bacteriol* **178**, 6579-6586(1996).
56. Pomposiello, P.J., Bennik, M.H. & Dimple, B. Genome-wide transcriptional profiling of the Escherichia coli responses to superoxide stress and sodium salicylate. *J. Bacteriol* **183**, 3890-3902(2001).
57. Helmann, J.D., Ballard, B.T. & Walsh, C.T. The MerR metalloregulatory protein binds mercuric ion as a tricoordinate, metal-bridged dimer. *Science* **247**, 946-948(1990).
58. Leitch, S. et al. Nickel-specific response in the transcriptional regulator, Escherichia coli NikR. *J. Am. Chem. Soc* **129**, 5085-5095(2007).
59. D'Aquino, J.A. et al. Mechanism of metal ion activation of the diphtheria toxin repressor DtxR. *Proc. Natl. Acad. Sci. U.S.A* **102**, 18408-18413(2005).
60. Jacquamet, L. et al. Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding. *Mol. Microbiol* (2009).doi:10.1111/j.1365-2958.2009.06753.x

61. Cavet, J.S. et al. A nickel-cobalt-sensing ArsR-SmtB family repressor. Contributions of cytosol and effector binding sites to metal selectivity. *J. Biol. Chem* **277**, 38441-38448(2002).
62. Brocklehurst, K.R. et al. ZntR is a Zn(II)-responsive MerR-like transcriptional regulator of zntA in Escherichia coli. *Mol. Microbiol* **31**, 893-902(1999).
63. Patzer, S.I. & Hantke, K. The ZnuABC high-affinity zinc uptake system and its regulator Zur in Escherichia coli. *Mol. Microbiol* **28**, 1199-1210(1998).
64. Outten, C.E. & O'Halloran, T.V. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488-2492(2001).
65. Hitomi, Y., Outten, C.E. & O'Halloran, T.V. Extreme zinc-binding thermodynamics of the metal sensor/regulator protein, ZntR. *J. Am. Chem. Soc* **123**, 8614-8615(2001).
66. Pruteanu, M., Neher, S.B. & Baker, T.A. Ligand-controlled proteolysis of the Escherichia coli transcriptional regulator ZntR. *J. Bacteriol* **189**, 3017-3025(2007).
67. Yamamoto, K. & Ishihama, A. Transcriptional response of Escherichia coli to external zinc. *J. Bacteriol* **187**, 6333-6340(2005).
68. Sawers, R.G., Ballantine, S.P. & Boxer, D.H. Differential expression of hydrogenase isoenzymes in Escherichia coli K-12: evidence for a third isoenzyme. *J. Bacteriol* **164**, 1324-1331(1985).
69. Mulrooney, S.B. & Hausinger, R.P. Nickel uptake and utilization by microorganisms. *FEMS Microbiol. Rev* **27**, 239-261(2003).

70. Lindahl, P.A. & Chang, B. The evolution of acetyl-CoA synthase. *Orig Life Evol Biosph* **31**, 403-434(2001).
71. Lubbers, M.W. et al. Purification and characterization of urease from *Schizosaccharomyces pombe*. *Can. J. Microbiol* **42**, 132-140(1996).
72. Dixon, N.E. et al. Letter: Jack bean urease (EC 3.5.1.5). A metalloenzyme. A simple biological role for nickel? *J. Am. Chem. Soc* **97**, 4131-4133(1975).
73. Nielsen, F.H. Nutritional requirements for boron, silicon, vanadium, nickel, and arsenic: current knowledge and speculation. *FASEB J* **5**, 2661-2667(1991).
74. Mertz, W. *Trace Elements in Human and Animal Nutrition : Volume 1 (Trace Elements in Human & Animal Nutrition)*. (Academic Press: 1987).
75. Ermler, U. et al. Crystal structure of methyl-coenzyme M reductase: the key enzyme of biological methane formation. *Science* **278**, 1457-1462(1997).
76. Volbeda, A. et al. Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. *Nature* **373**, 580-587(1995).
77. Fontecilla-Camps, J.C. et al. Structure/function relationships of [NiFe]- and [FeFe]-hydrogenases. *Chem. Rev* **107**, 4273-4303(2007).
78. Peters, J.W. et al. X-ray crystal structure of the Fe-only hydrogenase (Cpl) from *Clostridium pasteurianum* to 1.8 angstrom resolution. *Science* **282**, 1853-1858(1998).

79. Nicolet, Y. et al. Desulfovibrio desulfuricans iron hydrogenase: the structure shows unusual coordination to an active site Fe binuclear center. *Structure* **7**, 13-23(1999).
80. Culotta, V.C., Yang, M. & O'Halloran, T.V. Activation of superoxide dismutases: Putting the metal to the pedal. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1763**, 747-758(2006).
81. Youn, H.D. et al. A novel nickel-containing superoxide dismutase from *Streptomyces* spp. *Biochem. J* **318 (Pt 3)**, 889-896(1996).
82. Yost, F.J. & Fridovich, I. An iron-containing superoxide dismutase from *Escherichia coli*. *J. Biol. Chem* **248**, 4905-4908(1973).
83. Weisiger, R.A. & Fridovich, I. Superoxide dismutase. Organelle specificity. *J. Biol. Chem* **248**, 3582-3592(1973).
84. Vance, P.G., Keele, B.B. & Rajagopalan, K.V. Superoxide dismutase from *Streptococcus mutans*. Isolation and characterization of two forms of the enzyme. *J. Biol. Chem* **247**, 4782-4786(1972).
85. Keele, B.B., McCord, J.M. & Fridovich, I. Superoxide dismutase from *Escherichia coli* B. A new manganese-containing enzyme. *J. Biol. Chem* **245**, 6176-6181(1970).
86. Fee, J.A. Studies on the reconstitution of bovine erythrocyte superoxide dismutase. 3. Evidence for a strong interdependence between Cu^{2+} and Zn^{2+} binding in the expression of the spectroscopic properties of the native

- protein and for a close proximity of the Zn²⁺ and Cu²⁺ sites. *Biochim. Biophys. Acta* **295**, 107-116(1973).
87. McCord, J.M. & Fridovich, I. Superoxide dismutase. An enzymic function for erythrocyte hemocuprein (hemocuprein). *J. Biol. Chem* **244**, 6049-6055(1969).
88. Pearson, M.A. et al. Structures of Cys319 variants and acetohydroxamate-inhibited *Klebsiella aerogenes* urease. *Biochemistry* **36**, 8164-8172(1997).
89. Garcin, E. et al. The crystal structure of a reduced [NiFeSe] hydrogenase provides an image of the activated catalytic center. *Structure* **7**, 557-566(1999).
90. Pochapsky, T.C. et al. A refined model for the structure of acireductone dioxygenase from *Klebsiella* ATCC 8724 incorporating residual dipolar couplings. *J. Biomol. NMR* **34**, 117-127(2006).
91. Darnault, C. et al. Ni-Zn-[Fe₄-S₄] and Ni-Ni-[Fe₄-S₄] clusters in closed and open subunits of acetyl-CoA synthase/carbon monoxide dehydrogenase. *Nat. Struct. Biol* **10**, 271-279(2003).
92. Dobbek, H. et al. Carbon monoxide induced decomposition of the active site [Ni-4Fe-5S] cluster of CO dehydrogenase. *J. Am. Chem. Soc* **126**, 5382-5387(2004).
93. Barondeau, D.P. et al. Nickel superoxide dismutase structure and mechanism. *Biochemistry* **43**, 8038-8047(2004).

94. He, M.M. et al. Determination of the structure of Escherichia coli glyoxalase I suggests a structural basis for differential metal activation. *Biochemistry* **39**, 8719-8727(2000).
95. Ahn, B. et al. Nur, a nickel-responsive regulator of the Fur family, regulates superoxide dismutases and nickel transport in Streptomyces coelicolor. *Mol. Microbiol* **59**, 1848-1858(2006).
96. De Pina, K. et al. Isolation and characterization of the nikR gene encoding a nickel-responsive regulator in Escherichia coli. *J. Bacteriol* **181**, 670-674(1999).
97. Chivers, P.T. & Sauer, R.T. Regulation of high affinity nickel uptake in bacteria. Ni²⁺-Dependent interaction of NikR with wild-type and mutant operator sites. *J. Biol. Chem* **275**, 19735-19741(2000).
98. Campbell, D.R. et al. Mycobacterial cells have dual nickel-cobalt sensors: sequence relationships and metal sites of metal-responsive repressors are not congruent. *J. Biol. Chem* **282**, 32298-32310(2007).
99. Zhang, Y. et al. Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization. *BMC Genomics* **10**, 78(2009).
100. Andrews, S.C. et al. A 12-cistron Escherichia coli operon (hyf) encoding a putative proton-translocating formate hydrogenlyase system. *Microbiology (Reading, Engl.)* **143 (Pt 11)**, 3633-3647(1997).

101. Ballantine, S.P. & Boxer, D.H. Nickel-containing hydrogenase isoenzymes from anaerobically grown *Escherichia coli* K-12. *J. Bacteriol* **163**, 454-459(1985).
102. Clugston, S.L. et al. Overproduction and characterization of a dimeric non-zinc glyoxalase I from *Escherichia coli*: evidence for optimal activation by nickel ions. *Biochemistry* **37**, 8754-8763(1998).
103. Rossmann, R., Sawers, G. & Böck, A. Mechanism of regulation of the formate-hydrogenlyase pathway by oxygen, nitrate, and pH: definition of the formate regulon. *Mol. Microbiol* **5**, 2807-2814(1991).
104. Lutz, S. et al. Molecular characterization of an operon (hyp) necessary for the activity of the three hydrogenase isoenzymes in *Escherichia coli*. *Mol. Microbiol* **5**, 123-135(1991).
105. Zhang, J.W. et al. A role for SlyD in the *Escherichia coli* hydrogenase biosynthetic pathway. *J. Biol. Chem* **280**, 4360-4366(2005).
106. Atanassova, A. & Zamble, D.B. *Escherichia coli* HypA is a zinc metalloprotein with a weak affinity for nickel. *J. Bacteriol* **187**, 4689-4697(2005).
107. Leach, M.R. et al. Metal binding activity of the *Escherichia coli* hydrogenase maturation factor HypB. *Biochemistry* **44**, 12229-12238(2005).
108. Hottenrott, S. et al. The *Escherichia coli* SlyD is a metal ion-regulated peptidyl-prolyl cis/trans-isomerase. *J. Biol. Chem* **272**, 15697-15701(1997).

109. Maier, T., Lottspeich, F. & Böck, A. GTP hydrolysis by HypB is essential for nickel insertion into hydrogenases of *Escherichia coli*. *Eur. J. Biochem* **230**, 133-138(1995).
110. Navarro, C., Wu, L.F. & Mandrand-Berthelot, M.A. The nik operon of *Escherichia coli* encodes a periplasmic binding-protein-dependent transport system for nickel. *Mol. Microbiol* **9**, 1181-1191(1993).
111. Cherrier, M.V. et al. Crystallographic and spectroscopic evidence for high affinity binding of FeEDTA(H₂O)- to the periplasmic nickel transporter NikA. *J. Am. Chem. Soc* **127**, 10075-10082(2005).
112. Cherrier, M.V. et al. Structural characterization of a putative endogenous metal chelator in the periplasmic nickel transporter NikA. *Biochemistry* **47**, 9937-9943(2008).
113. Rowe, J.L., Starnes, G.L. & Chivers, P.T. Complex transcriptional control links NikABCDE-dependent nickel transport with hydrogenase expression in *Escherichia coli*. *J Bacteriol* **187**, 6317-23(2005).
114. Wu, L.F. et al. Antagonistic effect of nickel on the fermentative growth of *Escherichia coli* K-12 and comparison of nickel and cobalt toxicity on the aerobic and anaerobic growth. *Environ. Health Perspect* **102 Suppl 3**, 297-300(1994).
115. Richard, D.J. et al. Transcriptional regulation in response to oxygen and nitrate of the operons encoding the [NiFe] hydrogenases 1 and 2 of

- Escherichia coli. *Microbiology (Reading, Engl.)* **145 (Pt 10)**, 2903-2912(1999).
116. Schreiter, E.R. et al. Crystal structure of the nickel-responsive transcription factor NikR. *Nat. Struct. Biol* **10**, 794-799(2003).
117. Chivers, P.T. & Sauer, R.T. NikR is a ribbon-helix-helix DNA-binding protein. *Protein Sci* **8**, 2494-2500(1999).
118. Chivers, P.T. & Sauer, R.T. NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA. *Chem. Biol* **9**, 1141-1148(2002).
119. Bradley, M.J., Chivers, P.T. & Baker, N.A. Molecular dynamics simulation of the Escherichia coli NikR protein: equilibrium conformational fluctuations reveal interdomain allosteric communication pathways. *J. Mol. Biol* **378**, 1155-1173(2008).
120. Rodrigue, A., Effantin, G. & Mandrand-Berthelot, M. Identification of rcnA (yohM), a nickel and cobalt resistance gene in Escherichia coli. *J Bacteriol* **187**, 2912-6(2005).

Chapter 2
Identification of the RcnR Metalloregulator and Its Linkage to NikR
Function

This chapter is adapted from Iwig, J.S., Rowe, J.L. and Chivers, P.T. *Mol Microbiol.* 2006 62: 252-262. Ni accumulation and hydrogenase plate assays were carried out by Jessica Rowe.

Summary

The nickel physiology of *Escherichia coli* is dominated by its Ni-Fe hydrogenase isozymes, which are expressed under anaerobic growth conditions. Hydrogenase activity in *E. coli* requires the NikABCDE nickel transporter, which is transcriptionally repressed by NikR in the presence of excess nickel. Recently, a nickel and cobalt-efflux protein, RcnA, was identified in *E. coli*. This chapter examines the effect of RcnA on nickel homeostasis in *E. coli*. Under nickel-limiting conditions, deletion of *rcnA* increased NikR activity *in vivo*. Nickel and cobalt-dependent regulation of *rcnA* expression required the newly identified transcriptional repressor RcnR (formerly YohL). Deletion of *rcnR* results in constitutive *rcnA* expression and a corresponding decrease in NikR activity. Purified RcnR binds directly to the *rcnA* promoter DNA fragment and this interaction is inhibited by nickel and cobalt. Nickel accumulation is affected differently among deletion strains with impaired nickel homeostasis. Surprisingly, in low nickel growth conditions *rcnA* expression is required for nickel import via NikABCDE. The data support a model with two distinct pools of nickel ions in *E. coli*. NikR bridges these two pools by controlling the levels of the hydrogenase-associated pool based on the nickel levels in the second pool.

Introduction

All cells monitor their intracellular transition metal levels to allow for optimal growth¹. When the concentration of a particular metal is limiting, cells must engage systems that increase metal availability, such as the synthesis of import proteins, the release of metals from storage proteins, as well as the turnover of dispensable metalloproteins². Conversely, under conditions of metal excess, import systems must be repressed, while export and storage pathways are activated. Regulation is frequently achieved at the transcriptional level but post-transcriptional and post-translational regulation are also important, for example iron regulation via RNA stability³ and the post-translational regulation of the human copper transporter, hCTR1, by internalization and proteolysis⁴.

Previously, transition metal levels in aerobically grown *E. coli* were shown to be quite different, with Zn and Fe most abundant¹. These levels are controlled by metal-responsive machinery and the complement of intracellular metalloenzymes (Chapter 1). Transcriptional regulatory responses have been studied for a variety of first-row transition metals in *E. coli* (Figure 2.1), with each regulator having to discriminate between different metal ions despite the intrinsic metal affinities dictated by the Irving-Williams series⁵. All of these regulators belong to one of the six identified structural families of metalloregulators or the classic two-component regulator family, and have different mechanisms of linkage between metal and DNA binding. These examples illustrate the diversity of solutions to solve the problem of how to tightly regulate

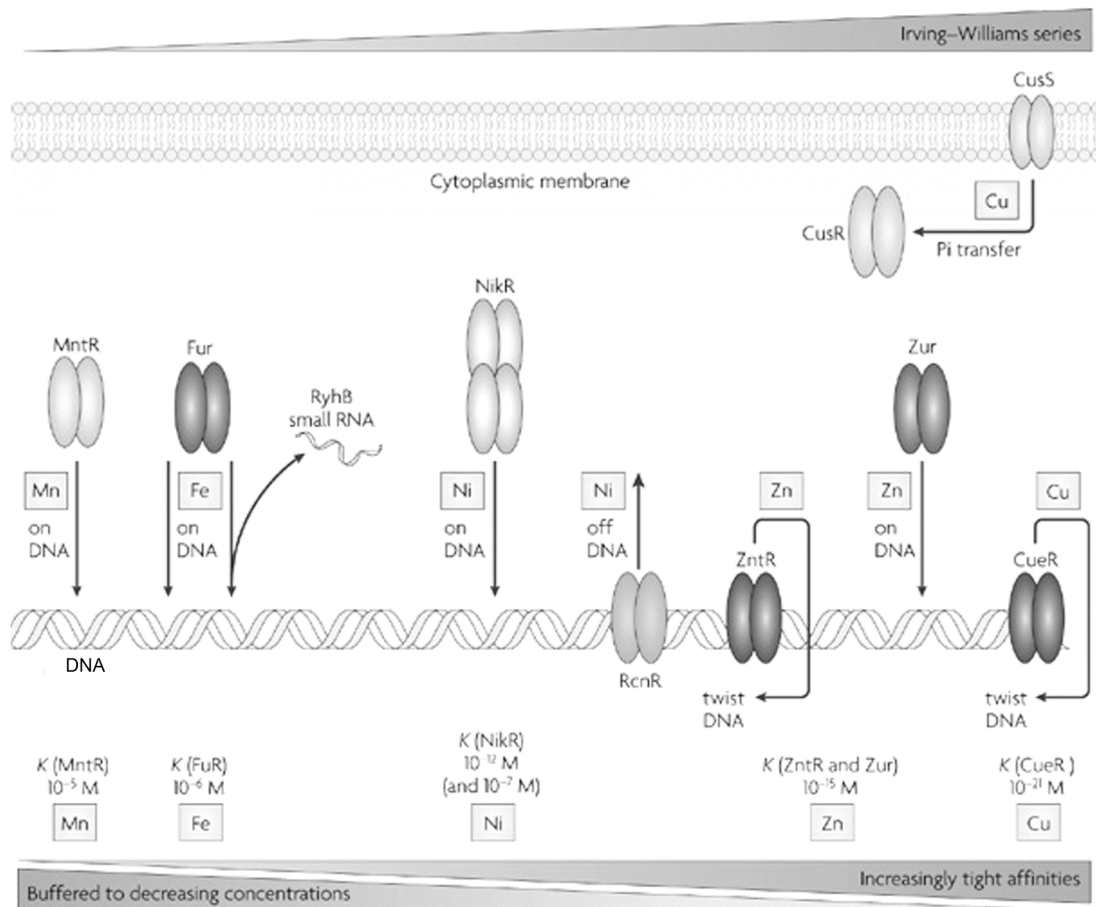


Figure 2.1. Metalloregulators of *E. coli* (modified from ⁵). The known first row transition metal regulators with approximate metal binding affinities are shown. The relative affinities for different metals correlate with the Irving-Williams series. Metal ions can either increase or decrease DNA binding affinity of the transcription factors. RcnR is identified in this chapter as a Ni- and Co-responsive protein.

intracellular metal levels in *E. coli*. In some cases a single global regulator is used (Fur), and in others complementary regulators are employed (ZntR and Zur). It is also clear that some regulatory components remain to be identified.

The recent discovery of the nickel- and cobalt-regulated RcnA efflux protein in *E. coli*⁶ suggests that nickel homeostasis also requires control of import and export pathways, and an unidentified regulator could be present. The bulk of nickel utilization in *E. coli* occurs during specific anaerobic growth conditions that require Ni-Fe hydrogenase isozyme expression^{7,8}. A hydrogenase-specific nickel transporter, NikABCDE, is synthesized under anaerobic growth conditions in accordance with the nickel-requirement, and is regulated by NikR⁹⁻¹¹. Assembly of the hydrogenase active site nickel-iron (Ni-Fe) metallocluster¹² must compete with NikR to ensure optimal *nikABCDE* expression for the prevailing growth condition, but it is not known how this occurs. Further details on Ni homeostasis are found in Chapter 1. It is not known whether NikR acts as a single global regulator of nickel ion levels in *E. coli*, like Fe-responsive Fur, or if multiple proteins are involved, as with zinc-homeostasis.

This chapter examines the coordination of RcnA expression and activity with NikR-dependent regulation of the NikABCDE transporter. *rcnA* expression is controlled by a newly identified nickel-dependent regulator RcnR (formerly YohL) in *E. coli*. RcnR is a member of a widespread, uncharacterized family of regulatory proteins found throughout the eubacterial kingdom. Interestingly,

RcnA affects NikR function and is necessary for hydrogenase activity in the presence of NikR. Nickel accumulation studies in different genetic backgrounds demonstrate that deletion of *rcnA* results in reduced nickel accumulation under certain conditions.

Results

Nickel-dependent regulation of P_{rcnA} and P_{nik} are sequential.

The nickel-dependent induction of *rcnA* expression⁶ raised the question of whether this regulation is linked to that of *nikABCDE* expression. Nickel-dependent transcriptional regulation of *rcnA* expression was examined using a *lacZ* reporter. A P_{rcnA} -*lacZ* fusion (pJl114) containing the 408-bp upstream of the *rcnA* start codon fused to *lacZ* was constructed and assayed under anaerobic growth conditions in M63 minimal media. Expression of P_{rcnA} was not induced until 0.5 μ M added nickel chloride. At 3 μ M added NiCl_2 , P_{rcnA} -*lacZ* expression was ~30-fold higher than uninduced levels (< 10 nM added NiCl_2). P_{rcnA} expression did not occur until P_{nik} expression was maximally repressed under identical growth conditions (Figure 2.2). Nickel concentrations greater than 4 μ M resulted in inhibition of cell growth in the M63 growth medium. The same patterns of *nik* repression and *rcnA* activation were also observed in rich growth medium (LB) except that the level of P_{nikA} -*lacZ* expression was ~2-fold less at low nickel concentrations (Figure 2.3) and the nickel-dependent transitions for both promoter-*lacZ* constructs were shifted to higher nickel concentrations (1/2 max LacZ for P_{nikA} occurred at ~ 10 μ M added NiCl_2). These shifts are due to changes in the nickel availability in the complex growth medium (JR and PTC, in preparation).

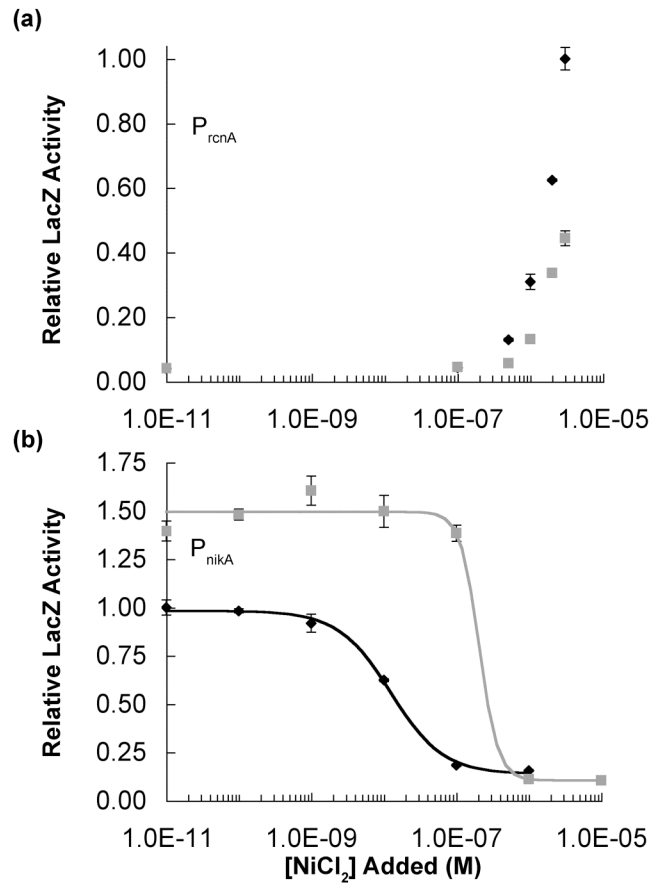


Figure 2.2. Nickel-dependent regulation of $P_{nikABCDE}$ -*lacZ* and P_{rcnA} -*lacZ* in RZ4500. (a) P_{rcnA} expression (pJI114) in M63 minimal media with glucose (black diamonds) or glucose and 10mM formate (gray squares). A relative LacZ activity of 1 corresponds to 200 Miller Units. (B) $P_{nikABCDE}$ expression (pPC181) in M63 minimal media with glucose (black diamonds) or glucose and 10mM formate (gray squares). Cells were grown anaerobically and also contained pNIK103⁹. A relative LacZ activity of 1 corresponds to ~4000 Miller Units.

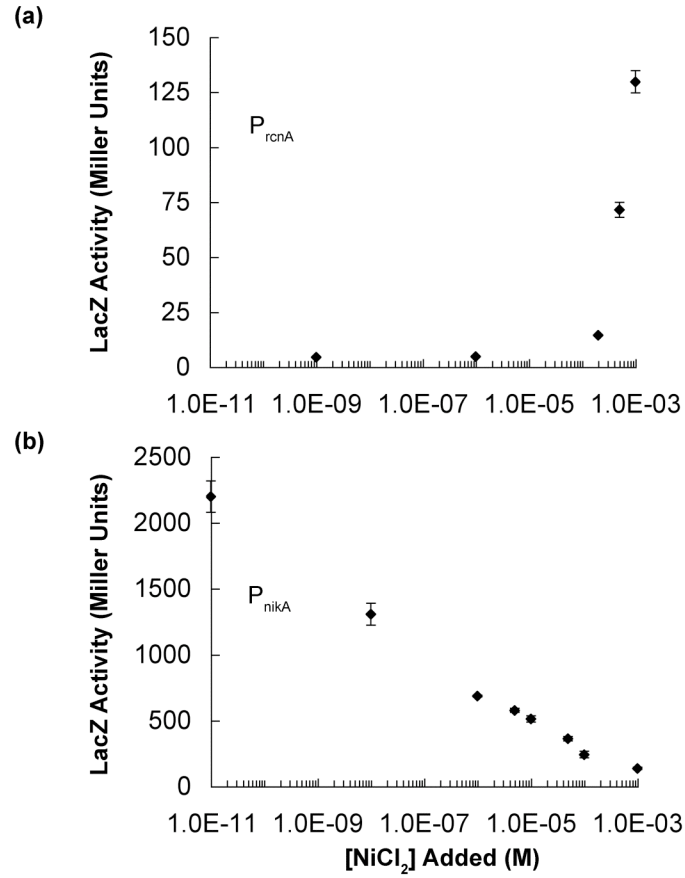


Figure 2.3. Regulation of $P_{nikABCDE-lacZ}$ and $P_{rcnA-lacZ}$ in LB media. (a) P_{rcnA} expression (pJI114) and (b) $P_{nikABCDE}$ expression (pPC181) of RZ4500 cells in LB media are shown in black diamonds. Cells were treated as in Figure 2.2.

The presence of formate, which activates the transcription factor FhlA, thereby increasing hydrogenase-3 expression^{8,13}, caused a shift in P_{rcnA} -*lacZ* expression to higher extracellular nickel concentrations, similar to that seen for the P_{nik} repression curve (Figure 2.2). Formate likely does not act directly to affect RcnR and NikR function. Their activity is instead attenuated due to the increasing nickel requirement for hydrogenase-3 activity. Growth in nitrate, which represses hydrogenase isozyme expression^{14,15}, did not affect the relative position of the repression/induction curves (Figure 2.4a,⁹). The nickel dependence of P_{rcnA} expression suggested that NikR does not regulate both P_{nik} and P_{rcnA} . Indeed, P_{rcnA} expression was unaltered in a $\Delta nikR$ strain (Figure 2.4b), strongly indicating that a second nickel-dependent regulator exists in *E. coli*.

Using bioinformatics to identify the regulator of *rcnA* expression

Proximal to *rcnA* (Figure 2.5a) is a small, divergently transcribed gene designated *yohL* (273 bp). Based on the results described below, we have renamed this protein RcnR, a repressor of *rcnA* expression. RcnR (Pfam accession #02583) contains a charged N-terminal region (amino acids 1-34), followed by a highly conserved stretch of primarily nonpolar residues (amino acids 35-64), and a variable segment consisting of the final 36 residues (Figure 2.5b). Secondary structure algorithms predict three helices, and no beta sheet structure (Figure 2.5b).

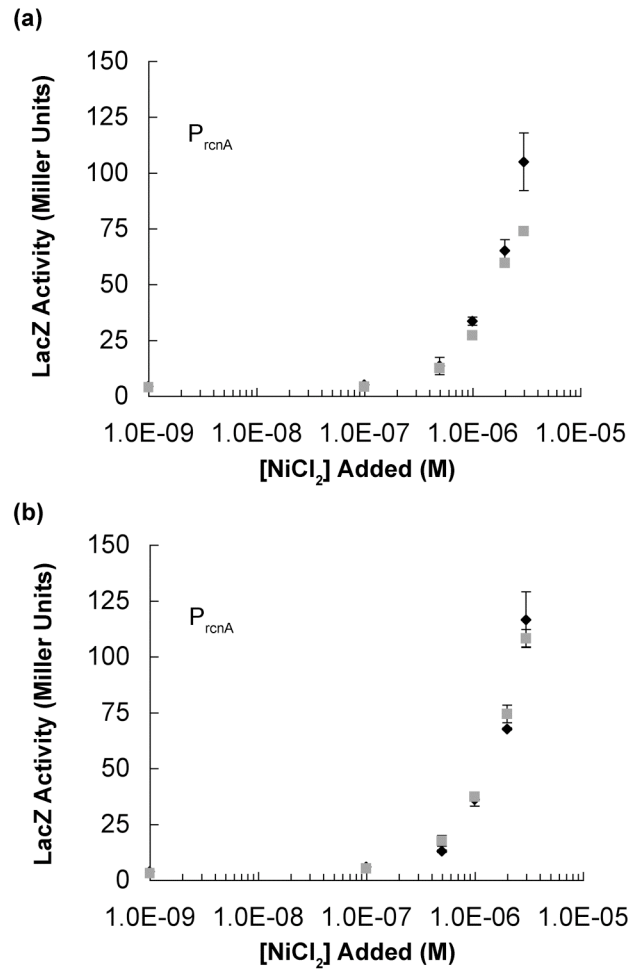


Figure 2.4. Neither nitrate nor NikR affect P_{rcnA} -*lacZ* expression. (a) RZ4500 cells with pJ114 were assayed for P_{rcnA} -*lacZ* activity in glucose alone (black diamonds) or with glucose plus 10 mM potassium nitrate (gray squares). (b) Wild type (black diamonds) or $\Delta nikR$ cells (b) were grown with pJ114 in M63 minimal media with glucose and assayed for P_{rcnA} -*lacZ* expression.

A BLAST search revealed many RcnR homologs in bacteria but no apparent sequence homology to known DNA-binding proteins. The genomic context of each of the 95 most similar RcnR-family members was analyzed to look for patterns in gene neighbors (Table 2.1). Any gene 200-bp upstream or downstream of the *rcnR*-related ORF was cataloged based on its conserved orthologous group [COG; <http://www.ncbi.nlm.nih.gov/COG>]. 51% of the *rcnR*-related ORFs are located next to a gene encoding putative metal efflux proteins. These ORFs include genes similar to *E. coli zntA*, *B. subtilis czcD* and *E. coli rcnA*. *rcnR*-related genes (23%) are also found adjacent to metalloenzymes including zinc-requiring alcohol dehydrogenases, while 13% of the genes are next to genes for copper chaperones. The remaining RcnR-family members (13%) are predominantly near genes for cardiolipin synthases or membrane proteins. Similar results were found using the STRING algorithm¹⁶ to look at gene neighborhoods. However, only co-transcribed genes are analyzed by this method. Because RcnR homologs are preferentially located near genes encoding metal requiring enzymes, or genes important in metal trafficking, they likely play a role in metal homeostasis. In an analysis of divergently transcribed gene pairs, Korbelt et al. predicted the product of the *rcnR* gene to be important in *rcnA* regulation based on the pattern of conservation of this gene pair¹⁷. Together, these data implicate RcnR as the nickel-dependent regulator of RcnA. RcnR homologs are likely to also act as

Table 2.1. *rcnR* gene neighbors.

COG	Protein	Number
262	Dihydrofolate Reductase ^a	1
265	Trypsin-like Serine Protease	1
491	Zn-dependent hydrolases	3
534	Na-driven Multi-drug efflux pump	1
542	ATPase with chaperone activity	4
762	Predicted integral membrane protein	1
793	Periplasmic protease	3
814	Amino acid permease	4
845	Membrane fusion protein	1
1062	ADH III	15
1179	molybdopterin & thiamine biosyn	1
1230	Co/Zn/Cd Efflux protein	27
1280	Threonine efflux protein	1
1502	Phosphatidylserine/Cardiolipin syn	9
1961	Site-specific Recombinase	1
1971	Membrane protein	1
2215	ABC-type permease component	9
2217	Cation transport ATPase	12
2226	Methylase	2
2236	Phosphoribosyltransferase	1
2608	Copper Chaperone	10
3118	Thioredoxin	1
3642	Mn-dependent kinase	1
3664	Beta-xylosidase	2
4232	Thiol-Disulfide interchange protein	1
4275	Uncharacterized	2
4347	Membrane protein	1
5455	Membrane protein	1
5584	Small secreted protein	2
5561	Metal-binding protein	1
	Metal Transport	48
	Metalloproteins	22
	Metallochaperones/binding proteins	12
	Other Membrane proteins	3
	Other/None	10
	Total	95

^aThe 95 proteins most similar to *E. coli* RcnR were identified with BLAST. Each gene was then catalogued based on the conserved orthologous group (COG) of gene neighbors within 200 bp. Results are summarized at the bottom of the table.

transcriptional regulators, although in some instances they will likely exhibit different signal specificities.

RcnR is a cobalt- and nickel-dependent regulator of RcnA expression

The original P_{rcnA} -*lacZ* construct described earlier contained the entire *rcnR* gene in the 403-bp P_{rcnA} fragment. To test whether *rcnR* has any effect on *rcnA* expression, the sixth codon of *rcnR* was mutated to a stop codon to create $rcnR_{stop}$ - P_{rcnA} -*lacZ* (pJI115). The activities of the wild-type plasmid and the *rcnR* stop mutant were compared (Figure 2.6a) in an *E. coli* strain lacking the chromosomal *rcnR* gene (PC563). The stop codon mutation resulted in high, constitutive LacZ activity under all nickel concentrations tested. The constitutive LacZ activity is similar to that seen with the wild-type *lacZ* construct at 3 μ M added NiCl₂ suggesting that *rcnA* is maximally induced at this nickel concentration and would be the same if it could be tested at higher nickel concentrations. Low-level constitutive expression of the *rcnR* gene from pRcnR restores Ni²⁺-dependent *rcnA* activation, albeit with slightly lower induction (4.4-fold at 3 μ M NiCl₂; 121.3 \pm 11.5 MU at 3 μ M Ni, 27.6 \pm 7.1 MU at 1x10⁻⁹ Ni). The incomplete complementation is likely due to low *rcnR* copy number effects, as the baseline expression at low nickel is ~6-fold higher than that of the wild-type construct. Similar results were obtained upon titration with CoCl₂ (Figure 2.6b), except that Co²⁺ did not induce *rcnA* expression as strongly as Ni²⁺, even at 5 μ M CoCl₂. Higher concentrations of CoCl₂ (\geq 10 μ M) resulted in a significant decrease in cell growth. The activity of $rcnR_{stop}$ - P_{rcnA} -*lacZ* in the presence of

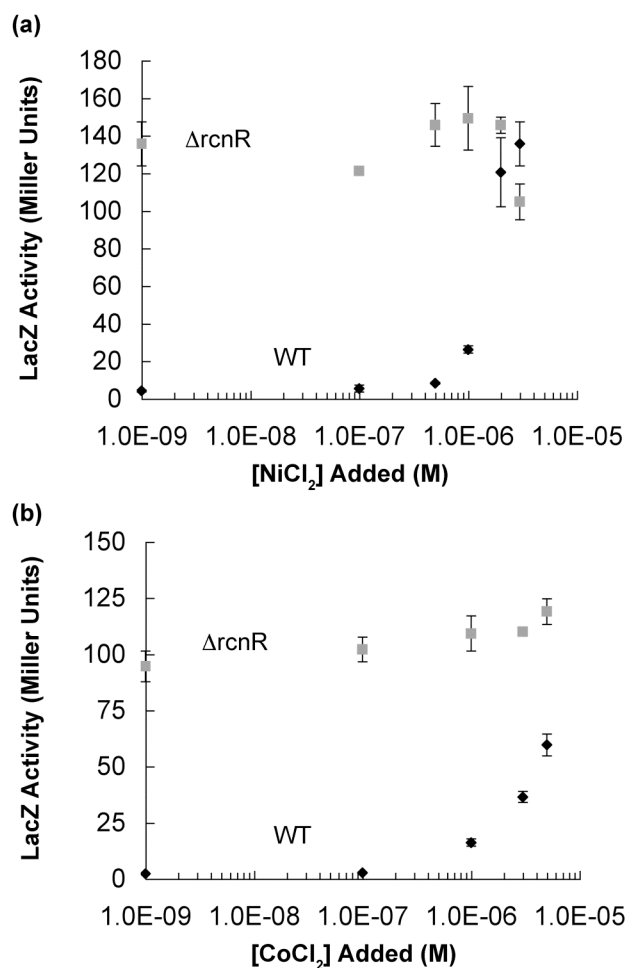


Figure 2.6. RcnR is necessary for Ni- and Co-dependent P_{rcnA} activation.

PC563 (RZ4500 $\Delta rcnR$) harboring either pJI114 or pJI115 grown anaerobically in M63 minimal medium with varying Ni(II) (a) or Co(II) (b) concentrations. Black diamonds indicate activity of pJI114, gray squares represent activity of pJI115 which contains a stop codon in $rcnR$, preventing translation of the full length protein.

Co²⁺ was high and independent of metal concentration, as observed for Ni²⁺. The experimental data support the bioinformatic identification of RcnR as a repressor of *rcnA* expression.

The nickel-dependence of P_{rcnA} expression is identical under anaerobic and aerobic conditions. Because O₂ inactivates FNR and prevents P_{nik} expression, the nickel ions required for RcnR inactivation must enter the cells in the absence of NikABCDE. Indeed *rcnR-P_{rcnA}-lacZ* assays in a *nikA* knockout strain had activity similar to that of the wild-type strain (Figure 2.7). Thus, the cell responds to nickel stress in the absence of a high nickel ion requirement in the bacterium, i.e., hydrogenase isozyme expression under anaerobic growth conditions.

Nickel controls the direct interaction of RcnR with P_{rcnA}

To determine if RcnR directly controls *rcnA* expression by promoter binding, electrophoretic mobility shift assays (EMSA) were performed with purified RcnR and a 160-bp DNA fragment spanning the intergenic region between *rcnA* and *rcnR*. A single shifted band was observed with increasing RcnR concentration when 1 mM EDTA was present in the binding buffer (Figure 2.8a). Half-maximal binding was observed at approximately 30 nM RcnR. This interaction was not observed when 100 μM NiCl₂ was present in the binding buffer instead of EDTA (Figure 2.8b). The same effect was observed in the presence of 100 μM CoCl₂ (Figure 2.9a). No shift was observed at the same protein concentrations with a 160 bp internal fragment of the *H. pylori* P_{nixA}

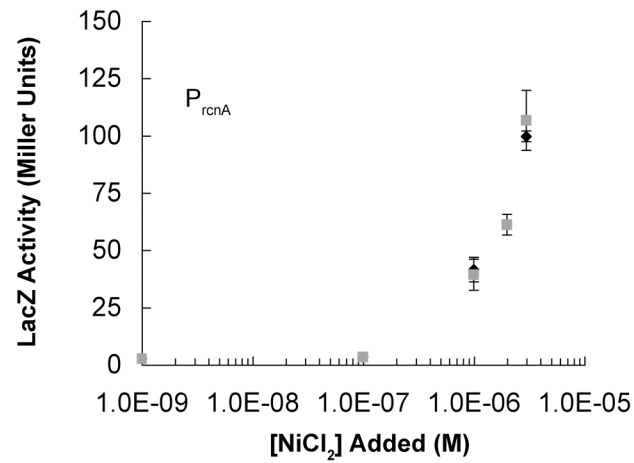


Figure 2.7. NikA does not affect P_{rcnA} -*lacZ* expression. Wild type (black diamonds) or $\Delta nikA$ cells (b) were grown with pJ1114 in M63 minimal media with glucose and assayed for P_{rcnA} -*lacZ* expression.

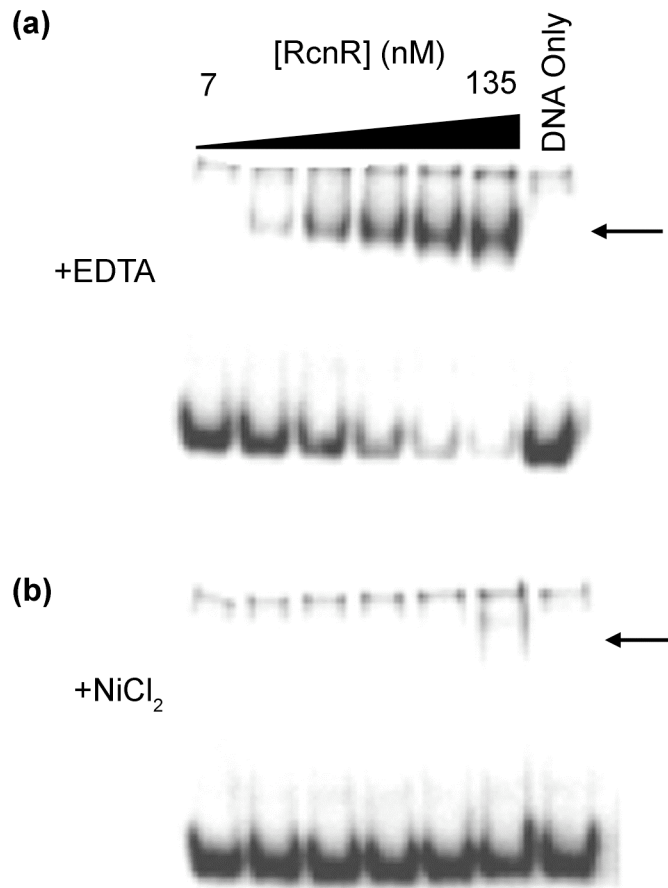


Figure 2.8. RcnR directly interacts with *rcnA* promoter region. Purified RcnR was titrated against a 160-bp ³²P-end-labelled *P_{rcnA}* fragment in binding buffer at 22 °C containing either 1 mM EDTA (A) or 100 μM NiCl₂ (B) prior to electrophoresis. The arrows indicate the migration of any shifted band. The protein concentrations were the same in each experiment. The right-hand lane was a DNA only control that shows the migration of the *P_{rcnA}* fragment in the absence of RcnR.

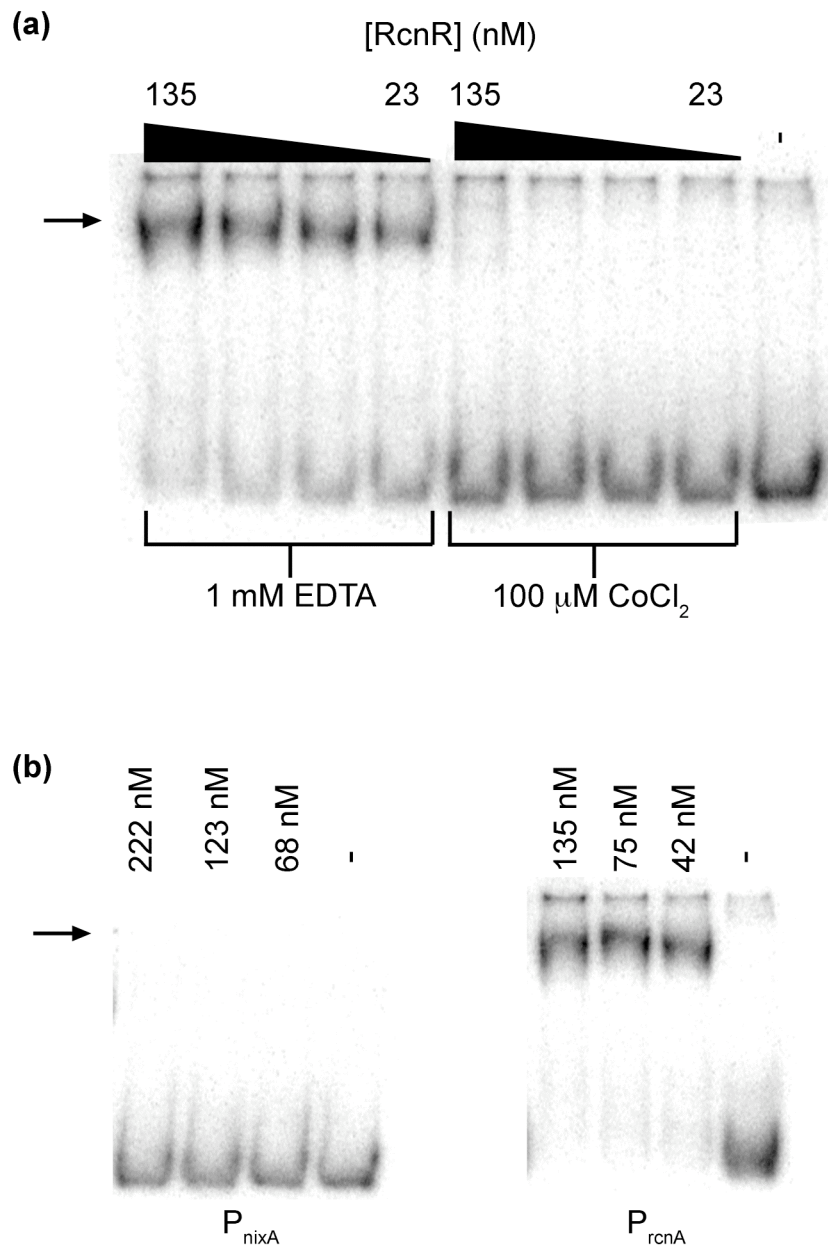


Figure 2.9. Cobalt inhibits specific RcnR DNA binding. (a) RcnR binding to a 160-bp ³²P-end-labelled P_{rcnA} fragment in binding buffer at 22 °C was assessed with 1 mM EDTA or 100 μ M CoCl₂ in the binding reaction. The arrow indicates the presence of a protein-DNA complex. The right-hand lane contains DNA alone.

(b) Various RcnR concentrations were incubated with P_{nixA} from *Helicobacter pylori* as a negative control or P_{rcnA} in parallel reactions in binding buffer with 1 mM EDTA and run on separate gels. DNA alone is present in the – lane.

promoter region (Figure 2.9b). Together, these *in vitro* experiments demonstrate that RcnR binds specifically to P_{rcnA} , and this interaction is negatively affected by the presence of nickel or cobalt ions.

Competition between nickel homeostasis pathways

The identification of the RcnA-RcnR nickel homeostasis module raised two important questions. Is RcnA activity coordinated with the NikR-regulated nickel import module? Do these modules compete for nickel ions within the cell? The effects on P_{nik} and P_{rcnA} expression of various gene deletions in these modules were determined by measuring LacZ activity. Deletion of *rcnA* had a striking effect on P_{nik} regulation (Figure 2.10), resulting in a 50% decrease in P_{nik} expression at low nickel (< 10 nM), which corresponds to an increase in NikR activity. Deletion of *rcnR* had the opposite effect (Figure 2.10), increasing P_{nik} expression by 50 % at low Ni^{2+} concentrations. Neither deletion affected the response of NikR to higher nickel concentrations. Addition of a plasmid-borne copy of *rcnA* (pRcnA) to the $\Delta rcnA$ strain restored P_{nik} activity to levels slightly higher than that seen in wild-type. Nickel-dependent repression by NikR was still observed. Complementation of the *rcnR* knockout with the low level expression from pRcnR also restored P_{nik} expression to wild-type levels.

Surprisingly, the *rcnA* deletion effect was observed at nickel concentrations that precede its induction. These results suggest that *rcnA* likely competes with NikR for nickel ions at low concentrations and *rcnA* expression

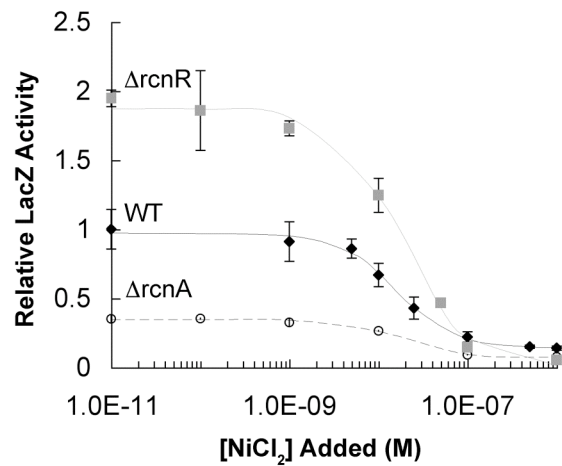


Figure 2.10. The effect of RcnA on P_{nik} expression. P_{nik} -*lacZ* reporter activity was measured in different *E. coli* deletion strains containing pPC181 and pNIK104 grown anaerobically in M63 minimal medium. Black diamonds, WT (RZ4500); circles, $\Delta rcnA$; gray squares, $\Delta rcnR$.

must occur at some basal level prior to the induction observed at higher nickel concentrations. In particular, the results suggest that *rcnA* deletion would have an impact on NikABCDE-dependent nickel uptake and hydrogenase activity due to an increase in NikR activity.

RcnA inactivation affects hydrogenase-3 function via an increase in NikR activity.

Hydrogenase-3 activity can be assayed qualitatively on solid media under anaerobic conditions using MacConkey agar base supplemented with glucose and formate¹⁸. Wild-type cells are white, while hydrogenase-3-deficient cells are small and red due the accumulation of protons from incomplete formate oxidation by the hydrogenase-3 linked formate dehydrogenase¹⁹. *rcnA*-deleted cells were red, indicating reduced hydrogenase-3 activity (Table 2.2). A similar phenotype was observed for deletion of *nikA*¹⁸ (Table 2.2). The data in Figure 2.10 suggest that hydrogenase-3 activity may be reduced because of impaired NikABCDE-dependent nickel uptake. Deletion of *nikR* in the *rcnA* deletion strain restored hydrogenase-3 activity (white colonies). The *nikR* deletion did not restore activity in the *nikA* deletion strain. The single and double mutant phenotypes suggest that *rcnA* is not directly involved in hydrogenase-3 assembly but is important for modulating NikR function. Deletion of *rcnR* did not affect hydrogenase-3 activity.

Table 2.2. Ni accumulation and hydrogenase activity in *E. coli* strains.

Strain	⁶³Ni content (atoms/cfu)	Hydrogenase-3 activity
Wild type	3067 ± 24	White
<i>ΔnikR</i>	5176 ± 58	White
<i>ΔrcnA</i>	1159 ± 79	Red
<i>ΔnikA</i>	68 ± 6	Red
<i>ΔnikR rcnA</i>	7923 ± 154	White
<i>ΔnikR nika</i>	80 ± 3	Red
<i>ΔrcnR</i>	2338 ± 198	White

These results indicate that without RcnA to export excess nickel ions, NikR limits nickel delivery to the hydrogenase assembly pathway by repressing *nikABCDE* expression.

Nickel homeostasis in strains defective for nickel regulation.

The interrelationship of RcnA and NikR was further tested by measuring nickel accumulation in anaerobically grown *E. coli* strains deleted for genes with different roles in nickel homeostasis (Table 2.2). ⁶³Ni-accumulation was tested at a low nickel concentration (10 nM in LB growth medium) in the plateau region of *nikABCDE* expression where the *rcnR* and *rcnA* deletion affected NikR function. Deletion of *nikA* blocked virtually all nickel accumulation within the cell, while deletion of *nikR* resulted in a 70% increase in nickel accumulation above wild-type levels. As predicted, deletion of *rcnA* resulted in a 60% decrease in intracellular nickel accumulation relative to the wild-type strain. Deletion of both *rcnA* and *nikR* resulted in the highest nickel accumulation of any of the strains tested and indicated that *nikR* is responsible for the decreased nickel accumulation in the *rcnA* deletion strain. Less nickel accumulated in the *rcnR*-deletion strain, which was surprising because of the higher $P_{\text{nik}}\text{-lacZ}$ expression level in this strain. This result suggests that high RcnA expression at pre-stress nickel concentrations can compete with hydrogenase assembly, although not at a level sufficient to affect the plate assay.

Discussion

The precise control of intracellular transition metal ion levels is critical for all cells. The coordinated regulation of metal import and export is a key component of this process that has been studied in limited detail in bacteria²⁰. Here, we identify RcnR as a repressor of *rcnA* expression, and thereby a critical component of the nickel regulatory network in *E. coli*. LacZ expression data and *in vitro* DNA binding experiments reveal that RcnR acts by binding to P_{rcnA} in the absence of Ni²⁺ (or Co²⁺) and repressing transcription. In the presence of one of these metals, RcnR has a weaker affinity for P_{rcnA} , and transcription can occur. This mode of regulation is also seen for the ArsR/SmtB family of metalloregulators²¹; however, RcnR shows little sequence or structural similarity with this family. SmtB proteins possess a mixed alpha-beta fold, while RcnR is predicted to be only α -helical. Additionally, the RcnR protein is about 30-residues shorter than most ArsR/SmtB family members. To our knowledge, these results provide the first evidence of DNA-binding by a member of the RcnR protein family. Because of the prevalence of similar proteins in dozens of microbial genomes, this discovery should aid in the understanding of metal homeostasis in other bacteria.

Further experiments will be required to identify an RcnR-binding site within the *rcnA* promoter as well as determining the stoichiometry of the RcnR-DNA complex. While a single shifted band is observed under our binding conditions, the stoichiometry of this complex is unresolved. There are two sets of perfect

inverted repeats and one with a single mismatch in the promoter fragment used in the EMSA experiments. Binding to more than one site could result in autoregulation of *rcnR* expression, which would provide a feedback loop to allow the cell to rapidly adapt to changing nickel ion concentrations in the environment. Aspects of DNA binding will be addressed in Chapter 4.

RcnR residues Cys35 and His60 are highly conserved among homologous proteins (Figure 2.5). Due to the location of RcnR-like proteins with other ORFs implicated in metal homeostasis, these residues are candidates for metal-binding ligands. The sequence variability of the RcnR-family at positions such as residue 64 may indicate altered metal specificity in the functions of these different homologs. Different positions could play roles in metal-specificity in RcnR homologs, as has been observed for the structurally related regulators MntR and DtxR²²⁻²⁵.

RcnR has 40% sequence identity with *E. coli* FrmR. Overexpression of FrmR results in a decrease in both *frmA*, encoding a zinc-requiring formaldehyde dehydrogenase, and *frmB*, encoding a putative S-formylglutathione hydrolase, levels²⁶. Consequently, it was hypothesized that FrmR is a transcriptional regulator of *frmA* and *frmB*. Expression of the *frmAB* operon is induced by formaldehyde, but it is unclear if FrmR mediates this response, or if it is responding to another signal. Fundamental aspects of FrmR function are investigated in Appendix 2.

The importance of RcnA and RcnR in *E. coli* cobalt homeostasis is not clear. A cobalt ion transporter and related repressor have not been identified. Cobalt ions are necessary for vitamin B₁₂-dependent enzymes and *E. coli* is known to import vitamin B₁₂ directly. The effect of excess added cobalt chloride on NikR function in wild-type or mutant *E. coli* strains is not consistent with an increase in intracellular nickel under defined minimal medium growth conditions²⁷. The interaction of Co²⁺ with RcnR may reflect the preference of nickel and cobalt ions for similar coordination environments. There may have been no pressure to select for an exclusively nickel-responsive efflux system.

RcnA likely competes with NikR for nickel ions and is necessary for full hydrogenase-3 activity at low nickel ion concentrations by preventing premature repression of *nikABCDE*. RcnA could also be important in the efflux of Ni²⁺ after the turnover of hydrogenase-3 large subunit, preventing access by NikR. RcnA contains a His-rich stretch that is predicted to form a cytoplasmic loop^{28,29}. Consequently, the protein could potentially act as both a nickel chelator and a nickel exporter, but the mechanistic importance of the His-rich region in RcnA function remains to be elucidated. The His-rich region in another metal efflux protein, NreB from *Achromobacter xylosoxidans*, is not essential for function but is hypothesized to be necessary for efficient metal transport³⁰.

The presence of two Ni²⁺-responsive regulators in *E. coli* raises the question of how Ni²⁺ partitions between NikR and RcnR so that repression of *nikABCDE* occurs before activation of *rcnA*. Coordination could be achieved by

tuning the nickel-binding affinities and/or kinetics of both repressors so that NikR binds Ni^{2+} preferentially. The RcnR- and NikR-mediated transcriptional responses at added NiCl_2 concentrations > 10 nM are independent of NikABCDE. Since *nikABCDE* expression is very low in the absence of FNR and *rcnA* activation is identical under aerobic and anaerobic growth conditions, nickel ions are probably entering the cell via another transport pathway.

The offset between the transcriptional regulation of nickel import and export is reminiscent of zinc ion regulation in *E. coli*, for which there are two different transcriptional regulators²⁰. The analogy is not perfect because ZntR acts as a transcriptional activator of the zinc exporter ZntA in the presence of excess zinc, whereas *rcnA* expression is derepressed by inactivation of RcnR in the presence of excess nickel. Regardless of these differences, the outcomes of the two pathways are identical.

The analysis of changes in nickel ion accumulation and *rcnA* and *nik* promoter activity in different genetic backgrounds and growth conditions, in combination with previous data on NikR function, suggests the existence of two separate nickel ion pools in *E. coli* (Figure 2.11). One pool (Pool A) is generated by NikABCDE and is used for assembly of active Ni-Fe hydrogenase isozymes. The second pool (Pool B) modulates the DNA-binding properties of RcnR and NikR and may supply nickel for Glx I activity. Pool B is established by a yet to be identified pathway, which may not necessarily be nickel-specific. ⁶³Ni-uptake

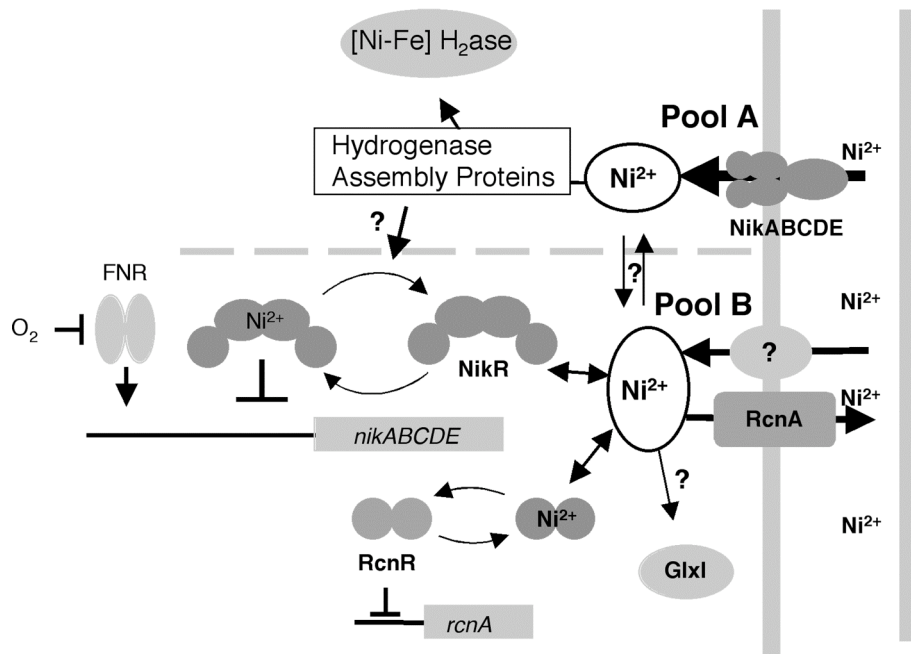


Figure 2.11. Model for nickel homeostasis in *E. coli*. Nickel ions targeted to pool A enter the cell through NikABCDE and enter the hydrogenase assembly pathway. Nickel ions in Pool B enter the cell through an unidentified pathway and are available to other Ni-requiring enzymes, such as GlxI, as well as the Ni²⁺-efflux protein RcnA, and the regulatory proteins RcnR and NikR. In the Ni²⁺-bound form, NikR represses transcription from *P_{nik}*, providing a link between the two metal pools. Question marks (?) indicate aspects of the model that are supported by published data but must still be tested.

data indicate that Pool A is substantially larger than Pool B under the growth conditions tested here. When hydrogenase activity is not required, i.e., no NikABCDE, both pools have low nickel. The nickel ions in these pools are probably mostly protein-associated (pool A, hydrogenase isozymes and assembly proteins; pool B, NikR and RcnR) with excess ions in complex with water or other small ligands.

NikR serves as a link between the two pools because its function depends at least partially on nickel in pool B, thus placing a limit on the size of pool A that is independent of NikABCDE activity. It is possible that the hydrogenase assembly pathway has access to the nickel in Pool B, however, high concentrations of extracellular nickel are required to observe any hydrogenase activity in a *nikA* deletion strain¹⁰. Nickel ions from Pool A may enter Pool B when hydrogenase activity is saturated. They may also enter when hydrogenase-3 large subunits are degraded and RcnA probably plays a role in nickel efflux after enzyme turnover. The low-level expression inferred from the effect of the *rcnA* deletion on NikR activity may be required for buffering effects of transient increases in intracellular nickel ions. The competition between RcnA and NikR suggests either a direct interaction between the two proteins or that free nickel ion levels (i.e. pool B) have a longer lifetime in an *rcnA* deletion strain, allowing NikR to bind these ions.

Studies on the kinetics of *nikABCDE* and *rcnA* transcription upon changes in Ni²⁺ levels will allow for a more detailed understanding of how nickel partitions

between these different pathways. It is unclear if other essential transition metal ions are partitioned in distinct pools that allow metal-requiring enzymes access to these cations. Prokaryotes clearly face challenges in intracellular metal partitioning and the solutions to these problems are only beginning to be understood.

Materials and Methods

Molecular biology.

Oligonucleotides, plasmids, and *E. coli* strains used in this study are listed in Table 2.3. Oligonucleotides were purchased from Operon Biotechnologies (Huntsville, AL). The sequences of all plasmids constructed were verified with primer extension sequencing by Seqwright (Houston, TX).

The *rcnR*-*P*_{*rcnA*}-*lacZ* reporter fusion was constructed from the pPC163 parent plasmid, which contains *lacZ* cloned into the *Sall* and *Bam*HI sites of pACYC184 and the *P*_{*nik*} promoter cloned into the *Eag*I and *Sall* sites³¹. The *P*_{*nik*} promoter was replaced with a 403-bp fragment upstream of the *rcnA* ATG start codon that was amplified from *E. coli* genomic DNA using primers JI132 and JI133 resulting in pJI114. The *rcnR*_{stop}-*P*_{*rcnA*}-*lacZ* reporter (pJI115), which changes codon six (CGT) of *rcnR* to a stop codon (TGA), was constructed by PCR mutagenesis of *rcnR*-*P*_{*rcnA*}-*lacZ* plasmid with primers JI145 and JI145r following the QuikChange protocol (Stratagene, La Jolla, CA).

In-frame gene deletions of *rcnA* (bases 166 to 753, primers JR118 and JR119) and *rcnR* (bases 1-324, primers JI137 and JI138) were constructed in strain PC113 (RZ4500) using the method of Datsenko and Wanner³². The numbering refers to base positions relative to the start codon of each gene.

The *rcnA* gene was PCR amplified from *E. coli* strain MG1655 genomic DNA with primers JI132 and JI133. The purified fragment was then cloned into the *Eco*RI and *Sall* sites of the pMPM-A5 vector³⁴, a generous gift from Matthias

TABLE 2.3. *E. coli* strains, plasmids and oligonucleotides.

Strain	Genotype	Parent	Reference
RZ4500	$\Delta lacZ$	MG1655	33
PC269	$\Delta nikR$	RZ4500	9
PC563	$\Delta rcnR$	RZ4500	This work
PC519	$\Delta rcnA$	RZ4500	This work
PC552	$\Delta rcnA nikR$	RZ4500	This work
PC237	$\Delta nikA$	RZ4500	9
PC379	$\Delta nikR$	RZ4500	9

Plasmid	Insert	Parent	Reference
pPC181	$P_{nik}-lacZ$	pACYC184 (cm^R)	9
pNik103	<i>nikR</i>	pET22-b (amp^R)	31
pRcnR	<i>rcnR</i>	pET22-b (amp^R)	This work
pRcnA	<i>rcnA</i>	pMPM-A5 (amp^R)	This work
pJI114	<i>rcnR</i> - P_{rcnA} - <i>lacZ</i>	pJI114 (cm^R)	This work
pJI115	<i>rcnR</i> _{stop} - P_{rcnA} - <i>lacZ</i>	pJI114 (cm^R)	This work
pJI119	P_{rcnA} (-137 to +23)	pBluescript II KS	This work

Oligo	Sequence	Purpose
JR118	ACGATGATGGCGGCGTTTATCATCGCCATCAAAGGCACCA TAAACAAGCGGTGATGCTCGTGTAGGCTGGAGCTGCTTC	<i>rcnA</i> deletion
JR119	GCCCATGAAGCCATGTACGCCATATACACACCGACTAAGCC AATCAACAGACTGGAAAAATTCCGGGGATCCGTGACC	<i>rcnA</i> deletion
JI130	CTATGTCATATGTCTCATAAATCCGTGATAAACAG	<i>rcnR</i> cloning
JI131	GTCATACTCGAGTTATTTGATATATGAATCCAGCAC	<i>rcnR</i> cloning
JI132	CTATGTGAATTCATGACCGAATTTACAACCTTTC	<i>rcnA</i> cloning
JI133	GTCATAGTCGACTTATCGCATTATGCCCATGAAG	<i>rcnA</i> cloning
JI134	CTATGTCGGCCGGATTGTATGAGACATGGCAACACC	P_{rcnA} cloning
JI135	GTCATACTCGAGAAGAGTTGTAAATTCGGTCATG	P_{rcnA} cloning
JI137	ATGTCTCATAAATCCGTGATAAACAGAACTGAAAGCGCGT	<i>rcnR</i>

	GCCAGTAAGATTCAGGGCGTGTAGGCTGGAGCTGCTTC	deletion
J1138	GCACGCGTTGTAAACGAGGAAAAACGGCAGTTTTACAATCGC	<i>rcnR</i>
	GTTAATAAATTATTTGATATTCCGGGGATCCGTCGACC	deletion
J1140	GTCATAGTCGACGATAATAATTCTTAGTATTAATTCGGC	<i>rcnR-P_{rcnA}-lacZ</i> cloning
J1145	GCCATGTCTCATACAATCTGAGATAAACAGAAACTGAAAGCG	<i>rcnR</i> codon six mutation
J1145r	CGCTTTCAGTTTCTGTTTATCTCAGATTGTATGAGACATGGC	<i>rcnR</i> codon six mutation
J1147	CTATGTCGGCCGTTAATAAATTATTTGATATATGAATCC	<i>rcnR-P_{rcnA}-lacZ</i> cloning

Mayer (University of Vienna), to create pRcnA. *rcnR* was cloned into pET-22b (Novagen, Madison, WI) at the *Nde*I and *Xho*I restriction sites using primers JI130 and JI131 to create pRcnR.

RcnR overexpression and electrophoretic mobility shift assay.

Native RcnR was purified from *E. coli* DL41 (DE3) cells containing pRcnR (*E. coli rcnR* in pET-22B) grown in LB at 37 °C. Expression was induced using 120 mg/L IPTG. Protein purification used a sequential cation-exchange (SP-Sepharose) and size exclusion chromatography (Superdex-75)) resulted in ~5 mg protein/L of growth. The complete RcnR purification procedure and protein characterization can be found in Chapter 3³⁵. Final RcnR concentrations in the binding reactions ranged from 7 nM to 135 nM.

Electrophoretic mobility shift assays were carried out in binding buffer [10 mM Hepes (pH 7.0) 150 mM NaCl, 2 mM MgCl₂ 5% glycerol, 2 µg/mL salmon sperm DNA and 10 µg/mL thioredoxin]. 1 mM EDTA or 100 µM NiCl₂ were added where indicated. The DNA fragment was generated by PCR using ³²P-end labeled oligonucleotide JI134 and oligonucleotide JI135 with pJI119 as a template (Table 2.3). The resulting 160-bp DNA fragment contained bases –137 to +23 relative to the RcnA translation start site. RcnR was incubated with DNA at 23 °C for 30 minutes before loading on 0.5X TB 6% polyacrylamide gel running at 120 V that had been pre-run for 20 minutes. The gel and running buffers were both 0.5X Tris-borate (25 mM Tris, 12.5 mM borate).

Hydrogenase-3 plate assay.

MacConkey agar base was used to prepare indicator plates for determining hydrogenase-3 activity¹⁸. The medium was supplemented with 20 mM sodium formate and 0.25% glucose was added as a carbon source. Plates were incubated overnight in a GasPak chamber³⁶ with a BBL GasPak Anaerobic System (H₂ + CO₂) envelope (Becton, Dickinson, and Company; Sparks, MD).

β-galactosidase assay.

LacZ activity assays for P_{nik} and P_{rcnA} promoter activity were performed as previously described⁹. Briefly, strains carrying pPC181 or pJI114 were first grown in minimal media under aerobic conditions then inoculated into defined minimal media to give OD₆₀₀ = 0.0001-0.0002 and grown 14 to 16 h at 37°C in capped microcentrifuge tubes. Strains also contained either pNIK103 or pRcnR, which are the *E. coli nikR* and *rcnR* genes cloned into pET22-b, respectively. These plasmids provide low-level constitutive expression of either protein even in the absence of T₇ DNA polymerase. M63 minimal medium was prepared as described⁹ with trace metals added at the following concentrations: 1 mM MgCl₂, 100 nM MnCl₂, 10 μM FeCl₂, 1 μM ZnCl₂, 100 nM (NH₄)₂MoO₄, and 100 nM NaSeO₃. Nickel was added by making serial 10-fold dilutions into minimal growth media from a 1 mM stock. Glucose (0.25%) was used as a carbon source and 10 mM potassium nitrate or sodium formate were added when indicated. Data represent measurements of three individual colonies. Similar results were obtained with triplicate measurements of the same colony.

⁶³Ni-accumulation.

E. coli strains were grown in the presence of 5 nM NiCl₂ and 5 nM ⁶³NiCl₂ (specific activity 9.87 mCi/mg; PerkinElmer, Boston, MA). Cultures were grown in capped 1.87 mL microcentrifuge tubes with LB medium supplemented with 20 mM sodium formate and sealed with parafilm. After 14 h of growth at 37 °C, a 100 μL aliquot of cells was removed to determine OD₆₀₀. The OD₆₀₀ was converted to colony forming units (cfu) based on a previously determined standard curve for the spectrophotometer and the growth medium. The remaining culture volume was spun in a microcentrifuge at 5000 rpm for 2 min at room temperature. The medium was removed and the cells washed once with 500 μL of 20 mM Tris-Cl (pH 8.0), 100 mM NaCl. The pellet was resuspended in 1 mL 10 mM acetic acid and added to 3 mL of ScintiSafe scintillation fluid. ⁶³Ni content was measured by scintillation counting on a Beckman LS7000 using a pre-programmed 10 min ¹⁴C acquisition window. The nickel counts per min were converted to atoms of nickel per cfu. Three biological replicates of each condition were measured and the standard error was reported. The ⁶³Ni atoms/cell represent a lower limit on Ni content because of trace amounts of cold nickel present in the growth medium. Under these media conditions, ICP-AES did not detect Ni atoms within the cell under aerobic conditions²⁰.

References

1. Finney, L.A. & O'Halloran, T.V. Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* **300**, 931-936(2003).
2. Nanamiya, H. et al. Zinc is a key factor in controlling alternation of two types of L31 protein in the Bacillus subtilis ribosome. *Mol. Microbiol* **52**, 273-283(2004).
3. Eisenstein, R.S. Iron regulatory proteins and the molecular control of mammalian iron metabolism. *Annu. Rev. Nutr* **20**, 627-662(2000).
4. Petris, M.J. et al. Copper-stimulated endocytosis and degradation of the human copper transporter, hCtr1. *J. Biol. Chem* **278**, 9639-9646(2003).
5. Waldron, K.J. & Robinson, N.J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol* **7**, 25-35(2009).
6. Rodrigue, A., Effantin, G. & Mandrand-Berthelot, M. Identification of rcnA (yohM), a nickel and cobalt resistance gene in Escherichia coli. *J Bacteriol* **187**, 2912-6(2005).
7. Ballantine, S.P. & Boxer, D.H. Nickel-containing hydrogenase isoenzymes from anaerobically grown Escherichia coli K-12. *J. Bacteriol* **163**, 454-459(1985).
8. Sawers, R.G., Ballantine, S.P. & Boxer, D.H. Differential expression of hydrogenase isoenzymes in Escherichia coli K-12: evidence for a third isoenzyme. *J. Bacteriol* **164**, 1324-1331(1985).

9. Rowe, J.L., Starnes, G.L. & Chivers, P.T. Complex transcriptional control links NikABCDE-dependent nickel transport with hydrogenase expression in *Escherichia coli*. *J Bacteriol* **187**, 6317-23(2005).
10. Wu, L.F. et al. Nickel deficiency gives rise to the defective hydrogenase phenotype of *hydC* and *fnr* mutants in *Escherichia coli*. *Mol. Microbiol* **3**, 1709-1718(1989).
11. Wu, L.F., Navarro, C. & Mandrand-Berthelot, M.A. The *hydC* region contains a multi-cistronic operon (*nik*) involved in nickel transport in *Escherichia coli*. *Gene* **107**, 37-42(1991).
12. Blokesch, M. et al. Metal insertion into NiFe-hydrogenases. *Biochem. Soc. Trans* **30**, 674-680(2002).
13. Rossmann, R., Sawers, G. & Böck, A. Mechanism of regulation of the formate-hydrogenlyase pathway by oxygen, nitrate, and pH: definition of the formate regulon. *Mol. Microbiol* **5**, 2807-2814(1991).
14. Richard, D.J. et al. Transcriptional regulation in response to oxygen and nitrate of the operons encoding the [NiFe] hydrogenases 1 and 2 of *Escherichia coli*. *Microbiology (Reading, Engl.)* **145 (Pt 10)**, 2903-2912(1999).
15. Stewart, V. & Berg, B.L. Influence of *nar* (nitrate reductase) genes on nitrate inhibition of formate-hydrogen lyase and fumarate reductase gene expression in *Escherichia coli* K-12. *J. Bacteriol* **170**, 4437-4444(1988).

16. von Mering, C. et al. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res* **33**, D433-437(2005).
17. Korbil, J.O. et al. Analysis of genomic context: prediction of functional associations from conserved bidirectionally transcribed gene pairs. *Nat. Biotechnol* **22**, 911-917(2004).
18. Wu, L.F. & Mandrand-Berthelot, M.A. Genetic and physiological characterization of new *Escherichia coli* mutants impaired in hydrogenase activity. *Biochimie* **68**, 167-179(1986).
19. Uden, G. & Bongaerts, J. Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. *Biochim. Biophys. Acta* **1320**, 217-234(1997).
20. Outten, C.E. & O'Halloran, T.V. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488-2492(2001).
21. Busenlehner, L.S., Pennella, M.A. & Giedroc, D.P. The SmtB/ArsR family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance. *FEMS Microbiol. Rev* **27**, 131-143(2003).
22. Ding, X. et al. Identification of the primary metal ion-activation sites of the diphtheria toxin repressor by X-ray crystallography and site-directed mutational analysis. *Nat. Struct. Biol* **3**, 382-387(1996).

23. Glasfeld, A. et al. Structure of the manganese-bound manganese transport regulator of *Bacillus subtilis*. *Nat. Struct. Biol* **10**, 652-657(2003).
24. Schiering, N. et al. Structures of the apo- and the metal ion-activated forms of the diphtheria tox repressor from *Corynebacterium diphtheriae*. *Proc. Natl. Acad. Sci. U.S.A* **92**, 9843-9850(1995).
25. White, A. et al. Structure of the metal-ion-activated diphtheria toxin repressor/tox operator complex. *Nature* **394**, 502-506(1998).
26. Herring, C.D. & Blattner, F.R. Global transcriptional effects of a suppressor tRNA and the inactivation of the regulator *frmR*. *J Bacteriol* **186**, 6714-20(2004).
27. Leitch, S. et al. Nickel-specific response in the transcriptional regulator, *Escherichia coli* NikR. *J. Am. Chem. Soc* **129**, 5085-5095(2007).
28. Tusnády, G.E. & Simon, I. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J. Mol. Biol* **283**, 489-506(1998).
29. Tusnády, G.E. & Simon, I. The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**, 849-850(2001).
30. Grass, G. et al. NreB from *Achromobacter xylosoxidans* 31A is a nickel-induced transporter conferring nickel resistance. *J. Bacteriol* **183**, 2803-2807(2001).

31. Chivers, P.T. & Sauer, R.T. Regulation of high affinity nickel uptake in bacteria. Ni²⁺-Dependent interaction of NikR with wild-type and mutant operator sites. *J. Biol. Chem* **275**, 19735-19741(2000).
32. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A* **97**, 6640-6645(2000).
33. Choe, M. & Reznikoff, W.S. Anaerobically expressed Escherichia coli genes identified by operon fusion techniques. *J. Bacteriol* **173**, 6139-6146(1991).
34. Mayer, M.P. A new set of useful cloning and expression vectors derived from pBlueScript. *Gene* **163**, 41-46(1995).
35. Iwig, J.S. et al. Ni(II) and Co(II) sensing by Escherichia coli RcnR. *J Am Chem Soc* **130**, 7592-606(2008).
36. Brewer, J.H. & Allgeier, D.L. Safe Self-contained Carbon Dioxide-Hydrogen Anaerobic System. *Appl Microbiol* **14**, 985-988(1966).
37. Cuff, J.A. et al. JPred: a consensus secondary structure prediction server. *Bioinformatics* **14**, 892-893(1998).

Chapter 3

Ni(II) and Co(II) Sensing by *E. coli* RcnR

This work is modified from Iwig, J.S., Leitch, S., Herbst, R., Maroney, M.J. and Chivers, P.T. *J Am Chem Soc.* 2008 130: 7592-7606. XAS and EPR data were obtained and analyzed in collaboration with Michael Maroney's lab by Sharon Leitch and Bob Herbst. These contributions are noted in figure legends.

Summary

E. coli RcnR and *M. tuberculosis* CsoR are the founding members of a recently identified, large family of bacterial metal-responsive DNA-binding proteins. Here, the interaction of Ni(II) and Co(II) with wild-type and mutant RcnR proteins is examined to understand how these metals function as allosteric effectors. Both metals bind to RcnR with nanomolar affinity at six-coordinate high-spin sites that contain a thiolate ligand. Experimental data support a tripartite N-terminal coordination motif (NH₂-Xaa-NH-His) that is common for both metals. However, the Ni(II)- and Co(II)-RcnR complexes are shown to differ in the remaining coordination environment. Each metal coordinates a conserved Cys ligand but with distinct M-S distances. Co(II)-thiolate coordination has not been previously observed in Ni(II)-/Co(II)-responsive metalloregulators. The ability of RcnR to recruit ligands from the N-terminal region of the protein distinguishes it from CsoR, which uses a lower coordination geometry to bind Cu(I). These studies facilitate comparisons between RcnR and NikR, the other Ni(II)-responsive transcriptional regulator in *E. coli*, to provide a better understanding of how different nickel levels are sensed in *E. coli*. The identification of the Ni(II)- and Co(II)-binding sites in RcnR, in combination with bioinformatics analysis of all RcnR/CsoR family members, identified a four amino acid fingerprint that defines ligand-binding specificity, leading to an emerging picture of the similarities and differences between different classes of RcnR/CsoR proteins.

Introduction

Metal-responsive DNA-binding proteins that control transporter and enzyme gene expression in bacteria are well-known¹. These metalloregulators have been divided into several structural families, named after their founding members: MerR, ArsR/SmtB, DtxR, Fur, NikR, CopY and the newly discovered RcnR/CsoR¹. Less well-understood are the structural bases for metal-sensing and metal-specific function by these proteins in the context of the complex intracellular milieu. Specificity is important to avoid promiscuous activation by noncognate metals. An additional structural challenge is present when two regulatory proteins that control opposing functional pathways must respond to the same metal ion, but at different concentrations, such as the zinc-responsive regulatory proteins Zur and ZntR² that control influx and efflux pathways, respectively. Structural characterization of the metal-binding sites provides insight into their thermodynamic and kinetic properties, which provide a basis for understanding their function in the cell.

The present understanding of metalloregulator structure-function relationships indicates that regulation of DNA-binding activity by the cognate metal is generally allosteric, so that the proper conformational change required for function is linked to the metal-binding specificity. The metal ion selectivity is at least partly achieved by the coordination number/geometry of the metal-protein complex³, and by ligand selection, as not all metals that bind *in vitro* elicit a biological response. These features have been demonstrated for the Ni(II)-

responsive transcriptional repressor, NikR, from *E. coli*, which has been shown to bind to a number of metals with a high affinity⁴ but responds in vivo only to the presence of Ni(II) ions⁵.

In the previous chapter I identified a nickel- and cobalt-responsive metalloregulator, *E. coli* RcnR⁶, that represses transcription of *rcnA*, which encodes a nickel and cobalt efflux protein⁷. RcnR exhibits only minimal sequence identity with the previously described families of transcriptional regulators, leading to the proposal of a new metalloregulator family⁶. RcnR homologues are found throughout the eubacterial kingdom, emphasizing their widespread role in regulating the cell physiology of microbes from diverse growth niches. The subsequent functional and structural studies of *M. tuberculosis* CsoR, a Cu(I)-responsive homologue of RcnR, confirmed the discovery of a new structural family of metalloregulators (RcnR/CsoR family) characterized by an all α -helical structure anchored by a four-helix bundle⁸. The discovery of this new family has been reviewed recently¹.

To better understand ligand recognition and allosteric regulation in this metalloregulator family, binding of the cognate metals, Ni(II) and Co(II) to RcnR was investigated. Ni(II) and Co(II) do not commonly adopt the trigonal planar geometry (two cysteines and one histidine) observed in the Cu(I)-CsoR structure⁸, which suggested that the RcnR protein must supply additional metal-binding ligands. In this chapter X-ray absorption (XAS) and UV-vis spectroscopies revealed a six-coordinate metal-binding site for Ni(II) or Co(II)

containing one S- and five N/O-donor atoms. Studies of RcnR mutants show that Ni(II) and Co(II) use overlapping but distinct binding sites to inhibit RcnR DNA binding. The sites use a similar set of ligands as Cu(I)-CsoR, but a N-terminal metal binding motif completes the coordination spheres.

Results

RcnR Metal Specificity in vivo

Previous studies in Chapter 2 demonstrated that Ni(II) and Co(II) induced transcription from a *P_{rcnA}-lacZ* fusion construct⁶. To test the effect of other divalent transition metal ions on *P_{rcnA}-lacZ* expression, cells were grown anaerobically in LB media containing the maximal metal ion concentrations that inhibited growth by <10%. As shown in Figure 3.1, Mn(II), Fe(II), Cd(II), Cu(II) and Zn(II) all had no detectable effect on *lacZ* expression. These results establish Ni(II) and Co(II) as the only two physiologically relevant metal ions for *rcnA* induction under these growth conditions and led to the study of their interactions with the RcnR protein in more detail.

RcnR Purification and Characterization

Native *E. coli* RcnR was purified to homogeneity (~5 mg/l culture) by a combination of anion exchange and gel filtration chromatographies (see Materials and Methods). The N-terminal amino acid sequence of the protein (SHTIR) revealed the expected cleavage of the N-terminal methionine,⁹ resulting in an 89-residue native protein (the numbering used here is based on the *rcnR* gene sequence; S2, H3, etc). Using the PAR assay for metal ion detection¹⁰, purified RcnR contained <0.04 metal equivalents per monomer. Thiol titration with DTNB indicated that the free sulfhydryl content of natively folded RcnR was 0.9 ± 0.1 (expected 1.0).

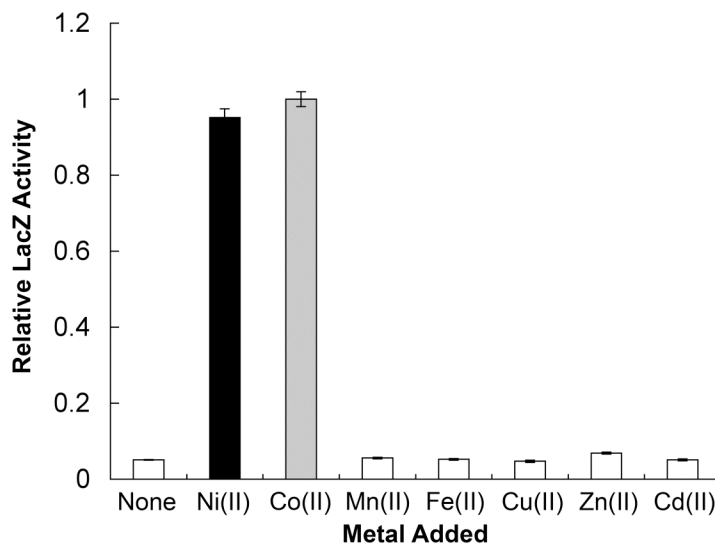


Figure 3.1. Divalent metal induction of P_{rcnA} expression. RZ4500 cells were grown with pJI114 in the presence of the indicated divalent metal ions as described in the Materials and Methods. LacZ activity was assayed as previously described⁶. Induction from P_{rcnA} only occurred in the presence of Ni or Co. Relative LacZ activity of 1 corresponds to 136 Miller Units. Error bars show the standard deviation of the LacZ activity of three separate samples.

The CD spectrum of purified RcnR (Figure 3.2a) showed minima at 222 and 208 nm consistent with a folded, α -helical protein. Conversion of the mean residue ellipticity to fraction helix gives 61% helix, in reasonable agreement with calculations of 68-78% helix from secondary structure prediction algorithms¹¹. The protein secondary structure showed little perturbation in the presence of NiCl₂ or CoCl₂ (Figure 3.2a), indicating a minimal net change in secondary structure due to metal binding.

The RcnR protein unfolded cooperatively as a function of increasing urea concentration (Figure 3.2b). RcnR showed a concentration-dependent stability, with transition midpoints of 2.7 M urea and 3.2 M urea for 2 μ M and 12 μ M protein, respectively, indicating that oligomer dissociation and RcnR unfolding are coupled (Figure 3.2c). Metal addition increased RcnR stability, shifting the midpoint of the denaturation curve from 2.7 M urea to 5.0 M urea for Ni, and 4.4 M urea for Co (Figure 3.2b).

Assembly State of RcnR

Sedimentation equilibrium experiments were performed to rigorously characterize the oligomeric state of the protein and determine whether metal binding caused any changes in RcnR assembly state. For the apo-protein, data sets were fit globally to a single species model, yielding a molecular mass of 39399 Da (Figure 3.3), consistent with an RcnR tetramer (expected mass 40010 Da). Residuals for this fit did not show systematic deviations and inclusion of a dimeric species in the model did not improve the quality of the fit because the

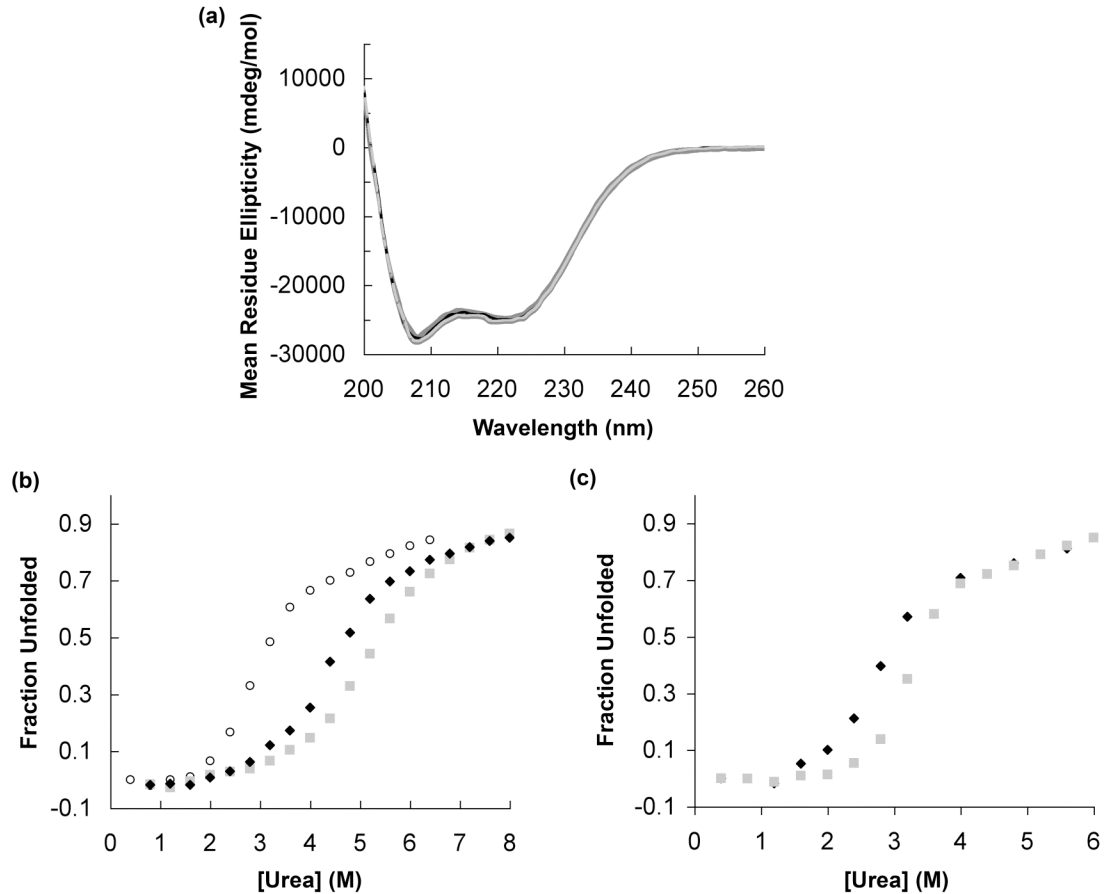


Figure 3.2. Secondary structure and stability of *E. coli* RcnR measured by CD spectroscopy. (a) Far UV CD spectra of 15 μM RcnR without metal (black line) or with 30 μM NiCl_2 (dark gray line) or CoCl_2 (gray line). Protein samples were in 10 mM Hepes (pH 7.0), 100 mM NaCl, 5% glycerol. Data were converted to mean residue ellipticity as described^{11,12}. (b) Urea denaturation (monitored at 230 nm) of apo- (white circles), Ni(II)- (black diamonds), or Co(II)-substituted (gray squares) RcnR (2.5 μM). NiCl_2 or CoCl_2 were added to 5 μM final concentration. Protein samples were in Buffer M. (c) Apo-RcnR melts were carried out at 2 μM (black diamonds) and 12 μM protein (gray squares).

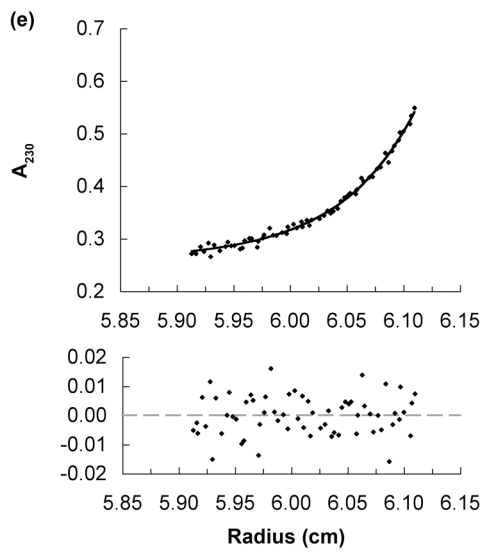
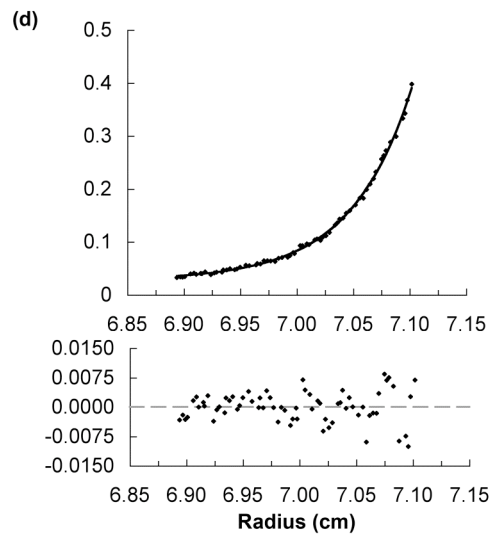
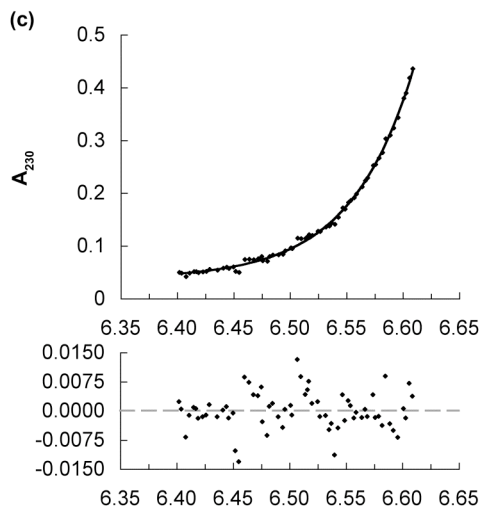
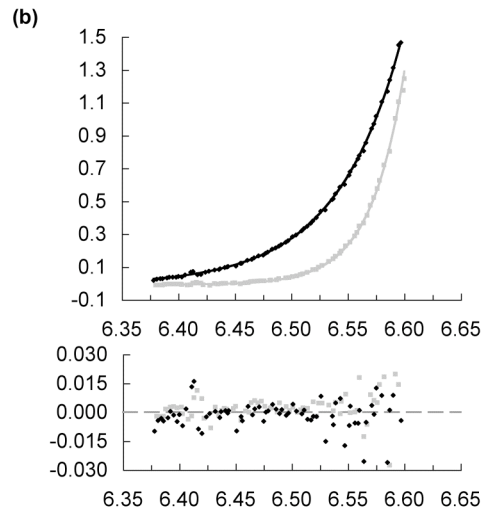
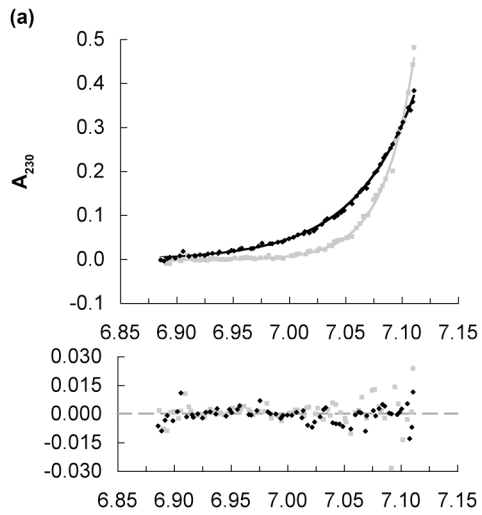


Figure 3.3. Oligomeric state of RcnR. Analytical ultracentrifugation data for 5 μM and 25 μM (a) were obtained at 20 $^{\circ}\text{C}$, at 25000 (black diamonds) and 35000 rpm (gray squares), monitoring A_{230} in 10 mM Hepes (pH 7.0), 150 mM NaCl, 200 μM EDTA, 5% glycerol. The solid lines are a global fit of a single species model to all four data sets, which yielded a molecular weight of 39399 Da (40100 Da expected for tetramer). Residuals for each fit are shown below the traces. Measurements were also obtained at 25000 rpm for 5 μM RcnR in 50 μM NiCl_2 (b), 50 μM CoCl_2 (c) (in the absence of EDTA) and 5 mM DTT (d), with the resulting individual fits giving a single species with masses of 41065 Da, 39502 Da and 36496 Da, respectively. Samples containing metal ions were found to have 0.7 SH equivalents (1.0 expected) after data collection as determined by DTNB assay.

best fit for the tetramer-dimer dissociation constant was > 2 orders of magnitude less than the protein concentrations used. Thus, negligible amounts of dimer (<10%) were present under these solution conditions. Likewise, the addition of an octameric species did not improve data fitting. Experiments were also carried out in the presence of 200 μM NiCl_2 or 200 μM CoCl_2 instead of EDTA and under these conditions, RcnR remained tetrameric (Figure 3.3b,c). Similarly, the addition of 1 mM DTT (Figure 3.3d) did not affect the species distribution.

Thus, under these experimental conditions, RcnR formed a stable tetramer and this tetrameric state was not perturbed by Ni(II) or Co(II) binding, suggesting that the tetrameric form is the active biological species. However, dimer or monomer species may occur under solution conditions different than those tested here.

Electronic Absorption Spectroscopy

The Ni(II)- and Co(II)-RcnR complexes were analyzed by UV-visible spectroscopy (Figure 3.4a) and exhibited features consistent with six-coordinate metal centers featuring thiolate ligation. The difference spectra (holo-apo) for 1:1 M(II)-RcnR (monomer equivalents) complexes exhibited intense absorptions for each metal [M = Co(II), $\epsilon_{314} = 2664 \text{ M}^{-1}\text{cm}^{-1}$ and $\epsilon_{238} = 4675 \text{ M}^{-1}\text{cm}^{-1}$; M = Ni(II), $\epsilon_{280} = 1778 \text{ M}^{-1}\text{cm}^{-1}$ and $\epsilon_{235} = 3937 \text{ M}^{-1}\text{cm}^{-1}$]. The feature at 310 nm in the Co(II)-RcnR spectrum was not present if RcnR was preincubated with the thiol modifying agent iodoacetamide (5 mM), while the peak at 235 nm was unchanged,

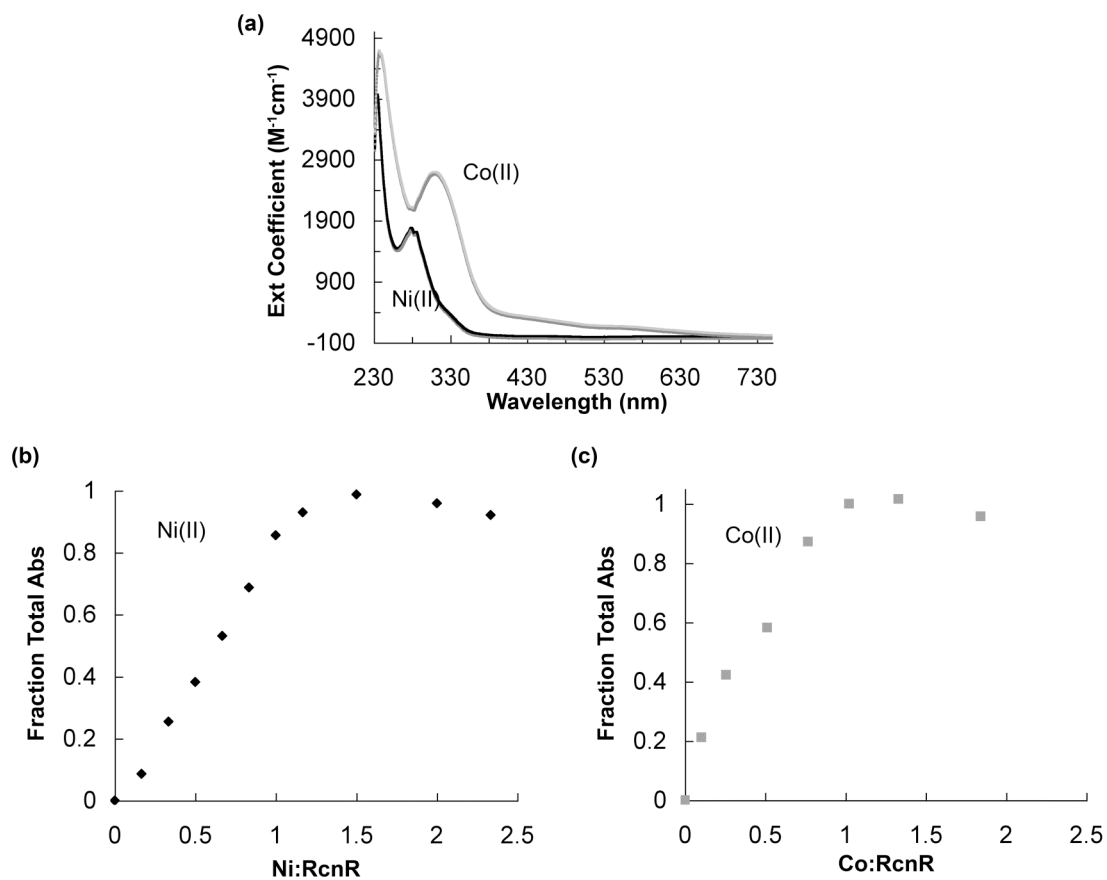


Figure 3.4. Electronic spectroscopy of Ni or Co-substituted RcnR. (a) UV-visible difference spectra of 1:1 Ni(II)-RcnR (60 μM) (black line) and 1:1 Co(II)-RcnR (85 μM) (gray line) were obtained in Buffer M. A metal-free sample was subtracted to obtain the difference spectrum. Data were collected every 0.5 nm. (b) Stoichiometric addition of NiCl₂ to RcnR (60 μM). (c) Stoichiometric addition of CoCl₂ to RcnR (88 μM). The absorbance signal saturated at 1.0 Co/monomer and 1.2 Ni/monomer under the same solution conditions as in (a).

indicating that the longer wavelength feature is due to thiolate to Co(II) ligand-to-metal charge transfer (LMCT) (Figure 3.5). Similarly, for Ni(II)-RcnR the 280 nm feature was nearly absent when RcnR is preincubated with iodoacetamide, while the 231 nm peak maintained a constant intensity, indicating that the peak at 280 nm was due to CysS \rightarrow Ni(II) LMCT (Figure 3.5). These assignments are also supported by EXAFS analyses of Co(II)-RcnR and Ni(II)-RcnR and mutagenesis data that confirm S-coordination (vide infra), as well as the high molar extinction coefficients observed for the UV-vis absorptions, and by similar absorptions in Co-substituted proteins^{13,14} and the few examples of six-coordinate Ni(II) thiolate complexes that are known^{15,16}. Thus, a preliminary interpretation of the observed LMCT spectra is that the Co(II)-RcnR complex exhibits a S \rightarrow M(II)_σ LMCT at 310 nm, while the S \rightarrow M(II)_{σ*} transition at higher energy is not observed. For Ni(II), the observed transition at 280 nm must be due to S \rightarrow M(II)_{σ*}, since no other transition is possible, and the energy relative to the Co(II) case reflects the higher energy of the transition moderated by the stabilization of the metal 3*d*-orbitals with respect to the sulfur 2*p*-orbitals that occurs with increasing effective nuclear charge on the metal.

Metal-Binding Stoichiometry of RcnR

Independent experimental methods were used to determine metal-binding stoichiometry. Titration of 60 μM RcnR with NiCl₂ and 88 μM with CoCl₂ revealed a saturation of the absorbance signals that corresponded to M(II)-RcnR complexes with stoichiometries of 1.2 Ni(II)/monomer (Figure 3.4b) and 1.0

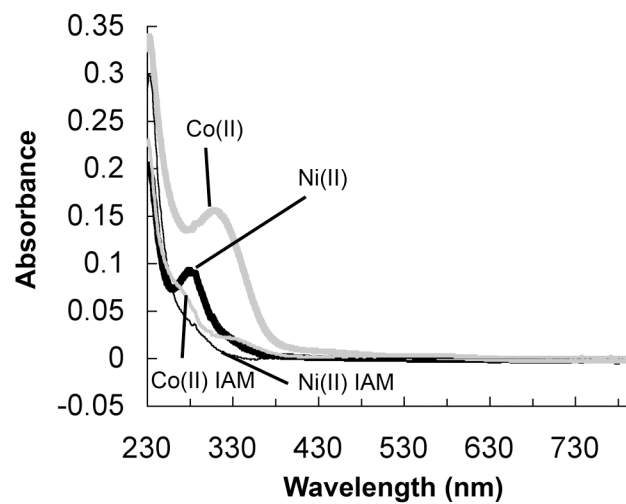


Figure 3.5. Iodoacetamide modified Ni- and Co-RcnR. 40 μ M RcnR was treated with 5 mM iodoacetamide (IAM) prior to addition of 1:1 Ni(II) (thin black) or Co(II) (thin gray) in Buffer M. Spectra of unmodified Ni(II)-RcnR (thick black) and Co(II)-RcnR (thick gray) are shown for reference.

Co/monomer (Figure 3.4c), or about four metals per tetramer. Because the absorbance change upon metal binding was linear at these concentrations, no accurate determination of binding affinities could be made. The number of metal-binding sites per RcnR monomer was also determined using the colorimetric divalent metal indicator, PAR (described in the Materials and Methods). The stoichiometries were found to be 0.7 ± 0.1 Co(II)/RcnR monomer and 1.2 ± 0.1 Ni(II)/RcnR monomer. The nickel content was also confirmed using atomic absorption spectroscopy (AAS), which yielded $1.03 \pm .01$ Ni(II)/RcnR monomer. These data are all consistent with one metal binding site per RcnR monomer, or four binding sites per tetramer.

Metal Affinities

To more accurately define metal-binding affinities, chelator competition assays were performed. It has been previously determined that mag-fura 2 fluorescence is quenched upon 1:1 complex formation with Ni(II)^{17,18}. To apply this approach to RcnR, the affinities for the dye itself were first determined by direct titration of 0.5 μ M mag-fura 2 with NiCl₂ in Buffer M. Data were fit to a single site model, giving an affinity of 330 nM for Ni(II), which is slightly higher than the published value of 130 nM, although this was carried out under different solution conditions¹⁷.

Titration curves were then performed with 0.5 μ M mag-fura 2 and 500 nM RcnR. The presence of RcnR resulted in a shift in the Ni-dependent fluorescence-quenching curve (Figure 3.6a), indicating competition for Ni-binding.

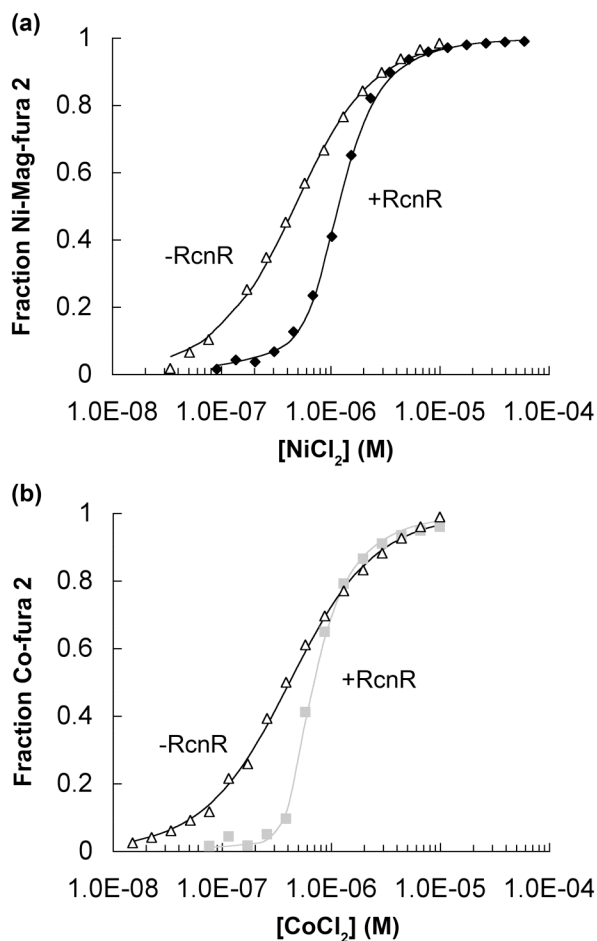


Figure 3.6. Ni(II)- and Co(II)-binding affinities determined by fluorescence competition assay. (a). Mag-fura 2 (0.5 μ M) fluorescence ($\lambda_{\text{ex}} = 380$ nm, $\lambda_{\text{em}} = 505$ nm) in the absence (white triangles) or presence of RcnR monomer (500 nM, black diamonds) and the indicated concentrations of NiCl₂. The solid line represents the fit to a model of four independent binding sites per RcnR tetramer ($K_d < 25 \pm 4$ nM). (b). Fura 2 (0.3 μ M) fluorescence ($\lambda_{\text{ex}} = 360$ nm, $\lambda_{\text{em}} = 500$ nm) in the absence (white triangles) or presence of RcnR monomer (360 nM, gray squares) and the indicated concentrations of CoCl₂. The solid gray line is a fit to a four independent binding site model ($K_d < 5 \pm 2$ nM).

However, the shape of the nickel-binding curve indicated that the protein was nearly fully saturated before quenching of mag-fura 2 began. Therefore, only a lower limit of the affinity could be determined from this experiment ($K_d < 25 \pm 4$ nM; four noninteracting sites model).

Similarly, to examine Co(II) interactions, CoCl_2 titrations were performed with 600 nM RcnR and 300 nM fura 2. Titration with fura 2 alone resulted in an affinity of 201 nM for the Co-fura 2 complex (1:1 complex expected)¹⁹. When RcnR was added, a noticeable shift in the cobalt titration curve was observed (Figure 3.6b), but again, only a lower limit of the affinity of RcnR for Co(II) (5 ± 2 nM) could be determined by this method.

Metal Site Structure

XANES

XAS spectroscopy of Ni(II)- and Co(II)-RcnR was used to obtain structural information about the metal coordination environment. Ni(II)- and Co(II)-substituted samples of a RcnRC35A were also studied to explore S-coordination at the metal site. Nickel and cobalt *K*-edge XANES analysis provides information about the coordination number and geometry at each site²⁰. In the WT Ni(II)-RcnR complex, the pre-edge XANES exhibits only one small pre-edge transition at ~8330 eV (peak area = 0.036(7) eV), associated with a $1s \rightarrow 3d$ electronic transition (Figure 3.7). In the absence of any other pre-edge features, this spectrum is diagnostic for a six-coordinate, pseudo-octahedral Ni(II) center²¹.

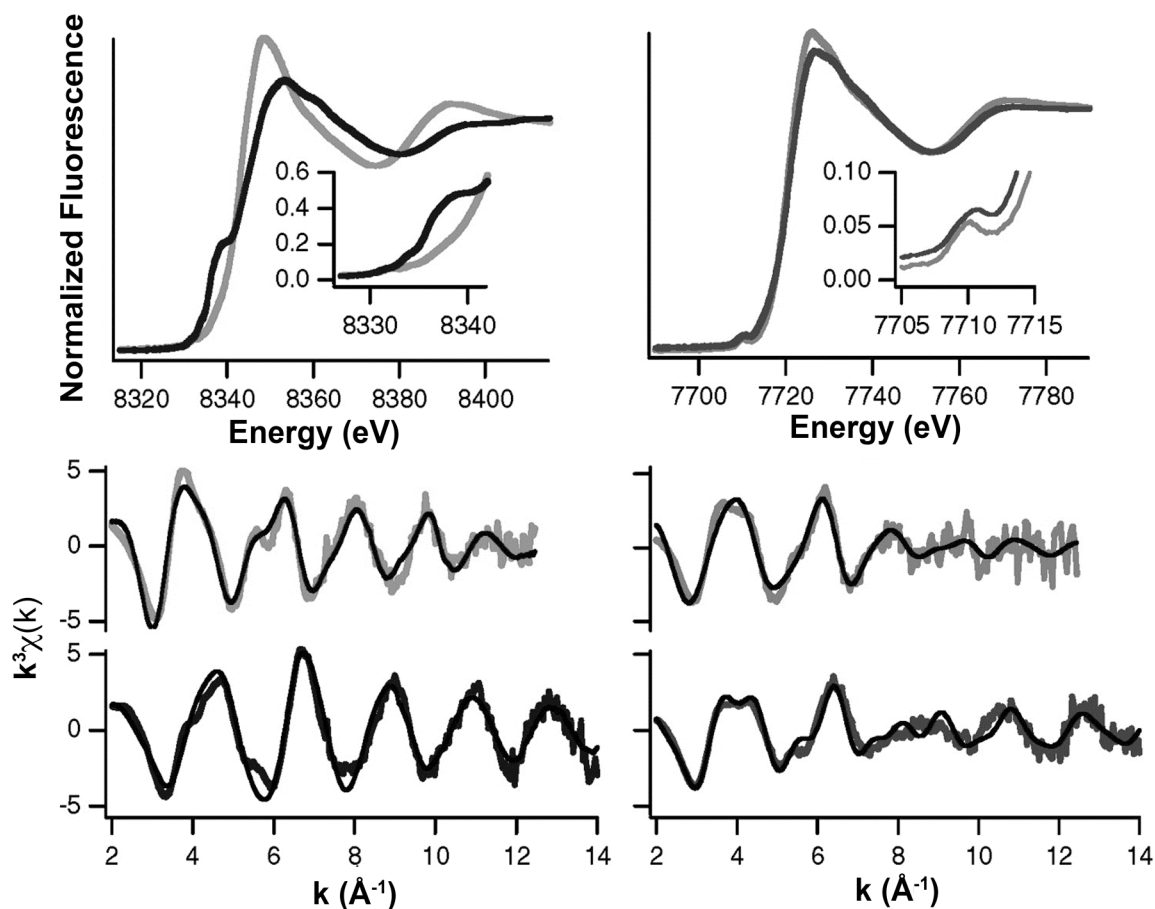


Figure 3.7. XAS of metal-substituted RcnR proteins. *K*-edge XAS spectra of Ni(II)-RcnR (left), wild type (gray) and Cys35Ala (dark gray) and Co(II)-RcnR (right), wild type (gray), and Cys35Ala (dark gray) in 300 mM NaBr buffer. Top: XANES spectra: insets are enlargements of the peaks associated with the $1s \rightarrow 3d$ and $1s \rightarrow 4p_z$ transitions. Bottom: Unfiltered EXAFS data (gray) and fits (black). The fits shown are from Table 3.1 and correspond to fits 5 and 10 for WT and Cys35Ala Ni(II)-RcnR, and fits 15 and 20 for WT and Cys35Ala Co(II)-RcnR, respectively. (Sharon Leitch)

The spectrum of Ni(II)-RcnRC35A showed a clear change in the intensity of the $1s \rightarrow 3d$ transition (peak area = 0.070(16) eV) and a new feature, a shoulder at ~ 8338 eV, associated with a $1s \rightarrow 4p_z$ transition (Figure 3.7). The increased intensity of the $1s \rightarrow 3d$ transition and the observation of a second pre-edge feature are both consistent with a five-coordinate, square pyramidal nickel center²¹.

The XAS spectrum of Ni(II)-RcnRC35A (and no other sample) exhibited a slight change with each sequential scan, indicating some exposure time-dependent radiation damage. The shift in the Ni *K*-edge energy and changes in the XANES and EXAFS spectra during scans 1-5 (used for the summed file) were too slight to interpret as a change in oxidation state or ligand environment (Figure 3.8). The peak area of the $1s \rightarrow 3d$ transition increases from 0.056(25) eV for scan 1 to 0.067(21) eV for scan 5, which are indistinguishable and lie within the range observed for 5-coordinate Ni(II)-centers²¹.

In the Co(II)-RcnR complex, the pre-edge XANES spectrum featured a single small pre-edge transition at ~ 7710 eV (peak area = 0.097(4) eV) that is associated with a $1s \rightarrow 3d$ electronic transition (Figure 3.7). The small peak area and the absence of any other pre-edge features are consistent with a six-coordinate, distorted octahedral geometry. While the peak area was slightly larger than usual for octahedral Co(II) complexes, it was significantly smaller than the peaks observed for five-coordinate or tetrahedral Co(II) complexes²². A slight increase in the $1s \rightarrow 3d$ peak area was also observed in other mixed-ligand,

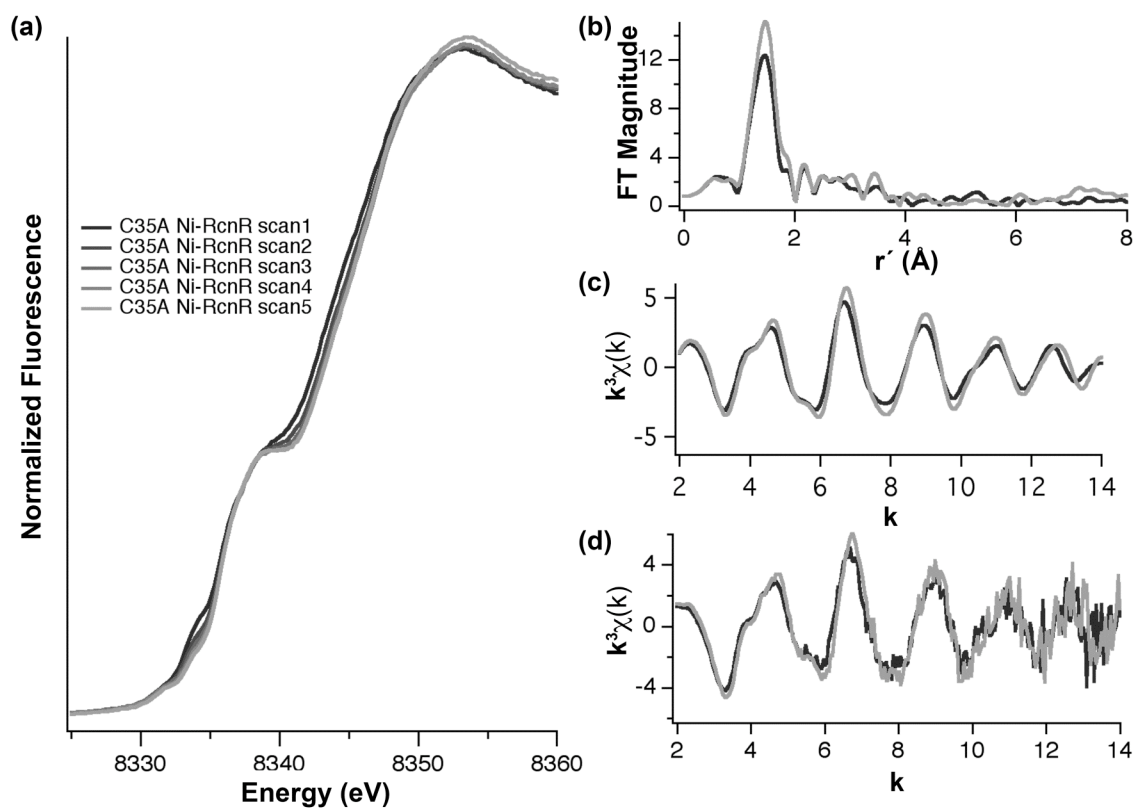


Figure 3.8. Effect of exposure time on XAS data for Ni-RcnRC35A. XANES spectra obtained from sequential scans of Ni-C35ARcnR are shown on in (a). EXAFS spectra for scan 1 (black) and scan 5 (gray) are shown in (b), with fourier-filtered data (c) and unfiltered data (d) below. (Sharon Leitch)

distorted octahedral Co(II) centers⁵. The XANES spectrum obtained for Co(II)-RcnRC35A was very similar to that obtained for the WT complex, with only one notable pre-edge feature at ~ 7710 eV (Figure 3.7). The peak area (0.075(6) eV) was slightly smaller than the corresponding feature in the spectrum of the WT cobalt center, and was also consistent with a six-coordinate site.

EXAFS

EXAFS analysis provides information about the types of scattering atoms and metric details about the complex²⁰. The results of EXAFS analyses on the RcnR samples are summarized in Figure 3.7 and Table 3.1, which contains five example fits for the EXAFS spectrum of each sample. Three of these fits represent the best fits of scattering arising from the first coordination sphere atoms (Fourier-filtered data with a back-transform window of 1.0 – 2.3 Å, not corrected for phase shifts), and include the best single shell fit and the best two shell fits with and without an S-donor ligand. In the fourth fit, data is included to allow the determination of the presence of histidine ligands and an estimate of the number of histidine imidazoles bound to be made (Fourier-filtered data with a back-transform window of 1.0 – 4.0 Å, not corrected for phase shifts). The fifth fit is a refinement against unfiltered EXAFS data. This data and the best fits to unfiltered data are shown in Figure 3.7. Additional fits of Fourier-filtered data are included in Figure 3.9 and tables in Appendix 3.

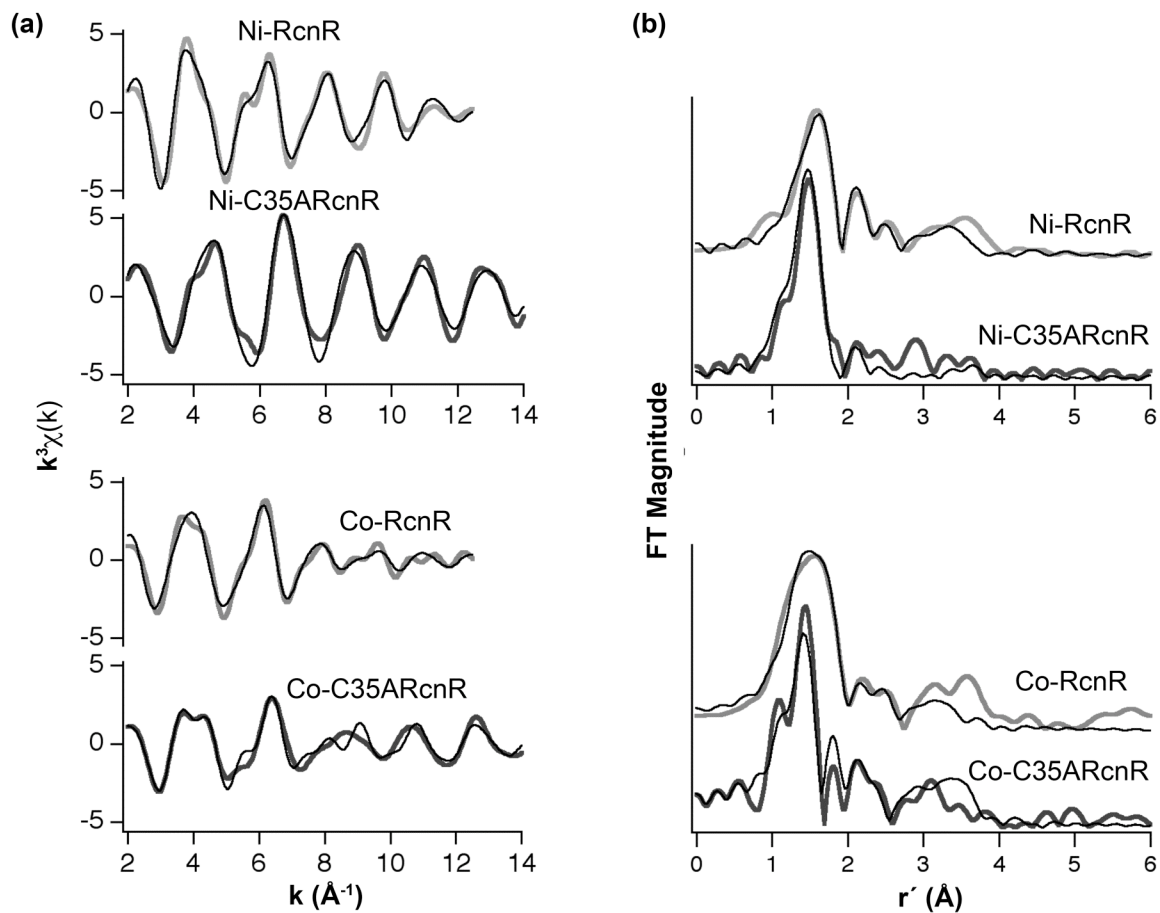


Figure 3.9. EXAFS comparison of M(II)-RcnR and M(II)-C35A RcnR. Fourier filtered (a) and Fourier-transformed (b) EXAFS data are shown for the indicated metal substituted RcnR and RcnRC35A with appropriate fits, which correspond to: Appendix 3 Table 1, fit 11; Table 2, fit 7; Table 4, fit 8; Table 5 fit 8, respectively. (Sharon Leitch)

Table 3.1. EXAFS fits for RcnR complexes.

Sample/Fit #	N ^a	r (Å) ^b	σ^2 ($\times 10^3 \text{Å}^{-2}$) ^c	ΔE_0 (eV)	GOF ^d
Ni(II)-RcnR	Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å				
1	6 Ni-N/O	2.05(2)	10.4(17)	7.6(21)	0.66
2	5 Ni-N/O	2.07(12)	7.7(12)	10.4(15)	0.48
	1 Ni-N/O	2.34(350)	[7.7] ^e	[10.4]	
3	5 Ni-N/O	2.06(1)	8.1(7)	9.5(9)	0.31
	1 Ni-S	2.57(2)	8.2(27)	[9.5]	
4	Fourier-Filtered fits with a Backtransform Window = 1-4 Å				
	2 Ni-N _{imd}	1.99(2)	3.4(19)	10.6(9)	0.55
	3 Ni-N/O	2.12(1)	[3.4]	[10.6]	
	1 Ni-S	2.55(2)	7.5(22)	[10.6]	
	Unfiltered EXAFS fit from 2-12.5 Å ⁻¹				
5	2 Ni-N_{imd}	1.986(5)	3.3(6)	9.8(3)	0.70
	3 Ni-N/O	2.112(3)	[3.3]	[9.8]	
	1 Ni-S	2.543(7)	7.4(8)	[9.8]	
C35A Ni(II)-RcnR	Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å				
6	5 Ni-N	1.87(1)	5.7(10)	2.5(27)	0.76
7	4 Ni-N/O	1.88(1)	4.1(5)	6.6(17)	0.43
	1 Ni-N/O	2.15(2)	[4.1]	[6.6]	
8	4 Ni/N/O				No fit
	1 Ni-S				
	Fourier-Filtered fits with a Backtransform Window = 1-4 Å				
9	1 Ni-N _{imd}	2.31(3)	4.1(5)	3.0(17)	0.65

	4 Ni-N/O	1.87(1)	[4.1]	[3.0]	
	Unfiltered EXAFS fit from 2-14.0 Å ⁻¹				
10	1 Ni-N_{imd}	2.316(8)	4.1(2)	4.4(4)	0.72
	4 Ni-N/O	1.876(2)	[4.1]	[4.4]	
<hr/>					
Co(II)-RcnR	Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å				
11	6 Co-N/O	2.05(3)	14.8(19)	-0.9(21)	0.44
12	5 Co-N/O	2.05(200)	14.8(32)	-0.9(27)	0.48
	1 Co-N/O	2.06(1000)	[14.8]	[-0.9]	
13	5 Co-N/O	1.97(2)	16.4(15)	-8.0(13)	0.16
	1 Co-S	2.25(1)	6.6(10)	[8.0]	
	Fourier-Filtered fits with a Backtransform Window = 1-4 Å				
14	2 Co-N _{imd}	1.94(3)	6.6(27)	-2.8(16)	0.51
	3 Co-N/O	2.11(3)	[6.6]	[-2.8]	
	1 Co-S	2.24(4)	13.1(60)	[-2.8]	
	4 Co-C _{imd}	3.01(4)	13.4(60)	[-2.8]	
	Unfiltered EXAFS fit from 2-12.5 Å ⁻¹				
15	2 Co-N_{imd}	1.945(12)	6.9(13)	-3.0(7)	0.76
	3 Co-N/O	2.113(14)	[6.9]	[-3.0]	
	1 Co-S	2.238(18)	12.6(28)	[-3.0]	
	4 Co-C_{imd}	3.009(19)	13.5(28)	[-3.0]	
<hr/>					
C35A Co(II)-RcnR	Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å				
16	6 Co-N/O	1.92(3)	15.6(29)	-5.2(48)	0.62

17	4 Co-N/O	1.92(2)	6.7(11)	-2.8(23)	0.43
	2 Co-N/O	2.12(3)	[6.7]		
18	5 Co-N/O				No fit
	1 Co-S				
Fourier-Filtered fits with a Backtransform Window = 1-4.0 Å					
19	2 Co-N _{imd}	1.87(7)	14.6(66)	-1.6(21)	0.66
	4 Co-N/O	1.97(4)	[14.6]	[-1.6]	
	4 Co-C _{imd}	2.89(5)	12.6(66)	[-1.6]	
Unfiltered EXAFS fit from 2-14.0 Å ⁻¹					
20	2 Co-N_{imd}	1.867(23)	14.3(22)	-1.6(6)	0.74
	4 Co-N/O	1.975(11)	[14.3]	[-1.6]	
	4 Co-C_{imd}	2.886(18)	13.2(24)	[-1.6]	

(Sharon Leitch)

^aN is the number of scattering atoms at a specific distance.

^br (Å) is the radial distance. ^cσ² is the root mean square disorder in the Ni-X distance, where X is the scattering atom.

^dGOF (goodness of fit) was calculated as described previously²³. The accuracy of distance determined by EXAFS for atoms in the first coordination sphere of the metal are limited to ± 0.02 Å by the theoretical phase parameters; numbers in parentheses represent standard deviations from the least-squares fits.

^eValues of parameters in [] were constrained to be equal to those of other scattering atoms in the same shell.

The EXAFS spectrum arising from the first coordination sphere of the metal in Ni(II)-RcnR was best modeled as a six-coordinate site composed of two shells of ligands, including one S-scattering atom (Figure 3.7; Table 3.1). The number of scattering atoms in the first coordination sphere was in agreement with the XANES analysis of the Ni(II) site (*vide supra*). Addition of a S-donor at a long distance (2.57(2) Å) improved the goodness of fit by 35% over the best fit composed of two shells of N/O scattering atoms. Data obtained from samples prepared in buffers containing 300 mM NaBr or 300 mM NaCl were essentially identical (Figure 3.10), confirming the assignment of a S-scattering atom and not a Cl⁻ ion from the buffer. Analysis of the EXAFS data including scattering atoms in the second and third coordination sphere confirmed the presence of histidine ligation (the best fit with histidine ligands has a GOF that is 15% better than the best fit obtained without imidazole parameters using the same back-transform limits), and an estimate of 2-3 histidine ligands. Refinement of the models against unfiltered data led to similar conclusions, with the fit obtained for two histidine ligands having the lowest GOF value (Appendix 3). Fits obtained using either Fourier-filtered or unfiltered data clearly showed a substantial improvement upon inclusion of a long Ni-S vector (Table 3.1 and Appendix 3). The EXAFS analysis is therefore consistent with a Ni(His)₂(Cys)(N/O)₃ site in WT Ni(II)-RcnR.

The EXAFS spectrum obtained for the Ni(II)-RcnRC35A further supported the inclusion of an S-donor in the coordination sphere of WT Ni(II)-RcnR, and identified the ligand involved as Cys35. The first coordination sphere EXAFS

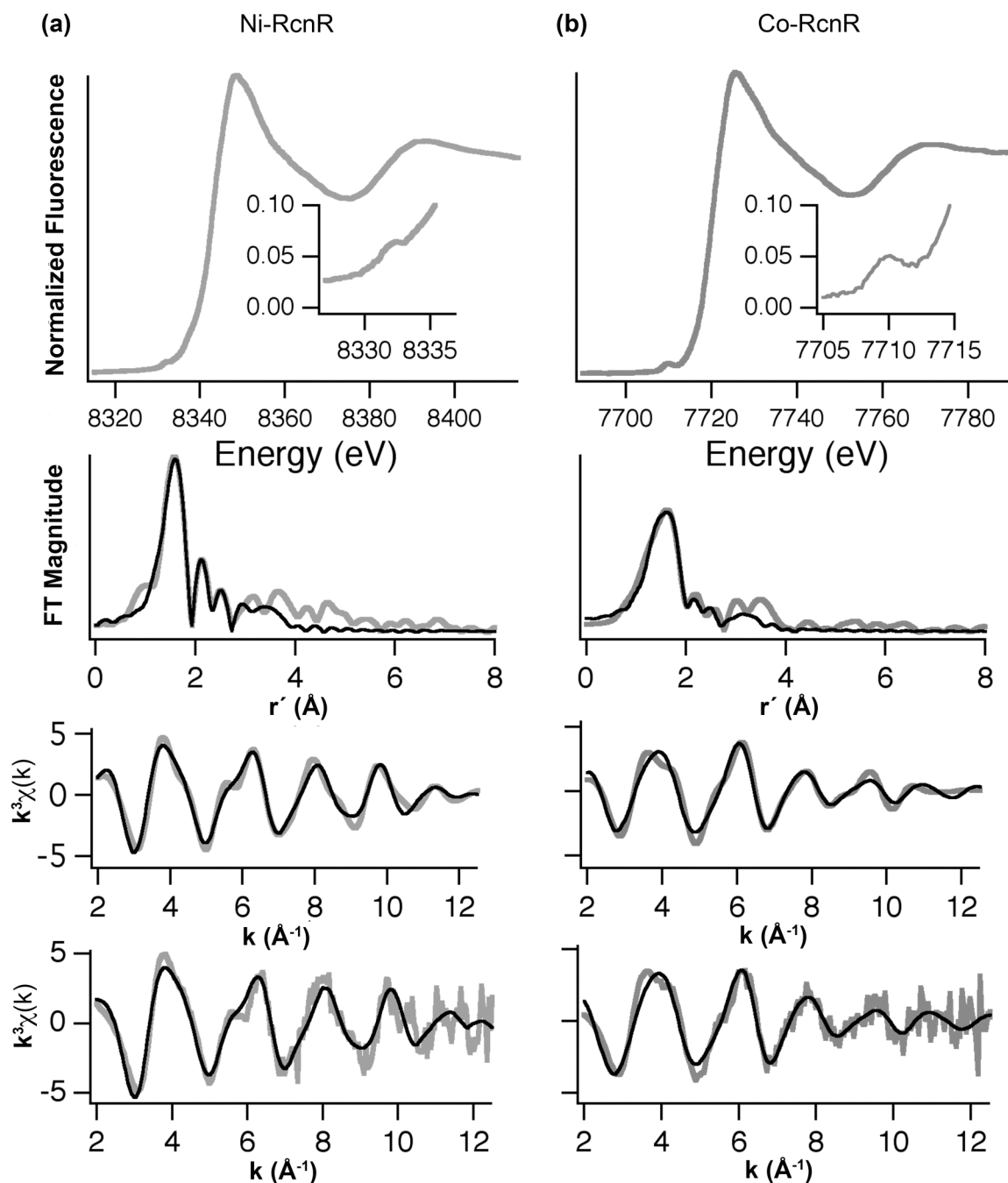


Figure 3.10. XAS of metal substituted RcnR proteins in NaCl. XAS spectra for Ni-RcnR (a) and Co-RcnR (b) with EXAFS fits in black, in 300 mM NaCl buffer. Top: XANES spectra; insets are enlargements of the peak associated with the

1s→3d transition. Second from top: Fourier-transformed EXAFS spectra ($k = 2$ -12.5 Å) and fits. Third from top: Fourier-filtered spectra (backtransform window = 1-4 Å) with fits. Bottom: unfiltered EXAFS spectra and fits. The fits shown are from: Appendix 3 Table 3 fits 11 and 20, for Ni-RcnR Fourier-filtered and unfiltered data, respectively; and Appendix 3 Table 6 fits 8 and 13, for Co-RcnR Fourier-filtered and unfiltered data, respectively. (Sharon Leitch)

data is best modeled with a single type of scatterer (N/O-donors) (Table 3.1). Splitting the coordination sphere into two shells improved the GOF by 43%. The number of N/O ligands determined for the first coordination sphere is five, which is also in agreement with the XANES analysis. Multiple-scattering analysis of data including second and third coordination sphere atoms was consistent with one histidine ligand, although the improvement over the fit obtained without imidazole parameters is modest (7%). There was no evidence in the data for a Br⁻ ion from the buffer coordinating to the Ni(II) center. The EXAFS analysis of Ni(II)-RcnRC35A was therefore consistent with a Ni(N/O)₅ site, where the number of histidine ligands is not clear. This might be due to the large distribution in Ni-NHis distances (static disorder) that occurs as a result of the loss of C35, a possibility that is consistent with the range of Ni-N/O distances observed (1.87-2.31 Å).

The analysis of first coordination sphere EXAFS data for Co(II)-RcnR also revealed the presence of a S-donor ligand. The addition of a single S-donor to a shell of five N/O-donors greatly improved the GOF compared to a fit containing six N/O donor ligands (67%, Table 3.1). Again, data obtained in 300 mM NaCl or NaBr containing buffers gave indistinguishable results, confirming the assignment of an S-donor ligand. The total number of ligands (six) agrees well with the XANES analysis (*vide supra*). Analysis of the spectrum using multiple-scattering parameters to account for histidine ligation is consistent with the presence of ~ 2 histidine ligands (15% improvement in GOF). The structure that

emerged from the EXAFS analysis for Co(II)-RcnR was $\text{Co}(\text{His})_2(\text{Cys})(\text{N/O})_3$, a result that is similar to the Ni site structure in Ni(II)-RcnR. Since there is only one Cys ligand present in *E. coli* RcnR, the two metals both bind Cys35 in six-coordinate sites that also contain at least two imidazole ligands. The largest difference apparent in the EXAFS spectra of the two metal sites was the M-S distance, which was short for Co(II) (2.24(2) Å) and long for Ni(II) (2.54(2) Å). The EXAFS from the first coordination sphere scattering atoms in Co(II)-RcnRC35A center was best modeled by six N/O-donors at two different distances (Table 3.1). Inclusion of a S-donor caused the fit to diverge. Multiple-scattering analysis of the Co(II) site was consistent with the presence of 2 ± 1 imidazole ligands. Thus, the perturbation in the metal sites for the Cys35Ala mutant was distinct. In the case of Ni(II), the site was five-coordinate, reflecting the loss of the S-donor. For Co(II), the site bound an unknown sixth N/O-donor ligand. Both Cys35Ala sites had poorly refined data for histidine ligation.

EPR

The analysis of the XAS data from both Ni(II)- and Co(II)-RcnR were consistent with six-coordinate, $\text{M}(\text{Cys})(\text{His})_2(\text{N/O})_3$ complexes, but differ strikingly in the M(II)-S distance found for the Cys ligand (2.24(2) Å for Co(II) and 2.54(2) Å for Ni(II)). Six-coordinate Ni(II) complexes are invariably high-spin, $S = 1$ complexes, and the M-ligand distances found for the Ni(II)-RcnR complex were consistent with this electronic configuration²⁴. Both high-spin ($S = 3/2$) and low-spin ($S = 1/2$) Co(II) complexes are common, but are easily distinguished by EPR

spectroscopy²⁵. The EPR spectrum of Co(II)-RcnR collected at 12 K shows features that were typical of an $S = 3/2$ Co(II) center (Figure 3.11). Integration of the EPR spectrum showed an absorption curve that lies on a flat baseline, indicating that no saturation effects were seen. The experimental spectrum could be simulated by a high-spin, $S = 3/2$ Co center with $M_s = 1/2$, $D \gg h\nu$ (50 cm^{-1} was used), $g_{x,y} = 2.210$, $g_z = 2.195$, and $E/D = 0.07$. These values are typical for high-spin 5- or 6-coordinate Co(II) in protein ligand environments²⁶.

Identification of Metal-Sensing Determinants by Amino Acid Substitutions

Site-directed mutagenesis of the *rcnR* gene (pJI114) was carried out to gain functional insight into the amino acids that form the metal-binding site, as suggested by spectroscopic measurements and sequence conservation. The regulation of *rcnA* expression by the RcnR mutants was tested in vivo using the previously described LacZ assay⁶ (Figure 3.12). A mutation that reduces metal binding should generate a constitutively active repressor and result in low LacZ activity, even in the presence of Ni(II) or Co(II). Importantly, any mutation that causes the protein to be poorly expressed, unstable, or unable to bind DNA would lead to high LacZ activity in both the presence and the absence of these metal ions. Mutated RcnR residues included all five histidine residues (three are highly conserved), a subset of conserved Glu and Asp residues, and the highly conserved Cys35. An alanine insertion mutation (A2*) resulted in an expressed protein with N-terminal sequence $\text{NH}_2\text{-Ala}_2\text{Ser}_3\text{His}_4$ after processing, instead of $\text{NH}_2\text{-Ser}_2\text{His}_3$, thereby disrupting the spacing between His3 and the $\text{NH}_2\text{-}$

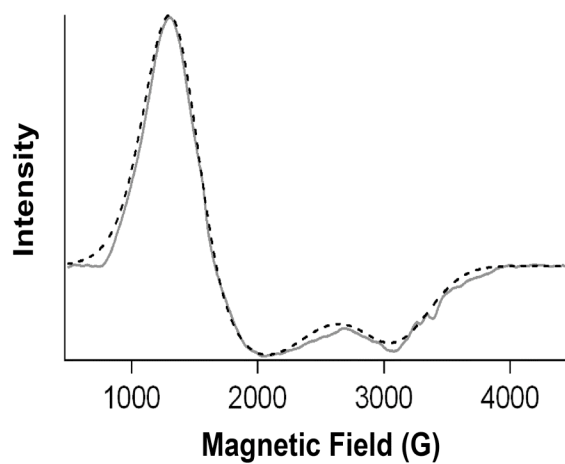


Figure 3.11. X-band EPR spectrum of Co(II)-RcnR. The spectrum (solid line) was obtained at 12 K. The parameters from the simulation (dashed line) are given in the text. (Bob Herbst)

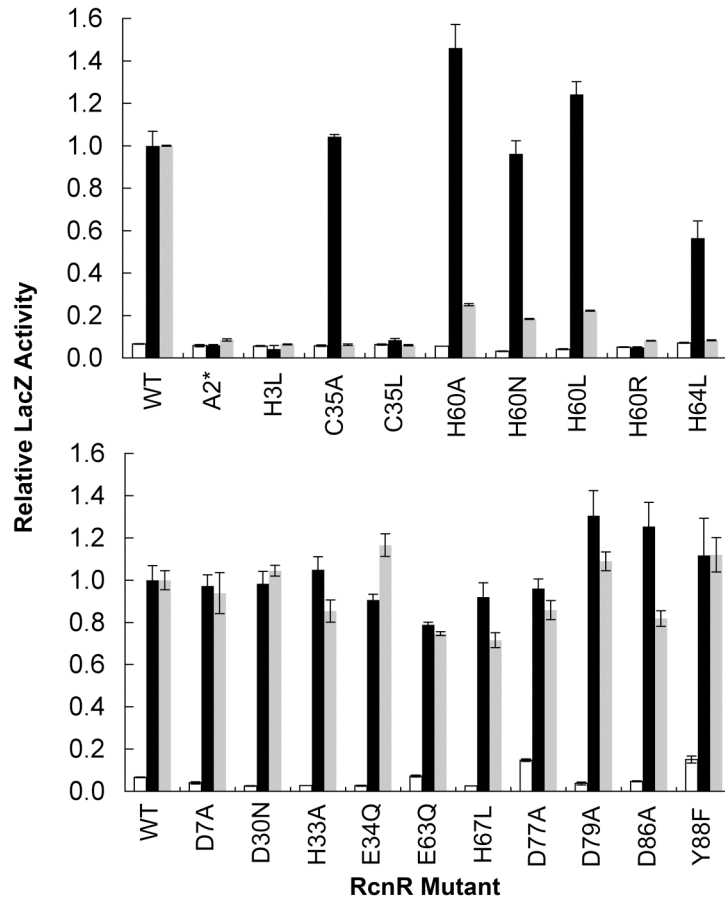


Figure 3.12. Effect of side-chain substitution of conserved residues on RcnR function in vivo. Top: Mutants that affect either Co(II) or Ni(II) responsiveness. Bottom: Mutants that do not affect Ni(II) and Co(II) responsiveness. β -galactosidase activity was measured for the P_{rcnA} - $lacZ$ reporter construct encoding wild-type or mutant $rcnR$ genes (Table 3.4) in *E. coli* strain PC563 grown anaerobically in LB media in the absence (white bars), or presence of 1 mM NiCl₂ (black bars) or 200 μ M CoCl₂ (gray bars). Each mutant is normalized to wild-type activity with either NiCl₂ or CoCl₂. A relative activity of 1 corresponds to about 168 Miller units with Ni(II) and 209 Miller units for Co(II). All data are

normalized to WT plus Ni(II) for ease of comparison. Error bars show the standard deviation of the LacZ activity of 3 separate samples. Mutations were constructed in pJI114 as described in the Materials and Methods.

terminus. This mutation was constructed to test the role of amino-terminal coordination in metal responsiveness based on the precedent of N-terminal Ni-coordination seen with Ni-SOD^{27,28}.

Ni(II) responsiveness. Two mutations (Ala2* and His3Leu) completely abolished Ni-responsive LacZ expression. A third mutation (His64Leu) showed only partial induction (44%) as compared to wild type. Mutants at two positions (Cys35Ala and His60Ala) that lost Co(II) responsiveness (see below), maintained sensitivity to Ni(II), suggesting a key structural difference in the metal-mediated inhibition of DNA binding by Ni(II) versus Co(II). The response of Cys35Ala to Ni(II) was unexpected based on the XAS analysis for this protein, which showed the loss of S-coordination. A Cys35Leu mutation was nonresponsive to Ni(II), which could result from steric perturbation of the Ni(II)-binding site. Additional mutations of His60 to Leu and Asn also maintained Ni sensitivity, suggesting that this His residue is not a Ni ligand, in contrast to the results for Co(II) (see below). Only the mutation of His60 to Arg completely abolished nickel responsiveness, most likely due to steric effects from the large, positively charged Arg side chain. Mutation of Asp7, Asp30, His33, Glu34, Glu63, His67, Asp77, Asp79, Asp86, or Tyr88 had no effect on Ni(II) responsiveness, eliminating these residues as possible Ni ligands.

Co(II) responsiveness. Mutant RcnR constructs were tested for LacZ activity in the presence of 200 μ M CoCl₂. LacZ activity increased ~20-fold under these experimental conditions when wild-type RcnR is present (Figure 3.1). The

Ala2*, His3Leu, Cys35Ala, His60Ala/Asn/Leu/Arg and His64Leu mutants were nonresponsive (<10% of wild-type) to added Co(II) and showed constitutive, low LacZ activity. These mutations potentially account for five of the six cobalt ligands determined by XAS. As with Ni(II), mutation of Asp7, Asp30, His33, Glu34, Glu63, His67, Asp77, Asp79, Asp86, or Tyr88 had little or no effect on Co(II)-dependent LacZ expression as compared to wild-type RcnR.

UV-visible Absorption Spectroscopy of RcnR Mutants

UV-visible spectroscopy of mutant RcnR proteins was used to explore the relationship between M(II) coordination and the LacZ assay data, in particular, the differences between Cys35, His60 and His64 in Ni(II) and Co(II) coordination. Individual mutations were constructed in the RcnR expression vector. Mutant proteins were overexpressed and purified the same as for wild-type RcnR, with similar yields. UV-visible spectra for M(II)-substituted RcnRs (1:1 metal:RcnR) are shown in the Figures 3.13 and 3.14. The extinction coefficients of the absorbance maxima for each M(II)-RcnR variant are listed in Table 3.2. The fully Ni(II)- and Co(II)-responsive D30N RcnR mutant (Figure 3.12) served as a control protein that exhibited little perturbation of the UV-visible spectra of the M(II)-substituted proteins.

Ni(II)-substituted mutants. WT Ni-RcnR exhibited minimal absorbance features between 400 – 700 nm, as expected of a 6-coordinate metal site that lacks *d-d* electronic transitions. Decreases in Ni(II)-coordination number are expected to increase absorbance in the visible range. An alteration in the

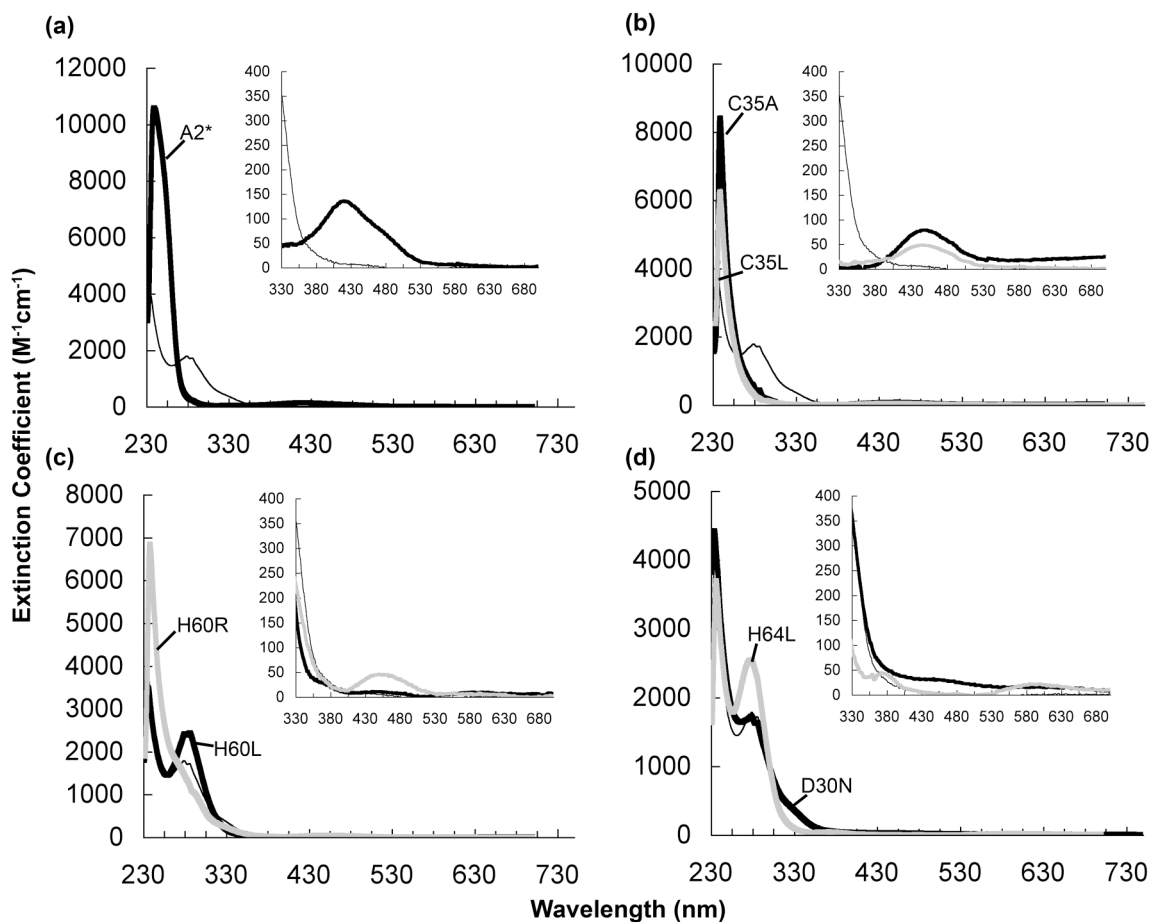


Figure 3.13. Electronic absorption spectra of Ni-substituted RcnR mutant proteins. Spectra for mutant RcnR proteins (110 μM) with 1:1 NiCl_2 in Buffer M as described in Materials and Methods for (a) A2* (thick black); (b) C35A (thick black), C35L (thick gray); (c) H60L (thick black), H60R (thick gray); (d) D30N (thick black), H64L (thick gray). WT Ni-RcnR is shown in thin black for reference.

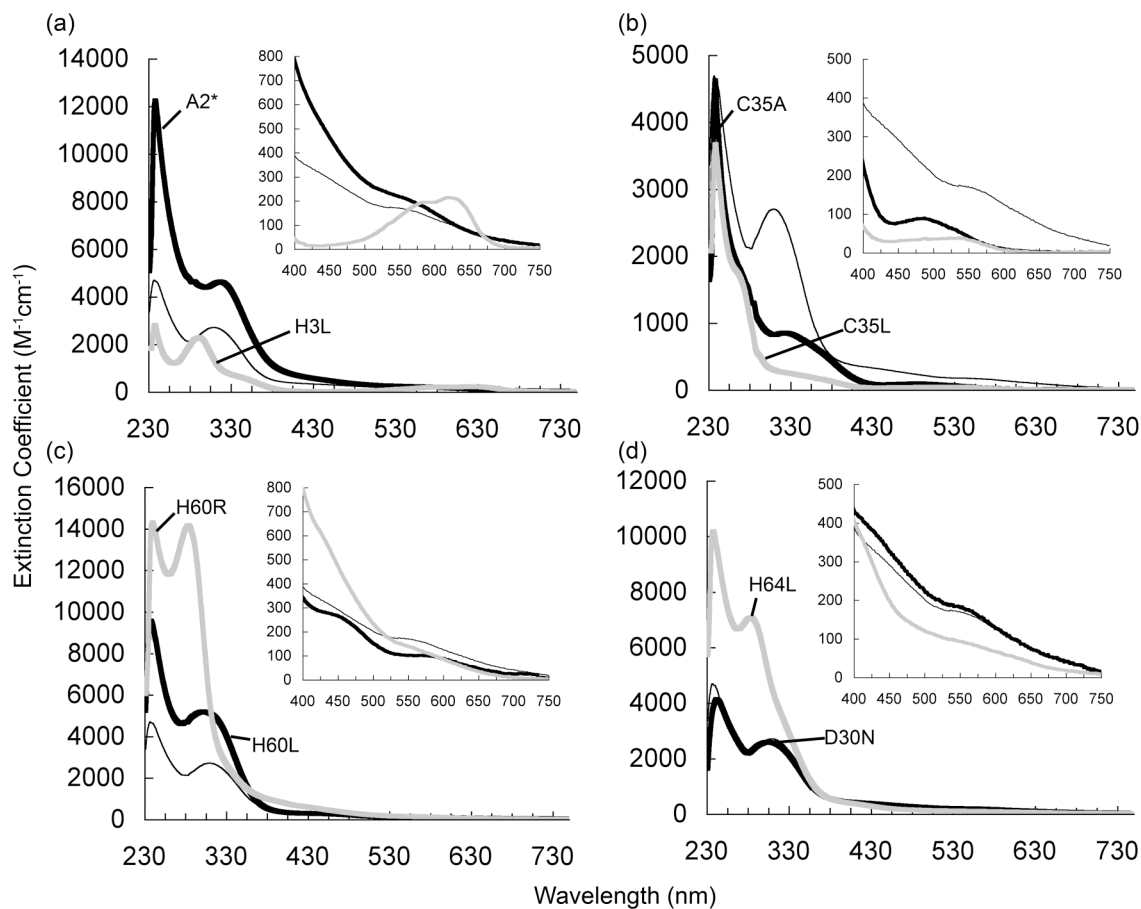


Figure 3.14. Electronic absorption spectra of Co-substituted RcnR mutant proteins. Spectra for mutant RcnR proteins (110 μM) with 1:1 CoCl_2 in Buffer M as described in Materials and Methods (except for D30N which was obtained in 10 mM Hepes (pH 7.0), 150 mM NaCl, 10% glycerol at 73 μM protein) for (a) A2* (thick black), H3L (thick gray); (b) C35A (thick black), C35L (thick gray); (c) H60L (thick black), H60R (thick gray); (d) D30N (thick black), H64L (thick gray). WT Ni-RcnR is shown in thin black for reference.

Table 3.2. UV-visible spectral features for M(II)-substituted RcnRs

RcnR variant	Ni(II)		Co(II)	
	λ (nm)	ϵ ($M^{-1} \text{ cm}^{-1}$)	λ (nm)	ϵ ($M^{-1} \text{ cm}^{-1}$)
Wild-type	235	3937	238	4675
			240	4609
	280	1778	314	2664
A2*	286	1713	549	shoulder
	240	10555	239	12200
	428	136	286	shoulder
			322	4681
H3L	none		557	shoulder
			238	2733
			296	2270
			597	194
D30N	234	4420	242	4100
	280	1709	312	2500
C35A	240	8308	238	4598
	456	76	272	1678
			325	850
C35L	240	6052	491	86
	452	48	239	3626
H60L	237	3462	346	shoulder
	281	2407	238	9539
	287	2429	307	5134
			450	shoulder
H60R			586	shoulder
	238	6840	240	14280
	290	shoulder	284	14145
H64L	450	45	550	shoulder
	237	3666	239	10148
	278	2546	289	6922
	380	40	550	shoulder
	600	19		

interaction with Cys35 is expected to shift the energy or magnitude of the 280 nm S→Ni LMCT peak. All RcnR mutants, except H3L, bound ~ 1 equivalent of Ni/RcnR monomer (Figure 3.15). H3L RcnR showed no features in the difference UV-visible spectrum indicative of bound Ni(II) atoms. For both C35A and C35L RcnR, the 280 nm S→Ni LMCT feature was absent, while the 235 nm feature remains. Additionally, both of these mutants have a feature at ~450 nm which can be attributed to a Ni(II) *d-d* transition. These data indicate that both mutants have 4- or 5-coordinate sites, the latter being consistent with the XAS analysis of C35A Ni(II)-RcnR. The Ni(II)-A2*-RcnR spectrum contained a feature at 420 nm ($\epsilon_{420}=136 \text{ M}^{-1} \text{ cm}^{-1}$) that is of similar intensity to one observed for the square-planar Ni(II)-site of *E. coli* NikR²⁹, and a similar geometry may be adopted for A2* Ni(II)-RcnR. This mutant also lacked the 280 nm S→M LMCT feature, suggesting loss of Cys-coordination. The H64L Ni(II)-RcnR spectrum contained two weak features at 380 nm and 600 nm not seen for WT. These features are consistent with a 6-coordinate Ni(II) site³⁰ that is clearly distinct from WT Ni(II)-RcnR. Except for A2* and H3L, each mutant shows a significant shift in the 280 nm S→M LMCT feature. Together, these data strongly indicate that the NH₂-group of the amino-terminus and the side-chains of His3, Cys35 and His64 coordinate nickel. When these residues were disrupted, the proteins adopted altered coordination geometries (4, A2*; 5, C35A; 6, H64L), or cannot bind Ni at all (H3L). In contrast, the H60L Ni(II)-RcnR spectrum showed neither an increase in the intensity of *d-d* features nor a significant perturbation of the 280

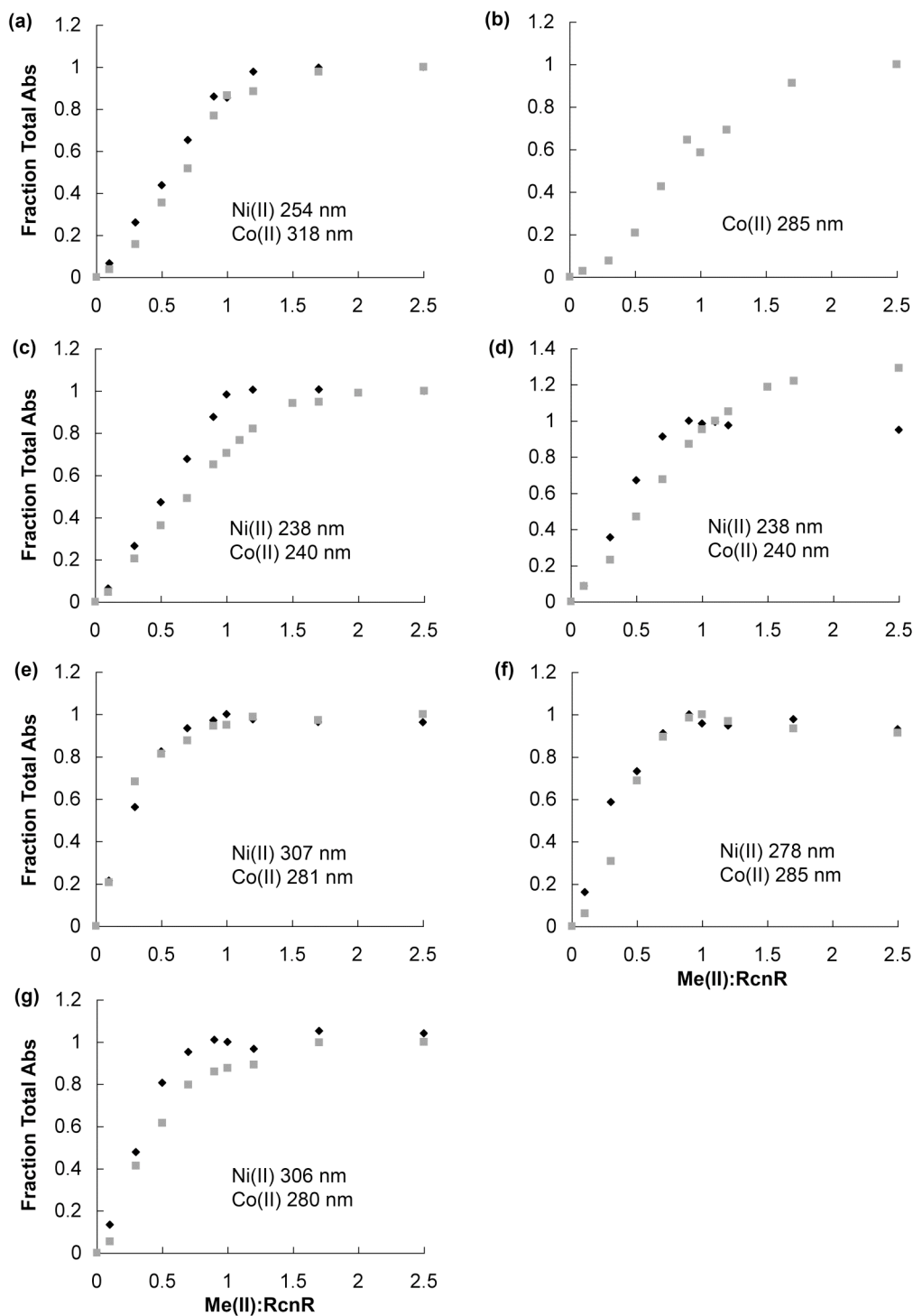


Figure 3.15. RcnR mutant metal titrations monitored by UV-visible spectroscopy. Ni(II) (black diamonds) and Co(II) (gray squares) titrations for

RcnR (a) A2* (b) H3L (c) C35A (d) C35L (e) H60L (f) H64L and (g) D30N were monitored by UV-visible spectroscopy at the indicated wavelengths as described in Materials and Methods. All mutants bound ~1 equivalent of Me(II) per monomer (0.8-1.4) except H3L which showed no spectral features for the Ni(II) complex, and displayed a sigmoidal curve with Co(II), reaching a plateau at ~2 Co/RcnR.

nm peak, suggesting that this residue is not involved in Ni coordination. These data are consistent with the LacZ data for the H60L/N mutants. Only when His60 was substituted with the bulky, charged Arg sidechain was the UV-visible Ni(II) spectrum affected, with features suggesting a five-coordinate metal-site. The effect of the H60R mutation on Ni(II)-responsiveness indicated this residue is near the metal-site but does not contribute to Ni(II) coordination.

Co(II)-substituted mutants. Unlike Ni(II)-RcnR UV-Vis spectra, the *d-d* transitions of wild-type and mutant Co(II)-RcnR proteins were difficult to resolve due to the extended tail of the S→Co LMCT (Figure 3.4; ~315 nm). However, other differences could be observed for Co(II)-substituted RcnR mutants (Figure 3.14). These changes, combined with alterations to the S→Co LMCT feature, are diagnostic of changes in Co(II) coordination. As with WT RcnR, all mutant proteins tested bound ~1 equivalent of Co(II) per RcnR monomer (Figure 3.15). The C35A and C35L Co(II)-RcnR spectra lacked the strong absorbance feature at 315 nm. These data confirmed the assignment of this peak as S→Co LMCT. They also provided further support for C35 as a Co ligand, in agreement with the XAS analysis of C35A Co(II)-RcnR. H3L RcnR still bound Co(II), although the Co(II) affinity appeared significantly weaker than for wild-type based on the sigmoidal shape of the titration curve. This mutant also showed distinct UV-visible spectral features in the *d-d* transition region with peaks at 597 nm ($\epsilon=194 \text{ M}^{-1} \text{ cm}^{-1}$) and 619 nm ($\epsilon=214 \text{ M}^{-1} \text{ cm}^{-1}$). These values are most consistent with a 5-coordinate Co(II) complex³¹.

Spectra for A2*, H60L, and H64L Co(II)-RcnR were not as different from WT, but each showed a large increase in absorbance for the S→Co transition and a shifts in energy for this transition (A2*, longer wavelength; H60L and H64L, shorter wavelengths). Changes were also observed between 500-700 nm, but these differences are difficult to resolve due to overlap with other features. It is likely that the mutant sites remain 6 coordinate, but with a variable sixth ligand as seen by XAS for C35A Co(II)-RcnR. While a precise characterization of these mutant sites cannot be made, it is clear that the mutations have altered the Co(II) site, and further supports the model of Co(II)-coordination by the NH₂-terminus, H3, C35, H60 and H64 suggested by the LacZ assay data.

Discussion

RcnR plays a pivotal role in intracellular sensing of Ni(II) and Co(II) ions in *E. coli*, where it must specifically interact with these metal ions when they are deemed to be in excess. The identification of the metal-binding ligands by a combination of XAS and mutagenesis experiments provides a starting point for understanding the allosteric regulation of RcnR activity by Ni(II) and Co(II). The characterization of the Ni(II)- and Co(II)-binding sites also permits a comparison of RcnR with NikR, the other nickel-responsive regulator in *E. coli*, as well as Ni(II)-/Co(II)-responsive regulators in other organisms. The results of this work, in combination with the prior study of Cu(I)-responsive CsoR⁸, establish a framework for understanding the ligand selectivity of other members of the RcnR/CsoR family of transcription factors.

Ni(II) and Co(II) bind to RcnR with a 1:1 stoichiometry and upper limits on the dissociation constants of 25 and 5 nM, respectively. These affinities make RcnR sensitive to low levels of free nickel and cobalt ions (lower limit ~ 1 nM) in *E. coli*. Both Ni(II) and Co(II) increase the stability of the RcnR tetramer without altering its oligomeric state or producing large-scale changes in secondary structure. The identity of the subunit interface required for RcnR tetramerization is unknown. The structure of *M. tuberculosis* CsoR revealed a dimeric species⁸, from which a reasonable structural homology model of RcnR can be constructed. The solution and crystal structure data presented for CsoR emphasized the dimeric form of the protein, but there was also evidence for a tetrameric species.

In particular, two distinct dimer-dimer interfaces are present in the unit cell, either of which may be relevant to tetramer formation by RcnR or may simply result from packing in the crystal lattice. CsoR also has a 30-amino acid C-terminal extension not visible in the crystal structure that may influence the multimeric state of the protein. RcnR dimers may occur in the cell, where its concentration is likely an order of magnitude less than those used in the analytical ultracentrifugation experiments. Differences in the oligomeric state between the two proteins may impact the mode of DNA binding and recognition, which is currently unknown. Nonetheless, the metal-binding properties of both CsoR and RcnR can be fully explained by examining only the dimeric assembly of the protein monomers.

The XAS analyses for Co(II)- and Ni(II)-RcnR reveal similar, but not identical, six-coordinate binding sites for each metal. XANES analyses indicate a pseudo-octahedral geometry for both metal ions, and EXAFS analyses are consistent with a similar set of six ligands: Cys(His)₂(N/O)₃ (Figure 3.7 and Table 3.1). The major difference seen between Co(II) and Ni(II) centers is the M-S distance, which is long for Ni(II)-RcnR (2.54 Å) and short for Co(II)-RcnR (2.24 Å). The 2.54 Å Ni-S distance is higher than the mean distance for unidentate S-donor ligands of all types (2.47 Å)²⁴ and at the high end of the distances found in the small number of structurally characterized high-spin, six-coordinate Ni(II) thiolate complexes that are known. Six-coordinate Ni(II) complexes are invariably high-spin ($S = 1$), which accounts for longer Ni-ligand distances in

these complexes. The lability of the high-spin $S = 1$ configuration also likely accounts for the fact that only a few high-spin nickel thiolates have been characterized, and these feature chelating thiolate ligands^{15,16}. For example, diaquabisdimethylaminothioquadratonickel(II) ($\text{NiS}_2\text{N}_2\text{O}_5\text{C}_{12}\text{H}_{20}$) contains an average Ni-S distance of 2.469(1) Å¹⁵, tetraethylammonium tris(2-thiopyridiniumnickelate(II)) ($\text{Ni}(\text{SC}_5\text{H}_4\text{N})_3$) contains Ni-S distances ranging from 2.518(1) to 2.541(1) Å¹⁶, and tetraphenylphosphonium tris(2-thiopyrrolidiniumnickelate(II)) ($\text{Ni}(\text{SC}_4\text{H}_3\text{N}_2)_3 \cdot \text{CH}_3\text{CN}$) contains Ni-S bond distances ranging from 2.460(2) to 2.545(2) Å¹⁶. The Ni-S distance in Ni(II)-RcnR is clearly distinct from the short Ni-S distances found in planar, $S = 0$, NiS_4 complexes (mean Ni-S = 2.18 Å)²⁴. The Ni-N average distances observed for Ni(II)-RcnR (1.98 and 2.11 Å) are also consistent with the high-spin configuration and fall in the range observed for structurally characterized high-spin Ni(II) N_6 complexes (mean Ni-N = 2.07 Å)²⁴. Thus, the metal-ligand distances found for the WT Ni(II)-RcnR complex are consistent with a high-spin ($S = 1$) complex. Co(II)-RcnR has an EPR spectrum that is typical of a high-spin, $S = 3/2$ center²⁶, and clearly different from the low-spin, $S = 1/2$, complexes that have been characterized³². The metal-ligand distances found for six-coordinate Co(II) thiolates are variable [2.338(1) to 2.585(1) Å]^{33,34}, a reflection of the $S = 3/2$, d^7 electronic configuration, which lacks an unfilled π -symmetry orbital in contrast to d^8 Ni(II). An example of a structurally characterized six-coordinate Co(II) thiolate complex with a confirmed $S = 3/2$ ground state ($\mu_{\text{eff}} = 4.83 \mu_{\text{B}}$) is dimethanolbis-

dimethylaminothioquadratocobalt(II) ($\text{CoS}_2\text{N}_2\text{O}_5\text{C}_{12}\text{H}_{20}$), which features an average Co(II)-S distance of $2.500(1) \text{ \AA}$ ³⁵. The Co-S distance found for Co(II)-RcnR lies at the lower end of the range of Co(II)-S distances observed. The difference between Co(II)-S and Ni(II)-S distances in RcnR suggests that either the Co(II)-S bond is supported by a π -interaction that is absent for Ni(II) or that the protein conformation allows stronger Co(II)-S binding. A likely consequence of the shorter Co(II)-S interaction is a different protein conformation resulting from Co(II) versus Ni(II) binding, which is discussed further below.

The structural characterization of the Ni(II)- and Co(II)-binding sites of RcnR allows a comparison with Ni(II) and Co(II) sites in other metalloregulators from *E. coli* and other bacteria. This study also completes structural characterization of the three known Ni(II)-binding sites involved in transcriptional regulation in *E. coli*. The two nickel-responsive transcriptional regulators from *E. coli*, RcnR and NikR, operate in a sequential fashion in response to increasing intracellular nickel⁶. NikR binds to DNA in the presence of Ni(II) and represses the transcription of the genes encoding the NikABCDE uptake permease. In contrast, RcnR-dependent repression of *rcnA* is maintained until *nikABCDE* expression is fully repressed, then increasing Ni(II) releases RcnR from DNA. NikR contains two Ni(II)-binding sites, termed high- and low-affinity. The high-affinity Ni(II)-binding site ($K_d \sim 1 \text{ pM}$)^{4,36} is a low-spin ($S = 0$), four-coordinate planar complex with His_3Cys coordination^{37,38} that is essential for biological function³⁹. Changes in high-affinity Ni(II)-site occupancy are thought to tune

nikABCDE expression according to Ni-enzyme expression levels^{6,39}. The low-affinity Ni-sites (K_d 5 nM)⁴⁰ are six-coordinate but do not employ thiolate coordination⁵. Low-affinity site occupancy increases DNA-binding affinity ($K_d \sim 5$ pM)^{36,40}, and results in the full repression of *nikABCDE* expression³⁹. The Ni(II) site in RcnR is high-spin and six-coordinate with an affinity similar to the six-coordinate site of NikR, but their relative numbers in the cell are probably quite different. The NikR low-affinity sites occur only when bound to DNA so there are two per cell (considering that there is only one known NikR DNA-binding site in the chromosome) while the number of RcnR sites is likely 1-2 orders of magnitude greater based on the range of intracellular concentrations observed for different repressor proteins in *E. coli*^{36,41-43}.

The differences in the electronic structures of the three sites have several ramifications for the physical and functional properties of the Ni(II) sites. First, the metal-ligand bond distances are quite different, seen most dramatically in the Ni-SCys distances for NikR³⁷ (2.13 Å) and RcnR (2.54 Å). These differences reflect the increased ionic radius of the high-spin ion that occurs when σ^* levels are populated. The mechanism for ligand exchange reactions for the two types of Ni(II) centers are also expected to differ. High-spin Ni(II) centers typically undergo ligand exchange via ligand-dissociation, which is favored by the high-spin electronic configuration. In contrast, four-coordinate planar complexes undergo ligand exchange via a mechanism involving association of the incoming ligand and therefore depend strongly on the nature of the nucleophile involved⁴⁴.

Thus, the NikR high-affinity site, which is not easily accessible in the protein, is not prone to easy nickel release once the metal is bound. A specific mechanism, such as a conformational change, might be required to release nickel. The occupancy of the high-affinity site of NikR is inversely correlated with nickel-enzyme expression^{6,39} and its function is consistent with the conservation of biosynthetic resources rather than stress. In contrast, the six-coordinate Ni(II) sites of NikR and RcnR appear to be designed for spontaneous metal release in response to decreasing metal ion concentration, as both of these sites function to respond to fluctuations in intracellular nickel concentrations that occur under stress conditions^{6,39}.

The use of a single cysteine thiolate ligation in NikR and RcnR would seem to be an ideal choice for a metal ion electronic structure sensor. The thiolate sulfur is both a σ - and a π -donor and is therefore able to discriminate between metal ions by sensing the occupancy of both sets of metal d -orbitals. Considering the Co(II)- and Ni(II)-substituted forms of RcnR and NikR, low spin Ni(II) has filled π -symmetry d -orbitals and the highest energy σ -orbital ($d_{x^2-y^2}$) is directed at the ligands and is unoccupied. This electronic structure leads to a very short in-plane Ni-S distance (2.13 Å). In contrast, for high-spin Ni(II), there is still no vacancy in the π -symmetry orbitals, but the σ -symmetry orbitals are both singly occupied, resulting in a much longer Ni(II)-SCys distance. For Co(II), the π -symmetry orbitals have one vacancy, opening up a possible π -interaction, resulting in an intermediate M-S distance (2.26 Å). Thus, the use of a single

cysteine ligand in metalloensors is a useful mechanism for interrogating the electronic structure of the bound metal ion, a critical aspect of generating the appropriate metal-specific allosteric response.

Mutation of RcnR identified specific residues involved in metal binding and support the ligand set determined from EXAFS analysis. The mutant data reveal that Ni(II) and Co(II) use distinct, but overlapping, ligand spheres to regulate RcnR DNA-binding activity. The mutation of His3 showed the most severe effects on metal binding, virtually eliminating UV-vis spectral features for Ni(II) and substantially decreasing Co(II)-binding affinity. The involvement of the N-terminus in the metal-binding site is supported by the CsoR structure, where the amino terminus is near the Cu(I)-binding site (Figure 3.16)⁸. The extension of the N-terminus by one amino acid (A2* mutant) also impaired Ni(II) and Co(II) binding and responsiveness. The involvement of the N-terminal amino group and the side chain of the second residues is similar to that observed for Ni-SOD^{27,28}. In this case, the deprotonated main chain amide of residue 2 also coordinates the Ni atom, suggesting that RcnR also uses main chain coordination by the amide of His3 for Ni(II) and Co(II) binding. Both Ni(II) and Co(II) are known to bind Gly-His and deprotonate the main chain amide to form a

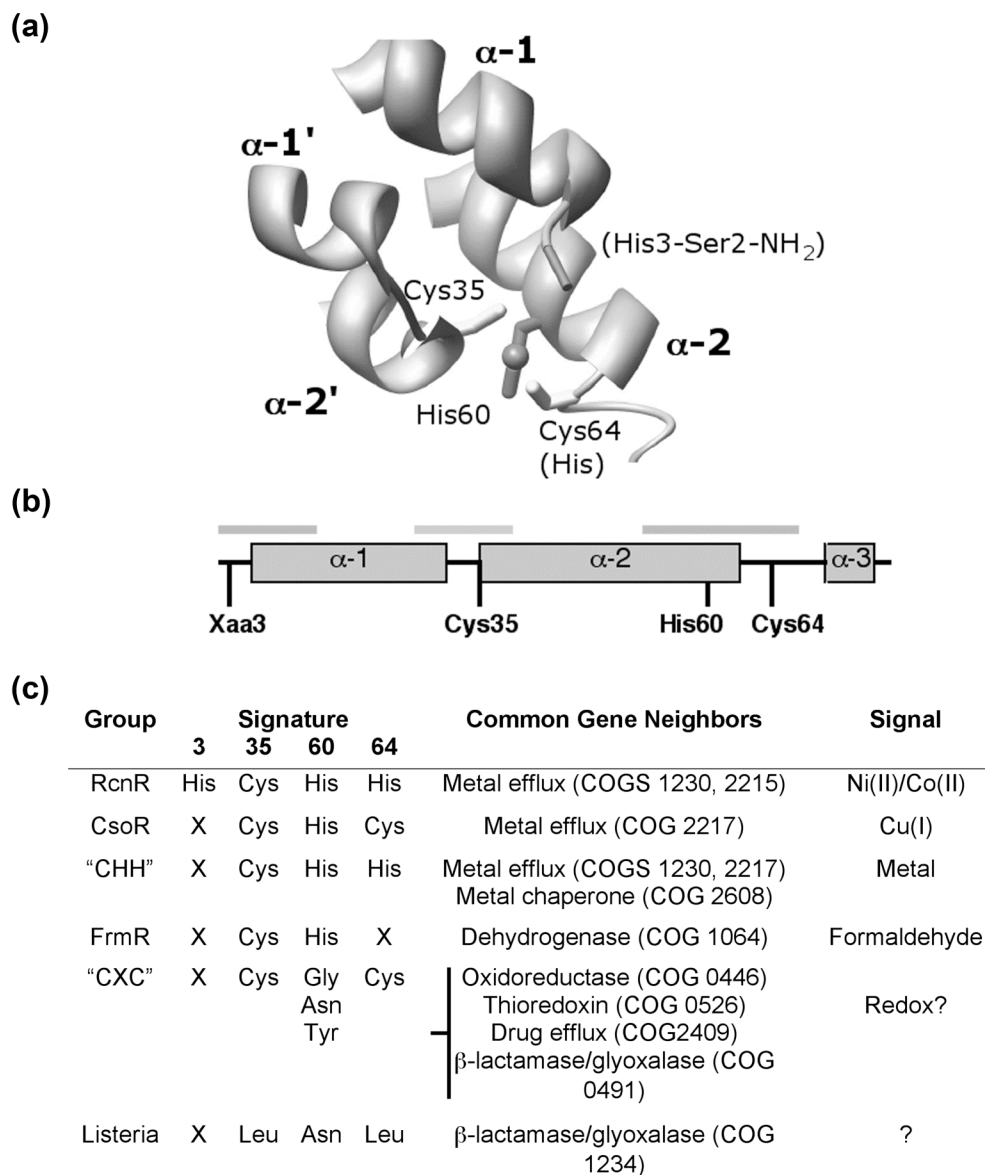


Figure 3.16. Fingerprint of the RcnR/CsoR family. (a) Image of Cu(I)-CsoR (PDB ID 2HH7) showing the residues (RcnR numbering) that coordinate the cuprous ion (gray sphere). The two different subunits of the CsoR dimer are shown. Text in parentheses refers to the identity of RcnR Ni/Co coordinating residues. (b) Scheme of the CsoR primary structure showing the positions of the fingerprint amino acids identified by bioinformatics analysis. The gray bars correspond to

the regions of structure shown in panel A. (c) Summary of the sequence fingerprint, gene neighborhood, and ligand predictions of different classes within the CsoR/RcnR family.

stable coordination complex^{45,46}, providing a chemical precedent for our model. The A2* mutation likely impairs Ni(II) responsiveness by forming a square-planar ATCUN motif (Xaa-Xaa-His)^{47,46}. The tripartite coordination of Ni(II) and Co(II) by the two N-terminal residues of RcnR is a key aspect of its selectivity toward metals with preferences for higher coordination numbers.

His64 also contributes to the responsiveness of Co(II) and Ni(II). The slightly different effect of the His64Leu mutation on the metal response for Ni(II) and Co(II) is consistent with expectations arising from differences in electronic structures of the metal-coordination complexes. This residue at this position in CsoR is Cys suggesting that changes in this residue position will be important for metal selectivity and responsiveness regardless of coordination number.

The differential contributions of Cys35 and His60 to Co(II) and Ni(II) responsiveness suggest that these residues play different roles in the structural response of the RcnR to metal binding. The spectroscopic studies of RcnR demonstrate that Cys35 is a ligand for both metals. The basis for the absence of an effect of the Cys35Ala mutation on Ni(II) responsiveness is currently unclear. Cys35 may play a more prominent role in metal selectivity rather than the allosteric response to metal binding. In contrast to Cys35, the mutation of His60 has little effect on the spectroscopic properties of the Ni(II) site. This position is invariant in RcnR and CsoR proteins, so its dispensability with regard to Ni(II) response is unexpected but might reflect another difference between the ligands used to bind Ni(II) and Co(II).

The minimal structure that emerges from the combination of XAS and mutagenesis analysis, $M(\text{Cys})(\text{His})_2(\text{NH}_2\text{-Ser})(\text{N/O})_2$, clearly identifies five metal ligands (the NH_2 -terminus, the His3 amide, and the sidechains of His3, Cys35, and His64) for both Ni(II) and Co(II), and a sixth, His60, for Co(II). The difference between EXAFS and mutagenesis, combined with structural homology, in assigning the number of His ligands is not unreasonable owing to the effect of His conformation on fitting these residues to EXAFS data⁴⁸. A sequence alignment of the predicted RcnR proteins (Figure 3.17) does not reveal a highly conserved sixth residue with a side chain likely to function in metal binding (Asp, Glu, His, and Cys). Mutation of several candidate residues, including those closest to the M(II)-binding site, did not cause any changes in Co(II) or Ni(II) responsiveness, leaving the identity of the sixth Ni(II) unresolved. The mutagenesis approach will only identify side chain ligands, so a Ni(II) coordination sphere completed by a second main chain amide will require additional experimental approaches for identification. Alternately, the coordination sphere could be completed by a ligand from the solvent (aqua or hydroxo) but not a halide ion, based on XAS data.

The regulation of DNA binding through allosteric changes communicated between metal- and DNA-binding sites is commonly observed in metalloregulator proteins¹. Giedroc and co-workers have shown for Zn-responsive CzrA and SmtB proteins that two residues of the tetrahedral metal-binding site are critical for transmitting the metal-binding state of the site to the rest of the protein^{49,50}.

	1	*	*	*	*
E.coli_RcnR	M S H T I R D K Q K L K A R A S K I Q G Q V V A L K K M L D E P - H E C A A V L Q Q I A A I				
A.tumefaciens	M S H T I R N K D K L L T R I R R L K G Q M E A V E R A L D D A - K P C G E V L Q L L A S V				
R. leguminosarum	M S H T T L Q K K K L V A R I S R L K G Q M E A V E R A L E A E - R P C G E I L Q L L A S I				
P. denitrificans	M S H T I R D K Q K L I A R V R R L R G Q I E G I E R A L E A E - K P C G E I L R Q L A S A				
D. acidovorans	M S H T S R Q K D K L V A R V R R I K G Q L E G I E R A L E S D - A A C V E V L R Q I A S V				
Acidovorax sp.	M S H T H G Q K D K L I A R V R R I K G Q M E G I E R A L E G D - A A C A E V L R Q L A S A				
N. hamburgensis	M S H T I K E K S K L I S R V R R I K G Q L E A V E R A L E S E - I G C A D V L M L V A S V				
Bradyrhizobium sp.	M S H T I K Q K S K L L A R V R R I K G Q I D A V E R A L E A E - L G C A D V L M L V A S V				
M. magnetotacticum	M S H T V K D K T K L L A R V R R I K G Q A E A V E R A L E A E - I G C T D V L M L V A S M				
X. autotrophicus	M A H T I K H K T K L L A R I R R I K G Q A E A V E R A L E A E - L G C T D V L M L V A S M				
Mesorhizobium sp.	M A H L Q K N N A K L V A R V R R L R G Q M E A V E R A L E G G - A P C G E V L H L V A S V				
R. sphaeroides	M A H L T T D K S K T L A R V R R L K G Q I E A V E R A I E A E - A D C A A L L Q L V A A V				
X. autotrophicus	M V H T I S N K D Q L I A R V R R I A G Q M A A I E K A I T E E - V G C S A V L H Q V A G A				
Nitrobacter sp.	M A H I K S S K D K L L G R V R R I S G Q L A A I E K A I D G E - A G C S A I L H Q V A G V				
S. alaskensis	M S H L A - A R P D L I A R V R R I A G Q V A A V E K A L A A D - A P C A T V L H R V A A V				
Sphingomonas sp.	- M H I V E D R E K L L A R T R R I A G Q I A A V E R Q L A G D - A D C S Q T L Q L V A S V				
S. wittichii	M G H I A K N R D Q L I A R V R R I A G Q V G A V E R S I Q A E - A G C S E T L H L V A A V				
P. bermudensis	M H L S Q N S D R L R A R L R R I T G Q L G A V D D A L G G E - A P C A A V L Q Q L A A A				
P. syringae	M T H V A S N K D D L L K R V K R I A G Q I Q A V E R A L E S D - H D C A K T L H L V A A T				
P. putida	M G H L A S N K E A L L K R V K R I A G Q V Q A I E K A L Q E E - E D C A K T L L L V A S A				
B. parapertussis	M S H L S D H K D E L L K R T Q R I A G Q V Q A I Q R S L E T E - A D C E S I L H L V A A T				
P. fluorescens	M S H T H E H K G E L L N R V R R I A G Q V Q A V E K A L E S E - A D C A K T L H L M A A I				
P. fluorescens	M S H T H A Q K D E L L K R V R R I L G Q V Q A V E R A L E S G - A D C G K T L H L V A A T				
Marinobacter sp.	M S H I T H N K A K L L G R V R R I R G Q V E A L E R A L D T D - K G C T E V L H Q I A A V				
X. campestris	M A H L N Q D K S K L L A R V R R I R G Q V E G L E K A L A Q D - V E C T A L L T Q V A A F				
X. campestris	M A H L T Q D K S K L L A R V R K I R G Q V E A L E K A L A Q D - V E C T A L L T Q V A A F				
V. cholerae	M S H T T R D Q K K L N T R V S K I Q G Q V N G L K K M L E E E - H E C H D V L Q Q I A A I				
R. solanacearum	M P H T I R D K A K L L A R V R R I Q G Q A A A L E K Q L N E E - G D C T E V L Q Q I A A I				
A. borkumensis	- M H T I K D K A K L L T R V R R I Q G T A A L E R Q L D E E - G D C T A V L Q Q I A A I				
C. violaceum	M S H T V R D K A K L L A R V R R L Q G Q A A A L E K Q L E N D - G D C S A I L Q Q I A A I				
C. psychrerythraea	M S H T I A S K K K L L S R V R R I K G Q S I A L E S V L E G E - P D C M A V L Q Q V A A I				
S. denitrificans	M S H I H K D N K K I L T R V R R I K G Q A E A L E K L L E T Q - P D C S Q V L Q Q I A A I				
P. aeruginosa	M A H T I H G K K K L L A R V R R I A G Q T R A L E Q A L E E G - S E C A A V L Q Q I A A I				
C. testosteroni	M S H T V S N Q K P L L A R I R R I K G Q T Q A L E R A L E E G - Q D C G V V L Q Q L A A V				
A. hydrophila	M S H T I H G K K K L L A R V R R I K G Q A G A L E S A L E Q E - S D C A A I L Q Q I A A I				
C. phaeobacteroides	M A H T K E G R K K L L N R V R R I K G Q I D A L E M A L E Q E I S D C S A V L Q Q I A S I				
gamma proteobacterium	M A H T I K Q R D K L L T R V R K I R G Q T E G L E K L L A K D - G E C S R I L Q Q I A A I				
C. limicola	M P H T K K D K K K L L T R I R R I K G Q T E A L E K A L E A G - V E C S G I L Q Q I A A I				
P. putida	M A H T L K S K K Q L L T R V R K I K G Q A A A L E T A L E Q D - K D C L A I L Q Q I A A V				
B. thailandensis	M S H T I R E K Q K L L N R V R R I K G Q V E A I E R A L E E E - C G C G D V L Q R I T S C				
B. xenovorans	M S H T V Q E K Q K L L N R V R R I K G Q V E A I E R A L E E E - H G C S D V L Q R I T S C				
B. cenocepacia	M S H T V R E K Q K L L N R V R R I K G Q V E A I E R A L E D E - R G C N D V L Q L I T S C				
B. vietnamiensis	M S H T I Q E K Q K L L N R V R R I K G Q V E A I E R A L E D E - R G C N E I L R L I T S G				
R. pickettii	M S H T I R E K Q K L L N R V R R I R G Q L D A I E R A L E E E - Q G C A P V L Q Q I T S C				
B. multivorans	M S H T I K E K Q K L L N R V R R I K G Q V E A I E R A L E E E - R S C S E I L Q L I T S G				
A. cryptum	M T H T V R E K R K L L A R V R R I R G Q V E A I E R A L E A E - A G C E T V M Q Q I A G V				
S. marcescens	M S H T H R D Q K K L L A R V R R I K G Q A E A L E R A L E S G - G E C S A V L Q Q I A A V				
A. baumannii	M S H I G Q D K K - I V N R V K R L K G Q I N S I E N A I E Q P E S S C I D I L Q Q V A A I				
consensus	M S H T I K D K K K L L A R V R R I K G Q V E A L E R A L E E E - A D C A A V L Q Q V A A V				

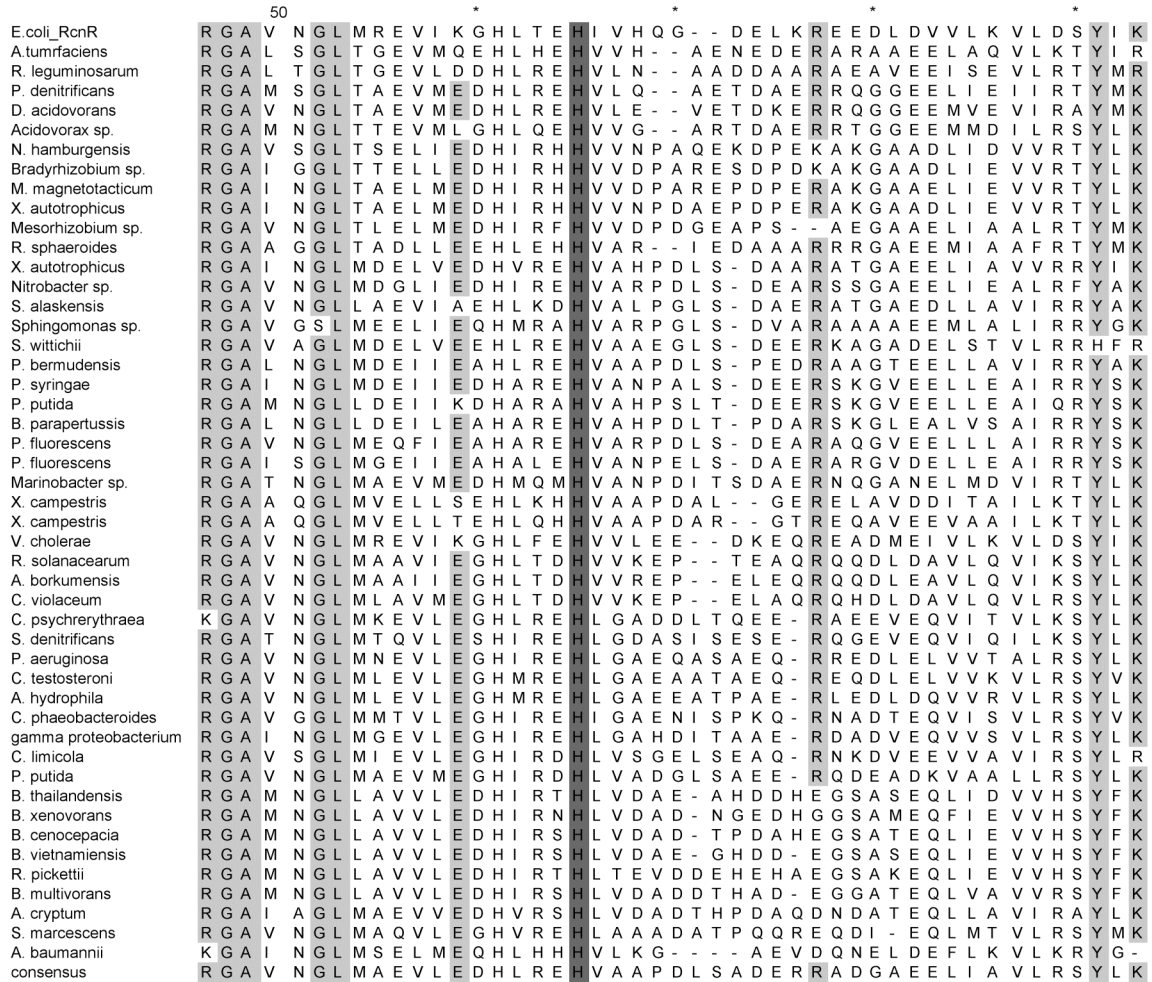


Figure 3.17. Alignment of predicted Ni-/Co-dependent regulators. 48

sequences retrieved using a BLAST search with *E. coli* RcnR, were found to have at least 30% sequence identity and have the sequence pattern of His3, Cys35, His60 and His64 (*E. coli* RcnR numbering; dark gray). Proteins with >90% identity with another sequence were also removed to limit redundancy. Positions with at least 75% identity in the alignment are shaded light gray. The sequence alignment was generated using ClustalW⁵¹.

The other two residues are primarily important for maintaining the tetrahedral coordination geometry. These observations suggest that within a particular structural class of metalloregulator, one or more master residues are critical for allostery. RcnR and CsoR share three conserved ligand positions (35, 60, and 64), suggesting that one or more of these residues may be essential for the allosteric response to metal binding. The regulation of RcnR by Co(II) and Ni(II) binding is coupled to a change in stability. For Co(II) and Ni(II), the ligand-binding sites utilize three distinct regions of the protein primary sequence and link these elements together (Figure 3.16). The stability change could include a rearrangement of secondary structure elements relative to the metal-free protein, or altering the orientation of monomers/domains relative to one another as seen for the iron-responsive regulator DtxR⁵². Either of these mechanisms is consistent with the absence of a change in RcnR secondary structure upon metal binding. The unaltered sensitivity of the Cys35Ala and His60 mutants to Ni(II) suggests either that these residues are not important for the allosteric response, or that more than one allosteric network exists in this protein family. The Cys35 mutant data suggest that Ni(II)-dependent allostery originates from intramonomer metal coordination, in contrast to Co(II)-RcnR and Cu(I)-CsoR where this residue is essential for activity. There is no functional requirement to preserve a single allosteric mechanism for the metal-dependent inhibition of DNA binding for a particular metalloregulator. The primary requirement of this class of metalloregulator is that metal binding reduces DNA-binding affinity. Thus, Ni(II)-

and Co(II)-RcnR are likely distinct structural entities whose common feature is an inability to bind DNA with appreciable affinity.

The requirement for thiolate coordination in Co(II) discrimination by RcnR is currently a singular occurrence. RcnR is functionally analogous to *M. tuberculosis* NmtR and KmtR, which also utilize six-coordinate geometry to bind Ni(II) and Co(II) between two monomer chains^{3,53,54}. However, KmtR and NmtR use only N and O ligands for metal coordination. Neither protein utilizes the NH₂-terminal, main-chain amide, and thiolate coordination environment proposed here for RcnR. Additionally, mutational studies of putative metal-binding ligands in both proteins have shown identical effects, or lack thereof, on Ni and Co responsiveness⁵⁴. KmtR only recently was identified⁵⁴, so the physiological requirement for two functionally similar Ni/Co responsive regulators in the same organism is not fully appreciated. However, the mutagenesis studies of NmtR and KmtR demonstrate that, despite their structural homology, the structural location of the metal-binding sites in each protein is distinct. This difference also manifests itself in the metal-binding affinities of the two proteins, wherein KmtR has a higher relative affinity for Ni(II) and Co(II) than NmtR⁵⁴.

NmtR and KmtR, like RcnR, are members of a large family of metalloregulators (ArsR/SmtB) that show distinct metal-binding specificities that correlate with differences in coordination number. Interestingly, the Zn-/Co-responsive *Staphylococcus aureus*³ CzrA protein uses a four-coordinate tetrahedral site. The higher coordination number observed in NmtR is achieved

by the recruitment of two His ligands from a short C-terminal extension in the protein that is not present in CzrA⁵³, similar to the recruitment of N-terminal residues for Ni/Co coordination by RcnR. The MerR family uses a variable loop region between two helices to increase the coordination number and alter metal specificity⁵⁵. These observations suggest that increases in metal site coordination number within a structural family generally require recruitment of ligands from distinct regions of the protein that are more tolerant to amino acid insertions and deletions, such as loops and the amino and carboxy termini.

The RcnR/CsoR family contains several distinct functional classes that likely respond to very different types of allosteric effectors. Ensuring ligand-binding specificity would require adaptation of the general binding site that has been identified from these studies of RcnR and those of CsoR. CsoR coordinates Cu(I) with three ligands (Cys35, His60 and Cys64, using RcnR numbering)^{8,56}, while RcnR responds biologically to metals that require higher coordination spheres by recruiting ligands from the N-terminal region, which is not used by CsoR for metal coordination, as well as a residue changes at position 64. The identification of four RcnR residues that when mutated drastically affect metal binding and metal responsiveness suggests that RcnR/CsoR-like sequences can be analyzed to make predictions about ligand-binding specificities. For example, sequences containing His/Cys/His/His residues at positions 3, 35, 60 and 64 (RcnR numbering) respectively, are likely Ni(II)- and Co(II)-responsive regulators (Figures 3.16, 3.18). Proteins with

Xaa/Cys/His/Cys residues (Xaa = nonconserved residue) are likely Cu(I) responsive regulators. Thus, the residues at these four positions provide a fingerprint that likely correlates with the ligand-binding specificity of a particular protein class in the largely uncharacterized RcnR/CsoR superfamily.

Using these minimal sequence criteria, ~450 unique sequences of the RcnR/CsoR family (COG 1937) were categorized into six different groups (Figure 3.16). CsoR-like regulators (Xaa/Cys/His/Cys) are the most abundant class, with 242 sequences. Of the remaining sequences, 72 (47 sequences with <95% identity; Figure 3.17) are predicted to be Ni/Co responsive transcriptional regulators. All sequences with this pattern are expected to express a protein with only one residue preceding His³, suggesting a strong bias toward the N-terminal coordination motif proposed here. All of the proposed *rcnR* genes have a neighboring gene that encodes a putative metal efflux protein (Figure 3.16), likely responsible for Ni/Co export. Position 3 in other groups varies widely and may not play a role in ligand binding in these other proteins. Thus, the RcnR class may be unique because it has recruited residues from the N-terminal region to aid in metal recognition.

Formaldehyde-responsive FrmRs⁵⁷ (12 sequences) can be singled out based on their fingerprint (Xaa/Cys/His/Xaa fingerprint) and proximity to formaldehyde dehydrogenase genes. The structural basis for formaldehyde responsiveness by FrmR is not known, (preliminary studies to address this are presented in Appendix 2) and the involvement of a metal-ligand has not been

explored. Additional sequence fingerprints for uncharacterized proteins were also detected, and predictions can be made about ligand specificities based on gene neighborhood analysis. For example, 92 genes encoding the Xaa/Cys/(Tyr, Asn, or Gly)/Cys proteins are preferentially near genes with functions related to antibiotic or oxidative stress responses, suggesting that not all RcnR/CsoR regulators are metal responsive. Some organisms, such as *Deinococcus geothermalis*⁵⁸, have multiple RcnR/CsoR family members, emphasizing the need for precise structural control of ligand-binding specificity. Further biochemical and biological studies will be required to identify the signals for each type of regulator. Once these signals are identified, the structural basis for ligand recognition and allosteric regulation can be investigated in detail.

Materials and Methods

Molecular Biology

All chemicals were purchased from Fisher Scientific (Pittsburgh, PA) except where indicated. Oligonucleotides (Operon Biotechnologies, Huntsville, AL), plasmids and bacterial strains used in this study are listed in Tables 3.3 and 3.4. Several *rcnR*-*P_{rcnA}*-*lacZ* mutations and pRcnR mutations were constructed using the following methods: (1) The QuikChange protocol (Stratagene, La Jolla, CA) was used with slight modification. For each 50 μ l reaction, 200 ng of plasmid DNA and 150 ng of each primer were used, with 2.5 units of *Pfu* Turbo DNA polymerase (Stratagene, La Jolla, CA). DNA was then subjected to digestion with 20 units of *Dpn* I (New England Biolabs, Beverly, MA) for 1.5 h at 37 °C. After heat inactivation for 10 min, 5 μ L of the reaction was transformed into *E. coli* JM109 cells following the TSS protocol⁵⁹. (2) Additional *rcnR*-*P_{rcnA}*-*lacZ* mutations and pRcnR mutations were constructed by overlap extension PCR²³ and ligation into pPC163 or pET-22b as described previously⁶. (3) Mutations near the 5' and 3' regions of *rcnR* were constructed by simple PCR amplification of the *rcnR* gene using a mutant primer (Table 3.3) followed by subcloning into the appropriate plasmid. All mutant constructs were sequenced (Seqwright, Houston, TX) to verify that only the desired mutation was present.

RcnR Overexpression and Purification

E. coli DL41 (DE3)⁶¹ cells harboring plasmid encoding wild-type or mutant RcnR proteins were grown in LB media containing ampicillin (200 μ g/ml, 538 μ M)

Table 3.3. Oligonucleotides.

Oligo	Purpose	Sequence
J1130	pRcnR	CTATGTCATATGTCTCATACAATCCGTGATAAACAG (ref ⁶)
J1131	pRcnR	GTCATACTCGAGTTATTTGATATATGAATCCAGCAC (ref ⁶) GTCATAGTCGACGATAATAATTCTTAGTATTAATTCGGC
J1140	rcnR-P _{rcnA} -lacZ	(ref ⁶) CTATGTGCGCCGTTAATAAATTATTTGATATATGAATCC (ref ⁶)
J1147	rcnR-P _{rcnA} -lacZ	
J1182	rcnRA2*-P _{rcnA} -lacZ	GTTAACCAGGTGTTGCCATGGCATCTCATACAATCCGTG
J1182r	rcnRA2*-P _{rcnA} -lacZ	CACGGATTGTATGAGATGCCATGGCAACACCTGGTTAAC
J1282	pRcnRA2*	CTATGTCATATGGCATCTCATACAATCCGTGATAAACAG
J1159	rcnRH3L-P _{rcnA} -lacZ	GGTGTGGCCATGTCTCTAACAATCCGTGATAAACAG
J1159r	rcnRH3L-P _{rcnA} -lacZ	CTGTTTATCACGGATTGTTAGAGACATGGCAACACC
J1189	pRcnRH3L	GGAGATATACATATGTCTCTAACAATCCGTGATAAACAG
J1189r	pRcnRH3L	CTGTTTATCACGGATTGTTAGAGACATATGTATATCTCC
J1305	rcnRD7A-P _{rcnA} -lacZ	ATGTCTCATACAATCCGTGCCAAACAGAACTGAAAGCG
J1305r	rcnRD7A-P _{rcnA} -lacZ	CGCTTTCAGTTTCTGTTTGGCACGGATTGTATGAGACAT
J1163	lacZ, pRcnRD30N rcnRD30N-P _{rcnA} -	GCGCTCAAGAAAATGCTCAACGAGCCGCACGAATGCGC
J1163r	lacZ, pRcnRD30N rcnRH33A-P _{rcnA} -	GCGCATTCGTGCGGCTCGTTGAGCATTTCCTTGAGCGC GAAAATGCTCGACGAGCCGGCGGAATGCGCTGCAGTTTT AC
J1303	lacZ rcnRH33A-P _{rcnA} -	GTAAAACCTGCAGCGCATTCCGCCGGCTCGTCGAGCATTTC
J1303r	lacZ rcnRE34Q-P _{rcnA} -	C
J1164	lacZ rcnRE34Q-P _{rcnA} -	GCTCGACGAGCCGCACCAATGCGCTGCAGTTTTAC
J1164r	lacZ rcnRC35A-P _{rcnA} -	GTAAAACCTGCAGCGCATTGGTGGCGGCTCGTCGAGC
J1136	lacZ, pRcnRC35A rcnRC35A-P _{rcnA} -	CTCGACGAGCCGCACGAAGCCGCTGCAGTTTTACAACAG
J1136r	lacZ, pRcnRC35A rcnRC35L-P _{rcnA} -	CTGTTGTAAAACCTGCAGCGGCTTCGTGCGGCTCGTCGAG
J1161	lacZ, pRcnRC35L rcnRC35L-P _{rcnA} -	CTCGACGAGCCGCACGAACCTCGCTGCAGTTTTACAACAG
J1161r	lacZ, pRcnRC35L rcnRH60R-P _{rcnA} -	CTGTTGTAAAACCTGCAGCGAGTTCGTGCGGCTCGTCGAG
J1141	lacZ rcnRH60R-P _{rcnA} -	CGGGAAGTGATTAAAGGTCTGCTGACGGAACACATCGTTC
J1141r	lacZ rcnRH60L-P _{rcnA} -	GAACGATGTGTTCCGTGACGACCTTTAATCACTTCCCG
J1298	lacZ, pRcnRH60L rcnRH60L-P _{rcnA} -	CGGGAAGTGATTAAAGGTCTACTGACGGAACACATCGTTC
J1298r	lacZ, pRcnRH60L rcnRH60N-P _{rcnA} -	GAACGATGTGTTCCGTGACGACCTTTAATCACTTCCCG
J1299	lacZ	CGGGAAGTGATTAAAGGTAACCTGACGGAACACATCGTTC

Jl299r	rcnRH60N-P _{rcnA} -lacZ	GAACGATGTGTTCCGTCAGGTTACCTTTAATCACTTCCCG GTGATTAAAGGTCATCTGACGCAACACATCGTTCACCAGG
Jl160	rcnRE63Q-P _{rcnA} -lacZ	G CCCTGGTGAACGATGTGTTGCGTCAGATGACCTTTAATCA
Jl160r	rcnRE63Q-P _{rcnA} -lacZ	C
Jl158	rcnRH64L-P _{rcnA} -lacZ, pRcnRH64L	GGTCATCTGACGGAACATAATCGTTCACCAGGGGGATGAG
Jl158r	rcnRH64L-P _{rcnA} -lacZ, pRcnRH64L	CTCATCCCCCTGGTGAACGATTAGTTCCGTCAGATGACC
Jl165	rcnRH67L-P _{rcnA} -lacZ	CTGACGGAACACATCGTTCTACAGGGGGATGAGCTAAAAC
Jl165r	rcnRH67L-P _{rcnA} -lacZ	GTTTTAGCTCATCCCCCTGTAGAACGATGTGTTCCGTCAG CTATGTCGGCCGTTAATAAATTATTTGATATATGAATCCAG C
Jl308	rcnRD77A- P _{rcnA} -lacZ	ACCTTCAGAACGACATCCAGAGCTTCTTCAC CTATGTCGGCCGTTAATAAATTATTTGATATATGAATCCAG C
Jl304	rcnRD79A- P _{rcnA} -lacZ	ACCTTCAGAACGACGGCCAGATC
Jl307	rcnRD86A- P _{rcnA} -lacZ	CTATGTCGGCCGTTAATAAATTATTTGATATATGAAGCCAG CACCTTC
Jl168	rcnRY88F-P _{rcnA} -lacZ	GAAGGTGCTGGATTCATTTATCAAATAATTTATTAACGGC
Jl168r	rcnRY88F-P _{rcnA} -lacZ	GCCGTTAATAAATTATTTGATAAATGAATCCAGCACCTTC

Table 3.4. Plasmids and bacterial strains.

Plasmid	Insert	Parent	Reference
pRcnR	rcnR	pET22b(amp ^R) pACYC163 (cm ^R)	6
pJI114	rcnR-P _{rcnA} -lacZ	pJI114(cm ^R)	6
pJI139	rcnRA2*-P _{rcnA} -lacZ	pRcnR(amp ^R)	This work
pJI155	rcnRA2*	pRcnR(amp ^R)	This work
pJI122	rcnRH3L-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI144	rcnRH3L	pRcnR(amp ^R)	This work
pJI163	rcnRD7A-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI131	rcnRD30N-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI167	rcnRD30N	pRcnR(amp ^R)	This work
pJI129	rcnRH33L-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI130	rcnRE34Q-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI118	rcnRC35A-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI117	rcnRC35A	pRcnR(amp ^R)	This work
pJI125	rcnRC35L-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI127	rcnRC35L	pRcnR(amp ^R)	This work
pJI116	rcnRH60R-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI161	rcnRH60N-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI160	rcnRH60L-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI159	rcnRH60L	pRcnR(amp ^R)	This work
pJI123	rcnRE63Q-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI124	rcnRH64L-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI142	rcnRH64L	pRcnR(amp ^R)	This work
pJI132	rcnRH67L-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI164	rcnRD77A-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI165	rcnRD79A-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI166	rcnRD86A-P _{rcnA} -lacZ	pJI114(cm ^R)	This work
pJI133	rcnRY88F-P _{rcnA} -lacZ	pJI114(cm ^R)	This work

Strain	Genotype	Parent	Reference
RZ4500	$\Delta lacZ$		60
PC563	$\Delta rcnR \Delta lacZ$	RZ4500	6
DL41	<i>metA</i> -	MG1655	61

at 37 °C to an OD₆₀₀ of ~1.0, prior to the addition of IPTG (120 mg/l, 0.5 mM) to induce RcnR expression. The cells were harvested after 3 h by centrifugation (4600 x g, 25 min), resuspended in residual media (~ 5 ml/l of culture) and frozen at -80 °C. Thawed cell pellets were lysed by sonication, at 4 °C in the presence of 5 mM Tris(2-carboxyethyl)-phosphine hydrochloride (TCEP) (Pierce) and 1.5 mM phenylmethylsulfonylfluoride (Sigma), a serine protease inhibitor. Subsequent steps were all performed at 20 °C. After centrifugation, the lysate supernatant was applied to an SP Sepharose column (18 ml) equilibrated with 20 mM Hepes (pH 7.0), 1.0 mM TCEP, 5.0 mM EDTA, 10% glycerol (Buffer A) with 50 mM NaCl. The column was then washed with 40 ml Buffer A with 50 mM NaCl followed by a linear gradient of 0.05 – 1 M NaCl in Buffer A (110 ml total volume; flow rate 2 ml/min; collected in 2 ml fractions). RcnR eluted at ~400 mM NaCl, and these fractions were pooled and concentrated at ~4000 x g to 1 ml using 2 ml, 5KDa MW cutoff Vivaspin concentrators (Vivascience, Hanover, Germany) and applied to a HiLoad 26/60 Superdex 200 column equilibrated with Buffer A containing 300 mM NaCl. RcnR eluted as a single peak, at a volume most consistent with a tetramer. Fractions containing RcnR were pooled and diluted to 25 ml and applied to an S6 ion exchange column (2 cm x 6 cm) (Biorad, Hercules, CA) running on a Biorad Biologic FPLC. The column was then washed with 20 ml Buffer A with 50 mM NaCl, and the protein was eluted using a linear gradient of 0.05 – 1 M NaCl (48 ml total volume, 4 ml/min flow rate). RcnR eluted between 490 and 620 mM NaCl. Purified RcnR was stored at 4 °C in 20

mM Hepes (pH 7.0), 500 mM NaCl, 1 mM TCEP, 1 mM EDTA, 5% glycerol for up to 4 weeks; thereafter, the DNA-binding specificity and affinity began to decrease. The identity of the purified protein was confirmed by N-terminal sequencing using Edman degradation (Midwest Analytical, St. Louis, MO). Prior to use in different experiments, RcnR samples were desalted two times into Buffer M [10 mM Hepes (pH 7.0), 150 mM NaCl, 5% glycerol] using Microbiospin 6 columns (Biorad) to remove EDTA and TCEP present in the storage buffer. Modifications to this approach are noted where relevant.

RcnR Characterization

An experimentally determined extinction coefficient of $\epsilon_{276} = 2530 \text{ M}^{-1}\text{cm}^{-1}$ was determined by measuring the absorbance of an aliquot of purified denatured RcnR and determining its concentration by amino acid analysis (Texas A&M Protein Chemistry Laboratory, College Station, TX).

The free sulfhydryl content of native RcnR was quantified by a standard thiol titration method⁶² using 5, 5'-dithio-bis(2-nitrobenzoic acid) (DTNB). Samples contained RcnR (5 -10 μM) in 100 mM sodium phosphate (pH 7.3). Quantification of the sulfhydryl concentration was obtained using a TNB extinction coefficient of $14150 \text{ M}^{-1}\text{cm}^{-1}$ at 412 nm using a Shimadzu UV-2401PC spectrophotometer⁶³.

Metal Content Assays

PAR assay.

For PAR assays, 3-4 monomer equivalents of NiCl₂ or CoCl₂ were added to RcnR protein previously desalted into Buffer M, and the samples were incubated at 20 °C for 2 h. Weakly bound and excess free metal ions were removed by a second passage over a Microbiospin 6 column equilibrated with Buffer M. 4-(2-pyridylazo) resorcinol (PAR) experiments followed an established protocol^{64,10} with minor modifications. Standard curves for Ni(II) and Co(II) in complex with PAR were generated from samples containing 1-9 μM metal, 20 μl of 1 mg/mL PAR in 120 μl total volume [50 mM Hepes (pH 7.3), 6 M guanidine hydrochloride]. The absorbance was measured for each sample (Ni, 505 nm and Co, 540 nm) after equilibration for 1 h at 20 °C. RcnR samples were prepared similarly, using two samples at two different protein concentrations between 1 and 9 μM. The reported metal content was the average of all four samples.

Atomic absorption spectroscopy (AAS).

Samples for AAS were prepared in a similar manner to those stated above except protein was desalted into Buffer M lacking glycerol to remove excess metal ions. Aliquots (100 μL) were then incubated overnight (20 °C) in 2% nitric acid, then diluted 10-fold with H₂O to a final volume of 1 mL. Samples were analyzed on a Perkin Elmer AAnalyst 600 graphite furnace atomic absorption spectrophotometer equipped with a nickel hollow cathode lamp (Department of Biology, Washington University). Nickel content was quantified by comparison to a standard curve of 2-100 μg/L nickel. Each sample was quantified three times and averaged.

UV-visible spectroscopy.

Metal-bound RcnR UV-visible spectra were measured on a Shimadzu UV-2401PC spectrophotometer. Metals were added from a stock solution to individual protein samples (previously desalted) for each nickel or cobalt concentration. Samples (100 μ l volume) were allowed to equilibrate for 15 h at 20 °C under aerobic conditions after metal addition before spectral collection (1 cm path length; 0.2 nm step size).

For iodoacetamide (IAM) (Sigma, St. Louis, MO) treated samples, 5 mM IAM was added to RcnR in storage buffer (2 h at 20 °C). Control samples with water added instead of IAM were prepared in parallel. The samples were then desalted twice into Buffer M before metal addition

Mag-Fura 2 and Fura 2 metal Titrations

Fluorescence measurements were all taken on a Cary Eclipse fluorescence spectrophotometer in a 150 μ l, 1 cm path length Hellma quartz cuvette. A range of concentrations of NiCl₂ or CoCl₂, was incubated with 0.5 μ M mag-fura 2 (Molecular Probes, Carlsbad, CA) or 0.3 μ M fura 2, respectively, for 3 h in the presence or absence of the indicated concentrations of RcnR.

Fluorescence measurements were made using $\lambda_{\text{ex}} = 380$ nm, $\lambda_{\text{em}} = 505$ nm for mag-fura 2 and $\lambda_{\text{ex}} = 360$ nm, $\lambda_{\text{em}} = 500$ nm for fura 2, with both excitation and emission slit widths set to 10 nm. Each reported value was the average of three readings. Data were fit to a four independent binding sites model using Scientist version 2.01 (Micromath, St. Louis, MO).

Analytical Ultracentrifugation

Sedimentation equilibrium data were obtained on a Beckman XL-A Ultracentrifuge using six-cell sample holders (1.2 cm pathlength). Samples of 110 μL RcnR (5 and 25 μM in Buffer M + 200 μM EDTA) were centrifuged at 25000 and 35000 rpm (20 °C) with a 120 μL buffer blank. A_{230} or A_{276} measurements were obtained after 10 h and every 2 h thereafter until three consecutive scans overlapped (20 h total) to ensure that equilibrium had been achieved. The partial specific volume for RcnR (0.7504 ml/g at 20 °C) was determined from the amino acid content using SEDENTERP⁶⁵. A buffer density of 1.019 g/ml was also calculated from SEDNTERP. The SEDPHAT software package was used to globally analyze data obtained at multiple rotor speeds and protein concentrations⁶⁶.

Circular Dichroism (CD)

CD experiments were carried out on a JASCO 710 spectropolarimeter using 1 cm (urea melts) or 1 mm (wavelength scans) path length cylindrical quartz cuvettes. Wavelength scans (20 °C, speed 20 nm/min, 1 nm step) were performed on samples containing 15 μM RcnR in 10 mM Hepes (pH 7.0), 100 mM NaCl, 5% glycerol. 30 μM Ni(II) or Co(II) was added where indicated. Samples were equilibrated for 3 h (20 °C) before scans. Each spectrum was the average of three accumulations. The raw, blank subtracted signal (mdeg) was converted to mean residue ellipticity (MRE) and fraction helix, as previously described^{11,12}.

Urea denaturation samples were performed as described previously³⁶. A freshly prepared stock of 10 M urea in Buffer M was used to make individual samples (900 μ l) containing different urea concentrations, which were allowed to equilibrate for 12-15 h at 20 °C. Ellipticity data for each urea concentration was collected for 60 s at 230 nm in 1 s intervals and averaged. Because the pathway for unfolding could not be unambiguously determined with respect to the oligomeric state, urea concentrations that cause half-maximal unfolding are reported based on fits to a model of folded monomer to unfolded monomer⁶⁷.

X-ray Absorption Spectroscopy

Samples of RcnR were desalted and metallated as described above. The samples were then concentrated to 1.5 – 3.2 mM in 20 mM Hepes (pH 7.0), containing either 300 mM NaCl or NaBr, prior to data collection. X-ray absorption spectroscopy (XAS) data for all RcnR samples were collected at beamline 9-3 at the Stanford Synchrotron Radiation Laboratory (SSRL) under dedicated ring conditions. Each sample was syringed into a polycarbonate cuvette that was wrapped in kapton tape and frozen in liquid nitrogen. Data were collected at 10 K using a liquid helium cryostat (Oxford Instruments). The ring conditions were 3 GeV and the current was 80-100 mA for each sample. The beamline optics consist of a Si(220) double crystal monochromator ($\phi = 90^\circ$) a flat rhodium coated mirror upstream of the monochromator for harmonic rejection and vertical collimation, and a toroidal focusing mirror after the monochromator. X-ray fluorescence was detected with a 30 element Ge detector (Canberra). The

energy of each K_{α} - edge was calibrated to the first inflection point of the corresponding foil: Co, 7709.5 and Ni, 8331.6 eV. X-ray absorption near edge spectroscopy (XANES) data were collected from ± 200 eV relative to the edge. Analysis of the XANES region was carried out by fitting a cubic function to the baseline in the pre-edge region of the spectrum and using a 75% Gaussian and 25% Lorentzian function to fit the rise in fluorescence occurring at the edge. Gaussian functions were added to fit each pre-edge transition, and the areas of the inserted Gaussians were taken to be the peak area⁵. Extended X-ray absorption fine structure (EXAFS) data were collected to 13.5 k above the absorption edge (E_0) for WT Co-RcnR and 16.3 k above the absorption edge energy (E_0) for both C35A RcnR complexes and WT each Ni-RcnR. XAS data reduction was carried out using EXAFS123 with graphics produced in Igor Pro (WaveMetrics ver. 5.0) as previously described⁵. For EXAFS analysis, four to seven scans were averaged to generate the summed EXAFS spectrum. Each summed spectrum was normalized and background corrected using a three-section cubic spline for the pre-edge and post-edge regions. The data were converted to k -space using the relationship $[2m_e(E - E_0) / \hbar^2]^{1/2}$ (where m_e is the mass of the electron and E_0 is the threshold energy above the absorption edge; 7723 eV for Co and 8340 eV for Ni). Least-squares fitting of EXAFS arising from the first coordination sphere was performed using Fourier-filtered spectra with transform limits of 2.0 - 12.5 \AA^{-1} for WT RcnR samples and 2.0 - 14.0 \AA^{-1} for C35A RcnR samples, as determined by the signal-to-noise in the data and to

allow appropriate comparisons. A back-transform window of 1.0 - 2.3 Å (uncorrected for phase shifts) was used to model scattering atoms in the first coordination sphere. Fits were calculated for coordination numbers four, five and six, holding the number of scattering atoms in a shell to integer values and using all combinations of S- and N-donors (see Appendix 3). Best fits were judged by minimizing the goodness of fit parameter (GOF) and deviation of σ^2 values from typical values. The back-transform window was expanded to 1.0 – 4.0 Å (uncorrected for phase shifts) to accommodate scattering from imidazole rings, which were fit using multiple-scattering parameters as previously described⁵. Histidine ligand counting was achieved by adding integer numbers of imidazole rings to the fit and observing the effect on GOF and σ^2 , and provides an estimate of the number of imidazole ligands in model compounds with an error of $\pm 25\%$ (about the same as the error in N, the number of first shell atoms). The final fits included in the text were selected from the best fits of Fourier-filtered data and were further refined using unfiltered data. For all fits, r (metal-scatterer distance), σ^2 (disorder parameter) and E_0 (phase shift) were allowed to refine; however they were constrained so that only one value of E_0 was used and only one value of σ^2 per type of scattering atom was employed.

EPR Spectroscopy

EPR spectra were obtained with a Bruker Elexsys 500 spectrometer equipped with an ER 4102 single cavity operating at X-band (9.46 Ghz). Measurements were performed at 12 K using an Oxford Instruments EPR-9000

liquid helium continuous flow cryostat. EPR spectra were simulated using the matrix diagonalization program Xsophe (Bruker Biospin). The spin Hamiltonian employed was: $H = \beta g \cdot H \cdot S + S \cdot D \cdot S$.

β -galactosidase Reporter Experiments

β -galactosidase reporter experiments were set up as described previously with the following modification⁶. Starter cultures of *E. coli* strain PC563 ($\Delta rcnR$) cells containing wild type or mutant *rcnR* genes on the *rcnR*-*P_{rcnA}*-*lacZ* plasmid (pJ1114) were grown aerobically in Luria-Bertani broth (LB) with chloramphenicol (34 mg/l, 0.1 mM). These cultures were used to inoculate triplicate cultures of 1.87 ml LB with chloramphenicol (34 mg/l, 0.1 mM) in capped microfuge tubes (starting OD₆₀₀ = 0.000001). Cultures were then grown for 12-15 hours at 37 °C and assayed as previously described⁶.

For metal induction experiments, cells were treated in a similar manner except pJ1114 was in the $\Delta lacZ$ *E. coli* strain RZ4500⁶⁰. Cells were grown anaerobically as above at the maximal concentration for each individual divalent metal that resulted in <10% inhibition of growth (measured by final OD₆₀₀), which was as follows: 30 μ M MnCl₂, 1.5 mM FeSO₄, 150 μ M CoCl₂, 500 μ M NiCl₂, 100 μ M CuCl₂, 300 μ M ZnCl₂ and 10 μ M CdSO₄.

References

1. Giedroc, D.P. & Arunkumar, A.I. Metal sensor proteins: nature's metalloregulated allosteric switches. *Dalton Trans* 3107-3120(2007).doi:10.1039/b706769k
2. Outten, C.E. & O'Halloran, T.V. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488-2492(2001).
3. Pennella, M.A. et al. Structural elements of metal selectivity in metal sensor proteins. *Proc. Natl. Acad. Sci. U.S.A* **100**, 3713-3718(2003).
4. Wang, S.C. et al. Selectivity of metal binding and metal-induced stability of Escherichia coli NikR. *Biochemistry* **43**, 10018-10028(2004).
5. Leitch, S. et al. Nickel-specific response in the transcriptional regulator, Escherichia coli NikR. *J. Am. Chem. Soc* **129**, 5085-5095(2007).
6. Iwig, J.S., Rowe, J.L. & Chivers, P.T. Nickel homeostasis in Escherichia coli - the rcnR-rcnA efflux pathway and its linkage to NikR function. *Mol Microbiol* **62**, 252-62(2006).
7. Rodrigue, A., Effantin, G. & Mandrand-Berthelot, M. Identification of rcnA (yohM), a nickel and cobalt resistance gene in Escherichia coli. *J Bacteriol* **187**, 2912-6(2005).
8. Liu, T. et al. CsoR is a novel Mycobacterium tuberculosis copper-sensing transcriptional regulator. *Nat Chem Biol* **3**, 60-8(2007).
9. Frottin, F. et al. The proteomics of N-terminal methionine cleavage. *Mol. Cell Proteomics* **5**, 2336-2349(2006).

10. Hunt, J.B., Neece, S.H. & Ginsburg, A. The use of 4-(2-pyridylazo)resorcinol in studies of zinc release from *Escherichia coli* aspartate transcarbamoylase. *Anal. Biochem* **146**, 150-157(1985).
11. Myers, J.K., Pace, C.N. & Scholtz, J.M. Helix propensities are identical in proteins and peptides. *Biochemistry* **36**, 10923-10929(1997).
12. Rohl, C.A., Chakrabarty, A. & Baldwin, R.L. Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci* **5**, 2623-2637(1996).
13. Brown, D.C. & Collins, K.D. Dihydroorotase from *Escherichia coli*. Substitution of Co(II) for the active site Zn(II). *J. Biol. Chem* **266**, 1597-1604(1991).
14. May, S.W. & Kuo, J.Y. Preparation and properties of cobalt(II) rubredoxin. *Biochemistry* **17**, 3333-3338(1978).
15. Ragunathan, K.G. & Bharadwaj, P.K. Nickel(II) complexes with tripodal ligands: synthesis, X-ray structural and spectroscopic studies. *J. Chem. Soc., Dalton Trans.* 2417-2422(1992).
16. Rosenfield, S.G. et al. New octahedral thiolato complexes of divalent nickel: syntheses, structures, and properties of (Et₄N)[Ni(SC₅H₄N)₃] and (Ph₄P)[Ni(SC₄H₃N₂)₃].CH₃CN. *Inorganic Chemistry* **26**, 2792-2797(1987).
17. Golynskiy, M.V. et al. Metal binding studies and EPR spectroscopy of the manganese transport regulator MntR. *Biochemistry* **45**, 15359-15372(2006).

18. Kwan, C.Y. & Putney, J.W. Uptake and intracellular sequestration of divalent cations in resting and methacholine-stimulated mouse lacrimal acinar cells. Dissociation by Sr²⁺ and Ba²⁺ of agonist-stimulated divalent cation entry from the refilling of the agonist-sensitive intracellular pool. *J. Biol. Chem* **265**, 678-684(1990).
19. Grynkiewicz, G., Poenie, M. & Tsien, R.Y. A new generation of Ca²⁺ indicators with greatly improved fluorescence properties. *J. Biol. Chem* **260**, 3440-3450(1985).
20. Scott, R.A. Measurement of metal-ligand distances by EXAFS. *Methods Enzymol* **117**, 414-459(1985).
21. Colpas, G.J. et al. X-ray spectroscopic studies of nickel complexes, with application to the structure of nickel sites in hydrogenases. *Inorganic Chemistry* **30**, 920-928(1991).
22. Padden, K.M. et al. Immobilized metal complexes in porous organic hosts: Development of a material for the selective and reversible binding of nitric oxide. *J. Am. Chem. Soc* **123**, 1072-1079(2001).
23. Ho, S.N. et al. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59(1989).
24. Melnik, M. et al. *Reviews in Inorganic Chemistry* **15**, 139-144(1995).
25. Bennett, B. *EPR of Cobalt-Substituted Zinc Enzymes. In Metals in Biology: Applications of High Resolution EPR to Metalloenzymes.* **29**, (Springer: New York, 2007).

26. Crawford, P.A. et al. Spectroscopic studies on cobalt(II)-substituted metallo-beta-lactamase ImiS from *Aeromonas veronii* bv. sobria. *Biochemistry* **44**, 5168-5176(2005).
27. Wuerges, J. et al. Crystal structure of nickel-containing superoxide dismutase reveals another type of active site. *Proc. Natl. Acad. Sci. U.S.A* **101**, 8569-8574(2004).
28. Barondeau, D.P. et al. Nickel superoxide dismutase structure and mechanism. *Biochemistry* **43**, 8038-8047(2004).
29. Chivers, P.T. & Sauer, R.T. Regulation of high affinity nickel uptake in bacteria. Ni²⁺-Dependent interaction of NikR with wild-type and mutant operator sites. *J. Biol. Chem* **275**, 19735-19741(2000).
30. Glennon, J.D. & Sarkar, B. Nickel(II) transport in human blood serum. Studies of nickel(II) binding to human albumin and to native-sequence peptide, and ternary-complex formation with L-histidine. *Biochem. J* **203**, 15-23(1982).
31. Outten, C.E. et al. Characterization of the metal receptor sites in *Escherichia coli* Zur, an ultrasensitive zinc(II) metalloregulatory protein. *Biochemistry* **40**, 10417-10423(2001).
32. Brennan, B.A. et al. Nitrile Hydratase from *Rhodococcus rhodochrous* J1 Contains a Non-Corrin Cobalt Ion with Two Sulfur Ligands. *Journal of the American Chemical Society* **118**, 9194-9195(1996).

33. Fun, H.K. et al. A neutral cobalt (II) complex of a Schiff base ligand containing mixed O₂S₂N₂ donors. *Acta Crystallographica Section C: Crystal Structure Communications* **55**, 896-899(1999).
34. Yu, P. et al. Tetraethylammonium Tris (pyridine-2-thiolato-S, N) cobalt (II),(Et₄N)[Co (2-S-C₅H₄N) ₃]. *Acta Crystallographica Section C: Crystal Structure Communications* **52**, 630-632(1996).
35. Heini, U. et al. Monomere und polymere Dimethylaminothioquadratato-Komplexe: Die Kristallstrukturen von Nickel(II)-, Cobalt(II)-, Silber(I)-, Platin(II)-, Gold(I)-, Quecksilber(II)- und Blei(II)-Dimethylaminothioquadrataten. *Zeitschrift für anorganische und allgemeine Chemie* **628**, 770-778(2002).
36. Chivers, P.T. & Sauer, R.T. NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA. *Chem. Biol* **9**, 1141-1148(2002).
37. Carrington, P.E. et al. Nickel coordination is regulated by the DNA-bound state of NikR. *Nat. Struct. Biol* **10**, 126-130(2003).
38. Schreiter, E.R. et al. Crystal structure of the nickel-responsive transcription factor NikR. *Nat. Struct. Biol* **10**, 794-799(2003).
39. Rowe, J.L., Starnes, G.L. & Chivers, P.T. Complex transcriptional control links NikABCDE-dependent nickel transport with hydrogenase expression in *Escherichia coli*. *J Bacteriol* **187**, 6317-23(2005).

40. Bloom, S.L. & Zamble, D.B. Metal-selective DNA-binding response of *Escherichia coli* NikR. *Biochemistry* **43**, 10029-10038(2004).
41. Gilbert, W. & Müller-Hill, B. ISOLATION OF THE LAC REPRESSOR. *Proc. Natl. Acad. Sci. U.S.A* **56**, 1891-1898(1966).
42. Gunsalus, R.P., Miguel, A.G. & Gunsalus, G.L. Intracellular Trp repressor levels in *Escherichia coli*. *J. Bacteriol* **167**, 272-278(1986).
43. Kolodrubetz, D. & Schleif, R. Identification of araC protein and two-dimensional gels, its in vivo instability and normal level. *J. Mol. Biol* **149**, 133-139(1981).
44. Wilkins, R.G. *Kinetics and mechanism of reactions of transition metal complexes*. (VCH Weinheim: New York, 1991).
45. Farkas, E., Sovago, I. & Gergely, A. Studies on transition-metal-peptide complexes. Part 8. Parent and mixed-ligand complexes of histidine-containing dipeptides. *J. Chem. Soc., Dalton Trans.* 1545-1551(1983).
46. Sóvágó, I. & Osz, K. Metal ion selectivity of oligopeptides. *Dalton Trans* 3841-3854(2006).doi:10.1039/b607515k
47. 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†. *Accounts of Chemical Research* **30**, 123-130(1997).
48. Ferreira, G.C. et al. Unraveling the substrate-metal binding site of ferrochelatase: an X-ray absorption spectroscopic study. *Biochemistry* **41**, 4809-4818(2002).

49. Eicken, C. et al. A metal-ligand-mediated intersubunit allosteric switch in related SmtB/ArsR zinc sensor proteins. *J. Mol. Biol* **333**, 683-695(2003).
50. Pennella, M.A., Arunkumar, A.I. & Giedroc, D.P. Individual metal ligands play distinct functional roles in the zinc sensor *Staphylococcus aureus* CzrA. *J. Mol. Biol* **356**, 1124-1136(2006).
51. Larkin, M.A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948(2007).
52. Pohl, E., Holmes, R.K. & Hol, W.G. Motion of the DNA-binding domain with respect to the core of the diphtheria toxin repressor (DtxR) revealed in the crystal structures of apo- and holo-DtxR. *J. Biol. Chem* **273**, 22420-22427(1998).
53. Cavet, J.S. et al. A nickel-cobalt-sensing ArsR-SmtB family repressor. Contributions of cytosol and effector binding sites to metal selectivity. *J. Biol. Chem* **277**, 38441-38448(2002).
54. Campbell, D.R. et al. Mycobacterial cells have dual nickel-cobalt sensors: sequence relationships and metal sites of metal-responsive repressors are not congruent. *J. Biol. Chem* **282**, 32298-32310(2007).
55. Changela, A. et al. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* **301**, 1383-1387(2003).
56. Smaldone, G.T. & Helmann, J.D. CsoR regulates the copper efflux operon *copZA* in *Bacillus subtilis*. *Microbiology* **153**, 4123-8(2007).

57. Herring, C.D. & Blattner, F.R. Global transcriptional effects of a suppressor tRNA and the inactivation of the regulator frmR. *J Bacteriol* **186**, 6714-20(2004).
58. Makarova, K.S. et al. Deinococcus geothermalis: the pool of extreme radiation resistance genes shrinks. *PLoS ONE* **2**, e955(2007).
59. Chung, C.T., Niemela, S.L. & Miller, R.H. One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U.S.A* **86**, 2172-2175(1989).
60. Choe, M. & Reznikoff, W.S. Anaerobically expressed Escherichia coli genes identified by operon fusion techniques. *J. Bacteriol* **173**, 6139-6146(1991).
61. Hendrickson, W.A., Horton, J.R. & LeMaster, D.M. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J* **9**, 1665-1672(1990).
62. Creighton, T.E. *Protein Structure: A Practical Approach*. (Oxford University Press, USA: 1997).
63. Riddles, P.W., Blakeley, R.L. & Zerner, B. Reassessment of Ellman's reagent. *Meth. Enzymol* **91**, 49-60(1983).
64. Siemann, S. et al. IMP-1 metallo-beta-lactamase: effect of chelators and assessment of metal requirement by electrospray mass spectrometry. *Biochim. Biophys. Acta* **1571**, 190-200(2002).

65. Laue, T.M. et al. Analytical Ultracentrifugation in Biochemistry and Polymer Science. Harding SE, Rowe AJ, Horton J C., editors. Cambridge, UK: R. Soc. Chem 90–125(1992).
66. Vistica, J. et al. Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Anal Biochem* **326**, 234-56(2004).
67. Santoro, M.M. & Bolen, D.W. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **27**, 8063-8068(1988).

Chapter 4

DNA Recognition and Wrapping by RcnR

This chapter is modified from a manuscript that is *in revision* for publication in the Journal of Molecular Biology.

Summary

The novel, all α -helical fold of RcnR/CsoR protein family lacks a well-characterized DNA-binding motif, necessitating a thorough investigation of the mechanisms of DNA binding by this protein family. In this chapter the biophysical properties of the RcnR-DNA interaction were probed using isothermal titration calorimetry (ITC) and footprinting techniques. One RcnR tetramer recognizes a TACT-G₆-N-AGTA motif, of which there are two in the *rcnA-rcnR* intergenic region. G-tracts are found in many predicted binding sites of other RcnR/CsoR proteins and here we show that they endow A-form DNA characteristics to the RcnR operator sites. Interestingly, RcnR also interacts non-specifically with the ~50 bp flanking the core binding site resulting in DNA wrapping and the introduction of a single negative supercoil into plasmid DNA. Comparisons to other RcnR/CsoR proteins reveal likely key differences in DNA binding amongst members of this family that result from variations in the number and sequence of operator sites.

Introduction

Bacteria routinely encounter environmental fluctuations that necessitate metabolic adaptations mediated by changes in gene expression. These events are commonly organized by transcription factors, which must couple intra- or extracellular inputs to site-specific DNA binding. To coordinate these functions transcription factors often use distinct, allosterically coupled domains. Current estimates predict upward of 50 different structural families of prokaryotic transcription factors¹. While the effector binding domains are often quite diverse, these families share a small subset of DNA-binding domains (winged helix, helix-turn-helix and ribbon-helix-helix), with well-characterized interactions with DNA.

Studies from David Giedroc's lab and work in Chapter 3 have shown that RcnR/CsoR proteins can directly respond to Cu(I)^{2,3} or Ni(II)/Co(II)^{4,5}, while others are predicted to sense formaldehyde^{4,6}, oxidative stress⁴ or other metal ions⁴. Unlike most microbial transcriptional regulators, RcnR/CsoR proteins do not possess one of the three common DNA-binding motifs listed above. A crystal structure of dimeric *Mycobacterium tuberculosis* CsoR (*MtCsoR*) revealed that RcnR/CsoR proteins fold into single-domain, antiparallel four-helix bundles ($\alpha 1$ and $\alpha 2$ of each monomer of the dimer) with two shorter $\alpha 3$ helices packed against one face of the bundle². *MtCsoR* contains two trigonal planar Cu(I) sites found at either end of the bundle², and the metal-binding site for other family members is likely at a similar location⁴. RcnR⁴ (which refers to the *E. coli* protein unless noted) and *Bacillus subtilis* CsoR (*BsCsoR*)⁷ are tetrameric, forming a

dimer of dimers, possibly through interaction of the $\alpha 3$ helices forming a donut-like structure², but this has not been experimentally confirmed. This unusual DNA-binding fold contains several conserved positively charged amino acids in $\alpha 1$ and $\alpha 2$, and two of these residues were shown to be necessary for high affinity DNA binding of *MtCsoR*².

While there has been substantial focus on understanding the evolution of signal sensing in RcnR/CsoR proteins, comparatively little is known about the DNA binding of this protein fold. Here, studies from Chapter 2 are extended using biophysical and biochemical techniques to probe the mechanism of DNA interaction. This chapter demonstrates that RcnR specifically recognizes a pair of TACT-G₆-N-AGTA operator sites likely through minor groove contacts, and nonspecifically binds to flanking DNA to promote DNA wrapping. The intervening G-tracts provide unique structural features that are important for protein binding. These studies provide key insights into how RcnR/CsoR proteins recognize their DNA-binding sites, and suggest that the protein-DNA complexes may differ significantly with respect to number of binding sites and DNA wrapping.

Results

DNase I Footprinting Reveals an Extended Binding Site for RcnR

In Chapter 2 RcnR was shown to form a specific complex with a 160 bp fragment of P_{rcnA} ⁵ by mobility shift assay. However, this experiment did not reveal a specific DNA motif that was recognized by the protein, as there are several different sets of inverted repeats in this region that could potentially interact with RcnR. To better define the RcnR binding site, DNase I footprinting was carried out on ~300 bp duplexes end-labeled on either the plus or minus strands of P_{rcnA} . RcnR protected a core region of 41 or 42 bases on each strand (Figure 4.1a and b; plus strand: -72→-32, minus strand: -79→-38 relative to *rcnA* ATG). This region encompasses two G-tracts of six and seven guanines that are each flanked by TACT inverted repeats. Outside the core binding site, hypersensitive sites every 10 bases separated by regions of protection were found for both strands, extending ~60 bp in either direction. This unexpected pattern in the flanking regions is indicative of DNA bending or wrapping by RcnR. Due to the regular spacing of the hypersensitive sites, they should all be located on one face of the DNA duplex. For practical reasons the location of the core site was closer to the ³²P-labeled ends of the duplexes used for plus and minus strand footprinting, so that binding to the flanking region was disproportionately visible on opposite sides of the core site for the plus strand (-31→+17) and the minus strand (-80→-128). Taking this into account, RcnR interacts with an

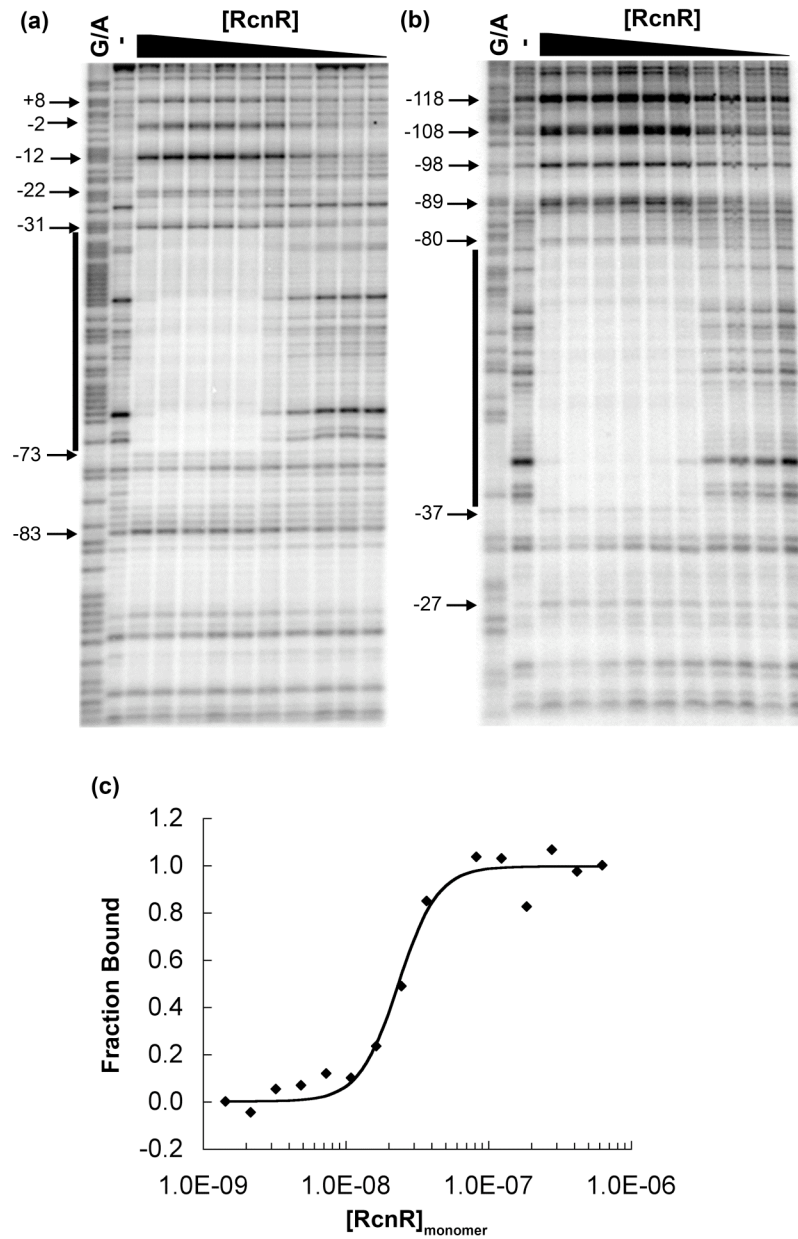


Figure 4.1. RcnR binds to an extended region of P_{rcnA} . DNase I footprinting of the plus (a) and minus (b) strands of P_{rcnA} . From left to right: G/A ladder, no RcnR, 250 to 7 nM RcnR. The core protected region (black bars) and hypersensitive sites (arrows) are indicated. Numbering is in relation to ATG start

codon of *rcnA*. (d) Quantitation of the core DNase I protection sites from separate gels with 1.5-fold dilutions were fit to the Hill equation.

extended region of ~150 bp of P_{rcnA} , which extends into the coding regions of both *rcnA* and *rcnR*.

The extension of the footprint into *rcnR* itself suggested that RcnR regulates its own expression. To test this, activity of a P_{rcnR} -*lacZ* fusion was measured with increasing Ni(II) concentrations in the presence and absence of RcnR (Figure 4.2). In the presence of RcnR, LacZ activity increases ~2-fold upon addition of 1 mM Ni(II), while constitutive, high activity is observed for a strain lacking RcnR. Although the induction is weak, RcnR clearly regulates transcription of its own gene in response to the same Ni(II) concentrations that control regulation of *rcnA*⁵.

Quantitation of the core binding site as a function of RcnR concentration and subsequent fitting using a Hill model demonstrated apparent positive cooperativity for DNA binding ($n=3.3$, Figure 4.1c), suggesting that multiple protomers bind to this DNA fragment. Although RcnR has been shown to be tetrameric between 5 to 25 μ M monomer⁴, the oligomeric state of the protein at the lower concentrations used in these experiments (7-250 nM), which are below the range accessible to sedimentation experiments, is unknown. Only an upper limit of 50 nM can be placed on the K_d of the dimer-tetramer equilibrium.

RcnR Binds to One Face of P_{rcnA}

Hydroxyl radical footprinting reveals the accessibility of deoxyribose sugars of the DNA backbone, and provides a small probe that can more precisely define DNA binding sites compared to the bulkier DNase I. For both duplex

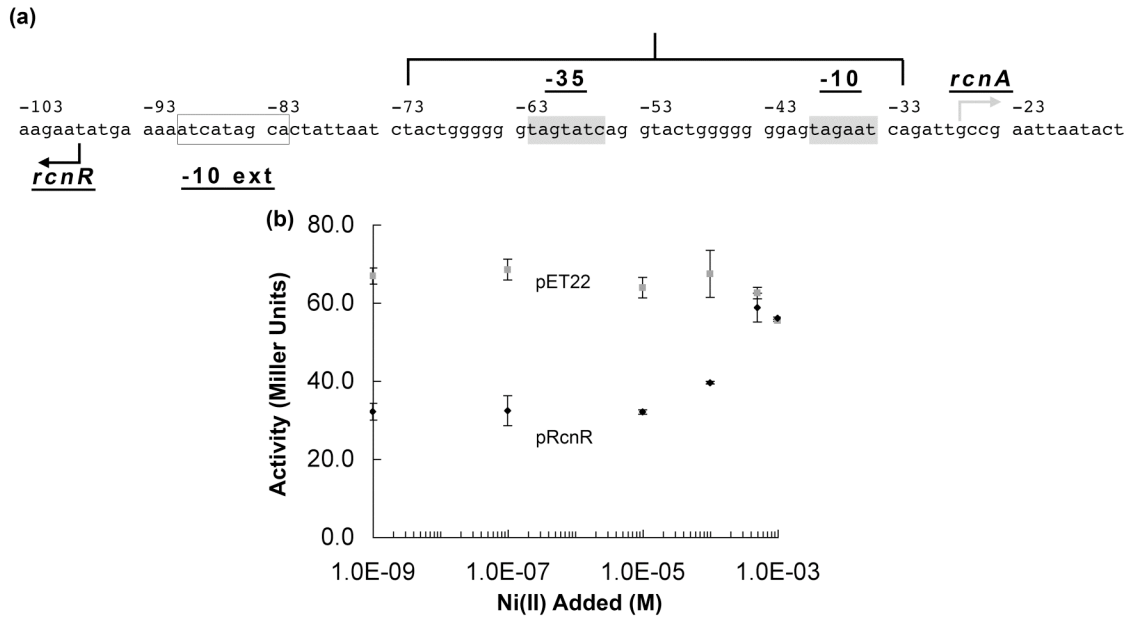


Figure 4.2. RcnR autoregulates its expression. (a) Transcription start sites were determined by primer extension (data not shown) and mapped onto the sequence with predicted -10 and -35 elements assigned based on the spacing and sequence of the consensus *E. coli* elements. The core RcnR footprint for the plus strand is denoted with a bracket. (b) P_{rcnR} -*lacZ* activity was assayed in a $\Delta rcnR$ strain with pRcnR (black diamonds) or pET22b (gray squares) as a control.

strands, RcnR binding lead to an extended footprint with protected regions of ~4 bases spaced evenly over most of the region visible on the gel (Figure 4.3a,b and c). The centers of protected regions were separated by 9 to 11 bases, demonstrating that RcnR binds to one face of the DNA duplex. Also, the protection within the core DNase I footprint region was indistinguishable from that outside the region in the hydroxyl radical footprinting experiment. Interestingly, the protected duplex face is opposite that of the DNase I hypersensitive sites. Thus, the DNA-binding surface recognized by RcnR is comprised of the minor groove of the TACT repeats and the major grooves of each G-tract.

RcnR Makes Non-specific Contacts with DNA Flanking the Core Site

The footprinting results demonstrated that RcnR-DNA contacts are made well outside of the core protection region, which suggests that RcnR binding affinity should change with DNA length. Mobility shift assays were employed with DNA fragments of increasing length, all centered on the core binding site. There was a ~4-fold increase in $K_{d\text{ app}}$ as the fragment length increased from 80 to 150 bp, while longer fragments showed minimal additional improvement in affinity (Figure 4.4). This places the optimal fragment size for RcnR binding at ≥ 150 bp, in good agreement with the size of the DNase I footprint.

DNA flanking the core binding site could contain additional RcnR binding sites or form specific or nonspecific contacts with repressors bound to the core site. To distinguish between these possibilities, a 47 bp fragment encompassing the core binding site (see Materials and Methods) was cloned into pBSII KS and

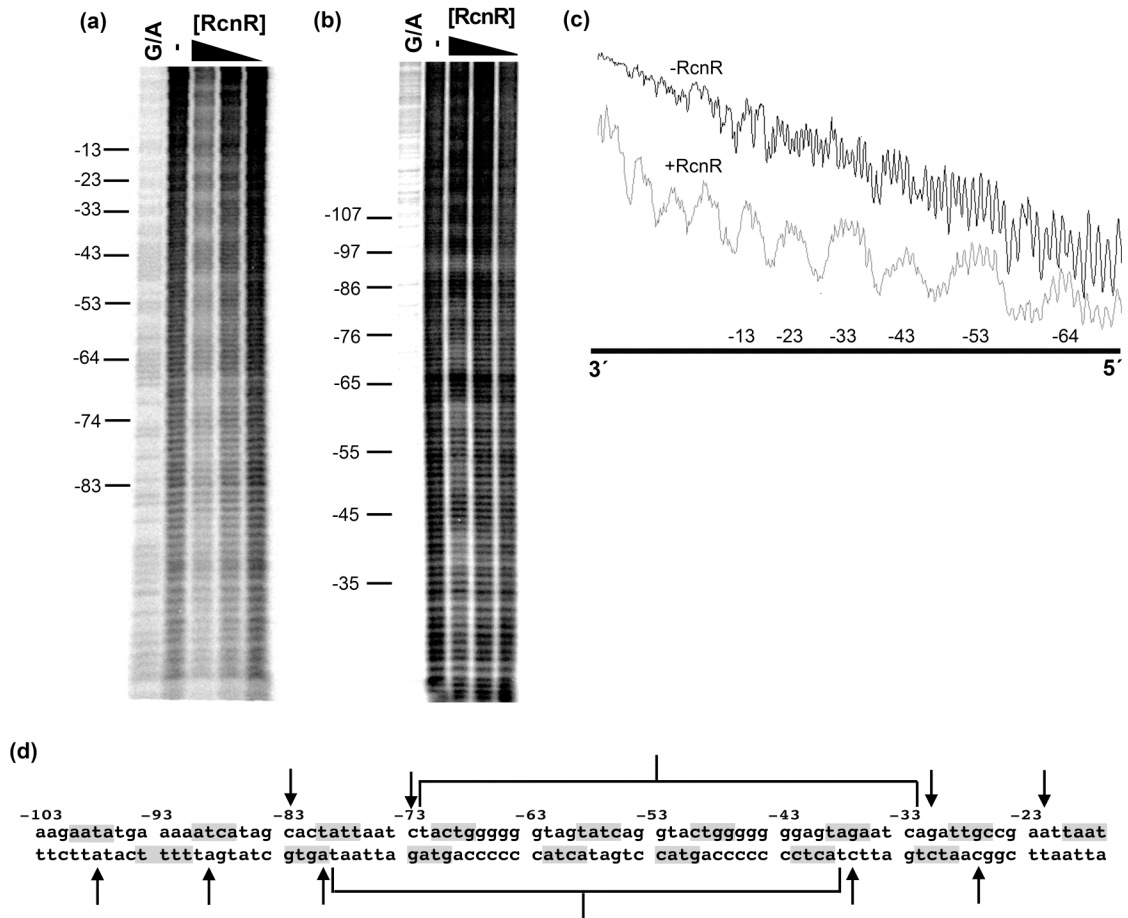


Figure 4.3. RcnR protects one face of the DNA duplex. Hydroxyl radical footprinting was conducted on the plus (a) and minus (b) strands of P_{rcnA} . From left to right, G/A ladder, no protein, 100 nM RcnR, 50 nM RcnR, 20 nM RcnR. The position relative to the ATG of *rcnA* is indicated to the left of the gel. (c) Densitometry scans of hydroxyl radical footprinting data showed an extended, and periodic protection pattern (DNA alone, black; 100 nM RcnR, gray). (d) A summary of the footprinting data for the central 80-bp portion of P_{rcnA} shows that DNase I hypersensitive sites (arrows) occurred opposite of hydroxyl radical protection (gray) for both strands. The core protection region is denoted with

brackets. Protection patterns that extended beyond this sequence are not shown.

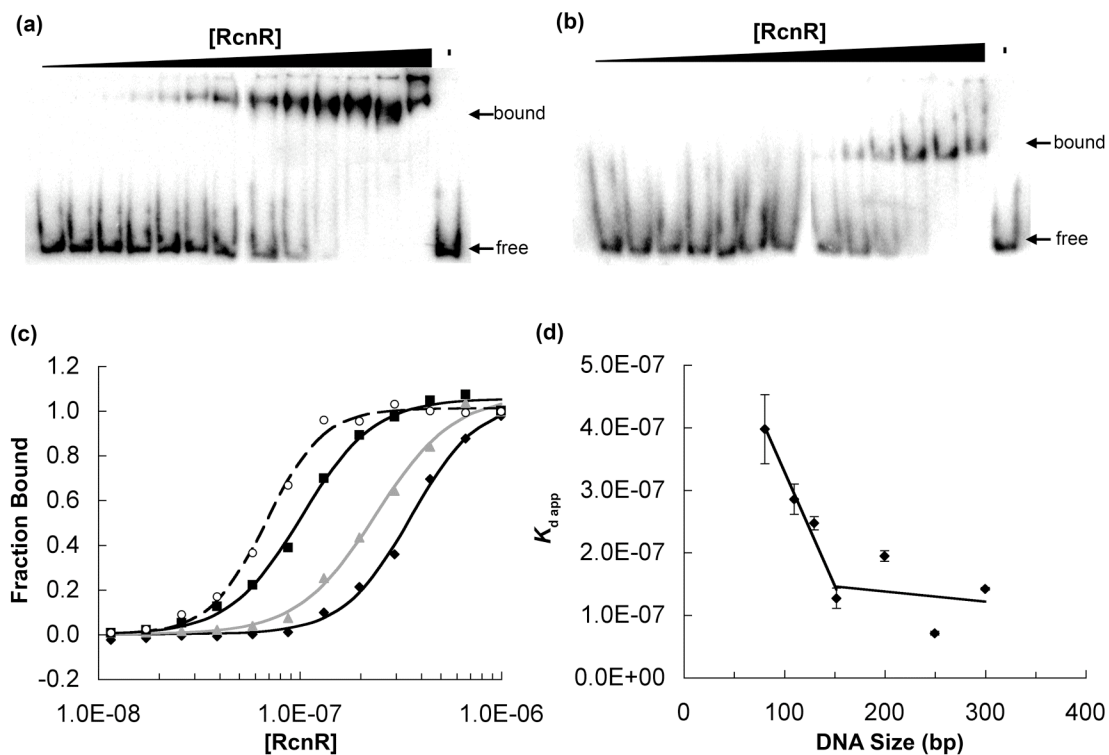


Figure 4.4. RcnR DNA affinity increases with length of flanking DNA. Mobility shift assays were conducted with DNA duplexes of different lengths, all centered on the core RcnR binding site. Titrations for 300 bp (a) and 81 bp (b) duplexes are shown. Protein concentrations ranged from 8 nM to 1 μM RcnR. Quantitation of four titrations are shown in (c) for 81 bp (black diamonds), 130 bp (gray triangles), 154 bp (black squares) and 250 bp (white circles). The average apparent dissociation constants for seven duplex lengths are plotted in (d).

primers were used to amplify a ~300 bp fragment that contained both the 47 bp core fragment and the flanking regions from the vector, which share minimal sequence identity with the wild type P_{rcnA} duplex used above. Remarkably, RcnR bound to this fragment with an affinity similar to the wild type P_{rcnA} fragment and DNase I hypersensitive sites separated by regions of protection were observed every ~10 bp, although they were slightly less intense than for the wild type fragment (Figure 4.5). Overall the cleavage patterns between the two DNA fragments are quite similar and indicate that the DNA distortion occurred for both duplexes. The small differences observed could indicate weak sequence preferences in these regions that affect RcnR-DNA interactions or local DNA flexibility.

RcnR Supercoils DNA

To gain more information about the structure of the RcnR-DNA complex and directly test for DNA wrapping as suggested by Figure 4.1, a DNA topology footprint assay was employed (see Materials and Methods). This assay, utilizing topoisomerase I, has been used to show topological changes induced by DNA-binding proteins including Orc I⁹ and enzymes such as lambda family recombinases¹⁰. A 297 bp fragment of P_{rcnA} was cloned into pBS and the resulting plasmid was preincubated with RcnR, then topoisomerase I was used to relax negative supercoils within the plasmid, and the topoisomers were resolved on a chloroquine agarose gel. The addition of RcnR to plasmid DNA containing a specific RcnR binding site shifted the topoisomer distribution to fewer positive

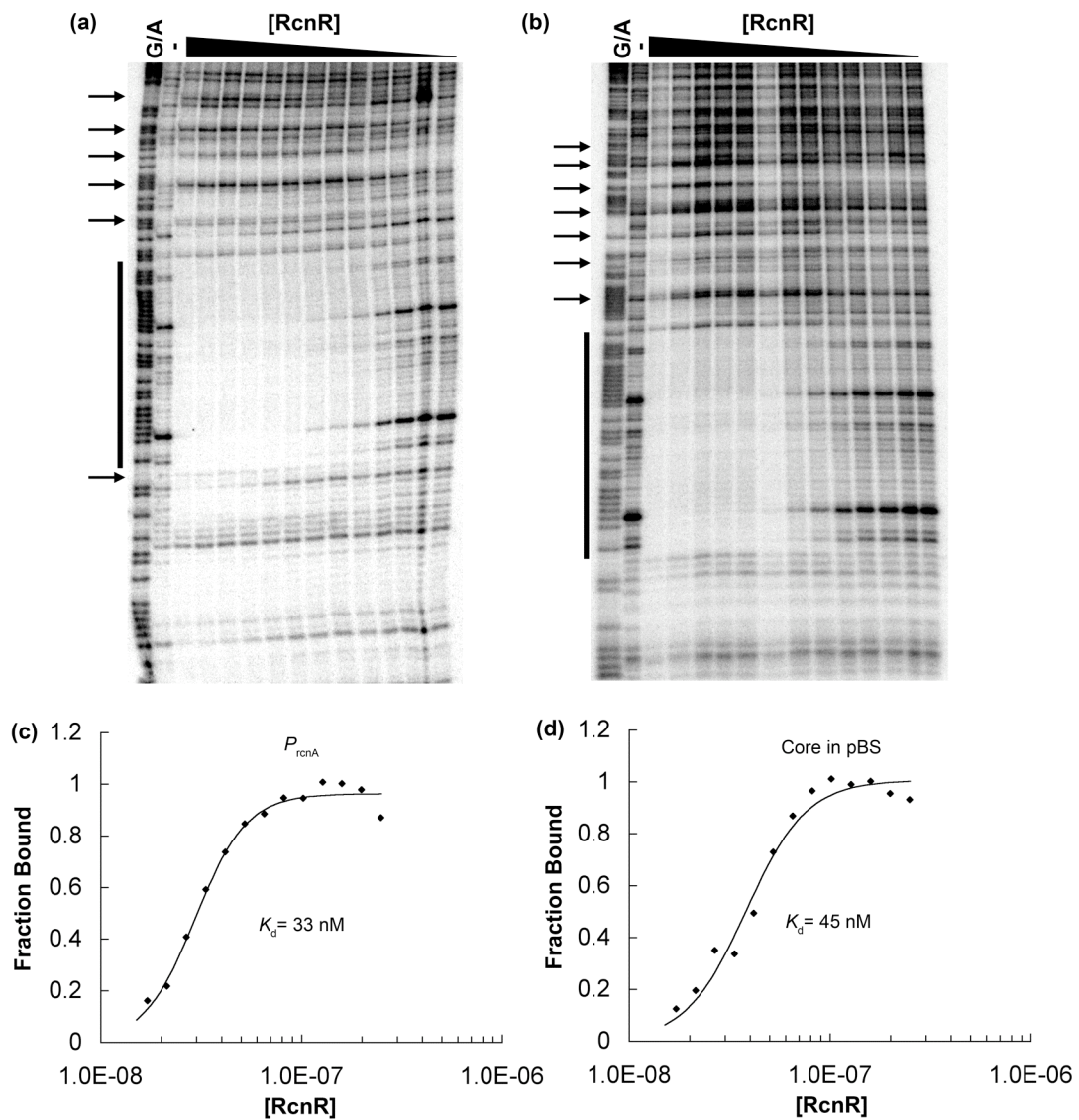


Figure 4.5. Flanking DNA sequence is not important for RcnR binding. DNase I footprinting was carried out on the plus strand of P_{rcnA} (a) and P_{rcnA} core in pBS (b) in Buffer P, and are quantitated in (c) and (d), respectively. Arrows indicate hypersensitive sites and a black bar delineates the core protection site.

supercoils (Figure 4.6). At 360 nM RcnR a change of 0.8 ± 0.1 supercoils was observed with P_{rcnA} , but not with vector DNA alone. Together with the size dependence of DNA binding affinity and extended footprinting patterns, these data are consistent with DNA wrapping around RcnR and crossing itself once, leading to the introduction of a single negative supercoil into the plasmid DNA.

A Single RcnR Tetramer Binds to Each Operator Site

The footprinting results identified a likely RcnR binding site, but did not reveal the stoichiometry of binding, nor did they precisely define the key bases necessary for the interaction. To test the importance of the TACT-G₆-N₁-AGTA motifs, isothermal titration calorimetry (ITC) was used to understand thermodynamic aspects of the RcnR-DNA interaction with short oligonucleotide duplexes. Based on the DNase I footprinting results, two 24 bp DNA fragments within the core DNase I protection region [Site 1(-76→-53) and Site 2 (-56→-33) relative to *rcnA* ATG] were first evaluated for RcnR binding (see Figure 4.3e). Preliminary studies indicated that the protein-DNA complex was poorly soluble in Buffer H, which necessitated the identification of a suitable substitute (Buffer P) that did not affect DNase I cleavage patterns, (Figure 4.5), apparent affinity, or protein oligomeric state (Figure 4.7) when compared to previous studies in Buffer H⁴ (Figure 4.1). Using ITC, titration of Site 2 into RcnR showed a signal change for an exothermic reaction, demonstrating that RcnR can indeed bind to this short DNA duplex (Figure 4.8a). Titrations with a scrambled 24 bp fragment with the same base composition as Site 2 showed no measurable heat absorbed or

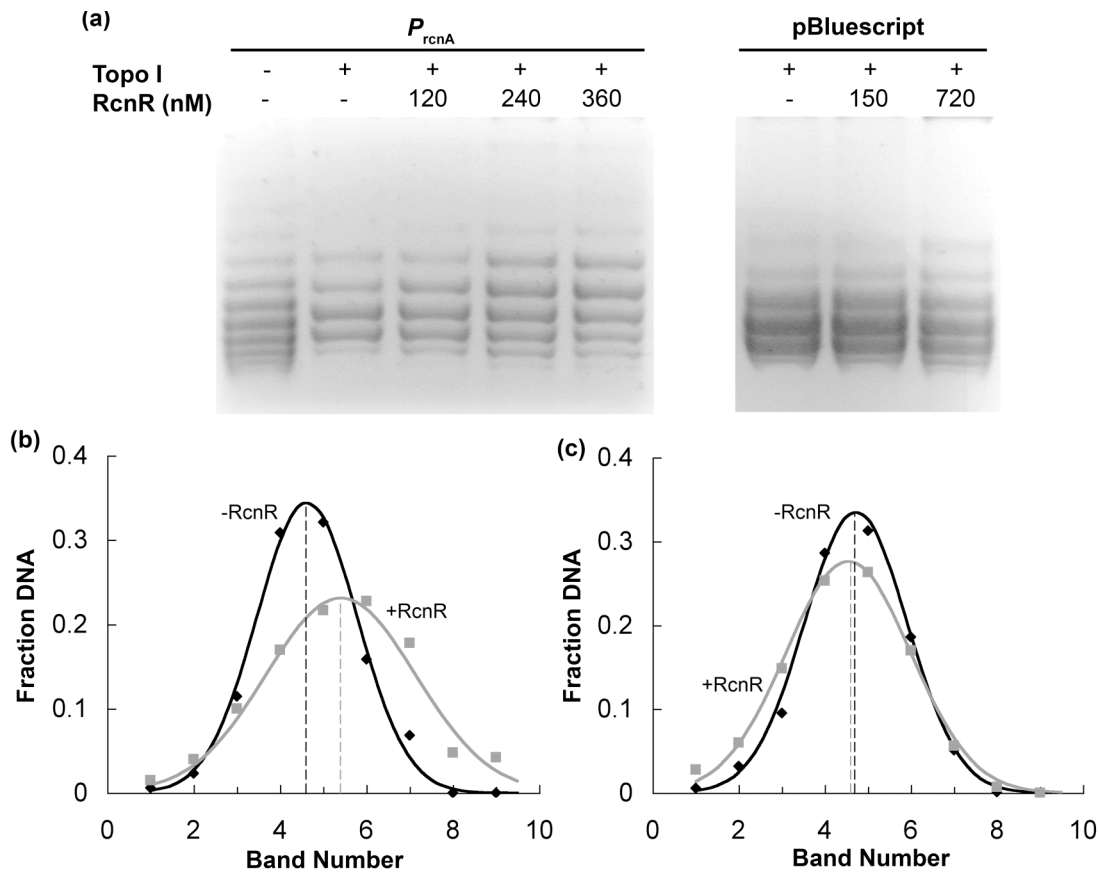


Figure 4.6. RcnR introduces a single negative supercoil into DNA. (a) RcnR binding reactions with plasmid DNA containing a 273-bp P_{rcnA} insert (P_{rcnA}) or vector alone (pBluescript) were incubated with *E. coli* topoisomerase I and run on chloroquine agarose gels to resolve each topoisomer. Relaxation times were long (4 h) so that the DNA bands represent positively supercoiled topoisomers. (b) DNA band intensities were quantitated for the 0 nM RcnR (black diamonds) and 360 nM RcnR lanes (gray squares) and fit to Gaussian curves. Dotted lines show the midpoints of each curve. The distribution shifted by -0.8 supercoils in the presence of RcnR, indicating that the protein introduced a single negative supercoil into the plasmid. (c) Experiments with the vector alone showed no

significant difference with (gray squares) or without (black diamonds) 720 nM RcnR.

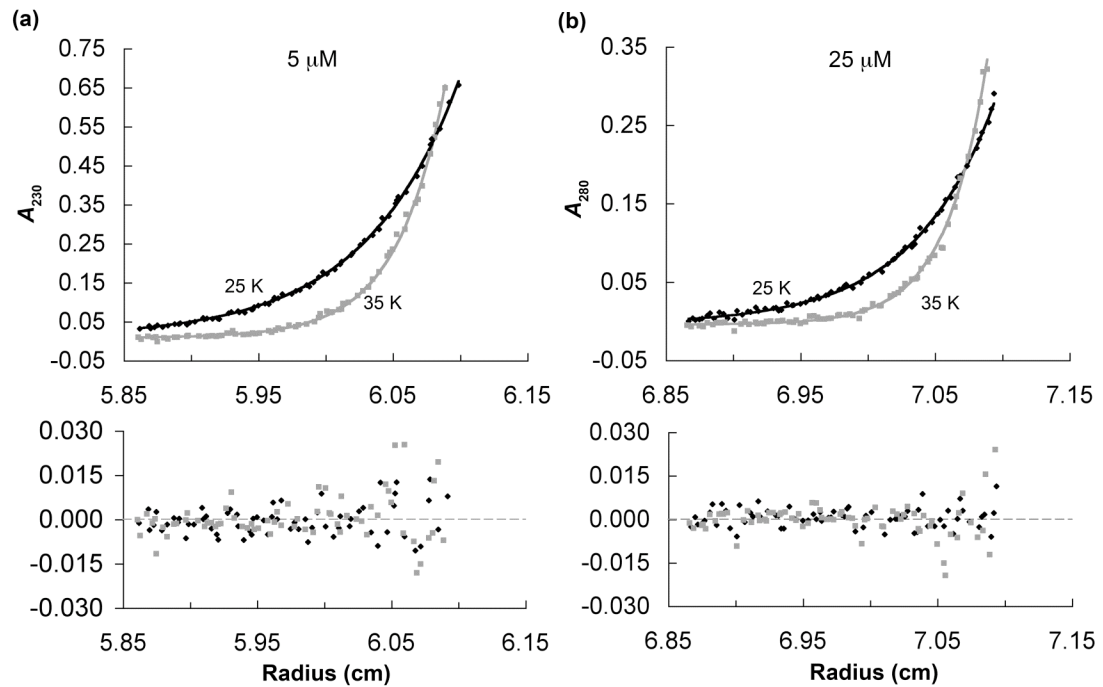


Figure 4.7. RcnR is tetrameric in Buffer P. Sedimentation equilibrium studies of 5 (a) and 25 μM (b) RcnR monomer in Buffer P at 25 000 (black diamonds) and 35 000 rpm (gray squares) at 20 °C. Absorbance was measured at 230 (5 mM) or 280 nm (25 μM), and the data were globally fit to a single species model giving a molecular weight of 38 955 Da (40 011 expected for a tetramer).

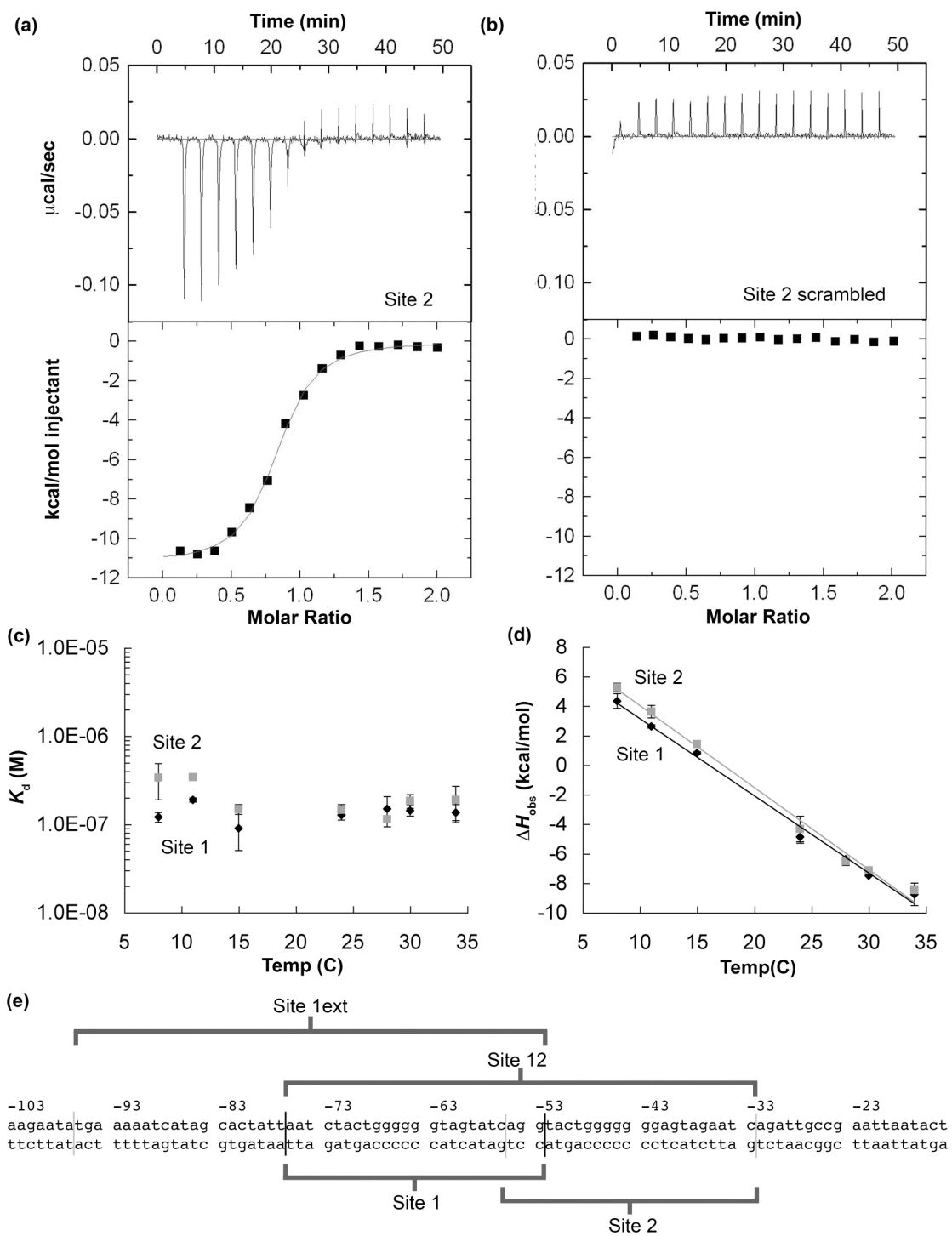


Figure 4.8. One RcnR tetramers binds to single site duplexes by ITC. (a) 80 μM Site 2 DNA was titrated into 8 μM RcnR tetramer in Buffer P at 34 $^{\circ}\text{C}$. The solid line is a fit to an n independent binding sites model with $n= 1.2$, $K_d= 245$ nM, and $\Delta H_{\text{obs}}= -8.97$ kcal/mol. (b) 100 μM of a 24 base pair duplex with the same base content of Site 2 was titrated into 10 μM RcnR tetramer under the same conditions as (a). No measurable response was observed at this temperature or at 8 $^{\circ}\text{C}$ or 30 $^{\circ}\text{C}$ (data not shown). Affinity (c) and enthalpy (d) measurements were made for Site 1 (black diamonds) and Site 2 (gray squares) from 8 $^{\circ}\text{C}$ to 34 $^{\circ}\text{C}$. The parameters for each site are similar with average parameters of $n= 1.3\pm 0.1$ and 1.2 ± 0.1 tetramers:DNA and $K_d= 126$ nM and 175 nM with $\Delta C_{p,\text{obs}}= -0.55\pm 0.03$ and -0.52 ± 0.01 , respectively. P_{rcnA} with sequence boundaries for different duplexes using in ITC experiments is shown in (e).

evolved at multiple temperatures tested, demonstrating that binding is specific for P_{rcnA} (Figure 4.8b). Each titration was fit to a model of n independent identical binding sites with fitting parameters n (stoichiometry), K_d (affinity) and ΔH_{obs} (enthalpy). RcnR bound Site 1 and Site 2 with a stoichiometry of ~ 1 tetramer per duplex (Site 1 = 1.3 ± 0.1 , Site 2 = 1.2 ± 0.1) and $\Delta C_{p,obs}$ values of -0.55 kcal/mol K and -0.52 kcal/mol K, over a temperature range from 8 °C to 34 °C. The DNA affinity did not significantly change over this range with average K_d values of 126 nM and 174 nM for Site 1 and Site 2, respectively. Therefore, RcnR interacted essentially identically with each operator site.

Two RcnR Tetramers Bind to the Full Core Binding Site

The interaction of RcnR with a 44 bp DNA duplex spanning both operator sites (Site 12, Figure 4.8e) was also examined using ITC to determine the stoichiometry of the complex and examine any potential cooperativity between bound tetramers (Figure 4.9). Analysis using a model of n independent identical binding sites revealed a stoichiometry of 1.9 ± 0.1 tetramers per DNA duplex. ΔH_{obs} values per site were more negative than that for either individual site alone, while the binding affinity was weaker. Titrations were also analyzed using a model of two sequential binding sites. However, values for ΔH_{obs} and K_d could not be precisely defined due to the increased number of variables in this model. It is clear through data fitting and simulations that the positive cooperativity seen in footprinting experiments with large DNA duplexes (~ 300 bp) does not exist for this significantly shorter fragment, and instead a slight negative cooperativity is

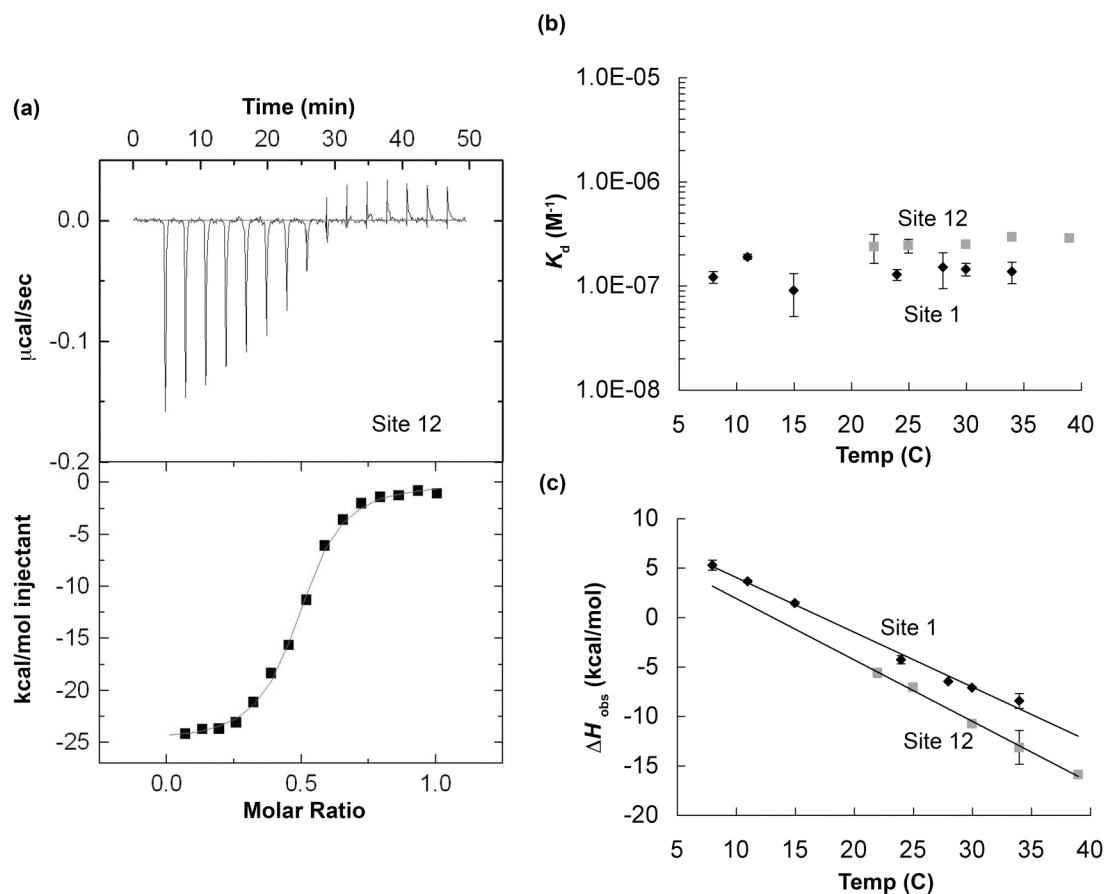


Figure 4.9. ITC with dual operator site duplex. (a) 50 μM Site 12 DNA was titrated into 10 μM RcnR tetramer in Buffer P at 34 °C. The solid line is a fit to an n independent binding sites model with $n= 2.1$, $K_d= 276$ nM, and $\Delta H_{\text{obs}}= -11.99$ kcal/mol, where the K_d and ΔH_{obs} are averages for each binding site. Affinity (b) and enthalpy (c) measurements were made from 8 °C to 39 °C (gray squares) and are shown with data for Site 1 (black diamonds). The average parameters for Site 12 are $n= 1.9\pm 0.1$ and $K_d= 255$ nM, with $\Delta C_{p,\text{obs}}= -0.62\pm 0.02$ kcal/mol K.

seen. Similar parameters were obtained during preliminary studies for binding to a 64 bp DNA fragment centered on Sites 1 and 2 ($n = 2.2$, $K_d = 2.92 \times 10^6$, $\Delta H = -13780$), although limited yields of longer oligonucleotides containing the G-tracts prohibited extensive studies of this fragment as well as larger duplexes. Binding to a 44 bp fragment that included Site 1 and the 20 bp upstream flanking region, denoted Site 1ext (-96→-53 relative to *rcnA* ATG; Figure 8e), showed a stoichiometry identical to that of Site 1 (Figure 4.10), suggesting that end-effects of the short fragments did not skew the stoichiometry measurements. Although the duplexes used in ITC were not as large as those used for footprinting, these experiments provide no evidence that more than two RcnR tetramers can bind to P_{rcnA} , instead demonstrating that one tetramer binds to each TACT-G₆N-AGTA motif.

Having established that there are two separable RcnR binding sites within P_{rcnA} (Site 1 and Site 2), we also tested the ability of RcnR to bind to a single site to distort DNA. Site 1 was cloned in pBS and the interaction of a ~300 bp duplex containing this site was tested for RcnR binding by DNase I footprinting (Figure 4.11). Interestingly, RcnR binds to this fragment with similar apparent affinity as the construct with both binding sites ($K_{app} = 5 \pm 1 \times 10^{-9}$ M and $14 \pm 2 \times 10^{-9}$ M for single and dual sites, respectively), but there was no extended footprint beyond the 22 bp protected region expected for Site 1, which is in contrast to the pattern with the dual site construct that contains hypersensitive sites outside the core

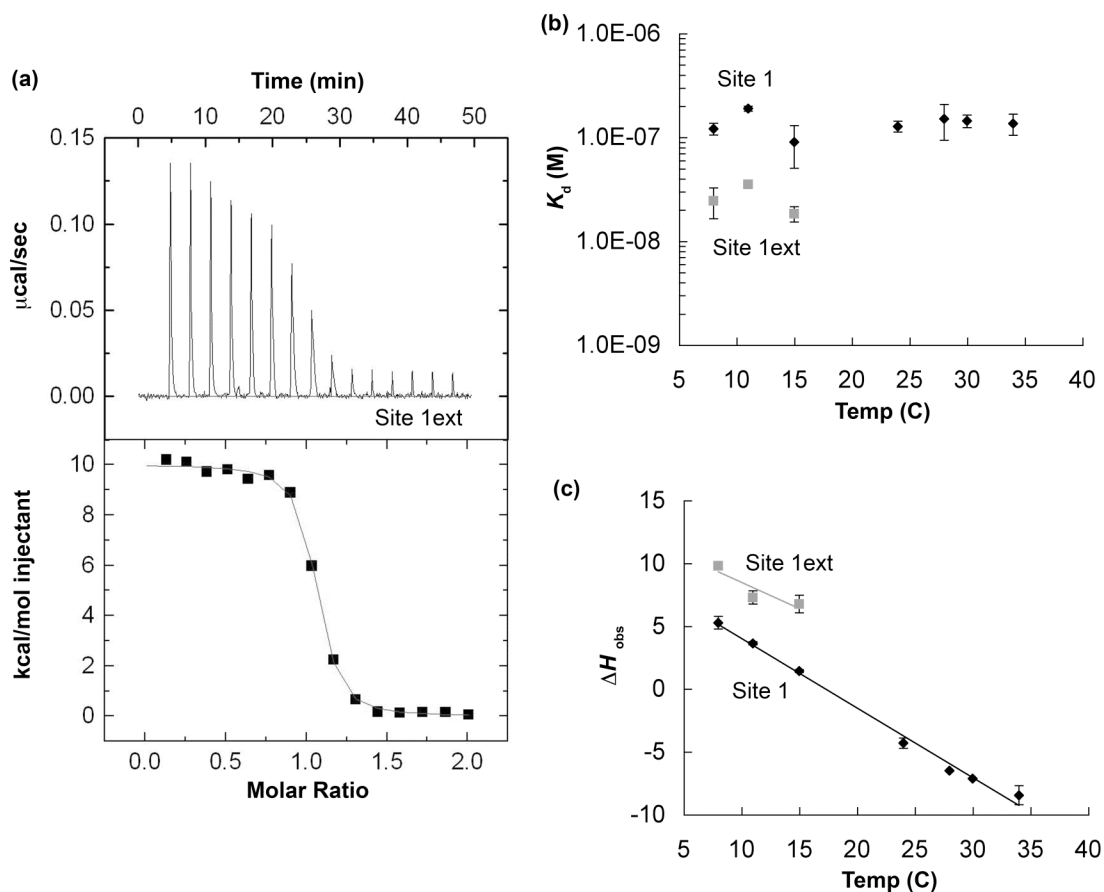


Figure 4.10. RcnR binds to an extended Site 1 with a 1:1 stoichiometry. (a) 80 μM DNA was titrated into 8 μM RcnR tetramer at 8 $^{\circ}\text{C}$. The titration was fit to an n independent binding sites model with $n= 1.0$ $K_d= 3.28 \times 10^{-8}$ $\Delta H_{\text{obs}}= 10.07$ kcal/mol . The K_d and ΔH_{obs} values from 8 to 15 $^{\circ}\text{C}$ are summarized in (b) and (c), respectively and compared to the values for Site 1.

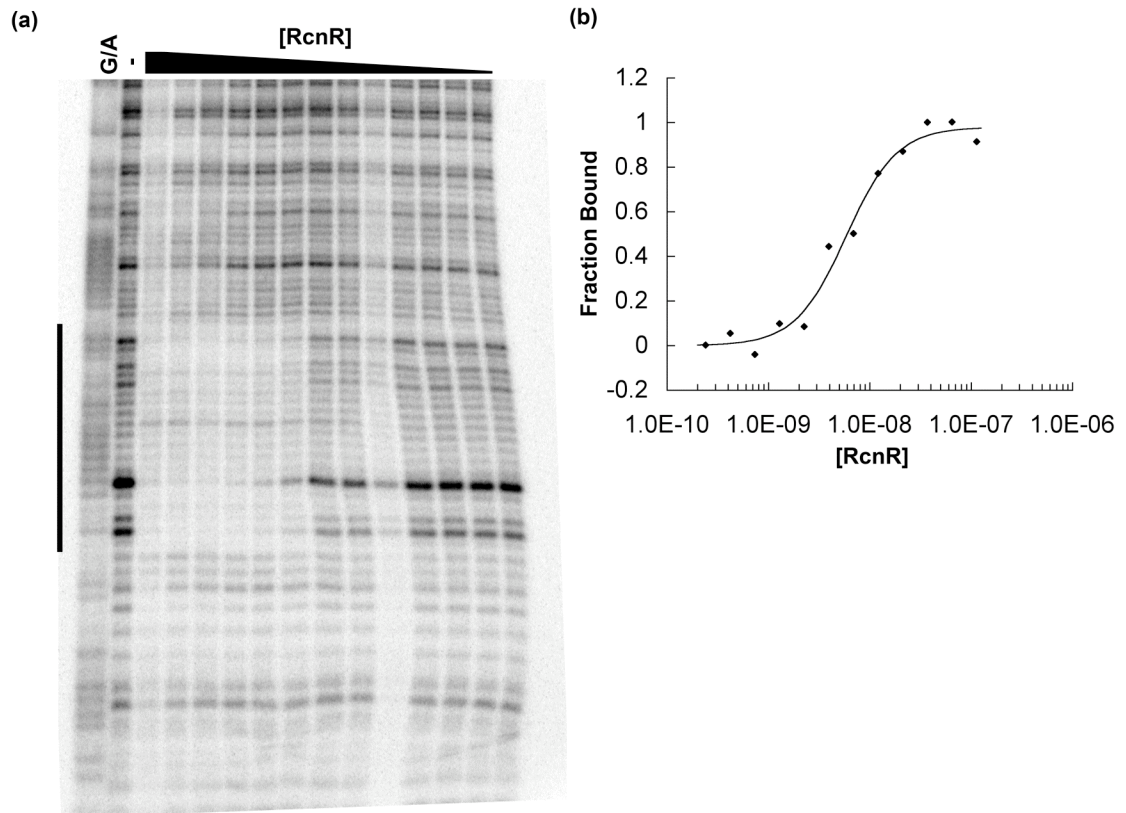


Figure 4.11. RcnR binding to a single site does not lead to DNA distortion. Site 1 was cloned into pBluescript and binding was measured as a function of RcnR concentration using DNase I footprinting. (a) Lane 1- G/A ladder, Lane 2- no protein, Lane 3 to 15- 1.25 fold protein dilutions, beginning at 200 nM protein. Binding does not result in an extended footprint with hypersensitive sites as seen when both sites are present. The protected region of 22 bases is denoted with a black bar. (b) Data points were fit to a Hill model with $K_{d\text{ app}} = 4.9 \pm 0.1$ nM.

protected region. Thus, the combination of Sites 1 and 2 is important for observing the extended footprint and DNA wrapping.

TACT Inverted Repeats Are Critical For High-Affinity DNA Binding by RcnR

The role of the T₁A₂C₃T₄ inverted repeats in RcnR binding was tested specifically by mutation of pairs of equivalent positions in the inverted repeat in Site 1 and monitoring protein binding by ITC (see Materials and Methods for constructs). Binding to each sequence variant duplex was examined at three different temperatures (Figure 4.12). All four sets of mutations had striking effects on RcnR binding, leading to decreases in affinity and increases in ΔH_{obs} over the full temperature range tested (8 to 30 °C). The most drastic shifts occurred when T₁ or C₃ were mutated, leading to an increase in $\Delta C_{p,\text{obs}}$ to 0.1±0.1 and 0.8±0.7 kcal/mol K and 32-fold and 25-fold decreases in affinity at 24°C, respectively. The A₂ and T₄ mutations had weaker, but significant decreases in affinity of 9-fold and 15-fold, respectively. This set of experiments indicates that the TACT inverted repeats are vital for high-affinity DNA binding by RcnR.

P_{rcnA} Operator Sites Have A-form DNA Characteristics

Experimental and computational studies have shown that GG basepair steps impart A-form DNA characteristics to DNA duplexes in solution^{11,12}. Differences between the characteristics of A- and B-form DNA include sugar pucker, major and minor groove dimensions, as well as hydration properties¹³, which all result in different available surfaces for protein binding. CD

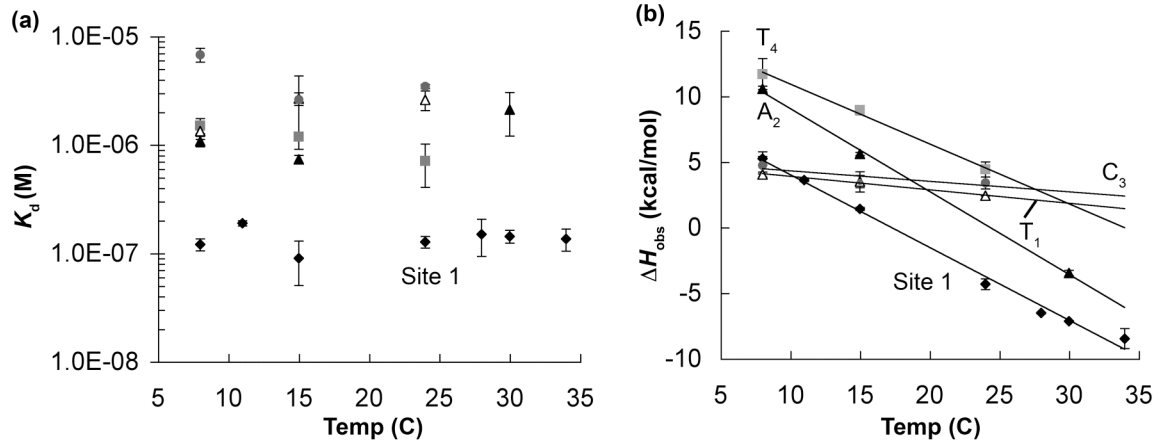


Figure 4.12. TACT inverted repeats are necessary for high-affinity DNA binding. 100 μ M Site 1 altered sequence DNA duplexes were titrated into 10 μ M RcnR tetramer in Buffer P between 8 °C to 34 °C. Affinity (a) and enthalpy (b) values determined from ITC titrations are shown for each mutant (T₁-white triangles, A₂-black triangles, C₃-gray circles, T₄-gray squares) relative to Site 1 (black diamonds).

spectroscopy has been used to probe DNA structure, and can distinguish between A- and B-DNA¹⁴. While a typical B-DNA spectrum shows a maximum from 275-285 nm, and a transition to negative ellipticity at ~270 nm, A-DNA spectra show a more intense maximum at ~270 nm, and a crossover to negative ellipticity between 240-250 nm^{11,14,15}. Comparison of the spectrum of the Site 2 duplex with that of a scrambled 24-bp duplex of the same base content but predicted to have B-DNA propensity showed differences consistent with Site 2 having A-DNA features (Figure 4.13a). The Site 2 spectrum had a maximum at 270 nm that was more intense than the scrambled sequence maximum at 278.5 nm as well as a crossover point shifted from 261.4 nm to 250.1 nm. The spectrum of Site 1 was very similar to that of Site 2, and changing the G-tract to an A-tract in the Site 1 background resulted in a spectrum more similar to that of the scrambled sequence, indicating that the G-tract is indeed responsible for inducing A-DNA features to these duplexes (Figure 4.14a). The G-tract to A-tract mutation also drastically inhibited RcnR binding as measured by ITC (Figure 4.14b, $K_d > 1 \times 10^{-5} \text{M}$).

B-form DNA is known to shift to A-form at high trifluoroethanol (TFE) concentrations¹⁴. DNA sequences with a propensity to adopt A-form structure in aqueous solution are expected to transition completely to A-form at lower TFE concentrations. To test if this is the case for P_{RcnA} Site 2, ϵ_{270} was monitored as a function of TFE concentration. As shown in Figure 4.13b, Site 2 had a much broader transition, with a midpoint at 68% TFE, compared to 72% TFE for the

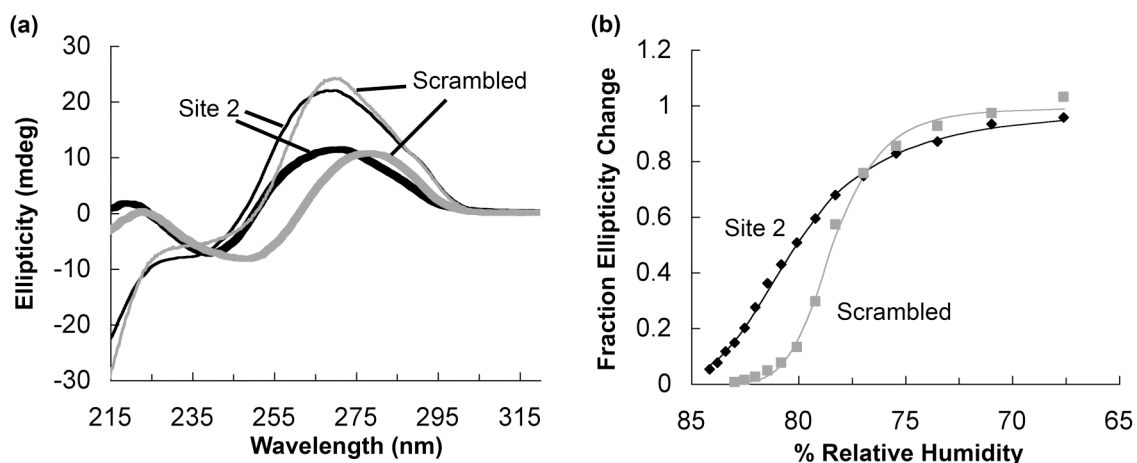


Figure 4.13. P_{rcnA} displays A-DNA properties. (a) Circular dichroism spectra of Site 2 were obtained in 0.5 mM Tris (pH 7.5) 1 mM NaCl 0.05 mM EDTA (thick black) or 84% TFE (thin black) at 22 °C. Reference spectra for a randomized duplex of the same size with B-form properties were also obtained in buffer (thick gray) or 84% TFE (thin gray) for comparison. (b) TFE titrations were obtained for both Site 2 (black diamonds) and random duplex (gray squares) by following the ellipticity at 270 nm.

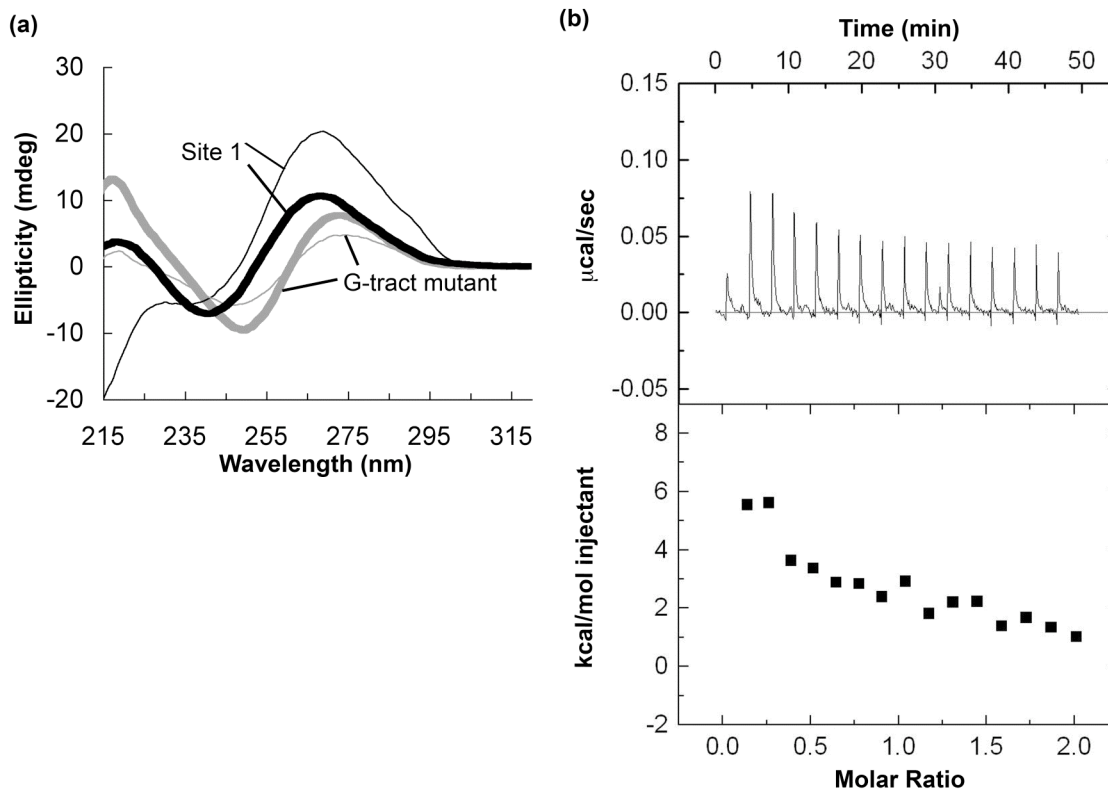


Figure 4.14. G-tracts impart A-form properties to P_{rcnA} and are required for RcnR binding. (a) Circular dichroism spectra of Site 1 were obtained in 0.5 mM Tris (pH 7.5) 1 mM NaCl 0.05 M EDTA (thick black) or 84 % TFE (thin black). Spectra for a duplex with mutation of the six Gs to As (JI371,r; Materials and Methods) are shown in buffer (thick gray) and in 84% TFE (thin gray). Site 1 showed features of A-form DNA in buffer alone and transitioned completely to A-form in 84% TFE, while the mutant duplex did not. (b) ITC titration of 100 μ M JI371,r duplex into 10 μ M RcnR tetramer showed minimal interaction at 34 $^{\circ}$ C.

random sequence, consistent with Site 2 having more A-form qualities. Broad transitions such as that seen for Site 2 are found for sequences known to form both A-form and B-form DNA in the same duplex¹⁶, such as the G-tracts (A-form) and flanking sequences (B-form) of P_{rcnA} .

Discussion

This study provides a detailed examination of the interaction between RcnR and its cognate DNA operator sites and provides the first insight into how the DNA-binding fold of RcnR/CsoR proteins recognizes DNA. RcnR specifically binds to the *rcnR-rcnA* intergenic region, resulting in the repression of both *rcnR* and *rcnA* at low Ni(II) levels. This autoregulation creates a negative feedback loop, allowing *E. coli* to more rapidly respond to fluctuations in intracellular metal levels. A model for the interaction between RcnR and P_{rcnA} must account for both the specific binding to a pair of TACT-G₆-N-AGTA motifs as well as the surprising, nonspecific interactions with the flanking DNA that lead to DNA wrapping. A search of regions upstream of known *E. coli* genes did not identify any other identical sequences, suggesting that RcnR only regulates expression of *rcnR* and *rcnA*. There are two similar sequences, TGCTCAGGGGGAGTA, upstream of *yfaX*, encoding a putative transcription factor, and TTATTGGGGGAAGTA upstream of *insH*, encoding a transposase, but it is not obvious how these could be linked to nickel physiology, or if RcnR could interact with these sites given the number of base substitutions.

ITC experiments revealed that two RcnR tetramers bind to the 44 bp core DNA binding site. Due to low DNA synthesis yields for longer oligonucleotides containing both G-tracts, the longest duplex that titrations could be conducted with was 64 bp, preventing thermodynamic analysis of the full (~150 bp) binding site and the contribution of wrapping to binding affinity. However, the 64 bp

duplex still showed a 2:1 stoichiometry, and binding to the 44 bp Site 1ext showed the same 1:1 stoichiometry as the 24 bp Site 1 alone. Therefore, all of the experimental data are consistent with two RcnR tetramers binding to P_{rcnA} , although the possibility of additional tetramers binding to large, wrapped DNA duplexes cannot be formally excluded. The positive cooperativity observed by footprinting is not observed for the shorter duplexes assayed by ITC. However, footprinting experiments require protein concentrations several orders of magnitude lower than those accessible by ultracentrifugation experiments. Therefore RcnR dimers may be present in these experiments, which could lead to positively cooperative binding as the dimer-dimer interface is formed. A second possibility is that tetramer binding and cooperativity or stoichiometry change with DNA duplex size.

Two sequence features were found to be important for high-affinity DNA binding by RcnR. The TACT inverted repeats appear likely to form sequence-specific interactions with RcnR. Hydroxyl radical footprinting showed that RcnR binds to the face of the DNA duplex on which the minor groove features of the inverted repeats are accessible. RcnR is an all-helical protein with poorly conserved loop sequences, which suggests that an α -helix of the protein binds in the DNA minor groove. The spacing of these half-sites (1 turn of A/B form DNA) is similar to the spacing of helical side chain residues in helix $\alpha 1$ (~34 Å) in two different monomers on the same face of the putative RcnR tetramer, which is consistent with the stoichiometry of the RcnR interaction with Site 1 and Site 2.

The minor groove offers fewer possibilities for hydrogen bond donor/acceptor interactions¹⁷ so that additional sequence or structural information may be present in the RcnR binding sites to facilitate specific, i.e. high-affinity, binding.

The G-tracts between the TACT inverted repeats likely provide a structural component to facilitate the RcnR-DNA interaction. Although the major grooves of the G-tracts of Site 1 and Site 2 are both accessible on the face of the DNA bound by RcnR, no protection of the N₇ position has been observed by DMS footprinting (Figure 4.15), indicating that RcnR is not making specific contacts with the G bases in the major groove, or at least not at this specific nitrogen. Our studies demonstrated that the G-tracts of the RcnR operator sites convey A-form-like properties to P_{rcnA} , suggesting that these sequences could provide a structural basis for DNA recognition. A-form DNA has dramatically different structural features compared to B-form DNA, including a C_{3'}-endo sugar conformation, a deep and narrow major groove, a wide, shallow minor groove and bases that are splayed out around the edge of the helical axis^{13,18}. These features combine to present a very different architecture and set of hydrogen bonding partners to proteins. Several studies suggest that A-DNA:B-DNA junctions form kinks in the DNA duplex, perhaps even in the absence of protein^{18,19}. This possibility seems quite likely for the TACT:G-tract junctions in the RcnR operator sequence, where intrinsic or RcnR-induced bending at these four junctions could help drive wrapping of the DNA around the protein. However, when only one binding site and two junctions are present (Site 1 footprint), no

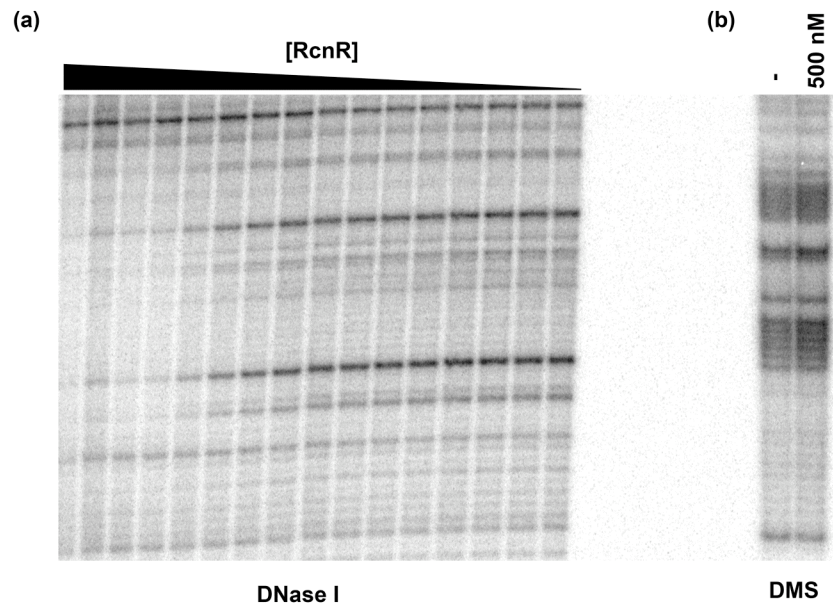


Figure 4.15. RcnR does not protect G-tracts from DMS. DNase I (a) and DMS (b) footprinting were conducted in parallel. No DMS protection for the plus strand of P_{rcnA} was observed with 500 nM RcnR, even though the protein binds at this concentration, as shown by DNase I footprinting. Black bars indicate G-tracts of P_{rcnA} .

wrapping occurs. A-form DNA does not require wholly G-C sequences¹¹, so the strong conservation of GC-rich sequences in other binding sites of RcnR/CsoR proteins^{2,7} suggests that there may be specific structural properties conveyed by these sequences that are not found in other A-forming sequences, or that some specific contacts are made with some of the bases between the inverted repeats.

There are precedents for the various aspects of the RcnR-DNA interaction suggested from the data presented here. Specific minor groove interactions are less common but have been observed for some proteins such as the helix-turn-helix containing PurR. In the co-crystal structure of this complex, sequence-specific contacts from two α -helix-minor groove interactions induce significant kinking of the DNA away from the protein and distortion of the minor groove²⁰. A-DNA has been reported to play a role in DNA binding by DNA polymerases¹⁸, small acid-soluble proteins (SASPs) in dormant *Bacillus* spores^{21,22} and catabolite repressor protein²³. Notably, the eukaryotic Tc3 transposase recognizes an A-DNA:B-DNA junction that depends upon a G₇ tract²⁴. Unlike RcnR, a helix-turn-helix motif binds in the DNA major groove near this junction, and the DNA bends sharply away from the protein. High-resolution structural studies of RcnR/CsoR proteins in complex with DNA will be necessary to completely understand the basis for sequence specific recognition.

The data indicate that wrapping is another feature of the RcnR-DNA complex, as the protein makes nonspecific contacts with the DNA over a ~50 base pair stretch on either side of the core binding site in addition to the specific

interactions made at each operator site. Swapping the flanking sequences with dissimilar plasmid DNA sequences did not significantly affect the binding affinity or protection pattern of RcnR as measured by DNase I footprinting. The DNase I and hydroxyl radical footprinting patterns of RcnR are very similar to those observed for *E. coli* DNA gyrase (GyrAB), which protects ~140 bp of DNase I cleavage, including a core region of ~40 bp, and periodic protected regions 10 bp apart in hydroxyl radical footprinting assays²⁵. DNA is thought to wrap around the C-terminal domain of GyrA utilizing a large basic patch on the outer edge of the protein, but this domain folds into a β -pinwheel structure²⁶ that is unlike that of *MtCsoR*². Like DNA gyrase, our studies indicate that the DNA wraps around RcnR instead of looping away from the protein. It is unclear if DNA wrapping by RcnR provides a physiological advantage for transcriptional repression or metal sensing. It is interesting to note that RcnR and RNA polymerase (RNAP) bind to opposite sides of the DNA helix at the *rcnA* promoter, and RcnR wraps DNA in the opposite direction necessary for RNAP binding²⁷. Therefore, the combination of DNA bending and wrapping may aid in the efficiency of repression. This higher order structure will also alter binding kinetics, and potentially the dynamics of the cell's response to changing metal levels. Intrinsic bending upstream of *rcnA* caused by the G-tracts may aid in RcnR binding, but also decrease the strength of the promoter.

DNA-binding data for other RcnR/CsoR proteins is limited to a small number of studies on *BsCsoR* and *MtCsoR* that explored the basis of protein-

DNA recognition. *MtCsoR* is thought to recognize a GC-rich 12-5-12 inverted repeat², while *BsCsoR* binds to an 8-bp inverted repeat separated by four G bases, although this could also be described as a 6-5-6 inverted repeat^{3,7} (Figure 4.16). More detailed studies of these interactions will be necessary to understand which bases are being contacted by each CsoR homolog.

Strikingly, these CsoR binding sites, the RcnR binding site described here, and the majority of predicted binding sites for many other RcnR/CsoR proteins all contain G-tracts of 3-8 bases, flanked by AT-rich inverted repeats of varying lengths (Figure 4.16). We hypothesize that proteins of this family recognize A-form:B-form junctions that are formed by these sequence elements. A single RcnR tetramer recognizes a site with a single G-tract and therefore both A:B junctions (one on each side). The predicted binding sites of other RcnR proteins all have G/C-tracts of at least four bases (Type 1) or two adjacent G/C-tracts separated by 2-4 bases (Type 2). This distinction is important as Type 2 sites have altered spacing between the flanking inverted repeats, placing the minor grooves on opposite sides of the DNA duplex and preventing a single tetramer from binding both half sites on one face of the DNA. This may explain why two *BsCsoR* tetramers bind to a single Type 2 site⁷. Unlike this single site, other promoters such as those in *V. cholera* have tandem binding sites similar to that in *E. coli*, which could be necessary for repression of both divergently transcribed genes, or fine-tuning the response to signal levels. It is also notable that all of

(a)

Type 1	A. tumefaciens	CCTTGTGCAAAAAATAACTGCAATATAGTATAC <u>CCCCCTATAGTATATGAGGTG</u> -4
	O. anthropi	GTGTGATTTGTAGCGCTCCAGAAAATAGTATAC <u>CCCCCTATTCTAATAGAGAGG</u> -9
	R. leguminosarum	CATTGAGCTCCGGCTTCACTCTTGATACCATAC <u>CCCCCTATGCATATATGCGAGC</u> -6
	R. solanacearum	CGAGGCGACGGGAGGGTAGCACGCAAA <u>TACTAGGGGGGAGTATAGTGTGGCACT</u> -26
	P. cryohalolentis	GACATTATTTTTATTGTCTCATT <u>TTTATACTGGGGGGGTAGCATATTTATAGGA</u> -17
Type 2	P. aeruginosa	TACTGTCTGTGAGAGCTCGAAAATATAC <u>TATACCCAGTATATTACCTTGAG</u> -10
	C. psychrerythraea	AATAGCATATATTTTAAACCTGAAATATAC <u>CCCCCTGTATATTTATTTTCA</u> -15
	C. phaeobacteroides	CCGCCTGTGAAAGTTTATTTAATTATAGTATAC <u>TCCCTTACTATAATTA</u> -14
	N. aromaticivorans	GCTGCCCTTCCCGCCCGCCAGATATAGGGTAGGGGGCTATGCCATGTACAT +9
	P. fluorescens	CGGAACAGCGACTATGCTCGCCATATAGGGGTAGGGGGTATAGGTATGTCCGAC +9
Type 2	R. sphaeroides	CATTGACGCTCCGCGCAGCGCAGGATACCCATAC <u>CCCCCTATAGGTATAAGGGGG</u> -6
	B. xenovorans	CTATGATATTCTCGCAGTGAACATATACCCATAGGGGTATTAGGAGGGTGCGA -1
	X. autotrophicus	GCTTCCTTTGGCCGTGATAATAGTATAGGGGAGTACCCTATGGTACACACCATA +15
	M. magnetotacticum	CGGAGCGGCAGTTTTGGGGCTTGATAGGGTAC <u>CCCCCTATATTATCAGATTGA</u> -8
	S. wittichii	CGATCGATTGCGCCCGTCTGTTGATATAGGGTAC <u>CCCCCTATGCATATGGGACATA</u> +10
B. parapturtussis	GGCTTCAGGATGTTATGATTTTCTATAC <u>CCCTAGGGGGTATGCATGTCACTCC</u> +10	
Tandem	E. coli K12	ATTAATCTACTGGGGGTAGTATCAGGTACTGGGGGGAGTAGAATCAGATTGC -26
	V. cholerae	TCTATCGTACTGGGGGTAGTATATTATAC <u>TGGGAGGGAGTATAGTTTGCCTCA</u> -27
	K. pneumoniae	TTTATCTACTGGGGGTAGTATTTTACTACTGGGGGGAGTAGGATAATCGCAC -20
	B. cepacia	ATATACCCCTCCCCCTATATN16ATATAGTCC <u>CCCCCGGTATATGAAGGGACT</u> -4
	X. campestris	CAGATACCCCCAGGGGTATTTN4ACCATCGGGGGGTATCTTAGTCAGGCCG -9
	P. syringae	TGCATAGGGGGGAAGGGTATGAN11TAGGTAAGGGGGGTATACATGACTCACG +10
	X. oryzae	AGGTACCCCTACAGGGTATCTTAAATACCCCTGGGGGGTATACGGCAGTGGCGC +7
	X. axonopodis	GTATACCCCCAGGGGTATTTAGGATAC <u>CCCTGTGGGGTACCTTAGCCAGAGCC</u> -8
B. subtilis (CsoR)	AATCATAAAGCGTTTTTTTATTGTAATAC <u>CCCTACGGGGGTATGGTAGGATGAAAA</u> -28	
M. tuberculosis (CsoR)	GTTTCGACTCCTTGGGTAGCCACCC <u>CCAGTGGGGTGGGATACCATGAACGGGTG</u> -19	
E. coli K12 (FrmR)	TATATACTATAGGGGGTATGCAN8TAGAATAC <u>CCCCCTATAGTATATATGCATG</u> -19	

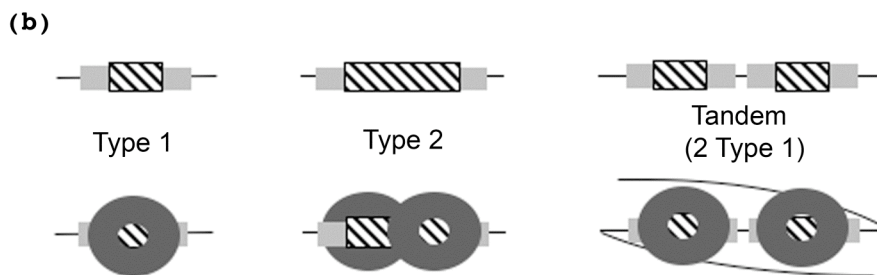


Figure 4.16. Common sequence features of DNA-binding sites of RcnR/CsoR proteins. (a) Promoter regions of genes predicted to be controlled by RcnR proteins (nickel/cobalt responsive) are shown from a variety of organisms. Inverted repeats are highlighted in gray, while G/C tracts are underlined. The sequence numbering is relative to the predicted translation start sites and refers

to the 3' end of each DNA sequence shown. The binding sites can be divided into three categories. *Type 1* sites have a single G/C tract flanked by an AT-rich inverted repeat. *Type 2* sites have two shorter G/C tracts with 2-4 intervening bases, and the AT-rich inverted repeat is outside of the two G/C tracts. The distance between the inside of each half of the inverted repeat is typically 11 bp, while that for *Type 1* sites is 6-9 bp. *Tandem* sites have two separate sites of either *type 1* or *type 2* or both. The spacing between the sites is variable. The binding sites for *BsCsoR*, *MtCsoR* and the likely binding site for *E. coli* FrmR are shown for comparison. The promoter sequences for *B. cepacia*, *X. campestris*, *P. syringae*, and *E. coli* FrmR contain large nucleotide inserts that are represented as N_x to maintain the alignment of the inverted repeats. (b) A schematic representation of the different types of binding sites described in (a) are shown with inverted repeats (light gray boxes), G/C tracts (hatched boxes) without (top) or with (bottom) protein tetramers (dark gray donuts) bound with the predicted stoichiometries based on *BsCsoR* and *E. coli* RcnR. The tandem site can be any combination of *Type 1* and *Type 2* sites but is shown with two *Type 1* sites such as that seen for *E. coli* RcnR, which induces DNA wrapping.

the inverted repeats in the predicted binding sites are AT-rich, which could be necessary to induce B-form DNA adjacent to the G-tracts.

It is likely that some, but not all proteins of the RcnR/CsoR family will also wrap DNA. Footprinting data is only available for *BsCsoR* and while it does show hypersensitive sites at the boundaries of the binding sites, which suggest bending of the DNA, these do not extend out from the core binding site as seen for RcnR. Footprinting with only a single site DNA construct does not show wrapping for RcnR, although binding still occurs, suggesting that other promoters with a single site cannot be wrapped. Our studies show that the spacing between the sites is also important. This indicates a fundamental difference in the DNA binding sites that could lead to distinct binding modes for different RcnR/CsoR proteins, which may also be reflected in changes in specific parts of the protein sequence.

The current biophysical studies paint a complex picture of base-specific, nonspecific and structural DNA features necessary for RcnR repression, while providing a clear basis for understanding future crystallographic experiments on the RcnR:operator structure as well as low-resolution structural studies on the larger, wrapped RcnR:DNA structure. By identifying specific sequence requirements for RcnR binding, the DNA recognition sites of RcnR proteins from other organisms have also been identified, revealing potentially distinct mechanisms of binding among this family due to single or dual binding sites as well as the sequence of the site itself.

Materials and Methods

All chemicals were purchased from Fisher Scientific (Pittsburg, PA) except where indicated. All oligonucleotides (Invitrogen, Carlsbad, CA), plasmids and bacterial strains used for this study are listed in the supplementary materials (Tables 4.1 and 4.2). RcnR was overexpressed in *E. coli* DL41 (DE 3)²⁸ cells and purified as previously described⁵. All nucleic acid modifying enzymes were purchased from New England Biolabs (Beverly, MA) except where noted.

Molecular Biology

The P_{rcnR} -*lacZ* fusion (pJI120) was constructed from the pPC163 parent plasmid²⁹, which contains the P_{nik} promoter and *lacZ* cloned into pACYC184. The 123 bp upstream of *rcnR* were amplified from *E. coli* genomic DNA with primers JI152 and JI153, digested with *EagI* and *Sall* and ligated (T4 DNA ligase) into the parent construct cut with the same enzymes. The *rcnR*-*rcnA* construct (pJI140) was made by PCR amplification of genomic DNA with JI133 and JI147, digestion with *EagI* and *Sall* and ligation into pBluescript KSII (pBS) (Stratagene, La Jolla, CA). Core RcnR binding site constructs were cloned into pBS using the oligonucleotide pairs JI350/JI351 (core) and JI369/JI370 (single site) which were purchased phosphorylated at the 5' end and contain overhangs equivalent to those generated by *EagI* and *XhoI*. DNA used for topoisomerase I relaxation assays (pJI141) contained the 273 base pairs immediately upstream from *rcnA* (primers JI139/JI188) cloned into the *EagI* and *EcoRI* sites of pBS.

DNase I Footprinting

Table 4.1. Oligonucleotides

Oligo	Purpose	Sequence
JI133	<i>rcnR-rcnA</i> cloning <i>P_{rcnA}</i> for topo assay,	gtcatagtcgacttatcgcattatgcccatgaag
JI139	WT footprinting	ctatgtcggccgaccggttaccgcgccacggatag
JI140	WT footprinting	gtcatagtcgacgataataattcttagtattaattcggc
JI144	WT footprinting	ctatgtcggccgggcaacacacctggtaacaagaatag
JI147	<i>rcnR-rcnA</i> cloning	ctatgtcggccgtaataaattattgatataatgaatcc
JI152	<i>P_{rcnR}-lacZ</i> cloning	ctatgtcggccggataataattcttagtattaattcgg
JI153	<i>P_{rcnR}-lacZ</i> cloning <i>rcnR</i> primer	gtcatagtcgacggcaacacacctggtaacaagaatag
JI173	extension	ctgtttatcacggattg
JI183	wt footprinting	gttgctgccagtcgagcatc
JI188	<i>P_{rcnA}</i> for Topo I assay	gtcatagaattcgataataattcttagtattaattcggc
JI197	gel shift 300 bp	cagtttctgtttatcacggattgtatgagac
JI200	gel shift 200 bp	aaccaggcgcttcctgtcgaagaagagttg
JI207	gel shift 152 bp	gaagagttgtaaattcggatcatgataataattc
JI210	site 2, ITC, CD	caggactgggggggagtagaatc
JI210r	site 2, ITC, CD	gatttactccccccagtagctg
JI241	gel shift 110 bp	gataataattcttagtattaattc
JI242	gel shift 110 bp	ggtaacaagaatagaaaatca
JI247	gel shift 81 bp	gaaaatcatagcactattaatctac
JI248	gel shift 81 bp	gtattaattcggcaatctgattc
JI249	gel shift 200 bp	ttcagtttctgtttatcacggattgta
JI250	gel shift 250 bp	gccctgaatctactggcacgcgcttt
JI251	gel shift 250 bp	caagtaagatggcgcctggggatgaaga
JI252	gel shift 130 bp	ggcaacacctggtaacaagaatag
JI253	gel shift 130 bp	attcggatcatgataataattcttagtatt
JI254	gel shift 152 bp	tatgagacatggcaacacctggtaac
JI256	gel shift 300 bp non-specific DNA,	aagaaccaggcgcttcctgtcgaagaag
JI284	ITC, CD non-specific DNA,	ggatgcagaaggctggatgtgacg
JI285	ITC, CD	cgtcacatccgaccttctgcatcc
JI335	site 1, ITC, CD	aatctactgggggtagtatcagg
JI335r	site 1, ITC, CD	cctgatactacccccagtagatt
JI336	site 1,2 ITC	aatctactgggggtagtatcaggtagtgggggggagtagaatc
JI336r	site 1,2 ITC	gatttactccccccagtagctgatactacccccagtagatt
JI347	footprinting with pBS	caacgcaattaatgtgagttagctcac
JI350	<i>P_{rcnA}</i> core	Phos- ggccgtaatctactgggggtagtatcaggtagtgggggggagtagaatcac

JI351	<i>P</i> _{rcnA} core	Phos- tcgagtgattctactccccccagtagctgatactacccccagtagattaac
JI355	site 1 T1 mutant	aatccactgggggtagtgcagg
JI355r	site 1 T1 mutant	cctgacactacccccagtgatt
JI356	site 1 C3 mutant	aatctattgggggtaatatcagg
JI356r	site 1 C3 mutant	cctgatattaccccccaatagatt
JI358	site 1 extended	tgaaaaatcatagcactattaatctactgggggtagtgcagg
JI358r	site 1 extended	cctgatactacccccagtagattaatagtgcctatgattttca
JI359	site 1 A2 mutant	aatctgctgggggtagcatcagg
JI359r	site 1 A2 mutant	cctgatgctacccccagcagatt
JI360	site 1 T4 mutant	aatctaccgggggtggtatcagg
JI360r	site 1 T4 mutant	cctgataccacccccggtagatt
JI369	<i>P</i> _{rcnA} core single site	ggccgtaatctactgggggtagtgcaggc
JI370	<i>P</i> _{rcnA} core single site	tcgagcctgatactacccccagtagattaac
JI371	site 1 G-tract mutant	aatctactaaaaaatagtagcagg
JI371r	site 1 G-tract mutant	cctgatactatttttagtagatt
JI372	Footprinting with pBS <i>rcnA</i> primer	ctatagggcgaattggagctc
JI373	extension	gaagagttgtaaaattcgggtcat

Table 4.2. Plasmids and bacterial strains

Plasmid	Insert	Parent	Reference
pRcnR	rcnR	pET22b(amp ^R)	5
pJI120	<i>P_{rcnR}-lacZ</i>	pACYC163(cm ^R)	This work
pJI140	<i>rcnR-rcnA</i>	pBluescript	This work
pJI141	<i>P_{rcnA}</i>	pBluescript	This work
pJI189	<i>P_{rcnA} core</i>	pBluescript	This work

Strain	Genotype	Parent	Reference
PC563	$\Delta rcnR \Delta lacZ$	RZ4500	5
DL41	<i>metA-</i>	MG1655	27

³²P-end-labeled DNA was generated by first labeling individual DNA primers with γ -³²P[dATP] and T4 polynucleotide kinase. The labeled primers were then used for PCR with reverse, unlabeled primers, using pJ1140 as a template. The resulting labeled products were then run on 6% acrylamide TBE gels. Bands were excised from the gel and incubated in 10 mM Tris (pH 7.5) 50 mM NaCl 1 mM EDTA 1 mM MgCl₂ 0.05% SDS at 37 °C overnight. The supernatant was removed and ethanol precipitated. The resulting DNA pellet was washed with 70% ethanol, then resuspended in ddH₂O. DNase I footprinting experiments were carried out in 50 μ l of Buffer H [10 mM Hepes (pH 7.0), 150 mM NaCl, 2 mM MgCl₂, 10% glycerol, 200 μ M EDTA] or Buffer P [10 mM Pipes (pH 7.0), 45 mM K₂SO₄, 1 mM MgSO₄, 0.5 mM TCEP, 15% glycerol, 200 μ M EDTA] where indicated. Labeled DNA was incubated with RcnR for 30 min at 22 °C before addition of 5 μ l of 3 μ g/ml DNase I (Sigma, St. Louis, MO; in Buffer H or P with 2.5 mM CaCl₂). Reactions were quenched with 110 μ l of 2.5 M ammonium acetate, 20 mM EDTA, 20 μ g/ml salmon sperm DNA after 1 min at 22 °C. DNA was then precipitated with 1 μ l linear acrylamide (Ambion, Austin, TX) and 300 μ l ethanol. After centrifugation and washing with 70% ethanol, pellets were resuspended in formamide loading dye and loaded on a pre-run 6% acrylamide urea gel, which was then run at 40 W for 2.5 hrs. Gel bands were quantitated using a Typhoon phosphorimager (GE Healthcare, Piscataway, NJ).

G/A ladders were generated by adding 25 μ l of 88% formic acid to 10 μ l ³²P-labeled DNA for 5 min at 22 °C, then quenching with 200 μ l 0.3 M sodium

acetate (pH 7.0), 0.1 M EDTA, 25 µg/ml salmon sperm DNA, and ethanol precipitated. After rinsing with 70% ethanol, the pellet was resuspended in 70 µl of 10% piperidine, and incubated at 90 °C for 30 min to cleave the DNA. After evaporation of the piperidine, samples were rinsed twice with 50 µl ddH₂O, dried and resuspended in formamide loading dye.

Hydroxyl Radical Footprinting

RcnR protein was desalted two times using Microbiospin 6 columns (Biorad, Hercules, CA) into 10 mM sodium cacodylate (pH 7.0), 150 mM NaCl, 2 mM MgCl₂ immediately prior to use to remove any glycerol from the storage buffer. RcnR was then incubated with ³²P-end-labeled DNA for 30 minutes at 22 °C in 50 µl binding reactions. Cleavage reactions were carried out by addition of 5 µl 2 mM FeSO₄/EDTA, 2.5 µl 20 mM sodium ascorbate and 2.5 µl 1.2% H₂O₂, and incubation at 22°C for 1 min. Reactions were terminated by addition of 10 µl of 0.1 M thiourea (Sigma) and 1 µl 0.5 M EDTA, and samples were precipitated and electrophoresed as for DNase I reactions

Mobility Shift Assays

Mobility shift binding assays were conducted in 50µl of 10 mM Hepes (pH 7.0), 150 mM NaCl, 2 mM MgCl₂, 5% glycerol, 1 µM EDTA, 2 µg/ml salmon sperm DNA, 10 µg/ml thioredoxin. Samples were prepared as previously described⁵. ³²P-labeled DNA was generated by including α-³²P[dATP] in PCR reactions with the appropriate primers (see Table 4.1) using pJ1140 as a template.

β -galactosidase Reporter Experiments

LacZ reporter assays were set up as previously described with several modifications^{5,29}. *E. coli* strain PC563 cells containing P_{rcnR} -*lacZ* (pJI120) and either pET22b or pRcnR (pJI111) were grown in Luria-Bertani broth (LB) with 34 mg/l chloramphenicol and 100 mg/l ampicillin. Cultures were used to inoculate 1.87 ml of LB containing the same antibiotics to a starting OD of 0.000001 in capped microfuge tubes in triplicate. These were then grown for 12-14 hrs at 37 °C, then assayed as described²⁹.

Isothermal Titration Calorimetry

DNA oligonucleotides for ITC were gel purified on 12 % acrylamide urea denaturing gels. DNA bands were excised from the gel, crushed, and soaked in 10 mM Tris (pH 8.0), 500 mM NaCl, 1 mM EDTA overnight. The extracted DNA was then desalted on Sep-Pak C18 cartridges (Waters, Milford, MA).

Oligonucleotides were eluted in 30% acetonitrile, and dialyzed extensively against Buffer P. Concentrations were determined by phosphodiesterase I (Worthington, Lakewood, NJ) digestion and absorbance measurements in 10 mM K_2PO_4 (pH 7.0) and 100 mM NaCl at 260 nm using previously determined extinction coefficients for each nucleotide³⁰. Equal amounts of top and bottom strands were then annealed by heating to 95 °C and slow cooling to 22 °C over several hours, followed by incubation on ice. RcnR was dialyzed against the stock of buffer solution to prevent mismatches in buffer composition.

ITC experiments were carried out on a MicroCal ITC 200 calorimeter (Northampton, MA) with the following settings: reference power = 2 μ cal/sec, stir speed = 1000 rpm, feedback = high, filter period = 5 sec, injection volume = 2.5 μ l x 15 injections, 180 s spacing between injections. For each titration 300 μ l of RcnR was placed in the cell with 40 μ l of DNA in the injection syringe. Protein and DNA concentrations are listed in Figure legends. Reference titrations of DNA into buffer were used to correct for heats of dilutions. Titrations were fit to the n independent binding sites model in Origin version 5.0 with a manual baseline correction. Errors reported for each temperature are standard deviations obtained from two or more titrations. Reported stoichiometries are averages for every titration for a given duplex.

Analytical Ultracentrifugation

Sedimentation equilibrium experiments were conducted on a Beckman XL-A ultracentrifuge with six-cell sample holders with a 1.2 cm path length in Buffer P as described previously for RcnR⁵. Data for 5 μ M and 25 μ M RcnR monomer at 25 000 and 35 000 rpm were globally analyzed using SEDPHAT³¹.

Circular Dichroism

CD measurements were obtained on a JASCO 710 spectropolarimeter with a 1 mm path length cylindrical cuvette at 20 °C. Wavelength scans are averages of five accumulations with a scan speed of 20 nm/min and 0.1 nm step size. 25 μ M DNA samples were made in 0.5 mM Tris (pH 7.5), 1 mM NaCl, 0.05 mM EDTA or in the same buffer with 84% trifluoroethanol (TFE). For TFE

titration samples, 20 μ M duplexes were made in 84% TFE and small aliquots of buffer were added, samples were scanned three times and the average ellipticity at 270 nm was recorded. Conversion to % relative humidity was based on the values for water/ethanol mixtures^{16,14}.

Topoisomerase I Relaxation

RcnR was desalted two times into 10 mM Tris (pH 7.9), 50 mM NaCl, 10 mM MgCl₂, 1 mM DTT, 5 % glycerol (Buffer T), and incubated with 20 nM plasmid DNA in Buffer T without glycerol and 0.1 mg/ml BSA (50 μ l). After incubation at 37 °C for 15 minutes 1.5 μ l *E. coli* topoisomerase I (New England Biolabs) was added and incubated for another four hours. Reactions were stopped by addition of 50 μ l of 1% SDS and 10 mM EDTA. After phenol:chloroform extraction and ethanol precipitation samples were loaded onto 1.5% agarose Tris-acetate-EDTA gels with 1.2 μ g/ml chloroquine. Gels were run at 38 V for 22-24 hr at 22 °C, and visualized with ethidium bromide.

References

1. Minezaki, Y., Homma, K. & Nishikawa, K. (2005). Genome-wide survey of transcription factors in prokaryotes reveals many bacteria-specific families not found in archaea. *DNA Res* **12**, 269-80.
2. Liu, T., Ramesh, A., Ma, Z., Ward, S.K., Zhang, L., George, G.N., Talaat, A.M., Sacchettini, J.C. & Giedroc, D.P. (2007). CsoR is a novel Mycobacterium tuberculosis copper-sensing transcriptional regulator. *Nat Chem Biol* **3**, 60-8.
3. Smaldone, G.T. & Helmann, J.D. (2007). CsoR regulates the copper efflux operon copZA in *Bacillus subtilis*. *Microbiology* **153**, 4123-8.
4. Iwig, J.S., Leitch, S., Herbst, R.W., Maroney, M.J. & Chivers, P.T. (2008). Ni(II) and Co(II) sensing by *Escherichia coli* RcnR. *J Am Chem Soc* **130**, 7592-606.
5. Iwig, J.S., Rowe, J.L. & Chivers, P.T. (2006). Nickel homeostasis in *Escherichia coli* - the rcnR-rcnA efflux pathway and its linkage to NikR function. *Mol Microbiol* **62**, 252-62.
6. Herring, C.D. & Blattner, F.R. (2004). Global transcriptional effects of a suppressor tRNA and the inactivation of the regulator frmR. *J Bacteriol* **186**, 6714-20.
7. Ma, Z., Cowart, D.M., Scott, R.A. & Giedroc, D.P. (2009). Molecular insights into the metal selectivity of the copper(I)-sensing repressor CsoR from *Bacillus subtilis*. *Biochemistry* **48**, 3325-34.

8. Rodrigue, A., Effantin, G. & Mandrand-Berthelot, M. (2005). Identification of rcnA (yohM), a nickel and cobalt resistance gene in *Escherichia coli*. *J Bacteriol* **187**, 2912-6.
9. Grainge, I., Gaudier, M., Schuwirth, B.S., Westcott, S.L., Sandall, J., Atanassova, N. & Wigley, D.B. (2006). Biochemical analysis of a DNA replication origin in the archaeon *Aeropyrum pernix*. *J Mol Biol* **363**, 355-69.
10. Grainge, I., Buck, D. & Jayaram, M. (2000). Geometry of site alignment during int family recombination: antiparallel synapsis by the Flp recombinase. *J Mol Biol* **298**, 749-64.
11. Basham, B., Schroth, G.P. & Ho, P.S. (1995). An A-DNA triplet code: thermodynamic rules for predicting A- and B-DNA. *Proc Natl Acad Sci U S A* **92**, 6464-8.
12. Peticolas, W.L., Wang, Y. & Thomas, G.A. (1988). Some rules for predicting the base-sequence dependence of DNA conformation. *Proc Natl Acad Sci U S A* **85**, 2579-83.
13. Sinden, R.R. (1994). *DNA Structure and Function*. Academic Press, San Diego.
14. Ivanov, V.I. & Krylov DYu. (1992). A-DNA in solution as studied by diverse approaches. *Methods Enzymol* **211**, 111-27.
15. Fairall, L., Martin, S. & Rhodes, D. (1989). The DNA binding site of the *Xenopus* transcription factor IIIA has a non-B-form structure. *EMBO J* **8**, 1809-17.

16. Ivanov, V.I., Minchenkova, L.E., Burckhardt, G., Birch-Hirschfeld, E., Fritzsche, H. & Zimmer, C. (1996). The detection of B-form/A-form junction in a deoxyribonucleotide duplex. *Biophys J* **71**, 3344-9.
17. Seeman, N.C., Rosenberg, J.M. & Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. U.S.A* **73**, 804-808.
18. Lu, X.J., Shakked, Z. & Olson, W.K. (2000). A-form conformational motifs in ligand-bound DNA structures. *J. Mol. Biol* **300**, 819-840.
19. Hud, N.V. & Plavec, J. (2003). A unified model for the origin of DNA sequence-directed curvature. *Biopolymers* **69**, 144-158.
20. Schumacher, M., Choi, K.Y., Zalkin, H. & Brennan, R.G. (1994). Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. *Science* **266**, 763-770.
21. Setlow, P. (1992). DNA in dormant spores of Bacillus species is in an A-like conformation. *Mol. Microbiol* **6**, 563-567.
22. Lee, K.S., Bumbaca, D., Kosman, J., Setlow, P. & Jedrzejewski, M.J. (2008). Structure of a protein–DNA complex essential for DNA protection in spores of Bacillus species. *Proceedings of the National Academy of Sciences* **105**, 2806-2811.
23. Ivanov, V.I., Minchenkova, L.E., Chernov, B.K., McPhie, P., Ryu, S., Garges, S., Barber, A.M., Zhurkin, V.B. & Adhya, S. (1995). CRP-DNA complexes:

- inducing the A-like form in the binding sites with an extended central spacer. *J Mol Biol* **245**, 228-40.
24. van Pouderooyen, G., Ketting, R.F., Perrakis, A., Plasterk, R.H. & Sixma, T.K. (1997). Crystal structure of the specific DNA-binding domain of Tc3 transposase of *C.elegans* in complex with transposon DNA. *EMBO J* **16**, 6044-6054.
25. Orphanides, G. & Maxwell, A. (1994). Evidence for a conformational change in the DNA gyrase-DNA complex from hydroxyl radical footprinting. *Nucleic Acids Res* **22**, 1567-75.
26. Corbett, K.D., Shultzaberger, R.K. & Berger, J.M. (2004). The C-terminal domain of DNA gyrase A adopts a DNA-bending beta-pinwheel fold. *Proc Natl Acad Sci U S A* **101**, 7293-8.
27. Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O. & Darst, S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285-1290.
28. Hendrickson, W.A., Horton, J.R. & LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J* **9**, 1665-1672.
29. Rowe, J.L., Starnes, G.L. & Chivers, P.T. (2005). Complex transcriptional control links NikABCDE-dependent nickel transport with hydrogenase expression in *Escherichia coli*. *J Bacteriol* **187**, 6317-23.

30. Gray, D.M., Hung, S.H. & Johnson, K.H. (1995). Absorption and circular dichroism spectroscopy of nucleic acid duplexes and triplexes. *Methods Enzymol* **246**, 19-34.
31. Vistica, J., Dam, D., Balbo, A., Yikilmaz, E., Mariuzza, R.A., Rouault, T.A. & Schuck, P. (2004). Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Anal Biochem* **326**, 234-56.

Chapter 5
Perspectives and Future Directions

In this thesis I have identified a novel family of metalloregulators that plays diverse roles in many bacteria. Investigation of the mechanisms of metal and DNA binding by RcnR has provided insights not only into Ni homeostasis in *E. coli* but also general features of this protein family, and paradigms of metal sensing in bacteria. Starting this project with an idea and a few primers and seeing it evolve to a significant understanding of the RcnR/CsoR protein family has been extremely rewarding, despite the challenges encountered along the way. These results have generated many exciting questions and future projects. The most important include: I. structural studies on RcnR to further investigate the basis for DNA wrapping and recognition of the unique G-rich binding sites in operator sequences, II. using comparisons between NikR and RcnR as a model to understand how differences in metal-binding properties translate to in vivo function and III. the examination of other RcnR/CsoR proteins to identify the mechanisms used to sense other signals and bind DNA.

I. Structural Studies on RcnR

A challenging part of my early work on RcnR was understanding aspects of protein structure/function without high-resolution structural information, or comparable systems, which created difficulties in putting the results in a broader context and applying them to additional transcription factors. In fact, the lack of any comparable systems led to a large number of interesting questions to explore, which became problematic at times and challenged my ability to focus

on one problem at a time. In many ways the publication of a *MtCsoR* crystal structure in late 2006¹ aided my work tremendously, and allowed for more refined arguments regarding RcnR metal binding as well a point of comparison with which to draw additional conclusions about the protein family. The CsoR structure was extremely helpful, but revealed very little about the DNA-binding mechanisms of this family. The DNA recognition sites are also quite unique compared to those of other transcription factor families, making the details of the RcnR-DNA interaction all the more interesting.

Future studies should be directed at solving the structure of an RcnR-DNA complex, which will reveal how RcnR interacts with the DNA minor groove, why the A-DNA-like G-tracts are important for binding, and what protein surface is responsible for the interaction. The G-tracts could play only a structural role for RcnR recognition, or specific contacts may also be made to these bases. My studies on RcnR indicate that DNA structure is a significant factor in DNA recognition and crystallography or NMR are the best way to determine how the recognition site deviates from B-form DNA. Preliminary data indicate a role of basic residues in helix 1 (Arg14, Lys17) of RcnR in DNA binding, but it is not known which DNA elements these residues bind. DNA binding with other mutant proteins could be carried out, but this cannot reveal why particular residues are important, and it is unclear which other residues should be targeted. Crystallographic or computational studies on the DNA alone may also be useful

to better understand the intrinsic structure of the duplex, and how this changes upon protein binding.

It is also unclear if RcnR and *MtCsoR* will have identical structures given their limited sequence identity and differences in oligomeric state, which further emphasizes the need for an RcnR structure. Comparison of DNA-bound, apo- and Ni-RcnR structures will also provide information on the allosteric changes induced by metal binding. The allosteric pathways for metal inactivation are not necessarily the same for RcnR and CsoR, as metal binding could induce different conformations, as long as they are not capable of DNA binding. The studies in Chapter 4 of this thesis identified biophysical and biochemical details of the RcnR-DNA complex, that will guide selection of DNA substrates for, and interpretation of a protein-DNA crystal structure.

These studies also reveal evidence for a larger, wrapped DNA complex that involves non-specific DNA contacts. To better understand the nature of this complex, lower resolution structural studies that allow for the analysis of large particles, such as small angle X-ray scattering and electron microscopy, which will provide information about the dimensions of these complexes, will be necessary. As I discovered in previous work, trying to make connections between studies with different size DNA duplexes will be challenging given the limitations on amounts of large DNA that can be synthesized. This has currently limited our understanding of the thermodynamics of wrapping. Wrapping is likely important for repression of *rcnA* as this type of complex will have different

dynamic properties than more conventional protein-DNA interactions, which may affect temporal aspects of gene expression.

II. A Tale of Two Regulators: NikR and RcnR

The metalloregulator field has developed rapidly over the past 20 years, but many challenging questions remain to be addressed. Paradigms for metal selectivity and allostery have been proposed, but in most cases these are oversimplified²⁻⁴. One key question is how the thermodynamics, kinetics and geometries of different metal binding sites translate to in vivo functions that are tuned to respond to specific conditions in the cell. Most current work focuses only on the thermodynamics of binding, paying less attention to the importance of coordination environments and dynamics of metal acquisition and exchange. Unfortunately, binding affinities are not necessarily meaningful in the absence of affinities for other proteins for the same organism.

The RcnR/NikR system is an extremely powerful tool to move the field in new directions, as the groundwork laid by this thesis opens up many opportunities not present in other previously described systems. I have shown that the in vivo responses of NikR and RcnR are sequential with respect to extracellular Ni(II) concentration, with NikR regulating nickel import and RcnR controlling nickel export, and this can be partially explained by the tighter Ni(II) affinity of NikR. However, the finding that RcnR has a six-coordinate, distorted octahedral Ni(II) site (versus the four-coordinate square planar NikR site)

indicates that the kinetics of metal binding will also differ for the two proteins, and will have an impact on the dynamics of gene regulation.

Previous experiments have focused on cellular responses to chronic Ni stress⁵, but complementary information can be obtained through analysis of the kinetics of regulation by NikR and RcnR in response to various amounts of Ni(II), or shifting cells from anaerobic to aerobic conditions, which would lead to degradation of hydrogenase and intracellular Ni release. While these experiments may be technically challenging, they would explore situations that more closely simulate environmental conditions encountered by *E. coli*. A complementary set of in vitro studies would examine the on and off rates of Ni binding to both NikR and RcnR using a stopped-flow UV-vis assay. The Chivers lab recently discovered that a Ni-His₂ complex is transported into the cell via NikABCDE, so it will be vital to understand how NikR and RcnR accept and release Ni in the presence of histidine or similar biologically relevant chelating agents. Currently only one study on the kinetics of metal binding of a metalloregulator has been published, however the results are difficult to interpret in the absence of kinetic information for other metal binding proteins, and without native chelators being used⁶. Binding kinetics may also play a key role for in vivo metal selectivity but this idea has not been thoroughly explored.

A second interesting line of comparison between NikR and RcnR will be to understand how sub-saturating Ni(II) concentrations affect protein function. Because both proteins are tetramers, there is the potential for cooperative

loading of metal sites as well as binding of sub-stoichiometric nickel ions inducing conformational changes, which may alter the repression capabilities of each regulator. All other known families of metalloregulators are dimeric⁷, but in the case of CzrA from the ArsR/SmtB family, binding of a single metal is sufficient to significantly decrease DNA binding⁸. To test this linkage, site-binding constants need to be measured for RcnR (these have recently been obtained for NikR), then DNA-binding experiments can be used as a functional readout with respect to the degree of Ni loading. Sub-saturating Ni concentrations may differentially affect NikR and RcnR due to their fundamental differences in function, as Ni must induce a specific conformation for NikR to bind DNA, while there are many potential conformations of Ni-RcnR that would be unable to bind DNA. By carrying out the above experiments on two complementary transcription factors there is the potential to truly understand metalloregulator function at a level previously not achieved, and learn how these functions evolve in the presence of a complementary transcription factor.

III. Untapped Potential of the RcnR/CsoR family

An intriguing direction for future research is investigating the functions of the many distinct proteins in the RcnR/CsoR family. Studies on RcnR and analysis of gene neighborhoods of other family members have led to a simple classification scheme consisting of six groups that can easily be tested through work on a small number of specific proteins that are distinct from RcnR or CsoR,

with the prediction that the conserved “fingerprint” amino acids are necessary to sense each signal. However, it is possible that similar to other metalloregulator families such as the ArsR/SmtB family⁴, RcnR/CsoR proteins may have evolved different signal-sensing locations on the same protein framework. While some RcnR/CsoR groups sense metal ions, others sense small molecule stressors. Metal ions could be necessary cofactors for these proteins, instead of acting as signals. Different family members will also be useful to probe the similarities and differences of DNA binding for this family, particularly related to stoichiometry of binding and DNA wrapping. The predicted binding sites for other RcnR/CsoR proteins are diverse, but share G-tract and AT-rich inverted repeat elements, and differ with respect to the number of sites and inverted repeat spacing, which will influence the stoichiometry and wrapping of each protein-DNA complex.

The further investigation of two particular RcnR/CsoR family members will broaden our knowledge of the functional aspects of the family that are discussed above, and are from genetically tractable organisms. *Bacillus subtilis* YrkD, from the “CXC” group of regulators, is predicted to sense redox stress, and is part of an operon that is induced by methylhydroquinone, which is similar to the ganomycin antibiotics⁹. The range of inducer compounds is unknown, but this suggests that antimicrobial compounds in the environment or used for human treatment of *Bacillus* infections could be detoxified by Yrk proteins. Recent evidence demonstrates that many classes of antibiotics act via a common oxidative stress pathway¹⁰, though the mechanistic details remain to be

elucidated¹¹. This leads to a prediction that YrkD uses the pair of “fingerprint” cysteines to sense oxidative stress through formation of a disulfide bond or direct modification of one or both residues. Understanding the mechanism of redox sensing will aid in understanding the allostery of DNA binding inhibition of the RcnR/CsoR family, and determine if proteins from different groups utilize equivalent residues for signal sensing. The *yrkE* promoter contains a single type 2 (Chapter 4) binding site that can be used to test stoichiometry and wrapping predictions based on results for RcnR. In addition, study of the *yrk* operon could form the basis for therapeutic advances against pathogenic *Bacillus* species that also contain this operon.

The second protein, the formaldehyde responsive *E. coli* FrmR, has already been subject to preliminary studies (see Appendix 1), and belongs to another distinct group of RcnR/CsoR proteins. FrmR represses expression of the *frm* operon, which includes genes encoding two formaldehyde detoxification proteins FrmA and FrmB. FrmR is easily purified but it is not known if solubility problems will arise with the protein-DNA complex as for RcnR. This posed significant difficulties in studies on RcnR, but the eventual success with RcnR suggests that exploring a variety of solution conditions may similarly overcome impediments with FrmR. In vivo experiments have revealed that small aldehydes induce P_{frm} expression, and DNA-binding studies identified a potential DNA binding site that is distinct from other experimentally studied family members in that it contains a much longer inverted repeat as well as a significantly longer

spacer between the two sites. The implications of this DNA-site architecture for DNA wrapping and stoichiometry are unclear, but could be determined using topoisomerase I relaxation and ITC assays.

Another intriguing line of experiments will be to elucidate the mechanism of formaldehyde sensing by first developing a solution-based DNA binding assay such as fluorescence or ITC to find conditions necessary to inactivate FrmR. Mass spectrometry and X-ray absorption spectroscopy could then be used to determine if the protein is covalently modified or the known metal-binding site has changed. While many studies in vitro have examined formaldehyde modifications on proteins, their relevance in the cell is unclear, so understanding the interaction of FrmR with aldehydes will also shed light on the mechanisms of aldehyde toxicity. In addition, FrmR is especially interesting as it is able to bind Zn(II), and this appears to act as a cofactor necessary for DNA binding, in direct contrast to metal-sensing RcnR and CsoR for which metal binding inhibits DNA binding. The precise role of the metal-binding site will need to be thoroughly probed. Overall, studies on YrkD and FrmR will identify signal-sensing mechanisms of two non-metal responsive groups of RcnR/CsoR proteins, and test many of the predictions made about the different modes of DNA binding by these proteins. Similar studies of the “CHH” proteins (likely metal responsive), will provide details on the mechanisms of metal specificity of RcnR/CsoR proteins, while study of the *Listeria* group may prove difficult, as the inducers of these operons is unclear.

IV. Relevance and Final Thoughts

Transition metal biochemistry is surprisingly at the heart of some critical current societal issues including the environment. Hydrogen fuel has shown promise to help curb rising atmospheric carbon dioxide levels, but is currently too expensive for many applications, leading some to investigate using hydrogenase-based systems, or hydrogenase-producing microbes for hydrogen production¹²⁻¹⁴. Another option to stem global warming is fertilizing the ocean with iron to encourage phytoplankton growth, CO₂ consumption, and subsequent CO₂ sequestration in ocean sediments^{15,16}. In both of these examples, a fundamental understanding of the physiological impacts of metal ions on microorganisms is of utmost importance. For hydrogen production, knowledge of the regulation of hydrogenase levels and trafficking of the necessary metal ions is vital to optimize enzymatic activity. Addition of iron to the ocean will affect all marine organisms, and understanding the mechanisms of metal tolerance and toxicity will be critical for adjustment of treatments to prevent unintended consequences. Research on fundamental mechanisms of metal transport, regulation and utilization will help guide these two important areas of research as well as those exploring human diseases of metal accumulation or deficiency and pathogenesis of Ni-utilizing bacteria. The basic chemical properties of transition metals are the same, regardless of the cell type, making the study of simplified bacterial models of great value.

Amazingly this project has followed a course very similar to what I originally envisioned, although the results have been surprising in many ways, particularly with regard to the RcnR-DNA interaction. The unexpected complexity on many different levels has made it very rewarding to have developed this project from its infancy. It has been exciting to not only identify a completely new protein family, but also discover the many unique properties of these proteins. With the recent burst of sequenced genomes it has become increasingly rare to find new bacterial protein families that are so widespread. Perhaps the most satisfying result is that these studies will lay a solid foundation for many diverse, stimulating projects, as there remains much to explore.

References

1. Liu, T. et al. CsoR is a novel Mycobacterium tuberculosis copper-sensing transcriptional regulator. *Nat Chem Biol* **3**, 60-8(2007).
2. Outten, C.E. & O'Halloran, T.V. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488-2492(2001).
3. Waldron, K.J. & Robinson, N.J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol* **7**, 25-35(2009).
4. Tottey, S., Harvie, D.R. & Robinson, N.J. Understanding how cells allocate metals using metal sensors and metallochaperones. *Acc. Chem. Res* **38**, 775-783(2005).
5. Rowe, J.L., Starnes, G.L. & Chivers, P.T. Complex transcriptional control links NikABCDE-dependent nickel transport with hydrogenase expression in *Escherichia coli*. *J Bacteriol* **187**, 6317-23(2005).
6. Busenlehner, L.S. & Giedroc, D.P. Kinetics of metal binding by the toxic metal-sensing transcriptional repressor *Staphylococcus aureus* pl258 CadC. *J. Inorg. Biochem* **100**, 1024-1034(2006).
7. Giedroc, D.P. & Arunkumar, A.I. Metal sensor proteins: nature's metalloregulated allosteric switches. *Dalton Trans* 3107-3120(2007).doi:10.1039/b706769k
8. Lee, S. et al. Structural insights into homo- and heterotropic allosteric coupling in the zinc sensor *S. aureus* CzrA from covalently fused dimers. *J. Am. Chem. Soc* **128**, 1937-1947(2006).

9. Nguyen, V.D. et al. Transcriptome and proteome analyses in response to 2-methylhydroquinone and 6-brom-2-vinyl-chroman-4-on reveal different degradation systems involved in the catabolism of aromatic compounds in *Bacillus subtilis*. *Proteomics* **7**, 1391-1408(2007).
10. Kohanski, M.A. et al. A common mechanism of cellular death induced by bactericidal antibiotics. *Cell* **130**, 797-810(2007).
11. Hassett, D.J. & Imlay, J.A. Bactericidal antibiotics and oxidative stress: a radical proposal. *ACS Chem. Biol* **2**, 708-710(2007).
12. Allakhverdiev, S.I. et al. Hydrogen photoproduction by use of photosynthetic organisms and biomimetic systems. *Photochem. Photobiol. Sci* **8**, 148-156(2009).
13. Kluwer, A.M. et al. Special Feature: Self-assembled biomimetic [2Fe2S]-hydrogenase-based photocatalyst for molecular hydrogen evolution. *Proc. Natl. Acad. Sci. U.S.A* (2009).doi:10.1073/pnas.0809666106
14. Maeda, T., Sanchez-Torres, V. & Wood, T.K. Enhanced hydrogen production from glucose by metabolically engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol* **77**, 879-890(2007).
15. Pollard, R.T. et al. Southern Ocean deep-water carbon export enhanced by natural iron fertilization. *Nature* **457**, 577-580(2009).
16. Smetacek, V. & Naqvi, S.W.A. The next generation of iron fertilization experiments in the Southern Ocean. *Philos Transact A Math Phys Eng Sci* **366**, 3947-3967(2008).

Appendix 1

Mechanistic Insights into Formaldehyde Sensing by the RcnR Homolog

FrmR

Introduction

The RcnR/CsoR family is thought to contain at least six different groups of proteins that each respond to diverse stimuli. While the Ni(II)/Co(II)-responsive and Cu(I)-responsive groups have been investigated, little is known about the signal sensing and DNA binding mechanisms of the remaining proteins. FrmRs are one group of proteins from this family and are distinguishable based on the proximity of these genes to those encoding formaldehyde detoxification proteins. All proteins contain conserved residues (C35 and H60) at two of the four fingerprint positions (see Chapter 3).

In *E. coli*, *frmR* is 34-bp upstream of *frmA*, a formaldehyde dehydrogenase, and *frmB*, a S-formylglutathione hydrolase. Upon entering the cell, formaldehyde spontaneously reacts with glutathione, giving S-hydroxymethyl glutathione, which is oxidized by FrmA to S-formylglutathione, before being hydrolyzed to formate and glutathione by FrmB. S-nitrosoglutathione, a product of nitrosyl stress is also a substrate for FrmA¹⁻³.

In examining global expression patterns caused by an amber suppressor mutant, transcript levels of the *frm* operon were found to increase when FrmR was translated with seven additional C-terminal amino acids⁴. Through this chance discovery, FrmR was proposed to regulate *frm* genes in response to formaldehyde, although this effect was not shown to be direct. This was the first proposed formaldehyde-responsive transcription factor, but AdhR from *B. subtilis* was subsequently identified as a formaldehyde- and methylglyoxal- dependent

regulator of the MerR family⁵ (Chapter 1). AdhR contains a conserved Cys residue that is important for the related regulator NmlR, which senses NO⁶, but the mechanism of interaction with aldehydes is unclear.

Formaldehyde is a very reactive and toxic molecule, capable of cross-linking proteins and modifying a variety of intracellular nucleophiles⁷⁻⁹.

Formaldehyde is a known by-product of methylated DNA repair by AlkB¹⁰, and *N*-methyltryptophan oxidase¹¹ in *E. coli*, so inducible mechanisms must be in place to prevent DNA damage. In this appendix the role of FrmR in the regulation of the *E. coli* response to aldehydes is examined. FrmR is found to directly bind to P_{frm} and repress transcription in the absence of aldehyde stress, and contain a metal cofactor that could be important for direct aldehyde sensing. Residues necessary for this response are also identified, shedding light on the mechanistic commonalities of stress sensing by the RcnR/CsoR family.

Results and Discussion

***P*_{frm} Is Induced by a Subset of Small Aldehydes**

Herring and Blattner have previously shown that FrmR is necessary for repression of *P*_{frm} in *E. coli*, and the operon is induced in the presence of formaldehyde⁴. To further explore the spectrum of molecules that activate expression, a *P*_{frm}-*frmR-lacZ* reporter fusion containing ~500 bp upstream of *frmA*, including *frmR* was constructed. Reporter activity was induced 28-fold (Figure A1.1a) in the presence of 700 μM formaldehyde. The role of FrmR in regulation was tested through mutation of codon six of *frmR* to a stop codon (*P*_{frm}-*frmR*_{stop}-*lacZ*) and measuring LacZ activity in a Δ *frmR* strain. In the absence of FrmR, constitutive high activity was observed, consistent with the idea that FrmR represses gene transcription in the absence of the proper signal (Figure A1.1b). Reporter activity without FrmR is significantly higher than maximal expression observed with formaldehyde, which could indicate that suboptimal conditions are being used for induction, or this could be an artifact of using a multi-copy reporter.

To better understand the specificity of *P*_{frm} induction, reporter activity was next assessed in the presence of a range of small aldehydes at the maximal concentrations that inhibited growth <15% to facilitate comparisons between molecules (Figure A1.2). Many of these compounds have been found to induce intracellular damage, indicating that they can cross the membrane¹²⁻¹⁴. Formaldehyde was found to induce the highest activity, but other aldehydes

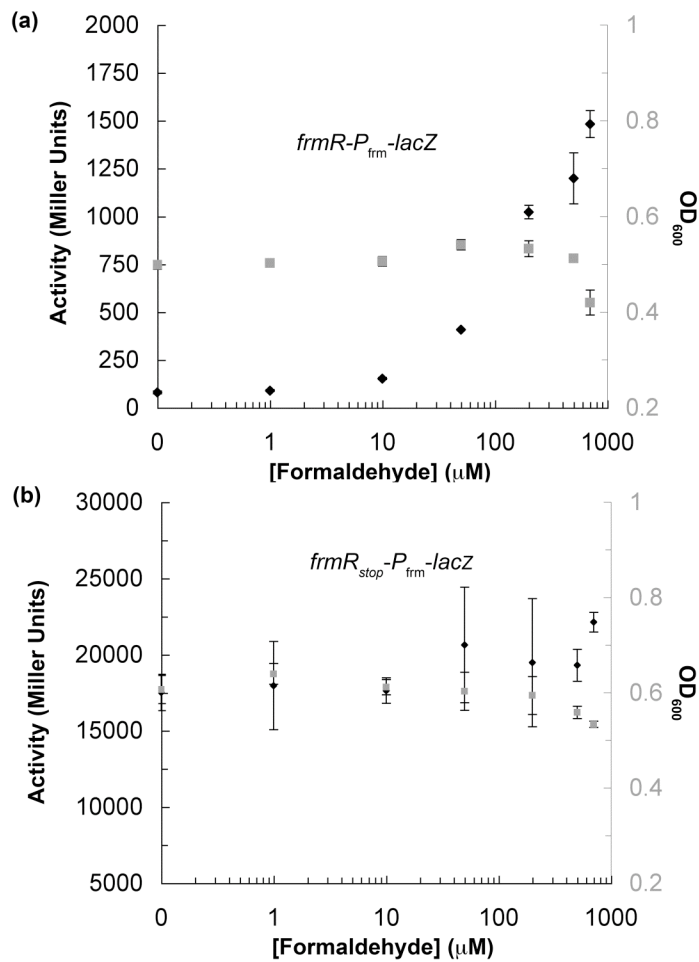


Figure A1.1. P_{frm} expression is FrmR dependent. P_{frm} -*frmR-lacZ* (a) or P_{frm} -*frmR_{stop}-lacZ* (b) expression was assayed anaerobically in $\Delta frmR$ cells. LacZ activity is shown in black diamonds, and the OD₆₀₀ is shown in gray squares. The wild type construct is induced 28-fold in the presence of 700 μ M formaldehyde, but the stop construct shows constitutive, high activity.

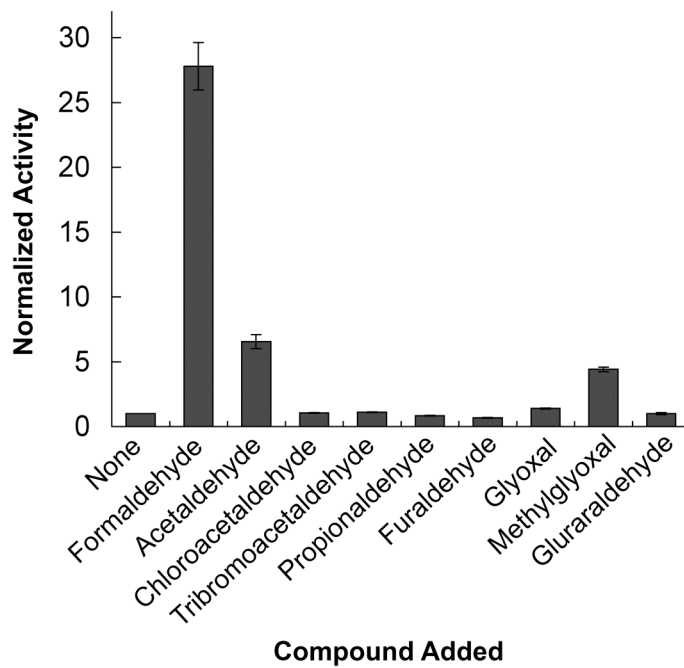


Figure A1.2. P_{frm} is induced by small aldehydes. P_{frm} -*frmR-lacZ* expression was assayed anaerobically as described in Materials and Methods. Activity was normalized to that observed in the absence of compound (73 Miller units). Compounds were present at maximal concentrations that inhibited growth by <15%. Error bars represent standard deviations.

showed a range of activities, with acetaldehyde, methylglyoxal, and glyoxal all inducing expression to lesser, but significant extents. In general bulky aldehydes such as furaldehyde and tribromoacetaldehyde were unable increase *lacZ* expression. While this inactivity could have many different sources, it suggests that there is a clear hierarchy for *frm* induction, and this could relate to the specificity of the detoxification machinery (FrmA and FrmB) as well as the sensory capabilities of the transcriptional regulators.

FrmR Directly Regulates P_{frm} Expression

To test if FrmR is directly regulating P_{frm} expression, FrmR was first overexpressed and purified by polyethyleneamine precipitation and cation exchange and gel filtration chromatographies. Purified FrmR was then assessed for DNA binding using a mobility shift assay with a ^{32}P -labeled DNA containing 189 bp upstream of *frmR*. Surprisingly, no binding was observed in 10 mM Hepes (pH 7.5), 150 mM NaCl, 2 mM MgCl_2 , 5% glycerol, 0.1% Nonidet P40, 2 $\mu\text{g/ml}$ salmon sperm DNA, 10 $\mu\text{g/ml}$ thioredoxin (Figure A1.3). The lack of activity could have been caused by use of the wrong DNA fragment, FrmR being unfolded or damaged during purification, or the absence of a necessary cofactor. Other members of the RcnR/CsoR family are known to bind metal ions, although in these cases metal binding inhibits DNA binding. FrmR has two conserved residues (Cys and His) at the fingerprint positions that were predicted to be involved in signal sensing for all RcnR/CsoR proteins, which suggested that FrmR could also bind metal ions. Intriguingly, with 10 μM ZnCl_2 in the gel,

binding buffer and running buffer FrmR formed a complex with P_{frm} with half-maximal activity observed at 56 nM FrmR. This surprising result suggested that Zn(II) can bind to FrmR and somehow induced a DNA-binding competent state. This is in direct contrast to RcnR, which only binds DNA in the absence of Ni(II)/Co(II). Given the physiological role of FrmR as an aldehyde sensor, one interpretation of this data is that Zn acts as an FrmR cofactor that somehow aids in signal sensing and is not a signal itself. Alternatively, Zn could prevent protein inactivation that occurs during electrophoresis by binding to specific residues. The metal specificity is unclear and should be probed in the future. The DNA fragment used in these studies contains a 19 bp inverted repeat containing six consecutive Gs (or Cs depending on the strand; Figure A1.3b). This is the likely site of FrmR interaction, given the prevalence of G-tracts in the binding sites of other RcnR/CsoR proteins (see Chapter 4), but the stoichiometry of the complex and the ability of FrmR to wrap DNA are not yet understood.

Glutathione is Not Necessary for FrmR Formaldehyde-sensing

The first step of formaldehyde detoxification is the spontaneous reaction of formaldehyde and glutathione. Because of the abundance of glutathione in the cell, this raises the question of if FrmR is responding to, formaldehyde itself, or a downstream product that is a glutathione adduct. To distinguish between these possibilities $P_{\text{frm}}\text{-frmR-lacZ}$ activity was assessed in a wild type and ΔgshA strain, which does not produce glutathione. If glutathione is necessary for FrmR

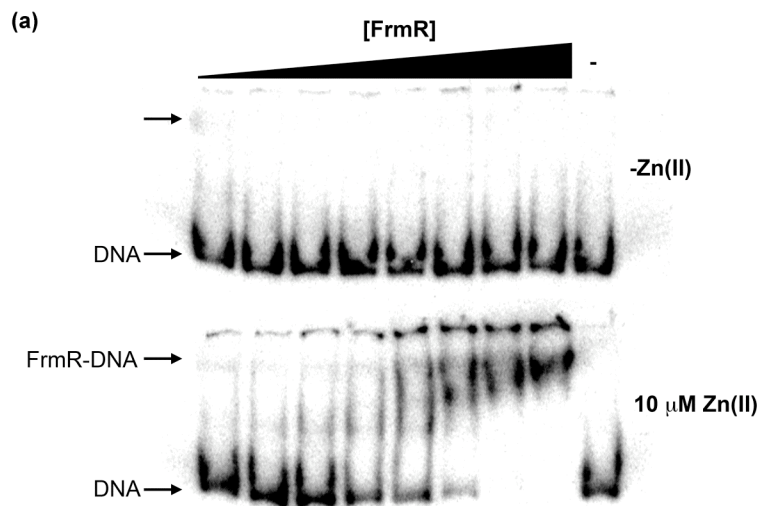


Figure A1.3. FrmR binds to P_{frm} in the presence of Zn(II). Mobility shift experiments were conducted with FrmR and a 189 bp fragment of P_{frm} (a). A shifted complex was only observed in the presence of 10 μ M ZnCl₂ in the binding reaction, gel, and running buffer. Protein concentrations ranged from 900 to 7 nM. A fragment of P_{frm} containing a possible G-rich binding site is shown in (b). The -10 and -35 elements are shown in gray boxes, the transcription start site is denoted with the arrow and inverted repeats are shown in white boxes.

inactivation, then LacZ activity should be diminished in the mutant strain in the presence of formaldehyde. However, as shown in Figure A1.4, activity is actually greater as a function of formaldehyde concentration when compared to the wild type strain. At 200 μ M formaldehyde the $\Delta gshA$ strain had 4.3-fold higher activity than wild type. The mutant strain was also more sensitive to formaldehyde as it did not grow in 500 μ M formaldehyde while wild type cells were not inhibited for growth. This experiment indicated that glutathione is not necessary for FrmR function and further suggested that FrmR may directly interact with formaldehyde, with glutathione acting as a competitor.

FrmR Directly Binds Zn(II)

Direct Zn(II) binding by FrmR was tested using a fluorescence competition assay with mag-fura 2, similar to the assay used to examine Ni(II) binding to RcnR. Mag-fura 2 binds to Zn(II) with a 1:1 stoichiometry which leads to a fluorescence enhancement at 480 nM. By following this signal in the presence and absence of FrmR, the protein affinity for Zn(II) can be measured. In the presence of 2 μ M FrmR the Zn(II) titration curve shifts to higher metal levels, indicating competition between the protein and the dye for Zn ions(Figure A1.5). Fitting the data to a 4-independent sites model yields an apparent affinity of $K_{d app} = 9 \pm 2$ nM using $K_d = 111$ nM for mag-fura 2. This affinity is significantly weaker when the titration is not done under reducing conditions ($K_{d app} = 146 \pm 32$ nM) suggesting that a cysteine residue is important for Zn binding. This residue must be extremely sensitive to oxidation as TCEP is removed from samples

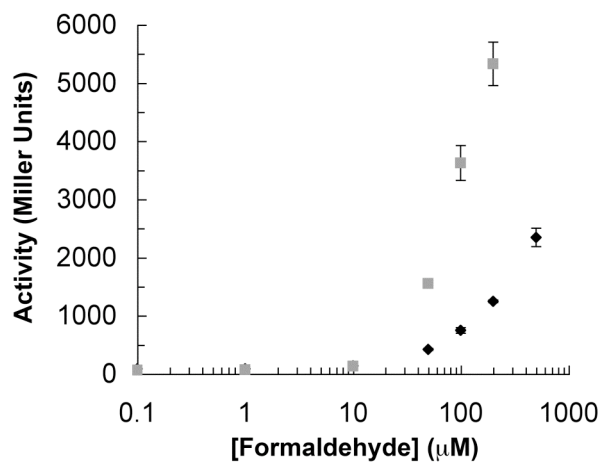


Figure A1.4. Glutathione is not necessary for FrmR formaldehyde sensing. BW25113 (black diamonds) or $\Delta gshA$ cells were grown anaerobically with varying formaldehyde concentrations and $P_{frm}-frmR-lacZ$ activity activity was assayed.

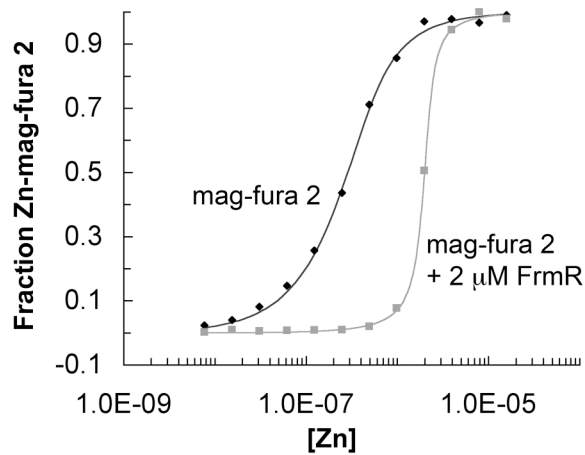


Figure A1.5. FrmR directly binds Zn(II). 500 nM mag-fura 2 (black diamonds) or 500 nM mag-fura 2 with 2 μ M FrmR (gray squares) were incubated with varying ZnCl₂ concentrations. The black line is a fit to a single binding site model with $K_d = 129$ nM. This number was used to obtain $K_{d,app} = 146 \pm 32$ nM for FrmR using a four independent sites model (gray).

immediately before the two hour incubation with Zn. The FrmR Zn affinity is weaker than those observed for other Zn proteins such as carbonic anhydrase ($K_d = 1 \text{ pM}$)¹⁵, but likely in the range of physiological relevance.

To gain structural insight into FrmR metal coordination, X-ray absorption spectroscopy (XAS) was carried out on Zn-FrmR in collaboration with Khadine Higgins of Michael Maroney's lab at University of Massachusetts-Amherst. Data was obtained and analyzed in a similar manner to Ni(II)/Co(II)-RcnR by the collaborator's lab and therefore are summarized here for simplicity. The best fit for Zn(II)-FrmR data was obtained with a model with a tetrahedral Zn(II) site with 2 N/O ligands at 2.00 Å (one of which is a His), 1 S ligand at 2.27 Å and 1 Br ligand from the solvent (Figure A1.6). Therefore, only three protein ligands form this metal site which is intriguingly similar to that of Zn-dependent enzymes such as human carbonic anhydrase II, which contains a tetrahedrally coordinated Zn(II) with three His ligands¹⁶. The fourth ligand is a hydroxide ion that is activated by the Zn(II) acting as a Lewis acid. This parallel suggests that the fourth ligand of FrmR will change depending on the solution conditions and could be important for interaction with formaldehyde and subsequent inhibition of DNA binding.

Identification of FrmR Residues Necessary for Formaldehyde Sensing

To identify FrmR residues necessary for formaldehyde sensing, site-directed mutants were made to the $P_{frm}\text{-}frmR\text{-}lacZ$ construct and tested for induction in the presence or absence of 250 μM formaldehyde. Under these

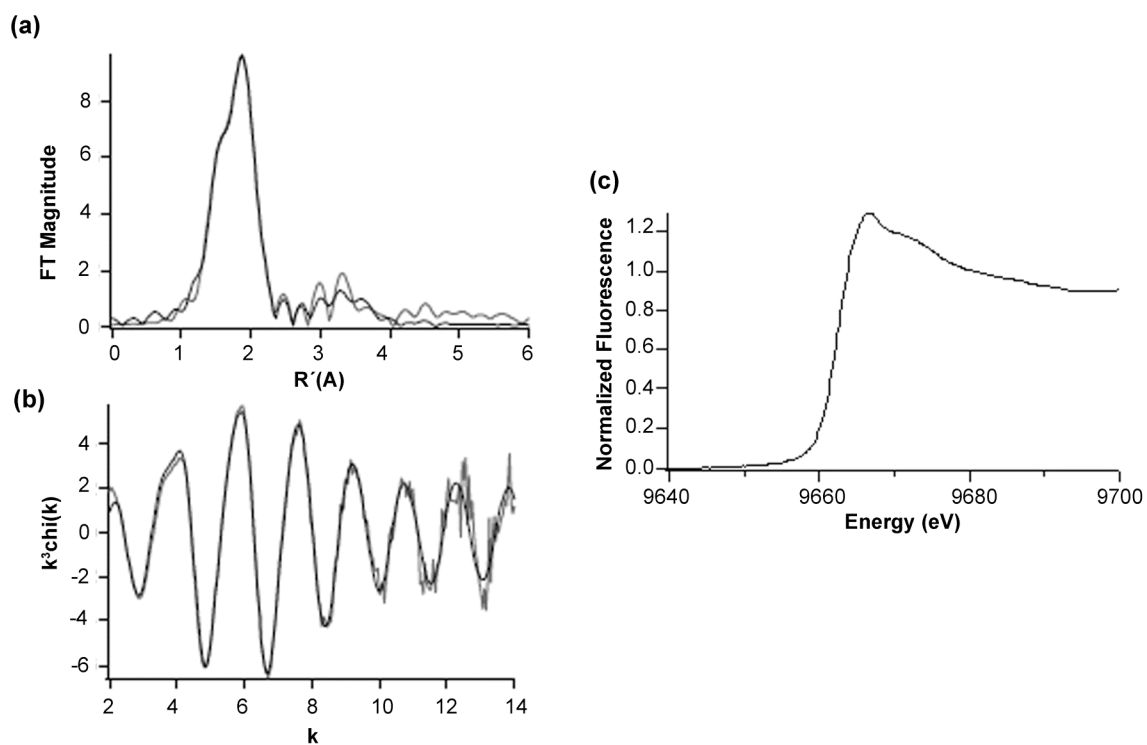


Figure A1.6. XAS of Zn-FrmR. Fourier transformed (a) and Fourier filtered (b) EXAFS spectrum of Zn-FrmR are shown in gray with best fits shown in black. The XANES spectrum is shown in (c). The spectra are consistent with a tetrahedral Zn site with 1S, 1NHis and 1N/O and 1Br ligand from the solvent.

conditions the wild type construct was induced 19-fold, while formaldehyde-insensitive strains should have constitutive, low LacZ activity. Three mutants demonstrated this behavior, A2* (Ala inserted at position 2), P2A and C35A (Figure A1.7). C35 is conserved in most RcnR/CsoR proteins and is necessary for Cu(I) sensing by CsoR and Co(II) sensing by RcnR. The N-terminal region was also important for RcnR, and appears to play a key role in FrmR function as well. Mutation of H60, another metal ligand for RcnR and CsoR, showed constitutive high activity, which could indicate that the protein is nonfunctional, but the cause of this is unclear, and other substitutions to probe the role of this residue will be required. The in vivo relationship between Zn(II) binding and aldehyde sensing is unclear, making it difficult to determine why these residues are important for protein function. It is tempting to speculate based on work on RcnR that C35, H60 and the N-terminus of the protein are the three Zn ligands, which is consistent with the XAS results presented above. C71 is the only other Cys in the protein, but mutation of this residue has little effect on function, suggesting that C35 accounts for the S-coordination observed by XAS. Mutation of T64, another “fingerprint” position, showed no effect on formaldehyde induction, indicating that not all of these key positions are important for FrmR function. The last “fingerprint” residue (S3) will be probed in the future.

Analysis of predicted bacterial FrmR proteins (those in operons with genes encoding formaldehyde dehydrogenases; 18 seqs) shows strict conservation of P2, C35 and H60. N-terminal sequencing analysis revealed that the N-terminal

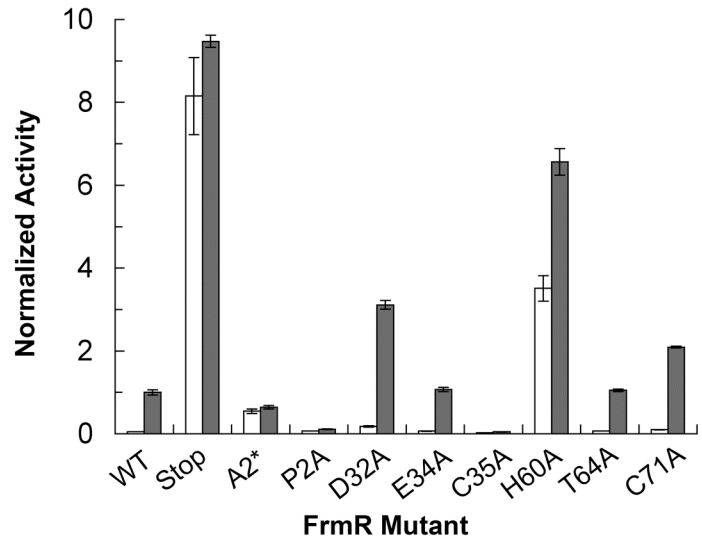
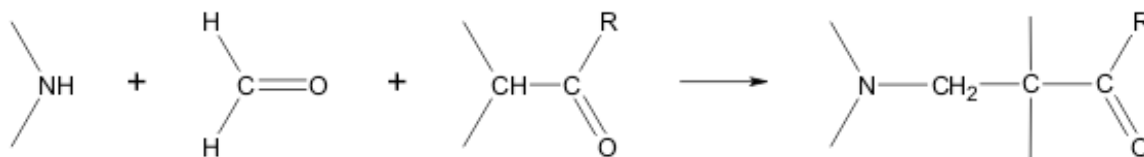


Figure A1.7. Identification of formaldehyde-insensitive FrmR mutants. The activity of $P_{frm}-frmR-lacZ$ or $P_{frm}-frmR_{mut}-lacZ$ were tested in $\Delta frmR$ cells without (white bars) or with (gray bars) 250 μ M formaldehyde as described in Materials and Methods.

methionine is cleaved from FrmR, leaving P2 as the first residue. In all other FrmR protein sequences, there are no N-terminal extensions, indicating that the first residue in each case is also likely to be Pro, and therefore the N-terminus itself could be important for protein function as observed for RcnR metal coordination. Combined with XAS, the mutagenesis and sequence conservation patterns suggest a model where Zn(II) is tetrahedrally coordinated by the N-terminus (Pro 2), C35, and possibly H60.

Overall these studies have provided a strong foundation to explore the mechanism of formaldehyde sensing by FrmR, and provide a system to further explore the similarities and differences in DNA binding observed for RcnR/CsoR proteins. The discovery that FrmR binds Zn(II) has uncertain implications for this mechanism, but suggests several models in which formaldehyde could change the metal coordination through covalent modification of a metal ligand, or be directly coordinated by Zn(II) at the position occupied by Br in the XAS samples. Several other stress sensors, including PerR and NorR are known to use metal ions as cofactors for signal sensing. PerR coordinates Fe(II), and in the presence of H₂O₂, Fenton chemistry produces localized OH· that oxidizes a coordinating His residue, decreasing DNA binding affinity and leading to expression of genes important for combating oxidative stress^{17,18}. Several studies with model peptides have shown N-termini, Lys, Arg, Trp and Cys can all be modified by formaldehyde, but this is sequence context-dependent, and subsequent cross-links can also be formed^{8,9}. In this context, it is possible that

for FrmR the N-terminus or C35 could be site of formaldehyde modification. A possible reaction mechanism for FrmR is suggested by the Mannich reaction^{19,20}, shown below.



First, a primary or secondary amine attacks the carbonyl group of formaldehyde, followed by dehydration to form a Schiff base. Under acidic conditions, an enol can then attack the iminium ion, forming a β -amino-carbonyl compound. For FrmR, Pro2 could be the secondary amine, but the identity of the functional group for the second nucleophilic attack is unclear.

Materials and Methods

All chemicals were purchased from Fisher Scientific (Pittsburg, PA) except where noted. Oligonucleotides were purchased from Invitrogen (Carlsbad, CA) or Operon Biotechnologies (Huntsville, AL), and are listed in Table A1.1. Plasmids and bacterial strains are listed in Table A1.2. All molecular biology reagents were from New England Biolabs (Beverly, MA).

Molecular Biology

The P_{frm} -*frmR-lacZ* plasmid was constructed from the pPC163 parent plasmid, which contains the P_{nik} promoter and *lacZ* cloned into pACYC184. The 499 bp upstream of *frmA*, including *frmR* were cloned into the *EagI*, *Sall* sites, replacing the P_{nik} cassette. *frmR* was amplified from *E. coli K12* genomic DNA with primers JI174 and JI175, and cloned into the pET22b at the *NdeI* and *NcoI* sites. The 700 bases upstream of *frmA* were amplified with JI180, JI181 primers, and cloned into pBSKSII at the *SacI*, *HindIII* sites, and used for subsequent amplification of FrmR binding substrates (pJI138). Mutants were constructed by multistep, overlapping PCR amplifications. For all cloning ligation steps, T4 DNA ligase was used.

FrmR Overexpression and Purification

FrmR was overexpressed in *E. coli* DL41 (DE3) cells with pFrmR grown in LB media containing 538 μ M ampicillin at 37 °C to an OD₆₀₀ of ~1.0. Cells were then induced with 0.5 M IPTG for ~3 hrs before centrifugation (4600g, 25 min) followed by resuspension in residual media and frozen at -80 °C. Thawed pellets

Table A1.1. Oligonucleotides.

Oligo	Purpose	Sequence
JI169	<i>P_{frm}-frmR-lacZ</i>	ctatgtcggccgattcctctgcgccgcccgtatccg
JI170	<i>P_{frm}-frmR-lacZ</i>	gtcatatgcgactctcgctcttctcaatatgg
JI174	pFrmR	ctatgtcatatgccagctactccggaagagaag
JI175	pFrmR	gtcataccatggctatttaagataggcacgaacc
JI176	<i>frmR</i> KO	gagaagaaaaaggtcctactcgagtcgctgtattcggggcagattgatgctct ggaagtgtaggctggagctgctc
JI177	<i>frmR</i> KO	cagttcaatagtgctgcaacggattggctgacttcgcggtgtagcagtcatttcg tcattccgggatccgctgacc
JI178	<i>P_{frm}-frmR_{stop}-lacZ</i>	gaaatgccagctactccgtaagagaagaaaaaggtccttac
JI178r	<i>P_{frm}-frmR_{stop}-lacZ</i>	gtaaggaccttttcttcttacggagtactggcatttc
JI179	FrmRC35A	ctggagggtgatgccgaagcccgctccatactccaacag
JI179r	FrmRC35A	ctgttgagatggcacgggctcggcatcacctccag
JI180	<i>P_{frm}</i>	ctatgtgagcttaagccggaacaggcttac
JI181	<i>P_{frm}</i>	gtcataaagcttatctcctccttctcaatatgg
JI203	FrmRC71A	cgttgaccgaaatgacgcctacagccggaagtacg
JI203r	FrmRC71A	gctgacttcgcggtgtaggcgtcatttcggtcaaacg
JI246	FrmRH60A	gcagaagtgctgaaagctgatccgggaaacggttgac
JI246r	FrmRH60A	gtcaaacgttcccggatcaagcttcaagcacttctgc
JI294	FrmRT64A	gaaagccatatccgggaagcgttgaccgaaatgactgc
JI294r	FrmRT64A	gcagtcatttcggtcaaacgctcccggatatggcttc
JI310	FrmRA2*	gatgagggtcgaaatggcggcagctactccggaag
JI310r	FrmRA2*	ctccggagtactggcgccatttcgcacctcatc
JI311	FrmRP2A	gatgagggtcgaaatggcagctactccggaagag
JI311r	FrmRP2A	ctctccggagtactggccatttcgcacctcatc
JI322	FrmRD32A	gaacggtcgctggagggtgctgccgaatgccgtgcc
JI322,r	FrmRD32A	ggcacggcattcggcagcaccctccagcaccgttc
JI323	FrmRE34A	ctggagggtgatccgcatgccgtgccatactc
JI323r	FrmRE34A	gagtatggcacggcatgcgcatcacctccag

Table A1.2. Plasmids and bacterial strains.

Plasmid	Insert	Parent	Reference
pFrmR	frmR	pET22b (amp ^R)	This work
pJI134	P _{frm} -frmR-lacZ	pACYC163 (cm ^R)	This work
pJI135	P _{frm} -frmR _{stop} -lacZ	pJI134 (cm ^K)	This work
pJI136	P _{frm} -frmRC35A-lacZ	pJI134 (cm ^K)	This work
pJI138	P _{frm}	pBluescript KSII (amp ^R)	This work
pJI156	P _{frm} -frmRH60A-lacZ	pJI134 (cm ^K)	This work
pJI157	P _{frm} -frmRC71A-lacZ	pJI134 (cm ^R)	This work
pJI175	P _{frm} -frmRP2A-lacZ	pJI134 (cm ^R)	This work
pJI176	P _{frm} -frmRE34A-lacZ	pJI134 (cm ^K)	This work
pJI177	P _{frm} -frmRA2*-lacZ	pJI134 (cm ^R)	This work
pJI178	P _{frm} -frmRD32A-lacZ	pJI134 (cm ^K)	This work
pJI179	P _{frm} -frmRT64A-lacZ	pJI134 (cm ^K)	This work

Strain	Genotype	Parent	Reference
RZ4500	ΔlacZ		21
PC677	ΔfrmRΔlacZ	RZ4500	This work
DL41	ΔmetA	MG1655	22
BW25113	ΔlacZ		23
ΔgshA	ΔgshA	BW25113	23

were lysed by sonication at 4 °C with 5 mM DTT and 1.5 mM PMSF. Next, to the clarified lysate, polyethyleneimine was added to 0.2% at 4 °C with stirring for 30 min. The pellet was harvested by centrifugation (27000g, 30 min) and resuspended in 20 mM Hepes (pH 8.0), 1 mM DTT, 1 mM EDTA, 5 % glycerol (Buffer F) with 500 mM NaCl at 4 °C by stirring for 30 min. The supernatant was then collected after centrifugation (27000g, 30 min) and precipitated with 75% ammonium sulfate. This pellet was harvested by centrifugation and resuspended in Buffer F with 50 mM NaCl, and dialyzed against the same buffer overnight at 20 °C to remove residual ammonium sulfate. The next day the dialysate was applied to a Q fast flow cation exchange column (3/4 in x 5 in) equilibrated with Buffer F with 50 mM NaCl. The column was then washed with 45 ml of the same buffer then proteins were eluted with a linear gradient of 50 mM to 1 M NaCl in Buffer F (120 ml total volume). FrmR eluted at ~450 mM NaCl, and these fractions were pooled, and applied to an 18 mL octylsepharose column equilibrated with Buffer F with 1 M NaCl. After washing with 40 ml of the same buffer, FrmR was eluted with 25 ml of Buffer F with 50 mM NaCl. NaCl was added to 300 mM to the elute, which was then concentrated to 1 ml with 2 ml 5 KDa MW cutoff Vivaspin concentrators (Vivascience, Hanover, Germany) by centrifugation (4000g) and applied to a HiLoad 26/60 Superdex 200 column equilibrated with Buffer F with 300 mM NaCl. Fractions containing FrmR were pooled and stored at 4 °C until further use.

β-Galactosidase Reporter Experiments

Starter cultures of PC677 ($\Delta frmR$) containing the appropriate reporter plasmid (wild type or mutants) were grown aerobically in LB media with chloramphenicol (0.1 mM) to an $OD_{600} \sim 0.5$. These cultures were then used to inoculate 1.87 ml of LB media with 0.1 mM chloramphenicol in capped microfuge tubes in triplicate with aldehyde compounds as indicated in figure legends. Cultures were grown overnight (12-15 hrs) and assayed as previously described. Maximal aldehyde concentrations used to inhibit growth by <15% were: 700 μ M formaldehyde, 7 mM acetaldehyde, 100 μ M chloroacetaldehyde, 11 μ M tribromoacetaldehyde, 4 mM propionaldehyde, 10 mM furaldehyde, 3.5 mM glyoxal, 350 μ M methylglyoxal and 500 μ M glutaraldehyde. Glutathione-dependent experiments were carried out in strain BW25113 or BW25113 $\Delta gshA$.

FrmR Mobility Shift Assays

Mobility shift binding reactions were set up in 10 mM HEPES (pH 7.5), 150 mM NaCl, 2 mM $MgCl_2$, 5% glycerol, 0.1% Nonidet P40, 2 μ g/ml salmon sperm DNA, 10 μ g/ml thioredoxin +/- 10 μ M $ZnCl_2$ where indicated and allowed to incubate for 30 min. Labeled DNA was generated by ^{32}P -labeling JI169 with T4 kinase and γ - ^{32}P [dATP] followed by PCR with JI171 using pJI138 as a template. Binding reactions were loaded on prerun 0.5x TB (25 mM Tris, 12.5 mM borate) 6% acrylamide gels running at 120 V with 0.5x TB running buffer. $ZnCl_2$ was present where indicated. Gels were visualized with a Storm phosphorimager (GE Healthcare, Piscataway, NJ).

FrmR Zn(II) Binding

Zn(II) affinity experiments were conducted in 10 mM Hepes (pH 7.5), 150 mM NaCl, 5% glycerol, with FrmR being desalted in the same buffer two times with Microbiospin 6 columns (Biorad) to remove DTT and EDTA. Separate samples contained 500 nM mag-fura 2 and the indicated Zn(II) concentrations with or without FrmR. Samples were allowed to equilibrate 3 h before scanning. Fluorescence measurements were obtained on a Cary Eclipse fluorescence spectrophotometer in a 150 μ l Hellma quartz cuvette, with slit widths = 5 nm, λ_{ex} = 320 nm and λ_{em} = 440-560 nm. Titrations were repeated 3 times.

References

1. Gonzalez, C.F. et al. Molecular basis of formaldehyde detoxification. Characterization of two S-formylglutathione hydrolases from *Escherichia coli*, FrmB and YeiG. *J. Biol. Chem* **281**, 14514-14522(2006).
2. Harms, N. et al. S-formylglutathione hydrolase of *Paracoccus denitrificans* is homologous to human esterase D: a universal pathway for formaldehyde detoxification? *J. Bacteriol* **178**, 6296-6299(1996).
3. Liu, L. et al. A metabolic enzyme for S-nitrosothiol conserved from bacteria to humans. *Nature* **410**, 490-494(2001).
4. Herring, C.D. & Blattner, F.R. Global transcriptional effects of a suppressor tRNA and the inactivation of the regulator frmR. *J Bacteriol* **186**, 6714-20(2004).
5. Nguyen, T.T.H. et al. Genome-wide responses to carbonyl electrophiles in *Bacillus subtilis*: control of the thiol-dependent formaldehyde dehydrogenase AdhA and cysteine proteinase YraA by the MerR-family regulator YraB (AdhR). *Mol. Microbiol* **71**, 876-894(2009).
6. Kidd, S.P. et al. NmlR of *Neisseria gonorrhoeae*: a novel redox responsive transcription factor from the MerR family. *Mol. Microbiol* **57**, 1676-1689(2005).
7. Craft, T.R., Bermudez, E. & Skopek, T.R. Formaldehyde mutagenesis and formation of DNA-protein crosslinks in human lymphoblasts in vitro. *Mutat. Res* **176**, 147-155(1987).

8. Toews, J. et al. Mass spectrometric identification of formaldehyde-induced peptide modifications under in vivo protein cross-linking conditions. *Anal. Chim. Acta* **618**, 168-183(2008).
9. Metz, B. et al. Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. *J. Biol. Chem* **279**, 6235-6243(2004).
10. Trewick, S.C. et al. Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. *Nature* **419**, 174-178(2002).
11. Koyama, Y. & Ohmori, H. Nucleotide sequence of the Escherichia coli solA gene encoding a sarcosine oxidase-like protein and characterization of its product. *Gene* **181**, 179-183(1996).
12. Dillon, D., Combes, R. & Zeiger, E. The effectiveness of Salmonella strains TA100, TA102 and TA104 for detecting mutagenicity of some aldehydes and peroxides. *Mutagenesis* **13**, 19-26(1998).
13. Mroczkowska, M.M., Kolasa, I.K. & Kusmierk, J.T. Chloroacetaldehyde-induced mutagenesis in Escherichia coli: specificity of mutations and modulation by induction of the adaptive response to alkylating agents. *Mutagenesis* **8**, 341-348(1993).
14. Zeiger, E., Gollapudi, B. & Spencer, P. Genetic toxicity and carcinogenicity studies of glutaraldehyde--a review. *Mutat. Res* **589**, 136-151(2005).
15. McCall, K.A. & Fierke, C.A. Probing determinants of the metal ion selectivity in carbonic anhydrase using mutagenesis. *Biochemistry* **43**, 3979-3986(2004).

16. Håkansson, K. et al. Structure of native and apo carbonic anhydrase II and structure of some of its anion-ligand complexes. *J. Mol. Biol* **227**, 1192-1204(1992).
17. Helmann, J.D. et al. The global transcriptional response of *Bacillus subtilis* to peroxide stress is coordinated by three transcription factors. *J. Bacteriol* **185**, 243-253(2003).
18. Lee, J. & Helmann, J.D. The PerR transcription factor senses H₂O₂ by metal-catalysed histidine oxidation. *Nature* **440**, 363-367(2006).
19. Cordova, A. et al. A Highly Enantioselective Route to Either Enantiomer of Both α - and β -Amino Acid Derivatives. *Journal of the American Chemical Society* **124**, 1866–1867(2002).
20. Mannich, C.;K. & Mannich, C.; Krosche, W. Ueber ein Kondensationsprodukt aus Formaldehyd, Ammoniak und Antipyrin. *Archiv der Pharmazie* **250**, 647–667(1912).
21. Choe, M. & Reznikoff, W.S. Anaerobically expressed *Escherichia coli* genes identified by operon fusion techniques. *J. Bacteriol* **173**, 6139-6146(1991).
22. Hendrickson, W.A., Horton, J.R. & LeMaster, D.M. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J* **9**, 1665-1672(1990).
23. Baba, T. et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol* **2**, 2006.0008(2006).

Appendix 2
Supplementary XAS Tables for Chapter 3

The following tables are in support of work in Chapter 3, and are the work of Sharon Leitch. They are published in Iwig, J.S. et al. *J Am Chem Soc.* 2008 130: 7592-7606.

Table A2.1. Selected EXAFS fits for Ni(II)-RcnR in NaBr.

Fit #	N ^a	r (Å) ^b	σ^2 (x10 ³ Å ⁻²) ^c	ΔE_0 (eV)	GOF ^d
Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å					
1	6 Ni-N/O	2.05(2)	10.4(17)	7.6(21)	0.66
2	6 Ni-S	2.15(4)	16.8(26)	-13.0(62)	1.00
3	5 Ni-N/O	2.07(12)	7.7(12)	10.4(15)	0.48
	1 Ni-N/O	2.34(350)	[7.7] ^e	[10.4]	
4	5 Ni-N/O	2.06(1)	8.1(7)	9.5(9)	0.31
	1 Ni-S	2.57(2)	8.2(27)	[9.5]	
5	4 Ni-N/O	2.05(1)	5.8(18)	10.4(16)	0.32
	1 Ni-N/O	2.17(8)	[5.9]	[10.4]	
	1 Ni-S	2.57(3)	8.4(28)	[10.4]	
Fourier-Filtered fits with a Backtransform Window = 1-4 Å					
6	5 Ni-N/O	2.06(1)	8.1(13)	9.7(14)	0.82
	1 Ni-S	2.55(3)	8.8(40)	[9.7]	
7	1 Ni-N _{imd}	1.69(16)	20.0(39)	4.4(27)	1.25
	5 Ni-N/O	2.02(3)	[20.0]	[4.4]	

8	2 Ni-N _{imd}	1.98(4)	6.7(33)	9.2(16)	0.87
	4 Ni-N/O	2.10(2)	[6.7]	[9.2]	
9	3 Ni-N _{imd}	2.01(2)	5.9(36)	10.7(14)	0.90
	3 Ni-N/O	2.14(3)	[5.9]	[10.7]	
10	1 Ni-N _{imd}	1.98(6)	6.2(23)	10.1(12)	0.69
	4 Ni-N/O	2.08(1)	[6.2]	[10.1]	
	1 Ni-S	2.55(3)	8.3(31)	[10.1]	
11	2 Ni-N _{imd}	1.99(2)	3.4(19)	10.6(9)	0.55
	3 Ni-N/O	2.12(1)	[3.4]	[10.6]	
	1 Ni-S	2.55(2)	7.5(22)	[10.6]	
12	3 Ni-N _{imd}	2.03(1)	2.8(14)	12.0(7)	0.56
	2 Ni-N/O	2.17(2)	[2.8]	[12.0]	
	1 Ni-S	2.56(2)	6.4(20)	[12.0]	
13	1 Ni-N _{imd}	1.88(6)	6.0(11)	8.1(17)	0.80
	4 Ni-N/O	2.05(1)	[6.0]		
	1 Ni-N/O	2.75(7)	[6.0]		
14	2 Ni-N _{imd}	1.88(3)	4.1(12)	7.9(18)	0.86
	3 Ni-N/O	2.05(1)	[4.1]	[7.9]	
	1 Ni-N/O	2.72(6)	[4.1]	[7.9]	

15	3 Ni-N _{imd}	2.02(1)	3.8(23)	10.7(12)	0.79
	2 Ni-N/O	2.15(3)	[3.8]	[10.7]	
	1 Ni-N/O	2.80(5)	[3.8]	[10.7]	
Unfiltered EXAFS fit from 2-12.5 Å ⁻¹					
16	1 Ni-N _{imd}	1.864(11)	6.7(4)	6.3(4)	0.87
	5 Ni-N/O	2.052(4)	[6.7]	[6.3]	
17	2 Ni-N _{imd}	1.970(9)	6.3(9)	8.6(4)	0.89
	4 Ni-N/O	2.095(5)	[6.3]	[8.6]	
18	3 Ni-N _{imd}	2.002(5)	6.1(4)	9.4(4)	0.94
	3 Ni-N/O	2.127(8)	[6.1]	[9.4]	
19	1 Ni-N _{imd}	1.961(14)	5.7(6)	8.9(4)	0.76
	4 Ni-N/O	2.077(3)	[5.7]	[8.9]	
	1 Ni-S	2.540(9)	8.5(11)	[8.9]	
20	2 Ni-N _{imd}	1.986(5)	3.3(6)	9.8(3)	0.70
	3 Ni-N/O	2.112(3)	[3.3]	[9.8]	
	1 Ni-S	2.543(7)	7.4(8)	[9.8]	
21	3 Ni-N _{imd}	2.024(3)	3.8(6)	11.1(3)	0.76
	2 Ni-N/O	2.155(8)	[3.8]	[11.1]	

	1 Ni-S	2.558(7)	6.7(8)	[11.1]	
22	1 Ni-N _{imd}	1.875(18)	6.1(4)	7.8(4)	0.82
	4 Ni-N/O	2.050(4)	[6.1]	[7.8]	
	1 Ni-N/O	2.748(21)	[6.1]	[7.8]	
23	2 Ni-N _{imd}	1.886(9)	4.2(4)	8.0(5)	0.90
	3 Ni-N/O	2.049(4)	[4.2]	[8.0]	
	1 Ni-N/O	2.721(18)	[4.2]	[8.0]	
24	3 Ni-N _{imd}	2.017(4)	4.5(8)	9.7(4)	0.88
	2 Ni-N/O	2.139(10)	[4.5]	[9.7]	
	1 Ni-N/O	2.782(19)	[4.5]	[9.7]	

^aN is the number of scattering atoms at a specific distance.

^br (Å) is the radial distance.

^cσ² is the root mean square disorder in the Ni-X distance, where X is the scattering atom.

^dGOF (goodness of fit) was calculated as described previously (Padden 2001) The accuracy of distance determined by EXAFS for atoms in the first coordination sphere of the metal are limited to ± 0.02 Å by the theoretical phase parameters; numbers in parentheses represent standard deviations from the least-squares fits.

^eValues of parameters in [] were constrained to be equal to those of other scattering atoms in the same shell.

Table A2.2. Selected EXAFS fits for C35A Ni(II)-RcnR in NaBr.

Fit #	N ^a	r (Å) ^b	σ^2 (x10 ³ Å ⁻²) ^c	ΔE_0 (eV)	GOF ^d
Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å					
1	5 Ni-N	1.87(1)	5.7(10)	2.5(27)	0.76
2	5 Ni-S	1.96(3)	11.7(16)	-22.8(64)	1.02
3	4 Ni-N/O	1.88(1)	4.1(5)	6.6(17)	0.43
	1 Ni-N/O	2.15(2)	[4.1] ^e	[6.6]	
4	3 Ni-N/O	1.85(23)	5.0(200)	2.7(48)	0.81
	2 Ni-N/O	1.90(48)	[5.0]	[2.7]	
5	4 Ni-N/O				No fit
	1 Ni-S				
Fourier-Filtered fits with a Backtransform Window = 1-4 Å					
6	4 Ni-N/O	1.88(1)	4.1(5)	6.7(17)	0.70
	1 Ni-N/O	2.15(3)	[4.1]	[6.7]	
7	1 Ni-N _{imd}	2.31(3)	4.1(5)	3.0(17)	0.65
	4 Ni-N/O	1.87(1)	[4.1]	[3.0]	
8	2 Ni-N _{imd}	1.83(3)	4.5(22)	4.8(18)	0.95
	3 Ni-N/O	1.90(2)	[4.5]	[4.8]	

9	1 Ni-N _{imd}				No fit
	3 Ni-N/O				
	1 Ni-Br				
Unfiltered EXAFS fit from 2-14.0 Å ⁻¹					
10	1 Ni-N _{imd}	2.316(8)	4.1(2)	4.4(4)	0.72
	4 Ni-N/O	1.876(2)	[4.1]	[4.4]	
11	2 Ni-N _{imd}	1.830(8)	4.5(6)	4.7(4)	1.00
	3 Ni-N/O	1.905(6)	[4.5]	[4.7]	
12	1 Ni-N _{imd}				No fit
	3 Ni-N/O				
	1 Ni-Br				

^aN is the number of scattering atoms at a specific distance.

^br (Å) is the radial distance.

^cσ² is the root mean square disorder in the Ni-X distance, where X is the scattering atom.

^dGOF (goodness of fit) was calculated as described previously (Padden 2001). The accuracy of distance determined by EXAFS for atoms in the first coordination sphere of the metal are limited to ± 0.02 Å by the theoretical phase parameters; numbers in parentheses represent standard deviations from the least-squares fits.

^eValues of parameters in [] were constrained to be equal to those of other scattering atoms in the same shell.

Table A2.3. Selected EXAFS fits for Ni(II)-RcnR in NaCl.

Fit #	N ^a	r (Å) ^b	σ^2 (x10 ³ Å ⁻²) ^c	ΔE_0 (eV)	GOF ^d
Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å					
1	6 Ni-N/O	2.05(2)	7.8(12)	7.5(20)	0.64
2	6 Ni-S	2.15(4)	16.2(24)	-13.4(59)	0.98
3	5 Ni-N/O	2.06(1)	7.4(12)	10.3(14)	0.47
	1 Ni-N/O	2.33(5)	[7.4] ^e	[10.3]	
4	5 Ni-N/O	2.06(1)	7.7(8)	9.3(11)	0.33
	1 Ni-S	2.56(2)	9.6(33)	[9.3]	
5	4 Ni-N/O	2.05(1)	5.7(22)	10.1(18)	0.36
	1 Ni-N/O	2.16(9)	[5.7]	[10.1]	
	1 Ni-S	2.57(3)	9.1(35)	[10.1]	
Fourier-Filtered fits with a Backtransform Window = 1-4 Å					
6	5 Ni-N/O	2.06(1)	7.8(11)	9.6(14)	0.79
	1 Ni-S	2.56(3)	10.1(47)	[9.6]	
7	1 Ni-N _{imd}	1.94(6)	7.5(2)	7.9(18)	0.85
	5 Ni-N/O	2.06(1)	[7.5]	[7.9]	
8	2 Ni-N _{imd}	1.98(3)	6.6(2)	8.9(16)	0.83
	4 Ni-N/O	2.09(2)	[6.6]	[8.9]	

9	3 Ni-N _{imd}	2.05(1)	5.0(2)	11.4(15)	0.87
	3 Ni-N/O	2.13(7)	[5.0]	[11.4]	
10	1 Ni-N _{imd}	1.98(5)	6.1(20)	9.9(12)	0.66
	4 Ni-N/O	2.08(1)	[6.1]	[9.9]	
	1 Ni-S	2.55(3)	9.6(27)	[9.9]	
11	2 Ni-N _{imd}	2.00(2)	4.4(16)	10.4(9)	0.55
	3 Ni-N/O	2.10(1)	[4.4]	[10.4]	
	1 Ni-S	2.55(2)	9.2(28)	[10.4]	
12	3 Ni-N _{imd}	2.03(1)	5.1(17)	11.3(9)	0.59
	2 Ni-N/O	2.13(2)	[5.1]	[11.3]	
	1 Ni-S	2.56(2)	8.7(29)	[11.3]	
13	1 Ni-N _{imd}	1.96(6)	5.2(2)	8.2(17)	0.75
	4 Ni-N/O	2.07(1)	[5.2]	[8.2]	
	1 Ni-N/O	2.79(7)	[5.2]	[8.2]	
14	2 Ni-N _{imd}	1.99(2)	3.7(2)	9.0(12)	0.67
	3 Ni-N/O	2.10(1)	[3.7]	[9.0]	
	1 Ni-N/O	2.80(5)	[3.7]	[9.0]	
15	3 Ni-N _{imd}				No fit

2 Ni-N/O

1 Ni-N/O

Unfiltered EXAFS fit from 2-12.5 Å⁻¹

16	1 Ni-N _{imd}	1.922(10)	7.1(6)	7.3(5)	1.03
	5 Ni-N/O	2.061(5)	[7.1]	[7.3]	
17	2 Ni-N _{imd}	1.972(10)	6.3(9)	8.2(5)	1.02
	4 Ni-N/O	2.087(5)	[6.3]	[8.2]	
18	3 Ni-N _{imd}	2.008(7)	7.2(11)	9.2(5)	1.06
	3 Ni-N/O	2.112(10)	[7.2]	[9.2]	
19	1 Ni-N _{imd}	1.967(17)	5.7(8)	10.6(4)	0.93
	4 Ni-N/O	2.072(4)	[5.7]	[10.6]	
	1 Ni-S	2.541(12)	8.6(13)	[10.6]	
20	2 Ni-N _{imd}	1.994(7)	4.3(7)	9.6(4)	0.89
	3 Ni-N/O	2.100(5)	[4.3]	[9.6]	
	1 Ni-S	2.545(11)	9.2(14)	[9.6]	
21	3 Ni-N _{imd}	2.029(5)	5.7(8)	10.6(4)	0.93
	2 Ni-N/O	2.124(11)	[5.7]	[10.6]	
	1 Ni-S	2.559(10)	[5.7]	[10.6]	

22	1 Ni-N _{imd}	1.946(16)	5.1(7)	7.3(5)	0.96
	4 Ni-N/O	2.065(4)	[5.1]	[7.3]	
	1 Ni-N/O	2.764(24)	[5.1]	[7.3]	
23	2 Ni-N _{imd}	1.984(7)	3.8(8)	8.4(4)	0.94
	3 Ni-N/O	2.096(5)	[3.8]	[8.4]	
	1 Ni-N/O	2.793(19)	[3.8]	[8.4]	
24	3 Ni-N _{imd}				No fit
	2 Ni-N/O				
	1 Ni-N/O				

^aN is the number of scattering atoms at a specific distance.

^br (Å) is the radial distance.

^cσ² is the root mean square disorder in the Ni-X distance, where X is the scattering atom.

^dGOF (goodness of fit) was calculated as described previously (Padden 2001.) The accuracy of distance determined by EXAFS for atoms in the first coordination sphere of the metal are limited to ± 0.02 Å by the theoretical phase parameters; numbers in parentheses represent standard deviations from the least-squares fits.

^eValues of parameters in [] were constrained to be equal to those of other scattering atoms in the same shell.

Table A2.4. Selected EXAFS fits for Co(II)-RcnR in NaBr.

Fit #	N ^a	r (Å) ^b	σ^2 (x10 ³ Å ⁻²) ^c	ΔE_0 (eV)	GOF ^d
Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å					
1	6 Co-N/O	2.05(3)	14.8(19)	-0.9(21)	0.44
2	6 Co-S	2.17(2)	20.5(14)	-16.2(27)	0.36
3	5 Co-N/O	2.05(200)	14.8(32)	-0.9(27)	0.48
	1 Co-N/O	2.06(1000)	[14.8] ^e	[-0.9]	
4	5 Co-N/O	1.97(2)	16.4(15)	-8.0(13)	0.16
	1 Co-S	2.25(1)	6.6(10)	[8.0]	
5	4 Co-N/O	1.97(83)	16.3(19)	-8.0(18)	0.19
	1 Co-N	1.98(330)	[16.3]	[-8.0]	
	1 Co-S	2.25(1)	6.6(11)	[-8.0]	
Fourier Filtered fits with a Backtransform Window = 1-4 Å					
6	5 Co-N/O	1.98(4)	16.5(30)	-7.3(31)	0.60
	1 Co-S	2.26(2)	6.9(23)	[-7.3]	
7	1 Co-N _{imd}	1.88(12)	12.7(77)	-5.4(31)	0.58
	4 Co-N/O	2.02(4)	[12.7]	[-5.4]	
	1 Co-S	2.27(3)	8.2(31)	[-5.4]	
	2 Co-C _{imd}	2.98(6)	7.8(65)	[5.4]	

8	2 Co-N _{imd}	1.94(3)	6.6(27)	-2.8(16)	0.51
	3 Co-N/O	2.11(3)	[6.6]	[-2.8]	
	1 Co-S	2.24(4)	13.1(60)	[-2.8]	
	4 Co-C _{imd}	3.01(4)	13.4(60)	[-2.8]	
9	3 Co-N _{imd}	1.97(2)	10.4(39)	-3.2(21)	0.59
	2 Co-N/O	2.15(8)	[10.4]	[-3.2]	
	1 Co-S	2.26(2)	7.1(28)	[-3.2]	
	6 Co-C _{imd}	3.01(5)	17.2(7)	[-3.2]	
10	1 Co-N _{imd}	1.90(6)	5.6(23)	-5.6(32)	0.59
	4 Co-N/O	2.10(1)	[5.6]	[-5.6]	
	1 Co-N/O	1.96(2)	[5.6]	[-5.6]	
	2 Co-C _{imd}	3.05(5)	7.6(5)	[-5.6]	
11	2 Co-N _{imd}	1.97(6)	6.2(51)	3.6(420)	0.67
	3 Co-N/O	2.12(2)	[6.2]	[3.6]	
	1 Co-N/O	2.30(15)	[6.2]	[3.6]	
	4 Co-C _{imd}	3.02(8)	5.1(50)	[3.6]	
Unfiltered EXAFS fit from 2-12.5 Å ⁻¹					
12	1 Co-N _{imd}	1.884(50)	12.7(31)	-5.2(12)	0.78
	4 Co-N/O	2.029(15)	[12.7]	[-5.2]	
	1 Co-S	2.270(13)	8.8(14)	[-5.2]	
	2 Co-C _{imd}	2.989(22)	7.6(26)	[-5.2]	

13	2 Co-N _{imd}	1.945(12)	6.9(13)	-3.0(7)	0.76
	3 Co-N/O	2.113(14)	[6.9]	[-3.0]	
	1 Co-S	2.238(18)	12.6(28)	[-3.0]	
	4 Co-C _{imd}	3.009(19)	13.5(28)	[-3.0]	
14	3 Co-N _{imd}	1.973(11)	9.9(14)	[-3.3]	0.80
	2 Co-N/O	2.157(29)	[9.4]	[-3.3]	
	1 Co-S	2.256(7)	7.0(13)	[-3.3]	
	6 Co-C _{imd}	3.009(18)	17.6(29)	[-3.3]	
15	1 Co-N _{imd}				No fit
	4 Co-N/O				
	1 Co-N/O				
	2 Co-C _{imd}				
16	2 Co-N _{imd}				No fit
	3 Co-N/O				
	1 Co-S				
	4 Co-C _{imd}				

^aN is the number of scattering atoms at a specific distance.

^br (Å) is the radial distance.

^cσ² is the root mean square disorder in the Ni-X distance, where X is the scattering atom.

^dGOF (goodness of fit) was calculated as described previously (Padden 2001). The accuracy of distance determined by EXAFS for atoms in the first coordination sphere of the metal are limited to ± 0.02 Å by the theoretical phase parameters; numbers in parentheses represent standard deviations from the least-squares fits.

^eValues of parameters in [] were constrained to be equal to those of other scattering atoms in the same shell.

Table A2.5. Selected EXAFS fits for C35A Co(II)-RcnR in NaBr.

Fit #	N ^a	r (Å) ^b	σ^2 (x10 ³ Å ⁻²) ^c	ΔE_0 (eV)	GOF ^d
Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å					
1	6 Co-N/O	1.92(3)	15.6(29)	-5.2(48)	0.62
2	6 Co-S	2.01(6)	20.5(34)	-30.6(110)	0.71
3	4 Co-N/O	1.92(2)	6.7(11)	-0.3(23)	0.43
	2 Co-N/O	2.12(3)	[6.7] ^e		
4	3 Co-N/O	2.13(2)	5.2(15)	4.3(25)	0.63
	3 Co-N/O	1.93(2)	[5.2]	[4.3]	
5	5 Co-N/O				No fit
	1 Co-S				
Fourier Filtered fits with a Backtransform Window = 1-4.0 Å					
6	4 Co-N/O	1.96(5)	15.6(60)	-2.2(8)	0.78
	2 Co-N/O	1.89(10)	[15.6]	[-2.2]	
7	1 Co-N _{ind}	1.83(11)	14.7(49)	-2.5(24)	0.66
	5 Co-N/O	1.95(2)	[14.7]	[-2.5]	
	2 Co-C _{ind}	2.89(5)	7.1(60)	[-2.5]	
8	2 Co-N _{ind}	1.87(7)	14.6(66)	-1.6(21)	0.66
	4 Co-N/O	1.97(4)	[14.6]	[-1.6]	

	4 Co-C _{imd}	2.89(5)	12.6(66)	[-1.6]	
9	3 Co-N _{imd}	1.89(4)	13.9(76)	-1.0(19)	0.68
	3 Co-N/O	2.10(6)	[13.9]	[-1.0]	
	6 Co-C _{imd}	2.90(5)	17.5(75)	[-1.0]	
Unfiltered EXAFS fit from 2-14.0 Å ⁻¹					
10	1 Co-N _{imd}	2.161(30)	12.1(6)	-0.8(6)	0.73
	5 Co-N/O	1.934(7)	[12.1]	[-0.8]	
	2 Co-C _{imd}	2.897(18)	7.6(23)	[-0.8]	
11	2 Co-N _{imd}	1.867(23)	14.3(22)	-1.6(6)	0.74
	4 Co-N/O	1.975(11)	[14.3]	[-1.6]	
	4 Co-C _{imd}	2.886(18)	13.2(24)	[-1.6]	
12	3 Co-N _{imd}	1.885(11)	12.8(23)	-1.2(6)	0.76
	3 Co-N/O	2.009(16)	[12.8]	[-1.2]	
	6 Co-C _{imd}	2.899(18)	18.9(29)	[-1.2]	

^aN is the number of scattering atoms at a specific distance.

^br (Å) is the radial distance.

^cσ² is the root mean square disorder in the Ni-X distance, where X is the scattering atom.

^dGOF (goodness of fit) was calculated as described previously (Padden 2001). The accuracy of distance determined by EXAFS for atoms in the first coordination sphere of the metal are limited to ± 0.02 Å by the theoretical phase parameters; numbers in parentheses represent standard deviations from the least-squares fits.

^eValues of parameters in [] were constrained to be equal to those of other scattering atoms in the same shell.

Table A2.6. Selected EXAFS fits for Co(II)-RcnR in NaCl.

Fit #	N ^a	r (Å) ^b	σ^2 (x10 ³ Å ⁻²) ^c	ΔE_0 (eV)	GOF ^d
Fourier-Filtered fits with a Backtransform Window = 1-2.3 Å					
1	6 Co-N/O	2.08(2)	12.7(18)	0.1(20)	0.50
2	6 Co-S	2.19(2)	18.6(14)	-16.3(13)	0.41
3	5 Co-N/O	2.09(2)	8.7(2)	-0.7(22)	0.47
	1 Co-N/O	1.92(6)	[8.7] ^e	[-0.7]	
4	5 Co-N/O	2.01(2)	16.1(15)	-6.5(17)	0.21
	1 Co-S	2.26(1)	5.6(10)	[-6.5]	
5	4 Co-N/O	2.00(2)	8.5(19)	-9.1(19)	0.20
	1 Co-N	1.82(4)	[8.5]	[-9.1]	
	1 Co-S	2.27(10)	4.3(14)	[-9.1]	
Fourier Filtered fits with a Backtransform Window = 1-4 Å					
6	5 Co-N/O	2.02(4)	16.0(28)	-5.7(31)	0.61
	1 Co-S	2.27(2)	6.1(2)	[-5.7]	
7	1 Co-N _{imd}	1.92(11)	11.4(60)	-3.6(31)	0.59
	4 Co-N/O	2.07(4)	[11.4]	[-3.6]	
	1 Co-S	2.28(3)	8.1(33)	[-3.6]	
	2 Co-C _{imd}	3.04(5)	6.9(59)	[-3.6]	

8	2 Co-N _{imd}	1.98(3)	6.9(33)	-1.7(19)	0.55
	3 Co-N/O	2.14(4)	[6.9]	[-1.7]	
	1 Co-S	2.25(3)	8.7(33)	[-1.7]	
	4 Co-C _{imd}	3.05(5)	14.1(70)	[-1.7]	
9	3 Co-N _{imd}	2.01(3)	9.6(43)	-1.8(21)	0.62
	2 Co-N/O	2.18(8)	[9.6]	[-1.8]	
	1 Co-S	2.27(1)	5.7(30)	[-1.8]	
	6 Co-C _{imd}	3.05(5)	18.6(80)	[-1.8]	
10	1 Co-N _{imd}	1.97(5)	5.8(23)	2.2(14)	0.70
	4 Co-N/O	2.12(1)	[5.8]	[2.2]	
	1 Co-N/O	2.30(15)	[5.8]	[2.2]	
	2 Co-C _{imd}	3.07(6)	3.4(39)	[2.2]	
11	2 Co-N _{imd}	2.02(2)	2.1(18)	2.8(11)	0.64
	3 Co-N/O	2.16(2)	[2.1]	[2.8]	
	1 Co-N/O	2.31(16)	[2.1]	[2.8]	
	4 Co-C _{imd}	3.00(8)	10.2(60)	[2.8]	
Unfiltered EXAFS fit from 2-12.5 Å ⁻¹					
12	1 Co-N _{imd}	1.902(44)	11.1(25)	-4.6(12)	0.82
	4 Co-N/O	2.057(18)	[11.1]	[-4.6]	
	1 Co-S	2.276(12)	7.5(12)	[-4.6]	
	2 Co-C _{imd}	3.026(24)	7.8(29)	[-4.6]	

13	2 Co-N _{imd}	1.968(14)	7.2(16)	-2.7(8)	0.81
	3 Co-N/O	2.131(17)	[7.2]	[-2.7]	
	1 Co-S	2.255(14)	8.7(14)	[-2.7]	
	4 Co-C _{imd}	3.033(22)	14.6(34)	[-2.7]	
14	3 Co-N _{imd}	1.999(12)	9.8(17)	-2.8(8)	0.85
	2 Co-N/O	2.169(30)	[9.8]	[-2.8]	
	1 Co-S	2.269(6)	5.8(11)	[-2.8]	
	6 Co-C _{imd}	3.037(22)	19.2(36)	[-2.8]	
15	1 Co-N _{imd}	1.971(18)	5.9(9)	2.2(6)	0.92
	4 Co-N/O	2.114(5)	[5.9]	[2.6]	
	1 Co-N/O	2.300(100)	[5.9]	[2.6]	
	2 Co-C _{imd}	3.060(23)	3.6(16)	-4.6(31)	
16	2 Co-N _{imd}				No fit
	3 Co-N/O				
	1 Co-N/O				
	4 Co-C _{imd}				

^aN is the number of scattering atoms at a specific distance.

^br (Å) is the radial distance.

^cσ² is the root mean square disorder in the Ni-X distance, where X is the scattering atom.

^dGOF (goodness of fit) was calculated as described previously (Padden 2001). The accuracy of distance determined by EXAFS for atoms in the first coordination sphere of the metal are limited to ± 0.02 Å by the theoretical phase parameters; numbers in parentheses represent standard deviations from the least-squares fits.

^eValues of parameters in [] were constrained to be equal to those of other scattering atoms in the same shell.

Appendix 3

Metal Binding by *Helicobacter pylori* HypA

A portion of this work was published as Kennedy, D.C., Herbst, R.W., Iwig, J.S., Chivers, P.T. and Maroney, M.J. A dynamic Zn site in *Helicobacter pylori* HypA: a potential mechanism for metal-specific protein activity. *J Am Chem Soc.* 2007 129: 16-17.

Results

H. pylori HypA is metallochaperone necessary for Ni insertion into both urease and Ni-Fe hydrogenase¹, which are required for colonization of the gastric mucosa in animal models²⁻⁴. To gain insight into how HypA specifically recognizes Ni(II) and delivers it to the necessary enzymes, HypA was purified and its Ni(II)-binding properties were investigated.

UV-Visible Spectroscopy

To test for metal binding by HypA, 150 μ M protein was incubated with one equivalent of metal chloride. As shown in Figure A3.1a, addition Ni(II) and Co(II) but not Mn(II) results in spectral features indicative of Me(II)-HypA complexes. The spectra for Ni(II) and Co(II) are indicative of a five- or six- coordinate site due to the extremely weak absorbance in the *d-d* transition region. Subsequent X-ray absorption spectroscopy on samples prepared under similar conditions revealed a six-coordinate site with 6 N/O ligands (1-2 of these are His) for Ni-HypA⁵. The absorbance feature at 269 nm for Ni-HypA was used to determine the stoichiometry of binding. Titration of NiCl₂ into 39 μ M HypA showed an increase in A_{269} that saturates at 0.8 Ni/HypA, indicating that the protein has 1 Ni-binding site per monomer (Figure A3.1b).

Fluorescence Spectroscopy

HypA has two tyrosine residues that can be used to probe conformational changes induced by metal binding. HypA (150 μ l of 17 μ M) was incubated for 2 h at 22 °C with increasing concentrations of metal chloride solution.

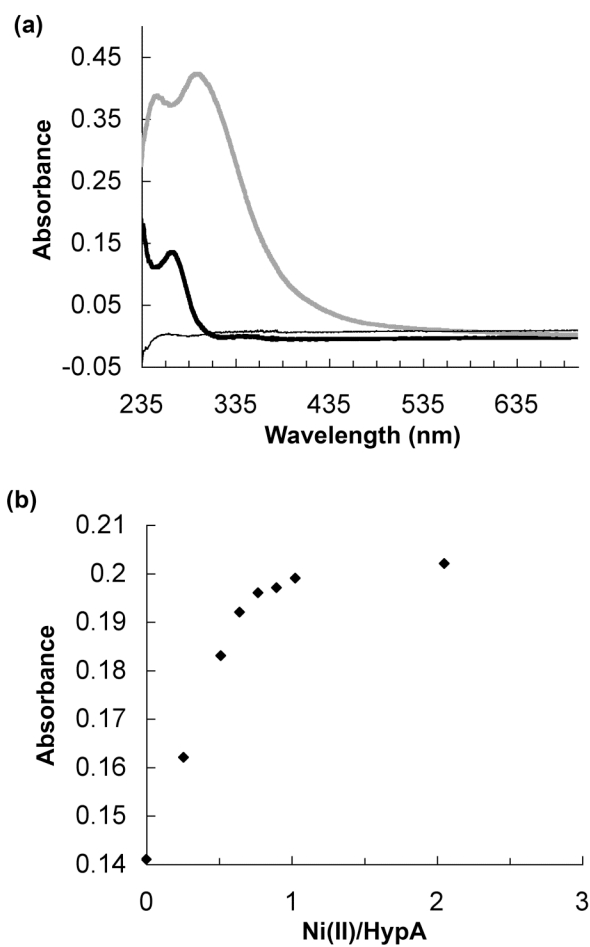


Figure A3.1. HypA binds metal ions. (a) HypA forms complexes with Ni(II) (thick black) and Co(II) (thick gray) but not Mn(II) (thin black). (b) The absorbance at 269 nm was followed after addition of Ni(II). The signal saturates near 1 Ni(II) equivalent.

Fluorescence spectra were obtained by excitation at 280 nm with the emission set at 302 nm. Ni(II) quenches the fluorescence by ~20% with the decrease saturating at a stoichiometry of ~1:1. Mn(II) does not lead to significant quenching (Figure A3.2).

HypA Stability

Protein stability was determined by monitoring the loss of secondary structure in the presence of urea by circular dichroism. Measurements used separate 900 μ l samples of 5 μ M protein in 20 mM Tris (pH 7.0), 150 mM NaCl with or without 5 μ M NiCl₂. Samples were incubated for 15 h at 22 °C before scanning. Ni binding led to a small increase in protein stability, as well as an increase in ellipticity in the absence of urea (Figure A3.3).

These preliminary studies showed that HypA directly binds one nickel ion per monomer at a six-coordinate site, and metal binding leads to a conformational change in the protein.

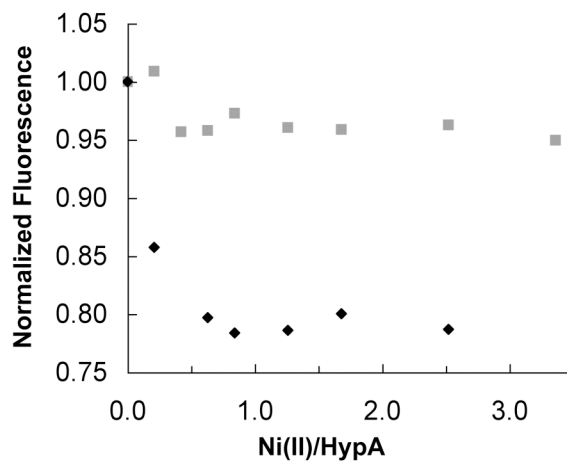


Figure A3.2. Ni(II) quenches intrinsic HypA fluorescence. Addition of Ni(II) (black diamonds) to 17 μ M HypA led to ~20% quenching of intrinsic tyrosine fluorescence. The fluorescence decreased until ~ 1 Ni(II) equivalent was added. Mn(II) did not lead to a signal change.

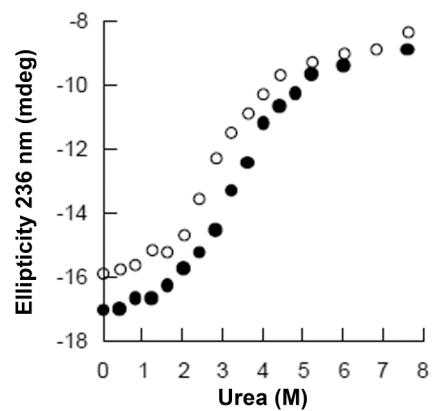


Figure A3.3. Ni(II) alters secondary structure and stability. 5 μ M HypA was incubated with (black) or without (white) 5 μ M NiCl₂ in urea, and the secondary structure content was assessed by circular dichroism.

Materials and Methods

Molecular Biology

The *hypA* gene was amplified from *H. pylori* 26695 genomic DNA using the primers JI101 and JI102 and *Klentaq* DNA polymerase. The resulting PCR fragment was then digested with restriction enzymes *Nde* I and *Xho* I and ligated into pET22b (Novagen, Madison, Wi) that had been digested with the same enzymes. The sequence of the *hypA* gene in the resulting plasmid was verified by DNA sequencing (Seqwright, Houston, TX).

JI101 CTATGTCATATGCATGAATACTCGGTCGTTTC

JI102 GTCATACTCGAGTTATTCCGCTAACATTTCTAAAG

Protein overexpression and purification

H. pylori HypA was overexpressed in *E. coli* DL41 (DE3) cells grown in LB to an OD of ~.9 before addition of 120 mg/l IPTG to induce gene expression. The cultures were then grown for an additional 3 h before harvesting by centrifugation (4600g, 25 min). Cell pellets were resuspended in the residual media and frozen at -80 C. Cell pellets were lysed by sonication in the presence of 5 mM DTT followed by centrifugation to remove the insoluble fractions. Initial purification was conducted using two ammonium sulfate cuts (25 and 75% w/v). In each case, the suspension was centrifuged after ammonium sulfate addition. After 25% saturation the supernatant was retained, and after 75% saturation the pellet

was harvested and dissolved in 20 mM Tris (pH 7.1) 50 mM NaCl 1 mM DTT. The resuspended pellet was dialyzed overnight against the same buffer at 4 C. The dialysate was applied to a Q Sepharose column equilibrated with 20 mM Tris (pH 7.1) 50 mM NaCl 1 mM DTT, and washed with four column volumes of the same buffer. For protein elution a linear gradient from 50 mM to 1 M NaCl was used, and HypA eluted at ~250 mM NaCl. Fractions containing HypA were concentrated to 2 ml using a 6 ml Vivaspin (Vivascience, Hanover, Germany) and run on a Sephacryl S-100 column using 20 mM Tris (pH 7.1) 150 mM NaCl 1 mM DTT. HypA eluted at the expected volume for a protein monomer.

References

1. Olson, J.W., Mehta, N.S. & Maier, R.J. Requirement of nickel metabolism proteins HypA and HypB for full activity of both hydrogenase and urease in *Helicobacter pylori*. *Mol. Microbiol* **39**, 176-182(2001).
2. Eaton, K.A. & Krakowka, S. Effect of gastric pH on urease-dependent colonization of gnotobiotic piglets by *Helicobacter pylori*. *Infect. Immun* **62**, 3604-3607(1994).
3. Eaton, K.A. et al. Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. *Infect. Immun* **59**, 2470-2475(1991).
4. Olson, J.W. & Maier, R.J. Molecular hydrogen as an energy source for *Helicobacter pylori*. *Science* **298**, 1788-1790(2002).
5. Kennedy, D.C. et al. A dynamic Zn site in *Helicobacter pylori* HypA: a potential mechanism for metal-specific protein activity. *J. Am. Chem. Soc* **129**, 16-17(2007).