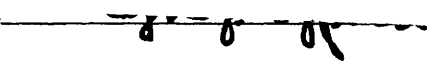





This doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:

Professor Stephen J. Lippard  \_\_\_\_\_  
Committee Chairman

Professor Catherine L. Drennan  \_\_\_\_\_  
Research Supervisor

Professor Robert T. Sauer  \_\_\_\_\_  
Committee Member

# Crystallographic Studies of the Metal-Responsive Transcription Factor NikR

By  
Eric R. Schreiter

Submitted to the Department of Chemistry in Partial Fulfillment of the Requirements for  
the Degree of Doctor of Philosophy in Biological Chemistry, June 2005

## ABSTRACT

Metal ion homeostasis is critical to the survival of all cells, because requirements for these essential nutrients must be balanced with their toxicity when present at elevated concentrations. Regulation of nickel concentrations in *Escherichia coli* is achieved via nickel-induced transcriptional repression of the nickel permease complex, NikABCDE, mediated by the NikR metalloregulatory transcription factor. To understand how this takes place at a molecular level, X-ray crystallography was used to determine three-dimensional structures of the NikR. Structures of apo-NikR and Ni<sup>2+</sup>-NikR show an interesting quaternary arrangement of the functional NikR tetramer, consisting of two dimeric Ribbon-Helix-Helix (RHH) DNA-binding domains separated by a tetrameric regulatory domain that binds metal ions (MBD). NikR is the only known metal-responsive member of the RHH family of transcription factors, and the structural relationships between NikR and other RHH proteins are detailed in this thesis. The high-affinity nickel-binding site of NikR is located at the tetramer interface of the MBD, with square-planar coordination by three histidines and a cysteine sidechain. A structure of NikR in complex with a 30 base-pair oligonucleotide containing the *nik* operator sequence illustrates the dramatic conformational changes associated with DNA binding, shows the nature of the specific and non-specific contacts between NikR and its operator, and reveals a second functionally relevant metal-binding site created at the interface of the RHH domain and MBD. Binding of nickel ions to the high-affinity site of NikR is proposed to increase affinity for the *nik* operator through a combination of mechanisms. Perhaps the most important of these is by ordering one face of the MBD of NikR, allowing productive interaction with the DNA phosphate backbone in the center of the operator. Additional metal ions further increase the affinity for operator by binding to the site situated at the RHH-MBD domain interface, and perhaps by binding to one of several sites on the surface of NikR identified by soaking crystals in excess Ni<sup>2+</sup>.

Thesis Supervisor: Catherine L. Drennan

Title: Associate Professor of Chemistry

## ACKNOWLEDGEMENTS

Many people contributed to my graduate education and helped to make the work described here possible:

First and foremost I would like to thank my advisor, Cathy Drennan, for her guidance, advice, and discussions, both scientific and otherwise. Thanks to her, my five years at MIT were enjoyable and memorable, and left me with a deeper passion to do meaningful science.

I would also like to thank my thesis committee, Prof. Steve Lippard and Prof. Bob Sauer, both of whom gave me helpful advice and provided forums to present my work for critical evaluation.

Thank you to all of the Drennan lab members, especially Luke Higgins for his friendship, intellectual discussions, and for making sure there was never a dull moment in the lab. Thanks to Mike Sintchak for his friendship and spending the time to teach me crystallography. Also thanks to Hector, Fred, Jess, Leah, Cintyu, Tzanko, Jeff, Yvain, Tina, Paul, Ainsley, and Stephanie for making the Drennan lab a fun and exciting place to come to work everyday.

My mother, Betsy, deserves many thanks for raising me well, believing in me, and supporting my dreams. Her pride in my accomplishments has always been a source of motivation for me.

Finally, I would like to thank the love of my life, Zarixia Zavala-Ruiz. Her friendship and support have opened many doors for me and made my graduate school experience far more than simply bearable. I look forward to a life with her filled with new experiences and joy.

TABLE OF CONTENTS	page
<b>Overview</b>	<b>8</b>
<b>Chapter I: Introduction</b>	<b>12</b>
<u>A. The importance of metals in biology</u>	12
<u>B. Nickel in biology</u>	13
<u>C. Nickel uptake in bacteria</u>	14
<u>D. The NikR protein</u>	15
<u>E. Structural mechanisms of ligand-regulated transcription factors</u>	19
<u>F. References</u>	21
<u>G. Figures</u>	25
<b>Chapter II: Crystal Structure of the Nickel-Responsive Transcription Factor NikR</b>	<b>32</b>
<u>A. Summary</u>	32
<u>B. Introduction</u>	33
<u>C. Results and Discussion</u>	34
1. Overall features of full-length apo-NikR	
2. Nickel binding	
3. MBD homology to other regulatory domains	
4. The Ribbon-Helix-Helix domain and operator recognition	
5. Implications for NikR function	
<u>D. Materials and Methods</u>	41
1. Protein expression and purification	
2. Crystallization and data collection	
3. Structure determination and model refinement	
<u>E. Acknowledgements</u>	43
<u>F. References</u>	44
<u>G. Tables and Figures</u>	47
<b>Chapter III: Crystal Structure of the NikR-Operator DNA Complex and the Mechanism of Repressor Activation by Metal Ions</b>	<b>56</b>
<u>A. Summary</u>	56
<u>B. Introduction</u>	57
<u>C. Results and Discussion</u>	59
1. Full-length nickel-activated NikR	
2. The NikR-operator complex	
3. NikR-operator interactions	
4. From zero to nanomolar affinity for DNA	
5. From nanomolar to picomolar affinity for DNA	

<u>D. Materials and Methods</u>	68
1. NikR-DNA complex formation	
2. Crystallization and data collection	
3. Structure solution and refinement	
4. NikR mutants and mobility shift assays	
<u>E. Acknowledgements</u>	73
<u>F. References</u>	73
<u>G. Tables and Figures</u>	77
<b>Chapter IV: Divalent transition metal ion binding to NikR</b>	<b>87</b>
<u>A. Summary</u>	87
<u>B. Introduction</u>	88
<u>C. Results and Discussion</u>	90
1. Crystal structure of Cu <sup>2+</sup> -MBD NikR	
2. Crystal structure of MBD NikR crystallized in the presence of Zn <sup>2+</sup> ions	
3. Additional Ni <sup>2+</sup> binding sites on NikR	
<u>D. Materials and Methods</u>	96
1. Crystal structure of Cu <sup>2+</sup> -MBD NikR	
2. Crystal structure of MBD NikR + Zn <sup>2+</sup>	
3. Excess Ni <sup>2+</sup> soak of Ni-NikR crystals	
4. Excess Ni <sup>2+</sup> soak of NikR-DNA crystals	
<u>E. Acknowledgements</u>	100
<u>F. References</u>	100
<u>G. Tables and Figures</u>	103
<b>Chapter V: NikR and the Ribbon-Helix-Helix Family</b>	<b>113</b>
<u>A. Summary</u>	113
<u>B. Introduction</u>	114
<u>C. Results and Discussion</u>	116
1. Sequence alignment of Ribbon-Helix-Helix proteins	
2. Structural alignment of Ribbon-Helix-Helix proteins	
3. DNA binding by Ribbon-Helix-Helix proteins	
<u>D. Materials and Methods</u>	121
1. Sequence alignment of Ribbon-Helix-Helix proteins	
2. Structural alignment of Ribbon-Helix-Helix proteins	
<u>E. References</u>	122
<u>F. Tables and Figures</u>	125
<b>Chapter VI: Afterword</b>	<b>137</b>
<u>A. Implications of this work</u>	137
<u>B. Future experiments</u>	137

<u>C. References</u>	139
<b>Chapter VII: Appendices</b>	<b>140</b>
<u>A. Protocols</u>	140
1. NikR plasmids	
2. Preparation of competent <i>Escherichia coli</i> cells and plasmid transformation	
3. Overexpression of NikR or MBD in <i>E. coli</i> strain DL41	
4. Overexpression of selenomethionine (SeMet) labeled NikR in DL41	
5. Purification of NikR or MBD for purification	
a. Immobilized nickel-affinity chromatography	
b. Size exclusion chromatography	
6. Quantitation of Ni <sup>2+</sup> :NikR ratio using PAR	
7. HPLC purification of deoxyoligonucleotides for co-crystallization with NikR	
<u>B. Crystal structure of the R33M NikR mutant</u>	149
1. Results and discussion	
2. Materials and methods	
<u>C. Two additional crystal forms of apo-NikR</u>	151
1. Results and discussion	
2. Materials and methods	
<u>D. References</u>	154
<u>E. Tables and Figures</u>	156
<u>E. Curriculum Vitae</u>	169

## Overview of this thesis

This thesis describes X-ray crystallographic experiments directed towards gaining a better understanding of the molecular mechanism of nickel-mediated repression of *nikABCDE* transcription by the *Escherichia coli* NikR protein. Chapter 1 provides the reader with a brief introduction to the importance of metal ions, specifically nickel, in biology. It also describes biochemical and biophysical characterizations of the NikR protein that have been carried out by other laboratories prior to and during our investigations. Finally, it gives a short overview of what is known regarding the structural mechanisms of ligand-regulated transcription factors.

In Chapter 2 we report the first two crystal structures of NikR to be determined, that of full-length apo-NikR and the nickel-bound isolated metal-binding domain (MBD). We learned that the NikR tetramer is composed of two dimeric Ribbon-Helix-Helix (RHH) DNA-binding domains (DBDs) flanking a central, tetrameric MBD, and that this separation of the two RHH domains explains how NikR can bind an operator sequence with sub-sites separated by two turns of DNA. The Ni<sup>2+</sup>-MBD structure also provided the first view of the high-affinity nickel-binding site of NikR, coordinated in a square-planar geometry by histidine and cysteine sidechains across the tetramer interface. This work was published, essentially as written here, in the October, 2003 issue of *Nature Structural Biology*, now called *Nature Structural and Molecular Biology*.

Chapter 3 describes two additional crystal structures of NikR, the full-length nickel-bound NikR tetramer and the complex between NikR and a 30 base-pair double-stranded oligonucleotide containing the *nik* operator sequence. The full-length Ni<sup>2+</sup>-NikR structure confirms that the structure of the MBD and the high-affinity nickel binding site are the same in the context of the full-length protein. It also shows a unique



conformation of the RHH domains with respect to the MBD, illustrating the conformational flexibility of NikR. In the structure of the NikR-operator DNA complex, the RHH domains have undergone a more dramatic rotation, pointing now towards the same face of NikR and creating an interface with the MBD where a putative potassium ion from the crystallization buffer is bound. This new metal-binding site involves coordination by two carboxylate sidechains of the RHH domain. Mutagenesis of these residues, done by Sheila Wang in Dr. Deborah Zamble's laboratory, supports the idea that this site is important for the operator-binding function of NikR. The NikR-DNA complex also reveals contacts to the slightly curved operator DNA from both the RHH domain and MBD. This work has been prepared in manuscript format and will be submitted for publication in the near future.

Chapter 4 represents a work in progress towards understanding the ability of NikR to bind divalent transition metal ions. A structure of the MBD of NikR in complex with  $\text{Cu}^{2+}$  ions shows the metal binding at the same position as, and with a similar coordination to  $\text{Ni}^{2+}$ . Another structure of the MBD, crystallized in the presence of  $\text{Zn}^{2+}$ , shows zinc ions putatively bound to a site adjacent to the  $\text{Ni}^{2+}/\text{Cu}^{2+}$  site, with two protein sidechain ligands in common. Further work will be required to confirm this assignment, however, as the low pH (5.6) and low occupancy of the zinc ion in the structure call its validity into question. Also described in this chapter, crystals of either  $\text{Ni}^{2+}$ -NikR or the NikR-DNA complex were soaked in solutions containing millimolar concentrations of  $\text{Ni}^{2+}$  ions in an attempt to observe a secondary nickel binding site on NikR that has been observed biochemically to substantially increase the affinity of NikR for its operator when occupied. Four unique nickel-binding sites were located in both crystals, and

provide a basis for future mutagenesis experiments to determine their relative importance. This work has been prepared as a manuscript such that it could be published pending additional work to characterize the metal ion binding sites that have been identified.

As mentioned previously, the amino-terminal DNA-binding domain of NikR belongs to the Ribbon-Helix-Helix (RHH) structural superfamily, which binds DNA via specific contacts from  $\beta$ -strands of the protein to the nucleotide bases exposed in a DNA major groove. [Chapter 5](#) reviews the important structural aspects of the RHH family and compares NikR to the other structurally characterized members. A structure-aided sequence alignment also allows comparison with putative RHH sequences for which no structural information is available. This chapter was written such that it could eventually be submitted for publication as a structural review of the RHH family.

A brief afterword is provided as [Chapter 6](#), to put into context for the reader what was accomplished in this work, what questions have been raised as a result, and what types of experiments could be done in the future to learn more about this system.

Finally, the appendices in [Chapter 7](#) are intended to provide future experimenters with detailed protocols that were used during this work, and to catalog works in progress that are not yet part of a larger story.

### **Some abbreviations used in this thesis:**

ABC: ATP-Binding Casette

ADA: N-(2-(Acetamido)imino)diacetic Acid

AdoMet: S-Adenosylmethionine

CD: Circular Dichroism

CSD: Cambridge Structure Database

EDTA: Ethylenediaminetetraacetic acid

EGTA: Ethylene glycol-bis(2-aminoethyl)-N,N,N',N'-tetraacetic acid

LB: Luria-Bertani media

MBD: Metal-Binding Domain of NikR

MSA: Gel-Mobility Shift Assay

PAR: 4-(2-Pyridylazo)resorcinol

PDB: Protein DataBank

PEG: Polyethylene Glycol

RHH: Ribbon-Helix-Helix

RMSD: Root Mean Square Deviation

SDS-PAGE: Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

TEA: Triethyl amine

XAS: X-ray Absorption Spectroscopy

## Chapter I: Introduction

### I.A. The importance of metals in biology

Metal ions play a variety of roles in biological systems. By themselves, they serve to carry charge, such as in the generation of an action potential in a neuron, and to maintain the osmotic balance of cells<sup>1</sup>. As cofactors of proteins and enzymes, metal ions serve to stabilize functional protein folds, reversibly bind and transport dioxygen, shuttle electrons, and promote chemical reactions that would otherwise be extremely unfavorable. I hope to illustrate in this thesis how metal ions can also serve as a specific signal to regulate the production of proteins at the transcriptional level.

Despite their many uses within cells, metal ions, particularly transition metal ions, can be harmful if present in greater than normal abundances<sup>1</sup>. They can catalyze the formation of hydrogen peroxide and other damaging reactive oxygen species, or can bind to non-specific sites on proteins and nucleic acids, either preventing the normal function of the molecule or creating a new reactivity that may be detrimental to the cell. Fortunately, a number of mechanisms exist to regulate the intracellular concentrations of metal ions such that they are maintained at productive levels. There are specialized metal-binding proteins such as the metallothioneins<sup>2</sup>, capable of chelating excess free metal ions, and integral membrane efflux proteins<sup>3</sup>, which pump metal ions out of the cell. Alternatively, some organisms control concentrations of specific metals by regulating production of the proteins responsible for their uptake from the environment. All of these homeostatic mechanisms can be controlled at the transcriptional level by metalloregulatory transcription factors, proteins capable of transducing the signal of

metal binding to an increase or decrease in gene expression. Structures of a number of metalloregulatory transcription factors are illustrated in Fig. I.1. Recently, strong arguments have been made against intracellular pools of “free” copper<sup>4</sup> or zinc<sup>5</sup> ions, both of which employ metalloregulatory transcription factors with exceptional sensitivity to their respective metals. It seems likely that other biologically important transition metal ions are maintained in a similar fashion.

Metals that do not usually have a use in our bodies, such as platinum, gold, lithium, technetium, and gadolinium, have found use as therapeutics or contrast agents for medical imaging<sup>1</sup>. The most talked about metal compound of recent is cisplatin, an anticancer drug that has received press because of its successful use in the treatment of American cyclist Lance Armstrong’s testicular cancer.

### **I.B. Nickel in biology**

Nickel is found ubiquitously throughout the environment at low to moderate concentrations, primarily in the Ni<sup>2+</sup> oxidation state<sup>6</sup>. Nickel is an essential trace element for the growth a number of bacteria, archaea, and plants, and studies have shown depressed growth rates of several mammals on nickel-deficient diets<sup>6</sup>. Although the reasons for the beneficial effects of nickel in mammals are not understood, the molecular details of the nickel requirements of bacteria have been well studied. To date, at least eight nickel-dependent enzymes have been identified and characterized from bacteria and archaea, representing a wide range of interesting chemistry<sup>7</sup>. Among them are carbon monoxide dehydrogenase (CODH), which is capable of interconverting the gases CO and CO<sub>2</sub>, acetyl-coenzyme A synthase (ACS), which can convert CO gas into a metabolically

available carbon source, NiFe-hydrogenase, which catalyzes the reversible oxidation of hydrogen gas, and methyl-coenzyme M reductase, which evolves methane. Considering these reactions, it is clear why bacteria and archaea that metabolize these gases have a nickel requirement for their growth. All eight of the characterized nickel-dependent enzymes have been studied structurally, either by X-ray crystallographic or NMR methods. The nickel cofactor-containing active site of each of these enzymes is represented in Fig. I.2. This figure shows that nickel cofactors can be complicated, sometimes requiring post-translational modification of the protein scaffold and assembly of a multinuclear metallocluster. The maturation processes of the urease and NiFe-hydrogenase enzymes are the most extensively studied and appear to include the involvement of nickel metallochaperones (UreE and HypB, respectively) capable of specifically inserting nickel into the correct position within the enzyme active site<sup>8</sup>, a feature they may have in common with many complex metalloenzymes.

### **I.C. Nickel uptake in bacteria**

Bacteria can acquire nickel via the low-affinity, high-capacity magnesium transport system, CorA, when nickel is present at very high concentrations<sup>9</sup>. However, under physiological conditions where magnesium concentrations are high and nickel concentrations are generally quite low (low nanomolar in sea and fresh water), nickel uptake via CorA is not relevant and more specific high-affinity transporters are required<sup>10</sup>. Two classes of high-affinity nickel importers have been identified in bacteria. The first is an ATP-Binding Cassette (ABC)-type transporter<sup>11</sup>, exemplified by the NikABCDE complex in *E. coli*<sup>12</sup>, which need nickel for incorporation into the three

NiFe-hydrogenase isozymes for growth under anaerobic conditions. This complex consists of a soluble periplasmic nickel-binding protein (NikA), two integral membrane subunits (NikB and NikC), and two cytoplasmic ATP-binding cassette subunits (NikD and NikE), which couple hydrolysis of ATP to a conformational change in the membrane-spanning subunits, allowing transport (Fig. I.3). Crystal structures of NikA were reported recently<sup>13</sup>, and NikBCDE can be thought of as generally similar on a structural level to the B<sub>12</sub> transporter complex BtuCD<sup>14</sup>. Transcriptional regulation of the *nik* operon in *E. coli* is mediated by the NikR protein in a nickel-dependent manner (Fig. I.3). The second class of transporter is represented by HoxN from *R. eutropha*, a single-component high-affinity nickel permease<sup>15,16</sup>. Nickel transport systems in bacteria have been reviewed in more detail recently<sup>7,17</sup>.

#### **I.D. The NikR protein**

When transposon mutants of *E. coli* with significantly reduced hydrogenase activity were isolated<sup>10,18</sup> and the disrupted *nik* locus was sequenced<sup>12</sup>, it was found to encode the ABC-type transporter complex NikABCDE, described earlier. The authors noticed that increased concentrations of nickel repressed transcription of the *nik* operon, and that an inverted repeat sequence flanked the putative -10 region of the *nik* promoter. These two pieces of information hinted at the existence of a transcription factor capable of regulating transcription of the *nik* operon dependent upon nickel concentrations. Subsequently, an open reading frame (ORF) immediately downstream of NikE was found to encode a 133 amino acid protein responsible for the nickel-dependent regulation of *nik* transcription, and was termed *nikR*<sup>19</sup>.

Since its discovery, the NikR protein from *E. coli* has been characterized extensively using biochemical and biophysical methods. An alignment of *E. coli* with homologous protein sequences is shown in Fig. I.4. NikR was identified as a member of the Ribbon-Helix-Helix (RHH) family of DNA-binding proteins using sequence profile-based database searching methods<sup>20</sup>. A detailed structural comparison of NikR with other RHH family members is provided in chapter V of this thesis. Treatment of purified NikR with the protease elastase produced a stable N-terminal fragment approximately the size of a RHH protein. Furthermore, this N-terminal domain was found to fold as a dimer, have the  $\alpha$ -helical content expected of a RHH protein, and to weakly bind DNA. Mutation of Arg3 to Ala, a residue predicted from homology with other RHH members to be important for DNA binding, abolished the DNA-binding activity of the N-terminal NikR fragment. The relation of NikR to other members of the RHH family is reviewed in detail in Chapter 5 of this thesis.

Sequence alignment of the C-terminal portion of NikR revealed seven conserved amino acid sidechains with the capability of coordinating nickel ions. Addition of nickel ions to purified NikR or the isolated C-terminal domain (hereafter referred to as the metal-binding domain, or MBD) revealed a 1:1 Ni<sup>2+</sup>:NikR binding stoichiometry and produced additional characteristic features at 302 nm and 460 nm in the UV-Vis absorption spectrum<sup>21</sup>. The nickel-bound protein is a faint yellow color when concentrated in solution. An X-ray absorption spectroscopy (XAS) study showed that this nickel binding site in NikR had a square-planar coordination geometry, with the nickel coordinated by one S and three N/O ligands<sup>22</sup>. Their results also suggest that this coordination geometry changes to a six-coordinate site with loss of thiolate ligation in the



presence of specific DNA. Although this result is not supported by our structural work, it is possible that differences in sample preparation could account for the discrepancy between the studies (see Chapter 3). Nickel binding to the MBD of NikR was associated with an increase in secondary structure content, as assayed by circular dichroism (CD), and a change in the shape of the protein, as monitored using analytical ultracentrifugation (AUC)<sup>23</sup>. Additionally, NikR is tetrameric, either in the absence or presence of nickel ions. By titrating nickel into a solution of NikR in the presence of EGTA to keep the free nickel ion concentration low, and monitoring the absorbance at 302 nm, a dissociation constant ( $K_d$ ) of nickel for NikR was measured to be 1-2 pM<sup>23,24</sup>. Other divalent transition metal ions such as Cu<sup>2+</sup>, Zn<sup>2+</sup>, Co<sup>2+</sup>, and Cd<sup>2+</sup> bind NikR with the same stoichiometry and similar affinities, implying that the high-affinity metal binding site of NikR is not selective for nickel<sup>24</sup>.

DNase footprinting of an oligonucleotide containing the sequence upstream of the *nikA* gene in the presence of NikR and stoichiometric nickel ions revealed that NikR protects a region encompassing the inverted repeat sequence and -10 region of the *nik* promoter mentioned earlier<sup>23,25</sup>. Titration of Ni-NikR into the footprinting assay solution gave a  $K_d$  of 5-30 nM for the *nik* operator. Oddly, despite this high affinity for the *nik* operator, no shift of this DNA sequence is observed in a gel-mobility shift assay (MSA) after incubation with Ni-NikR. A shift is seen, however, if the gel and running buffer contain 35  $\mu$ M NiCl<sub>2</sub>. Titrating NikR into this MSA reveals a  $K_d$  of 10-20 pM for the *nik* operator in the presence of excess nickel ions, a value significantly tighter than that obtained from the DNase footprinting assay with stoichiometric Ni-NikR. This result hints at the existence of secondary nickel site(s) on NikR in the presence of DNA.

Indeed, two independent techniques were successful at determining a  $K_d$  of  $\text{Ni}^{2+}$  for this secondary site, although the values obtained differ significantly. In one study, an extended footprint was seen in the DNase assay in the presence of excess nickel, and this feature was exploited to calculate a  $K_d$  of  $\sim 30 \mu\text{M}$  for this secondary site<sup>23</sup>. A second study, however, failed to reproduce this extended footprint but instead lowered the concentration of  $\text{Ni}^{2+}$ -NikR below the  $K_d$  with stoichiometric nickel and titrated in excess  $\text{NiCl}_2$  to regain the original DNase footprint<sup>25</sup>. Using this method, the authors determined a  $K_d$  for the secondary nickel site of 30 nM, three orders of magnitude different from the value obtained in the original study. Chapter III of this thesis describes structural evidence for the existence of a secondary metal binding site on NikR in the presence of operator DNA. Stoichiometric amounts of divalent transition metal ions such as  $\text{Cd}^{2+}$ ,  $\text{Co}^{2+}$ , and  $\text{Cu}^{2+}$  are capable of activating NikR to bind specifically to DNA with an affinity similar to that of  $\text{Ni}^{2+}$ . The response of NikR to excess metal ions is somewhat specific for nickel, however, because excess nickel produced the highest affinity for specific DNA relative to other metals<sup>25</sup>.

Finally, it was estimated that under anaerobic conditions in Luria-Bertani (LB) media, *E. coli* contain  $\sim 125$  molecules of NikR per cell, corresponding to an intracellular concentration of  $\sim 200 \text{ nM}$  using a cell volume of  $1 \times 10^{-15} \text{ L}$ <sup>23</sup>. This concentration of NikR with stoichiometric nickel bound would result in a substantial occupancy of the *nik* operator, given the measured  $K_d$  of 5-30 nM. It remains to be seen whether occupancy of both nickel binding sites and the associated increased operator affinity is physiologically relevant.

Some recent work has focused on characterization of the roles of the NikR ortholog from the human gastric pathogen, *Helicobacter pylori* (HpNikR). This organism requires significant amounts of nickel for incorporation into the active sites of both NiFe-hydrogenase and urease. The latter catalyzes the hydrolysis of urea to ammonia, increasing the pH of the bacterium's local environment and allowing it to colonize the acidic lining of the human stomach. The result is an increased risk of developing gastric ulcers.

First, it was demonstrated that HpNikR was responsible for inducing urease expression in that organism<sup>26</sup>, a role that contrasts with the repressor function performed by the NikR from *E. coli*. Recently, transcriptome analysis of a deletion mutant revealed that *H. pylori* NikR regulates transcription of not only the nickel importer and urease enzyme, but also other regulatory networks, through repression of the gene encoding the ferric uptake regulator (Fur) protein<sup>27,28</sup>.

### **I.E. Structural mechanisms of ligand-regulated transcription factors**

Ligand-regulated transcription factors represent a large and functionally important class of DNA-binding proteins that are capable of regulating metabolic pathways in response to the presence or absence of a small molecule (usually) ligand. NikR belongs to the metalloregulatory sub-class of this group and is capable of regulating nickel uptake in a nickel-dependent manner. Our results described in this thesis, including structures of NikR in the apo, ligand-bound, and DNA-bound states, have allowed us to propose a detailed mechanism of ligand-activated transcriptional repression (see chapter 3) for this

system. For only a small number of these ligand-regulated transcription factors is an equivalent amount of structural information available (see below).

The only other metalloregulatory transcription factor for which structures of the apo, metal-bound, and DNA bound states are available is the DtxR/IdeR family (Fig. I.1.A)<sup>29-33</sup>. For this family of iron-activated regulators, binding of metal ions to three unique sites is able to stabilize the dimeric form of the protein and produce a subtle conformational change that places the two DNA-binding domains (DBDs) of the dimer at the correct spacing to productively interact with two operator sub-sites along a DNA molecule. This seems to be a general theme among dimeric ligand-regulated transcription factors that bind operator DNA sequences containing two sub-sites. Similar mechanisms have been proposed for the TetR<sup>34</sup>, FadR<sup>35</sup>, and Trp<sup>36,37</sup> repressor transcription factors, where binding of the ligand forces a conformational change in the dimeric protein that alters the orientation of the DBDs relative to each other and leads to a strong increase or decrease in affinity for the operator DNA. In the case of the well-studied lactose repressor, LacI, ligand binding clearly reorganizes the interface between the DNA-proximal subdomains of the dimer, but exactly how that change is propagated to the DBDs remains a matter of debate as the DBDs are generally disordered in the absence of DNA<sup>38-40</sup>. Interestingly, ligand binding to PurR, a structural homolog of LacI, at the same structural location has exactly the opposite effect on operator affinity<sup>41</sup>.

Finally, the methionine repressor, MetJ, which has the same RHH DNA-binding domain fold as NikR, is activated by the co-repressor S-adenosylmethionine (AdoMet) via what appears to be a purely electrostatic mechanism. Structures of the apo, ligand-bound, and DNA-bound forms of the repressor show no significant difference in protein

conformation<sup>42,43</sup>. Furthermore, computational studies demonstrated that the charge increase at the DNA-binding surface of MetJ due to AdoMet binding could account for the 1000-fold increase in affinity of this repressor for its operator<sup>44</sup>.

## I.F. References

1. Lippard, S.J. & Berg, J.M. *Principles of bioinorganic chemistry*, xvii, 411 (University Science Books, Mill Valley, Calif., 1994).
2. Henkel, G. & Krebs, B. Metallothioneins: zinc, cadmium, mercury, and copper thiolates and selenolates mimicking protein active site features--structural aspects and biological implications. *Chem Rev* **104**, 801-24 (2004).
3. Rosen, B.P. Transport and detoxification systems for transition metals, heavy metals and metalloids in eukaryotic and prokaryotic microbes. *Comp Biochem Physiol A Mol Integr Physiol* **133**, 689-93 (2002).
4. Changela, A. et al. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* **301**, 1383-7 (2003).
5. Outten, C.E. & O'Halloran, T.V. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488-92 (2001).
6. Ankel-Fuchs, D. & Thauer, R.K. Nickel in Biology: Nickel as an Essential Trace Element. in *The Bioinorganic chemistry of nickel* (ed. Lancaster, J.R.) 93-110 (VCH, New York, N.Y., 1988).
7. Mulrooney, S.B. & Hausinger, R.P. Nickel uptake and utilization by microorganisms. *FEMS Microbiol Rev* **27**, 239-61 (2003).
8. Kuchar, J. & Hausinger, R.P. Biosynthesis of metal sites. *Chem Rev* **104**, 509-25 (2004).
9. Smith, R.L. & Maguire, M.E. Microbial magnesium transport: unusual transporters searching for identity. *Mol Microbiol* **28**, 217-26 (1998).
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**, 1709-18 (1989).
11. Locher, K.P. Structure and mechanism of ABC transporters. *Curr Opin Struct Biol* **14**, 426-31 (2004).

12. 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. *Molecular Microbiology* **9**, 1181-91 (1993).
13. Heddle, J., Scott, D.J., Unzai, S., Park, S.Y. & Tame, J.R. Crystal structures of the liganded and unliganded nickel-binding protein NikA from *Escherichia coli*. *J Biol Chem* **278**, 50322-9 (2003).
14. Locher, K.P., Lee, A.T. & Rees, D.C. The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* **296**, 1091-8 (2002).
15. Eitinger, T. & Friedrich, B. Cloning, nucleotide sequence, and heterologous expression of a high-affinity nickel transport gene from *Alcaligenes eutrophus*. *J Biol Chem* **266**, 3222-7 (1991).
16. Eberz, G., Eitinger, T. & Friedrich, B. Genetic determinants of a nickel-specific transport system are part of the plasmid-encoded hydrogenase gene cluster in *Alcaligenes eutrophus*. *J Bacteriol* **171**, 1340-5 (1989).
17. Eitinger, T. & Mandrand-Berthelot, M.A. Nickel transport systems in microorganisms. *Arch Microbiol* **173**, 1-9 (2000).
18. Wu, L.F. & Mandrand-Berthelot, M.A. Genetic and physiological characterization of new *Escherichia coli* mutants impaired in hydrogenase activity. *Biochimie* **68**, 167-79 (1986).
19. De Pina, K., Desjardin, V., Mandrand-Berthelot, M.A., Giordano, G. & Wu, L.F. Isolation and characterization of the nikR gene encoding a nickel-responsive regulator in *Escherichia coli*. *Journal of Bacteriology* **181**, 670-4 (1999).
20. Chivers, P.T. & Sauer, R.T. NikR is a ribbon-helix-helix DNA-binding protein. *Protein Sci* **8**, 2494-500 (1999).
21. Chivers, P.T. & Sauer, R.T. Regulation of high affinity nickel uptake in bacteria. Ni<sup>2+</sup>-Dependent interaction of NikR with wild-type and mutant operator sites. *Journal of Biological Chemistry* **275**, 19735-41 (2000).
22. Carrington, P.E., Chivers, P.T., Al-Mjeni, F., Sauer, R.T. & Maroney, M.J. Nickel coordination is regulated by the DNA-bound state of NikR. *Nature Structural Biology* **10**, 126-130 (2003).
23. Chivers, P.T. & Sauer, R.T. NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA. *Chemistry and Biology* **9**, 1141-8 (2002).
24. Wang, S.C., Dias, A.V., Bloom, S.L. & Zamble, D.B. Selectivity of metal binding and metal-induced stability of *Escherichia coli* NikR. *Biochemistry* **43**, 10018-28 (2004).

25. Bloom, S.L. & Zamble, D.B. Metal-selective DNA-binding response of *Escherichia coli* NikR. *Biochemistry* **43**, 10029-38 (2004).
26. van Vliet, A.H. et al. NikR mediates nickel-responsive transcriptional induction of urease expression in *Helicobacter pylori*. *Infect Immun* **70**, 2846-52 (2002).
27. Contreras, M., Thiberge, J.M., Mandrand-Berthelot, M.A. & Labigne, A. Characterization of the roles of NikR, a nickel-responsive pleiotropic autoregulator of *Helicobacter pylori*. *Mol Microbiol* **49**, 947-63 (2003).
28. van Vliet, A.H.M., Ernst, F.D. & Kusters, J.G. NikR-mediated regulation of *Helicobacter pylori* acid adaptation. *Trends in Microbiology* **12**, 489-494 (2004).
29. Wisedchaisri, G., Holmes, R.K. & Hol, W.G. Crystal structure of an IdeR-DNA complex reveals a conformational change in activated IdeR for base-specific interactions. *J Mol Biol* **342**, 1155-69 (2004).
30. Pohl, E., Holmes, R.K. & Hol, W.G. Crystal structure of a cobalt-activated diphtheria toxin repressor-DNA complex reveals a metal-binding SH3-like domain. *J Mol Biol* **292**, 653-67 (1999).
31. Feese, M.D., Ingason, B.P., Goranson-Siekierke, J., Holmes, R.K. & Hol, W.G. Crystal structure of the iron-dependent regulator from *Mycobacterium tuberculosis* at 2.0-Å resolution reveals the Src homology domain 3-like fold and metal binding function of the third domain. *J Biol Chem* **276**, 5959-66 (2001).
32. White, A., Ding, X., vanderSpek, J.C., Murphy, J.R. & Ringe, D. Structure of the metal-ion-activated diphtheria toxin repressor/tox operator complex. *Nature* **394**, 502-6 (1998).
33. 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**, 22420-7 (1998).
34. Orth, P., Schnappinger, D., Hillen, W., Saenger, W. & Hinrichs, W. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat Struct Biol* **7**, 215-9 (2000).
35. van Aalten, D.M., DiRusso, C.C. & Knudsen, J. The structural basis of acyl coenzyme A-dependent regulation of the transcription factor FadR. *Embo J* **20**, 2041-50 (2001).
36. Zhang, H. et al. The solution structures of the trp repressor-operator DNA complex. *J Mol Biol* **238**, 592-614 (1994).
37. Zhang, R.G. et al. The crystal structure of trp aporepressor at 1.8 Å shows how binding tryptophan enhances DNA affinity. *Nature* **327**, 591-7 (1987).

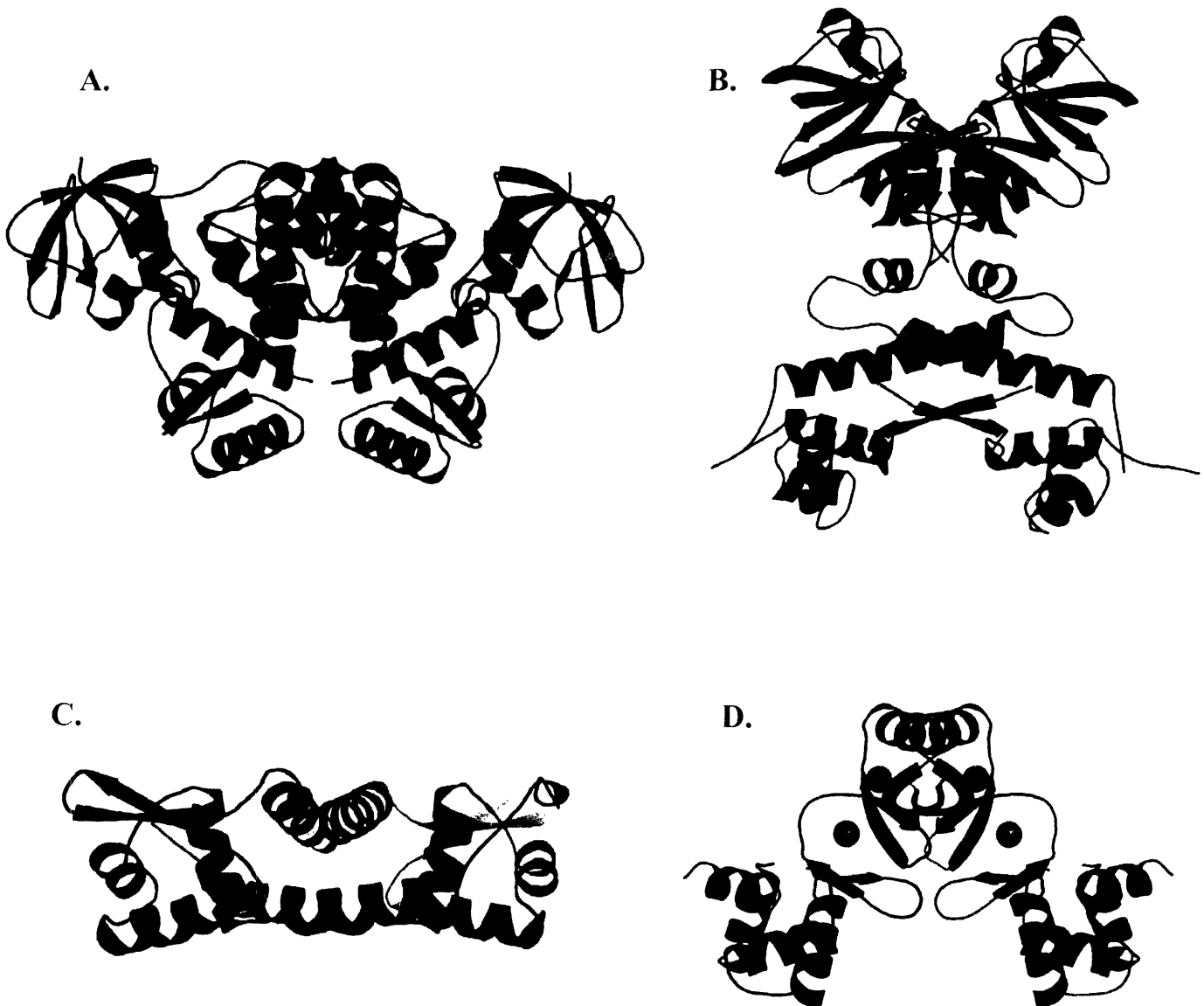
38. Sauer, R.T. Lac repressor at last. *Structure* **4**, 219-22 (1996).
39. Lewis, M. et al. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**, 1247-54 (1996).
40. Bell, C.E. & Lewis, M. The Lac repressor: a second generation of structural and functional studies. *Curr Opin Struct Biol* **11**, 19-25 (2001).
41. Schumacher, M.A., Choi, K.Y., Zalkin, H. & Brennan, R.G. Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. *Science* **266**, 763-70 (1994).
42. Somers, W.S. & Phillips, S.E. Crystal structure of the met repressor-operator complex at 2.8 Å resolution reveals DNA recognition by beta-strands. *Nature* **359**, 387-93 (1992).
43. Rafferty, J.B., Somers, W.S., Saint-Girons, I. & Phillips, S.E. Three-dimensional crystal structures of Escherichia coli met repressor with and without corepressor. *Nature* **341**, 705-10 (1989).
44. Phillips, K. & Phillips, S.E. Electrostatic activation of Escherichia coli methionine repressor. *Structure* **2**, 309-16 (1994).



## I.G. Figures

**Figure I.1.** Structures of prokaryotic metal-responsive transcription factors. Proteins are represented as ribbons colored by subunit. Nickel ions are shown as cyan spheres, other metal ions as green spheres. The name of each transcription factor and physiological functions regulated are given below. Notably, each is a dimer that uses a Helix-Turn-Helix (HTH) motif to specifically interact with DNA except NikR. Figure produced using PyMOL.

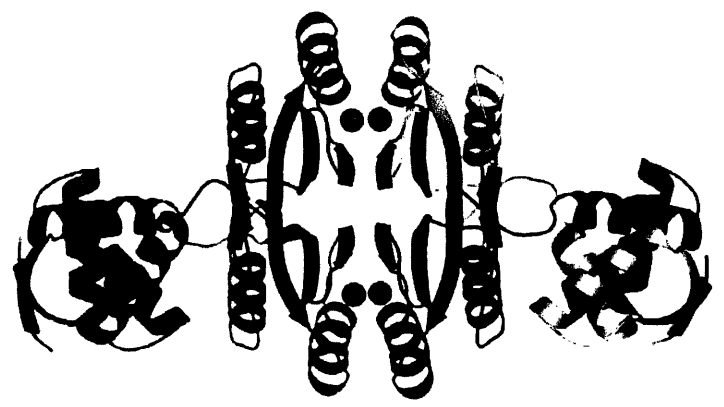
**A.** The DtxR/IdeR/MntR family: metal uptake, oxidative stress, toxin production **B.** ModE: molybdate uptake **C.** ArsR/SmtB family: heavy metal resistance **D.** Fur/Zur family: iron metabolism, zinc uptake **E.** MerR/CueR/ZntR family: mercury, copper, zinc efflux. **F.** NikR: nickel uptake.



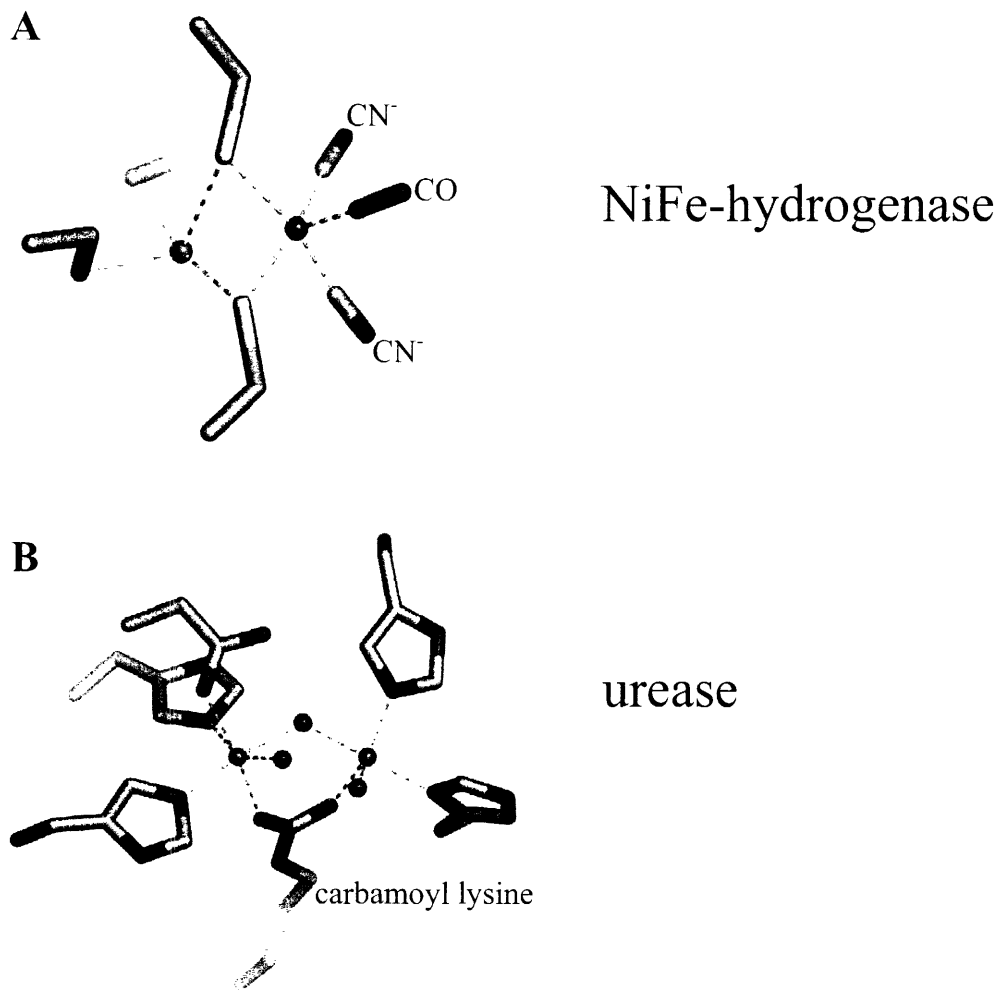
E.

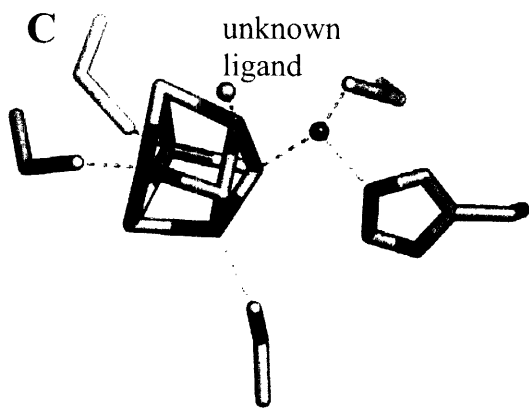


F.

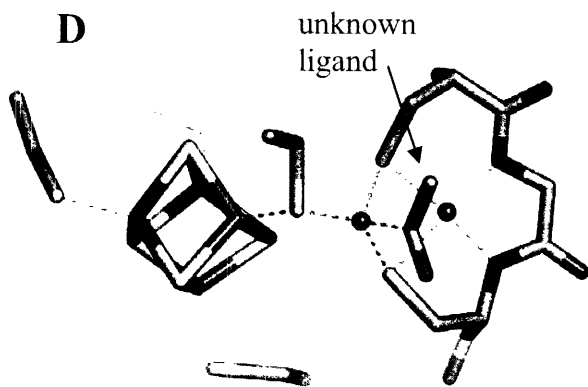


**Figure I.2.** Nickel-containing cofactors of nickel-dependent enzymes. Nickel is green, iron is brown, sulfide is yellow, and water molecules are red spheres. Metal ligation is by unmodified protein sidechains and backbone except where noted. **A.** Reduced, active hydrogenase from *D. norvegicum* (PDB code 1CC1). Similar structures in 2FRV, 1H2A, 1FRF, and 1E3D. **B.** Urease active site from *K aerogenes*, 1FWJ. Similar structures in 2UBP and 1E9Z. **C.** CODH active site from *R. rubrum*, 1JQK. Similar structures in 1JJY, 1MJG, and 1OAO. **D.** ACS active site from *M. thermoacetica*, 1MJG. Similar structure in 1OAO. **E.** Methyl-coenzyme M reductase active site from *M. marburgensis*, 1MR0. Similar structures in 1HBN and 1E6V. **F.** Oxidized Ni-SOD active site from *S. coelicolor*, 1T6U. Other structures in 1Q0D, 1Q0F, 1Q0G, 1Q0K, and 1Q0M. **G.** Active site of *E. coli* glyoxylase, 1F9Z. **H.** Active site of *K. pneumoniae* aci-reductone dioxxygenase, 1M4O. Figures were prepared using PyMOL.

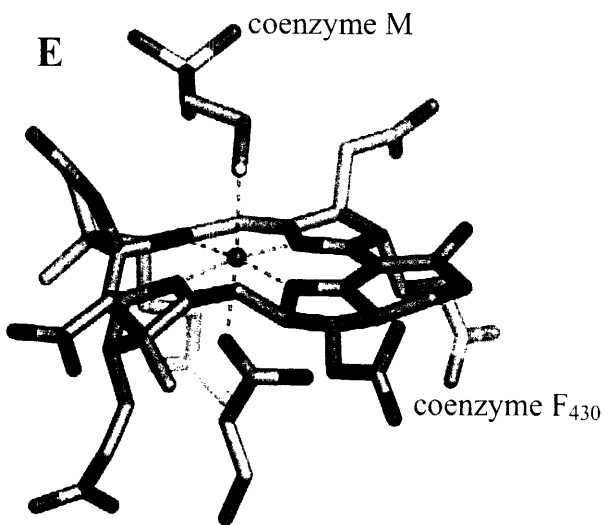




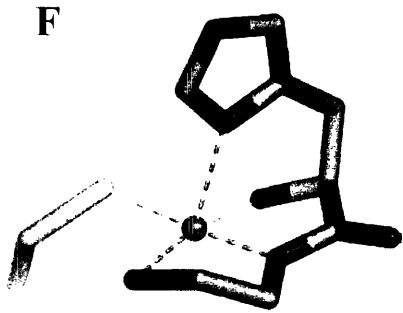
carbon monoxide  
dehydrogenase (CODH)



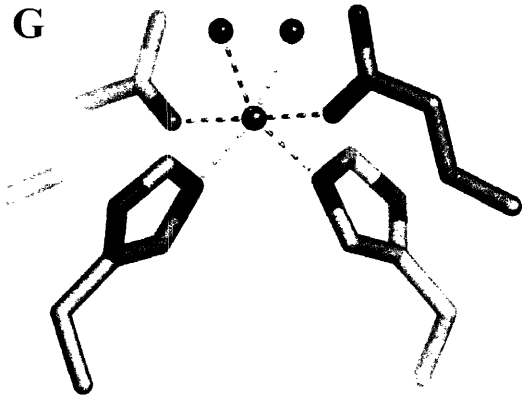
acetyl-CoA synthase (ACS)



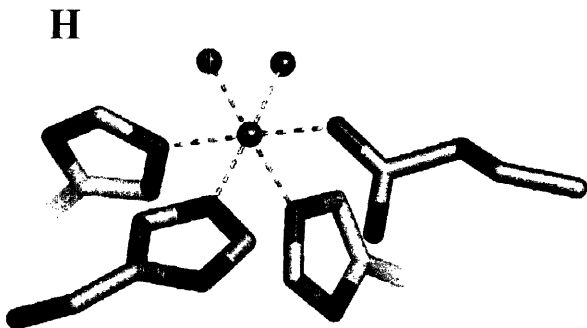
methyl-CoM reductase



Ni-superoxide dismutase  
(Ni-SOD)

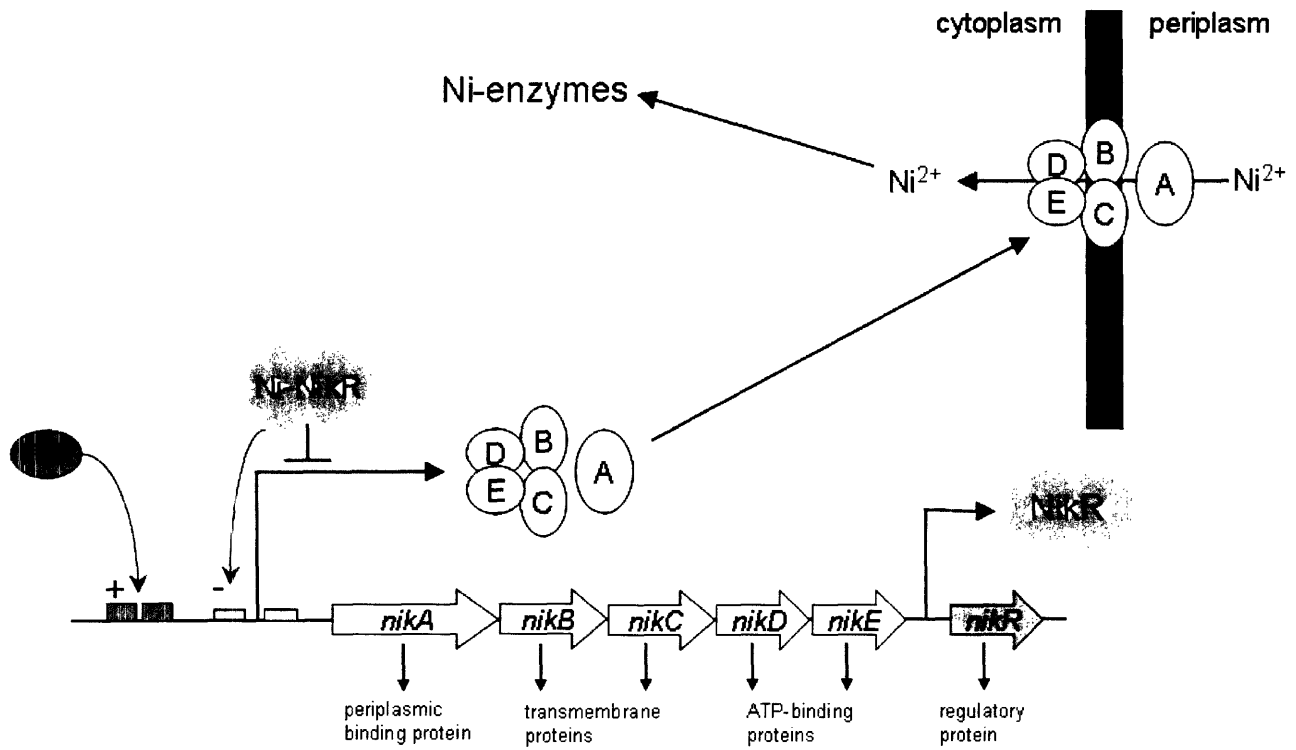


Ni-glyoxylase I



Ni-aci-reductone  
dioxygenase

**Figure I.3.** Schematic representation of the regulation of nickel uptake in *Escherichia coli*.





## Chapter II: Crystal Structure of the Nickel-Responsive Transcription Factor NikR

### II.A. Summary

NikR is a metal-responsive transcription factor that controls nickel uptake in *Escherichia coli* by regulating expression of a nickel-specific ABC transporter. We have determined two structures of NikR: the full-length apo-repressor at 2.3 Å resolution and the nickel-bound C-terminal regulatory domain at 1.4 Å resolution. NikR is the only known metal-responsive member of the ribbon-helix-helix family of transcription factors, and its structure displays an interesting quaternary arrangement consisting of two dimeric DNA-binding domains separated by a tetrameric regulatory domain that binds nickel. The position of the C-terminal regulatory domain enforces a large spacing between the contacts that each NikR DNA-binding domain can make with the *nik* operator. The regulatory domain of NikR contains four nickel-binding sites at the tetramer interface, each displaying a novel square-planar coordination by three histidines and one cysteine side chain. Differences between the apo-NikR and nickel-binding domain structures suggest mechanisms of DNA-binding activation upon nickel binding, and contribute to our understanding of intracellular metal regulation.

**Note:** The results presented in this chapter have been published in: Schreiter, E.R., Sintchak, M.D., Guo, Y., Chivers, P.T., Sauer, R.T. & Drennan, C.L. Crystal structure of the nickel-responsive transcription factor NikR. *Nature Structural Biology* **10**, 794-799 (2003).



## II.B. Introduction

Nickel is essential for anaerobic metabolism in bacteria, which use high-affinity uptake systems to obtain sufficient quantities of this metal for growth<sup>1</sup>. Nickel uptake must be tightly regulated, however, because too much nickel is toxic. In *Escherichia coli*, nickel concentrations are controlled by transcriptional repression of the nickel-specific ATP-binding cassette (ABC) transporter complex encoded by the *nikABCDE* operon<sup>2,3</sup>. In the presence of excess intracellular nickel, the nickel-dependent repressor NikR binds to an operator sequence within the *nikABCDE* promoter, turning off expression of the transporter and decreasing nickel import. NikR orthologs are found in Gram-negative bacteria and archaea, suggesting common mechanisms of nickel regulation across a diverse set of microorganisms.

NikR is the only known metal-responsive member of the ribbon-helix-helix (RHH) family of transcription factors. This 133 residue protein is a tetramer that shows no detectable DNA-binding in the absence of nickel<sup>4</sup>. There are two classes of nickel-binding sites in NikR<sup>4</sup>. One set of sites (four per tetramer) has picomolar affinity for nickel, and occupancy of these sites results in NikR binding to a region spanning roughly 40 operator base pairs (bp) with an affinity in the mid nanomolar range. Further addition of nickel fills low-affinity metal binding site(s) (apparent  $K_d = 30 \mu\text{M}$ ), extending the NikR-DNA contact region to ~65 bp and increasing DNA-binding affinity to 15 pM. The low-affinity nickel binding site(s) may only exist in the protein-DNA complex, and the mechanism(s) by which nickel binding to either set of sites modulates NikR's affinity for and contacts with operator DNA are not well understood. There is, however, evidence that DNA binding causes a change from four- to six-coordinate geometry in the high-affinity nickel site of NikR as assayed by X-ray absorption spectroscopy (XAS)<sup>5</sup>. *E. coli* NikR contains a dimeric N-terminal domain (Met1–His47) that binds operator DNA weakly and is

homologous to RHH transcription factors, including the Arc, Mnt, CopG, MetJ, TraY, and  $\omega$  repressors<sup>6</sup>. The C-terminal regulatory domain, or metal binding domain (MBD), of *E. coli* NikR (Gly48-Asp133) is tetrameric and contains the high affinity nickel-binding sites<sup>4</sup>.

To gain insight into how nickel regulates NikR function, thereby regulating its own uptake, we have determined the crystal structures of two forms of *E. coli* NikR: full-length apo-NikR at 2.3 Å resolution and the nickel-bound C-terminal regulatory domain at 1.4 Å resolution. These structures establish basic architectural features of NikR, including the relative positions of the DNA-binding domains, the nickel-binding sites, and the tetramerization interface. The structures explain the relatively large operator sub-site spacing, reveal a novel biological nickel coordination, and show how changes in nickel coordination geometry might be effected with minimal structural rearrangement.

## **II.C. Results and Discussion**

### *II.C.1. Overall features of full-length apo-NikR*

The NikR structure reveals a simple modular design, with dimeric RHH DNA-binding domains attached at either end of the tetrameric C-terminal regulatory domain (Fig. II.1.a). The N-terminal sequences of two NikR subunits intertwine to form each C<sub>2</sub>-symmetric RHH domain, as has been observed for other family members<sup>7-11</sup>. The C-terminal regulatory domain forms a tetramer with approximate D<sub>2</sub> (dimer of dimers) symmetry. We will refer to the two MBD subunits that are attached to each DNA binding domain as the core dimer. Each subunit in the MBD has a ferredoxin-like fold ( $\beta$ - $\alpha$ - $\beta$ - $\beta$ - $\alpha$ - $\beta$ ). The MBD subunits in each core dimer form an extensive hydrophobic interface, composed predominantly of branched hydrophobic side chains, between their respective  $\beta$ -sheets. The tetramerization interface between core dimers is located

at the end of the  $\beta$ -sheet distal to the DNA-binding domains. This interface results in an eight-stranded antiparallel  $\beta$ -sandwich structure with  $\sim 90^\circ$  of twist across its length. Several interactions appear to contribute to the formation of a stable tetramer, including a hydrophobic cluster formed by Leu85 from each MBD subunit and hydrogen bonding between the  $\beta$ -strands from adjacent subunits. The symmetry axes of the MBD tetramer are not coincident with the symmetry axis of either DNA-binding domain, making the overall structure of apo-NikR asymmetric.

### *II.C.2. Nickel binding*

In the absence of a structure of full-length NikR with nickel bound (see Methods Section), the structure of the isolated MBD has allowed us to locate the high-affinity binding site for nickel, and to compare apo and nickel-bound forms of this domain (Fig. II.1.b). The four nickel-binding sites in this structure are located at the tetramer interface. In each site, the nickel ligands are arranged in a novel square-planar coordination consisting of the side chains of His87, His89, and Cys95 from one NikR subunit and the side chain of His76 from the adjacent subunit across the tetramerization interface (Fig. II.2.a). The detailed geometry of the site is described in Table IV.1. The square-planar nature of this high-affinity binding site may confer specificity for nickel over other biologically relevant metal ions that do not commonly adopt this geometry. The nickel-binding sites of NikR are not equivalent to any of the catalytic nickel sites that have been identified in enzyme structures<sup>1</sup>. The square-planar nickel site in the anaerobic acetyl-CoA synthase (ACS) enzyme is the most similar, but employs two cysteine sulfur atoms and two protein backbone amide nitrogens to coordinate nickel<sup>12</sup>.

The overall fold of the MBD is largely the same with and without bound nickel, but superposition of the nickel-free and nickel-bound tetramers (1.6 Å rmsd for common  $C_{\alpha}$  atoms) reveals some notable differences (Fig. II.1.b). The most significant structural changes affect helix  $\alpha 3$  and the loop between strands  $\beta 3$  and  $\beta 4$  near the nickel-binding site (labeled with an asterisk in Fig. II.1.b). Helix  $\alpha 3$  is ordered in just one of the four subunits of the nickel-free structure and, even in this subunit, is shifted relative to the nickel-bound structure. This  $\alpha$ -helix has a significantly higher average B-factor than the rest of the apo-NikR model ( $88.8 \text{ \AA}^2$  vs.  $65.1 \text{ \AA}^2$ , respectively). His76 in helix  $\alpha 3$  serves as one of the nickel ligands (Fig. II.2.a), making it likely that the position and folding of this helix are stabilized by bound nickel. Indeed, nickel binding increases the  $\alpha$ -helical content of the MBD and full-length NikR in solution<sup>4</sup>. The remaining three nickel ligands in strands  $\beta 3$  and  $\beta 4$  are connected by a loop that moves closer to the nickel site in the nickel-bound structure, resulting in a shift of  $\sim 2.6 \text{ \AA}$  in the side-chain position of His89, one of the nickel ligands (Fig. II.3.a). The His87 and Cys95 nickel ligands adopt significantly different sidechain rotamers in the two structures, and the movement of these nickel ligands coincides with additional movements in the side chain positions of several nearby residues. These movements of the nickel ligands and neighboring residues between the nickel-free and nickel-bound structures suggest that the nickel-binding site is not pre-organized.

The nickel coordination observed in the NikR MBD structure is similar in most details to that predicted by XAS studies and site-directed mutagenesis experiments, with Ni-N and Ni-S bond lengths of 1.90 and 2.18 Å, respectively<sup>5</sup>. Square-planar geometry and the use of His87, His89, and Cys95 as nickel ligands are both observed in the structure and predicted by solution studies. The solution studies, however, predict that Glu97 is the fourth nickel ligand, whereas we observe His76 in this role. Glu97 hydrogen bonds to one of the nickel ligands in the MBD

structure (His87) and is part of an interesting hydrogen-bond network linking adjacent nickel sites, as shown in Fig. II.3.b. Mutations at this position would therefore be expected to destabilize nickel binding. However, the XAS spectra of the nickel-bound MBD and full-length NikR molecules are not identical, suggesting that several different nickel coordination geometries may be possible. Moreover, XAS experiments show that the nickel site changes from four-coordinate to six-coordinate upon binding to operator DNA<sup>5</sup>. In the MBD structure, several nearby side chains (5 to 6.5 Å from the nickel ion) are conserved in NikR orthologs (Figs. II.2.a,b) and might serve as new nickel ligands if the protein conformation changed modestly upon DNA binding.

### *II.C.3. MBD homology to other regulatory domains*

The MBD of NikR is structurally homologous to several classes of proteins containing an “ACT-like” domain<sup>13</sup> and is most similar to the regulatory domains of the metabolic enzymes phosphoglycerate dehydrogenase (PGDH)<sup>14</sup>, ATP phosphoribosyltransferase (ATP-PRTase)<sup>15</sup>, and phenylalanine hydroxylase (PheOH)<sup>16</sup>. PGDH and the MBD of NikR share a common topology, a similar quaternary structure for two protein subunits, and nearly identical positions for binding of their respective allosteric ligands, serine and nickel (Figs. II.1.c,d). PheOH has a different quaternary organization but its ligand (phenylalanine) has been proposed to bind in a similar location to the serine in PGDH. The topology and tertiary organization of the NikR MBD also matches the regulatory domain of LrpA<sup>17</sup>, thought to be representative of the leucine-responsive regulatory protein (Lrp) family of transcription factors. The LrpA C-terminus serves as the archetype for the recently categorized Regulation of Amino acid Metabolism (RAM) domain<sup>18</sup>. The similarities between these regulatory domains suggest that a common mechanism

may be employed to propagate ligand-binding signals to adjacent protein domains, but the molecular details remain to be elucidated.

#### *II.C.4. The Ribbon-Helix-Helix domain and operator recognition*

The DNA-binding domain of NikR has a dimeric structure similar to other proteins in the RHH family. This domain is structurally most similar to the Mnt, Arc, MetJ, CopG, and  $\omega$  repressors, in order of increasing rmsd values (Table II.1). Mutagenesis experiments and homology with other RHH proteins suggest that residues on the surface of the anti-parallel  $\beta$ -sheet in the NikR DNA-binding domain should contact the major groove in each *nik* operator sub-site<sup>6</sup>. RHH repressors also use a few residues from the second  $\alpha$ -helix to contact the DNA phosphate backbone. Interestingly, NikR homologues show considerably more sequence conservation throughout this second  $\alpha$ -helix than would be expected based on DNA-binding considerations. Since NikR is the only RHH transcriptional repressor that is known to be metal-responsive, it is possible that this helix is important for the allosteric regulation of NikR by nickel.

Most RHH proteins bind as tetramers to operator sequences containing tandem binding sites, usually arranged as an inverted repeat, for each dimeric DNA-binding domain. The center-to-center distance between the tandem operator sub-sites is typically one turn of the DNA helix or less (Figs. II.4.c,d). NikR is unusual in this regard, as its operator sub-sites are separated by a center-to-center spacing of more than two complete turns of the DNA helix. The structure of NikR reveals that the regulatory MBD tetramer enforces this larger spacing of operator sub-sites. Although apo-NikR is approximately the correct size to span the larger operator, closer inspection of Fig. II.4.b shows that the DNA-binding domains are not positioned to allow both to

bind simultaneously to a linear B-form DNA molecule. A conformational change in both the protein and DNA would be needed to allow bivalent interactions. Specifically, we expect a rigid-body rotation of the RHH domains relative to the MBD tetramer that brings the base-contacting  $\beta$ -sheets closer together, as well as a curvature of the DNA to prevent steric clashes with the regulatory domain helices (Figs. II.4.a,b). The DNA-binding domains of NikR are linked to the tetrameric MBD by residues that are partially disordered in one subunit of the apo-NikR structure and have different conformations in the others. Difference distance matrix analysis using ESCET<sup>19</sup> and  $\Delta\Phi, \Delta\Psi$  plots<sup>20</sup> of the four apo-NikR molecules support the notion that this poorly conserved linker serves as an inter-domain hinge, with vertices at residues 41 and 48-49. Linker flexibility coupled with nickel-induced allosteric changes in the MBD presumably allows the DNA-binding domains to assume conformations compatible with strong operator binding.

#### *II.C.5. Implications for NikR function*

The structures presented here are consistent with several possible models for how nickel binding could regulate NikR function. First, nickel binding creates or stabilizes a form of NikR in which the RHH domains are positioned correctly for operator sub-site binding. As mentioned above, the structure of the full-length protein in its apo-form is not capable of binding DNA, suggesting that the role of nickel could be to induce or stabilize a protein conformational change. Analytical ultracentrifugation (AUC) experiments<sup>4</sup>, XAS data<sup>5</sup>, and comparison of the structures presented here are all consistent with the idea of conformational changes in NikR. The ordering of helix  $\alpha 3$ , as well as movement of a loop and conserved side chains near the nickel binding site could all be part of a signal propagation to the RHH domains. The high degree of conservation

of the residues of the second helix of the RHH among NikR homologs suggests that this helix could be a region to which a signal is propagated. This communication would be a long-range allosteric effect since the nickel site is 25-30 Å from the center of this helix.

In a second model of NikR activation, nickel would exert its influence on gene transcription by formation or stabilization of the tetrameric form of NikR that is required for DNA binding. *In vitro*, however, nickel is not essential for tetramer formation. Our structure of apo full-length NikR is tetrameric, and AUC studies show that NikR is a tetramer with and without nickel<sup>4</sup>. The caveat for both crystallography and AUC is that these techniques require higher concentrations ( $\geq 50 \mu\text{M}$ ) of NikR than one would expect to find in an *E. coli* cell ( $\sim 200 \text{ nM}$ )<sup>4</sup>. The position of nickel ions bound at the dimer-dimer interface, coordinated by residues from two protein subunits, also suggests a role in tetramer stabilization. Perhaps tetrameric NikR can form *in vivo* and *in vitro* without nickel, but is not stable as a tetramer in the presence of DNA without nickel, and thus is ineffective as a repressor. This idea is consistent with denaturation experiments that show that nickel increases the stability of the MBD to the presence of urea<sup>4</sup>, but the stability of the tetramer in the presence of DNA has not been measured.

Finally, nickel binding could act to increase the affinity of NikR for operator DNA by decreasing the net negative charge of the regulatory MBD tetramer, making an interaction with DNA more electrostatically favorable. If the MBD tetramer does indeed contact operator DNA, it is likely to bind via the face of the MBD oriented toward the DNA model in Fig. II.4.b. An electrostatic potential surface generated in GRASP<sup>21</sup> (Fig. II.5) shows a stripe of neutral to positive charge amid an otherwise negatively charged MBD surface. It should be noted that these models of nickel-dependent regulation of NikR function are not mutually exclusive. It



seems likely that nickel plays multiple roles in activating NikR for operator binding, and that a combination of changes in repressor properties is required for function.

Metal-responsive transcription factors play a key role in intracellular physiology. These proteins must respond efficiently and specifically to excess metal, preventing toxicity while allowing the appropriate levels of trace metals required for cell growth and proliferation. The structure of NikR reveals the architecture of a protein capable of this level of regulation. Interestingly, the regulatory domains of NikR are similar in structure to regulatory domains of allosteric enzymes such as PGDH. Indeed, the nickel-binding sites of two NikR subunits can nearly be superimposed upon two PGDH allosteric effector sites. Thus, NikR uses a well-studied DNA-binding fold, a regulatory fold found in enzymes and transcription factors, and a unique metal-binding site that should exhibit specificity for nickel. This modification of a classical RHH repressor endows the NikR subclass of this family with novel metal-responsive capabilities. The presence of NikR homologues in organisms ranging from bacteria to archaea suggests that this architectural motif will be common to proteins that regulate nickel uptake through changes in gene expression.

## **II.D. Materials and Methods**

### *II.D.1. Protein Expression and Purification*

Full-length NikR and MBD were overexpressed as described<sup>2,4</sup>. To produce full-length selenomethionine (SeMet) NikR, cells were grown in LeMaster medium<sup>22</sup> supplemented with 50 mg/L SeMet. Proteins were purified by Ni-NTA chromatography (Qiagen) followed by size-exclusion chromatography on either Sephacryl S-100 HR 16/60 or Superdex 75 16/60 (Pharmacia). Protein samples were then concentrated using Centriplus (Millipore) concentrators

for crystallization. Nickel content of NikR samples was assayed using the divalent metal indicator 4-(2-pyridylazo)resorcinol (PAR)<sup>23</sup>.

#### *II.D.2. Crystallization and data collection*

Crystals of apo-NikR or SeMet apo-NikR were obtained by hanging-drop vapor diffusion after one week at room temperature by mixing 2  $\mu$ L protein (8 mg/ml in 20 mM Tris [pH 8.0], 300mM NaCl) with 2  $\mu$ L precipitant solution (100 mM N-(2-(acetamido)imino)diacetic acid (ADA) [pH 6.5], 6% w/v PEG 6000, 50 mM MgCl<sub>2</sub>). Although the protein was purified with one nickel ion bound per NikR subunit and was yellow in color, crystals of full-length NikR were colorless, presumably because the ADA buffer chelated the nickel at the concentrations used for crystallization. Apo and apo-SeMet NikR crystals belonged to space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> and were cryoprotected using Paratone-N oil. Yellow crystals of nickel-bound MBD were obtained by hanging drop vapor diffusion using 1.5  $\mu$ L protein (10 mg/ml in 20 mM Tris [pH 8.0], 300mM NaCl) and 1.5  $\mu$ L precipitant solution (0.2M di-sodium tartrate, 20% w/v PEG 3350) at room temperature after 4-5 days. MBD crystals also belonging to space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> were cryoprotected in the precipitant solution supplemented with 20% ethylene glycol by using sequential soaks. All crystals were flash-cooled in a gaseous N<sub>2</sub> stream at 100 K and all data were collected at this temperature (Table II.2). Diffraction data were reduced to intensities using the HKL program suite<sup>24</sup>.

#### *II.D.3. Structure determination and model refinement*

The nickel-bound MBD structure was phased by multi-wavelength anomalous dispersion (MAD) techniques using nickel peak and remote wavelengths (Table II.2). Nickel sites were

located and initial phases were generated using CNS<sup>25</sup>. Solvent flattening, phase extension and 4-fold NCS averaging in DM<sup>26</sup> dramatically improved the quality of the electron-density map. An initial model was manually built into these maps using Xfit<sup>27</sup> and refined using CNS against a high-resolution native dataset. Waters were added using wARP<sup>28</sup> and manually checked against  $2F_o-F_c$  and  $F_o-F_c$  electron density maps. After several rounds of iterative manual rebuilding, the model was further refined using SHELXL<sup>20</sup> to include anisotropic thermal displacement parameters for all atoms, and checked against a composite omit electron density map calculated in CNS<sup>25</sup>. Amino acids having obvious electron density for two side chain conformations were modeled at half-occupancy.

Selenium peak data was used to phase apo-NikR by Se single-wavelength anomalous dispersion (SAD). SOLVE<sup>29</sup> was used to locate four well-ordered Se atoms (Met106 of each molecule) and one lower occupancy site (Met1 from one molecule) and to generate initial phases. Solvent flattening in DM allowed placement of the MBD tetramer as well as two DNA-binding domains into the flanking RHH domain electron density. Phase extension and four-fold NCS averaging improved the experimental electron-density map and allowed for manual refitting of domain positions. This model was then refined against a high-resolution (2.3 Å) native dataset, and checked against a composite omit electron density map calculated in CNS<sup>25</sup>. Iterative rounds of model building and refinement were conducted using Xfit and CNS, respectively. Helix  $\alpha 3$  was removed in three apo-NikR molecules because of high B-factors and a lack of interpretable electron density in omit maps.

Coordinates and structure factors for the Apo-NikR and Ni-bound regulatory domain structures have been deposited in the Protein Data Bank (accession codes 1Q5V and 1Q5Y, respectively).

## II.E. Acknowledgements

This research was supported in part by the National Institutes of Health (RTS, CLD), Searle Scholars Program (CLD), Cecil and Ida Green Career Development Fund (CLD), Lester Wolfe Predoctoral Fellowship (ERS), and the Gray Fund for Undergraduate Research (YG). Data were collected at the NSLS, ALS, APS, and SSRL synchrotrons. Synchrotron facilities are funded by the U.S. Department of Energy (ALS 5.0.2, NSLS X25, SSRL), NIH National Center of Research Resources (APS NE-CAT 8BM, NSLS X25), and the National Institute of General Medical Sciences (NSLS X25).

## II.F. References

1. Maroney, M.J. Structure/function relationships in nickel metallobiochemistry. *Current Opinion in Chemical Biology* **3**, 188-99 (1999).
2. Chivers, P.T. & Sauer, R.T. Regulation of high affinity nickel uptake in bacteria. Ni<sup>2+</sup>-Dependent interaction of NikR with wild-type and mutant operator sites. *Journal of Biological Chemistry* **275**, 19735-41 (2000).
3. 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. *Molecular Microbiology* **9**, 1181-91 (1993).
4. Chivers, P.T. & Sauer, R.T. NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA. *Chemistry and Biology* **9**, 1141-8 (2002).
5. Carrington, P.E., Chivers, P.T., Al-Mjeni, F., Sauer, R.T. & Maroney, M.J. Nickel coordination is regulated by the DNA-bound state of NikR. *Nature Structural Biology* **10**, 126-130 (2003).
6. Chivers, P.T. & Sauer, R.T. NikR is a ribbon-helix-helix DNA-binding protein. *Protein Sci* **8**, 2494-500 (1999).
7. Breg, J.N., van Opheusden, J.H., Burgering, M.J., Boelens, R. & Kaptein, R. Structure of Arc repressor in solution: evidence for a family of beta-sheet DNA-binding proteins. *Nature* **346**, 586-9 (1990).

8. Burgering, M.J. et al. Solution structure of dimeric Mnt repressor (1-76). *Biochemistry* **33**, 15036-45 (1994).
9. Gomis-Ruth, F.X. et al. The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO Journal* **17**, 7404-15 (1998).
10. Rafferty, J.B., Somers, W.S., Saint-Girons, I. & Phillips, S.E. Three-dimensional crystal structures of Escherichia coli met repressor with and without corepressor. *Nature* **341**, 705-10 (1989).
11. Murayama, K., Orth, P., de la Hoz, A.B., Alonso, J.C. & Saenger, W. Crystal structure of omega transcriptional repressor encoded by Streptococcus pyogenes plasmid pSM19035 at 1.5 Å resolution. *Journal of Molecular Biology* **314**, 789-96 (2001).
12. Doukov, T.I., Iverson, T.M., Seravalli, J., Ragsdale, S.W. & Drennan, C.L. A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. *Science* **298**, 567-72 (2002).
13. Chipman, D.M. & Shaanan, B. The ACT domain family. *Current Opinion in Structural Biology* **11**, 694-700 (2001).
14. Schuller, D.J., Grant, G.A. & Banaszak, L.J. The allosteric ligand site in the Vmax-type cooperative enzyme phosphoglycerate dehydrogenase. *Nature Structural Biology* **2**, 69-76 (1995).
15. Cho, Y., Sharma, V. & Sacchettini, J.C. Crystal structure of ATP phosphoribosyltransferase from Mycobacterium tuberculosis. *Journal of Biological Chemistry* **278**, 8333-9 (2003).
16. Kobe, B. et al. Structural basis of autoregulation of phenylalanine hydroxylase. *Nature Structural Biology* **6**, 442-8 (1999).
17. Leonard, P.M. et al. Crystal structure of the Lrp-like transcriptional regulator from the archaeon Pyrococcus furiosus. *EMBO Journal* **20**, 990-7 (2001).
18. Ettema, T.J., Brinkman, A.B., Tani, T.H., Rafferty, J.B. & Van Der Oost, J. A novel ligand-binding domain involved in regulation of amino acid metabolism in prokaryotes. *Journal of Biological Chemistry* **277**, 37464-8 (2002).
19. Schneider, T.R. A genetic algorithm for the identification of conformationally invariant regions in protein molecules. *Acta Crystallographica. Section D: Biological Crystallography* **58**, 195-208 (2002).
20. Sheldrick, G.M. & Schneider, T.R. SHELXL: high-resolution refinement. *Methods in Enzymology* **277**, 319-343 (1997).
21. Nicholls, A., Sharp, K.A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281-96 (1991).

22. LeMaster, D.M. & Richards, F.M.  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NMR studies of Escherichia coli thioredoxin in samples isotopically labeled by residue type. *Biochemistry* **24**, 7263-8 (1985).
23. 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. *Analytical Biochemistry* **146**, 150-7 (1985).
24. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326 (1997).
25. Brünger, A.T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica. Section D: Biological Crystallography* **54**, 905-21 (1998).
26. Cowtan, K.D. & Main, P. Phase combination and cross validation in iterated density-modification calculations. *Acta Crystallographica. Section D: Biological Crystallography* **52**, 43-48 (1996).
27. McRee, D.E. XtalView/Xfit - A Versatile Program for Manipulating Atomic Coordinates and Electron Density. *Journal of Structural Biology* **125**, 156-165 (1999).
28. Lamzin, V.S. & Wilson, K.S. Automated refinement of protein models. *Acta Crystallographica. Section D: Biological Crystallography* **49**, 129-147 (1993).
29. Terwilliger, T.C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallographica. Section D: Biological Crystallography* **55**, 1174-1178 (1999).

## II.G. Tables and Figures

**Table II.1.** Structural comparison of the NikR RHH domain with other RHH domains.

<b>Structure</b>	<b>PDB code</b>	<b>Rmsd<sup>1</sup></b>	<b>N<sub>algn</sub><sup>2</sup></b>
Arc-DNA complex	1BDT	1.6	81
Mnt repressor(N-term domain)	1MNT	1.6	76
CopG-DNA complex	1B01	1.7	81
Arc repressor	1ARR	1.7	80
MetJ repressor	1CMB	1.8	77
CopG repressor	2CPG	2.0	78
MetJ-DNA complex	1CMA	2.2	76
$\omega$ repressor	1IRQ	2.6	76

<sup>1</sup>Calculated from a best C <sub>$\alpha$</sub>  alignment using the Secondary Structure Matching (SSM) server at (<http://www.ebi.ac.uk/msd-srv/ssm/ssmstart.html>) with residues 1-42 of two subunits of NikR corresponding to one RHH domain as a search model.

<sup>2</sup>N<sub>algn</sub> = number of C <sub>$\alpha$</sub>  atoms aligned for both molecules of each dimeric RHH domain.

**Table II.2.** Data collection, phase determination and refinement statistics.

<b>Data Collection</b>					
<b>Crystal</b>	WT NikR	Se-Met NikR	MBD 1	MBD 2	
<b>Data Set</b>	Native	Se-Peak	Native	Ni-Peak	Ni-Remote
<b>Space Group</b>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
<b>Cell Dimensions (Å)</b>					
<b>a</b>	46.34	46.29	45.85	45.89	45.87
<b>b</b>	69.61	69.54	78.29	78.36	78.35
<b>c</b>	159.10	159.32	81.39	81.75	81.73
<b>Beam Line</b>	X25	8BM	X25	5.0.2	5.0.2
<b>Temperature (K)</b>	100	100	100	100	100
<b>Wavelength (Å)</b>	1.1000	0.9791	1.1000	1.4852	0.9500
<b>Resolution Range (Å)</b>	50-2.3	50-2.5	50-1.4	50-2.7	50-1.4
<b>Observations</b>	250494	356862	360689	63581	221766
<b>Unique Reflections</b>	23458	17369	57926	7589	53354
<b>Completeness (%)<sup>1</sup></b>	99.2(97.4)	98.8(96.4)	99.2(94.2)	88.2(58.2)	90.6(60.8)
<b><i>I</i> / <math>\sigma(I)</math><sup>1</sup></b>	27.0(5.8)	23.6(5.7)	17.5(3.7)	33.5(19.3)	18.7(2.6)
<b><i>R</i><sub>sym</sub> (%)<sup>1,2</sup></b>	5.2(35.9)	7.1(31.9)	6.2(31.0)	6.1(8.3)	5.1(40.9)
<b>Phasing</b>					
<b>Number of sites/asu</b>		5		4	
<b>Resolution Range (Å)</b>		44.3-3.0		36.2-2.7	
<b>Figure of Merit</b>		0.39		0.65	
<b>Refinement</b>					
<b><i>R</i><sub>cryst</sub> (<i>R</i><sub>free</sub>) (%)<sup>3</sup></b>	24.7(30.5)		16.0(22.0)		
<b>Resolution Range (Å)</b>	50-2.3		10-1.4		
<b>Number of protein atoms</b>	3626		2528		
<b>Number of nickel ions</b>	-		4		
<b>Number of glycerol atoms</b>	-		20		
<b>Number of water molecules</b>	22		242		
<b>R.m.s. deviations</b>					
<b>Bond lengths (Å)</b>	0.009		0.010		
<b>Bond angles (°)</b>	1.30		-		
<b>Bond distances (Å)</b>	-		0.029		
<b>Ramachandran (%)</b>					
<b>Most favored</b>	91.1		90.8		
<b>Additionally Allowed</b>	8.2		9.2		
<b>Generously Allowed</b>	0.7		0.0		
<b>Disallowed</b>	0.0		0.0		
<b>Average <i>B</i>-factor (Å<sup>2</sup>)</b>	65.8		27.9		

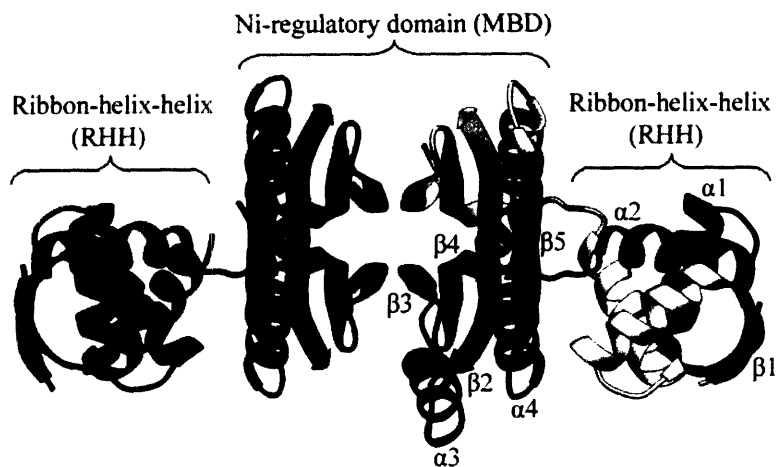
<sup>1</sup>The number in parentheses is for the highest resolution shell.

<sup>2</sup> $R_{\text{sym}} = \frac{\sum_{hkl} |I_{(hkl)}^i - \langle I_{(hkl)} \rangle|}{\sum_{hkl} \langle I_{(hkl)} \rangle}$ , where  $I_{(hkl)}^i$  is the  $i^{\text{th}}$  measured diffraction intensity and  $\langle I_{(hkl)} \rangle$  is the mean of the intensity for the miller index ( $hkl$ ).

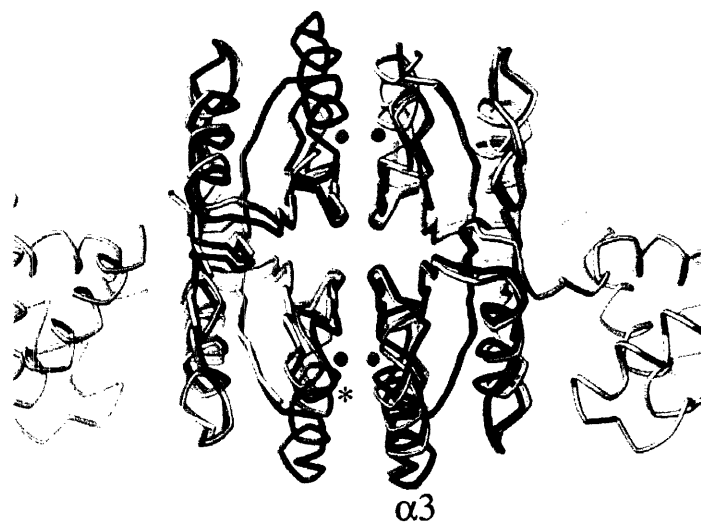
<sup>3</sup> $R_{\text{cryst}} = \frac{\sum_{hkl} ||F_o(hkl)| - |F_c(hkl)||}{\sum_{hkl} |F_o(hkl)|}$ .  $R_{\text{free}} = R_{\text{cryst}}$  for a test set of reflections (8% for WT NikR, 5% for MBD) not included in refinement.



**Figure II.1.** NikR structures and structural comparisons.



*Figure II.1.a.* Ribbon diagram of the apo-NikR tetramer colored by protein subunit. Domains are bracketed and labeled as either ribbon-helix-helix (DNA binding domain) or Ni-Regulatory domain (MBD). Secondary structure elements of the blue NikR subunit are labeled.



*Figure II.1.b.* Superimposition of the  $C_{\alpha}$  traces of the regulatory domain of apo-NikR (light grey) and nickel-bound MBD (colored by protein subunit) based on all common  $C_{\alpha}$  positions. Nickel atoms from the nickel-bound MBD structure are shown as cyan spheres. The most similar regions are the  $\beta$ -strands of the central  $\beta$ -sheets, while the most significant differences are in the loop (labeled with an asterisk near the green subunit) and helix (labeled  $\alpha 3$ ) nearest each nickel-binding site. Note the absence of this helix in the three other apo-NikR molecules.

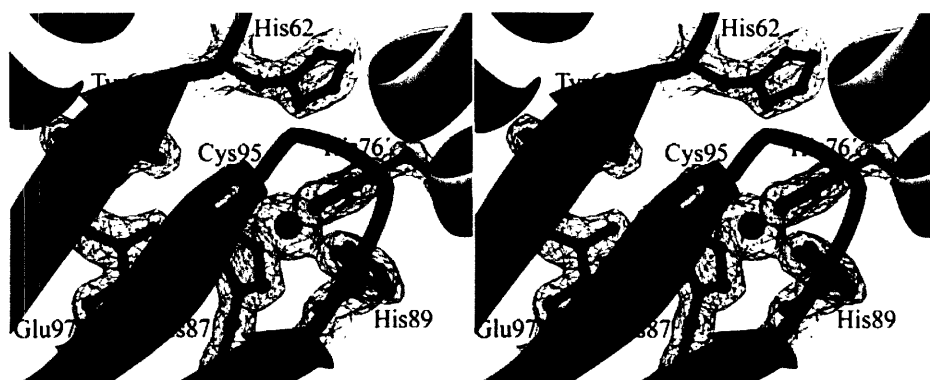


*Figure II.1.c.* Two subunits of the nickel-bound MBD tetramer displayed as ribbon diagrams and colored by molecule. Amino acid side chains ligating nickel are shown in ball-and-stick and colored by atom type (carbon atoms from the two subunits are green and grey, respectively) and nickel is shown as a cyan sphere.

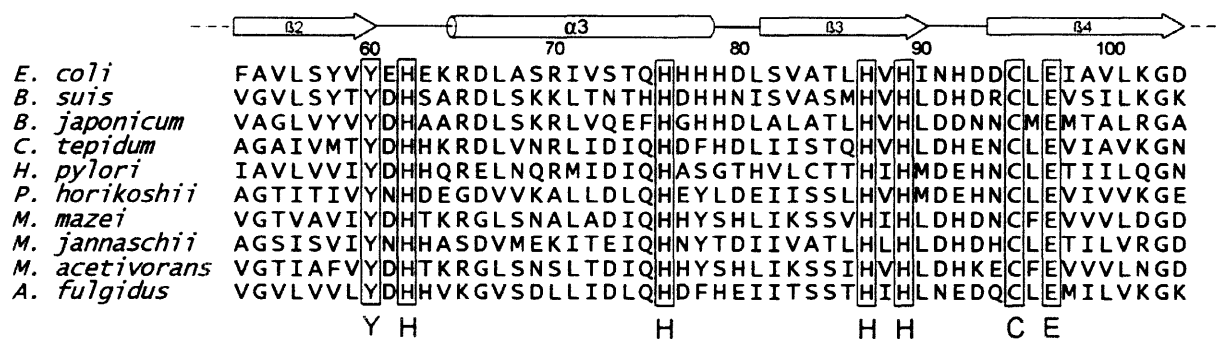


*Figure II.1.d.* Regulatory domains of two subunits of PGDH (from 1PSD), displayed as in Figure II.1.c. Bound serine is shown in cyan ball-and-stick. Note the similarity in location of bound allosteric ligand and the contribution to binding of each ligand from two protein subunits across their interface. All figures were prepared using Ribbons<sup>30</sup>.

**Figure II.2.** Strictly conserved amino acids at the high-affinity nickel-binding site of NikR.

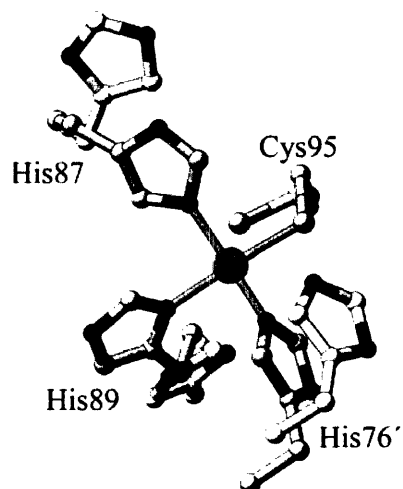


*Figure II.2.a.* Stereo view of the nickel-binding site in the Ni-bound MBD structure. Ribbons are colored as in Figure II.1. by protein subunit, conserved residues surrounding the nickel-site are shown in ball-and-stick and colored by atom type, and nickel is shown as a cyan sphere. A  $2F_o - F_c$  electron density map contoured at  $1.2\sigma$  is displayed around the nickel and conserved residues.

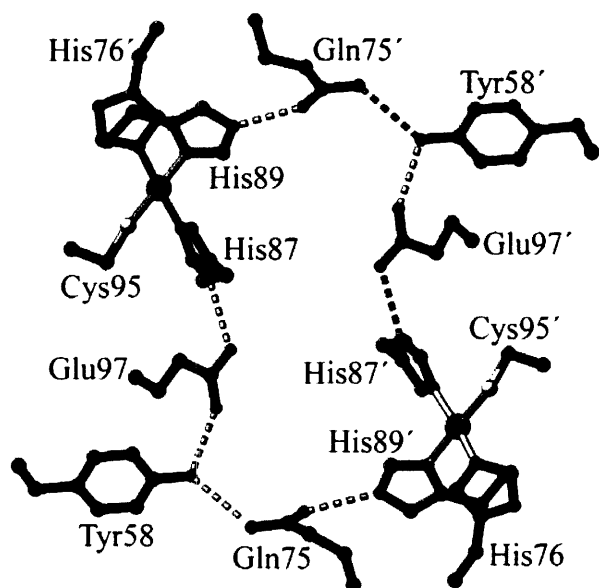


*Figure II.2.b.* Sequence alignment of NikR proteins from bacterial and archaeal species with numbering and secondary structure elements from the *E. coli* NikR structure. Invariant residues that are potential nickel ligands within the regulatory domain are boxed. Residues colored in blue are seen as nickel ligands in the nickel-bound MBD structure, those in red are within 4.8-6.5Å of the nickel.

**Figure II.3.** Side chain movements and interactions at the high-affinity nickel site.



*Figure II.3.a.* Comparison of the nickel ligands between apo-NikR and nickel-bound MBD after superposition of all common  $C_{\alpha}$  atoms of each tetramer. Amino acid sidechains of the nickel ligands are shown in ball-and-stick with carbon atoms in green for the nickel-bound MBD structure and white for the apo-NikR structure. Nickel ions are depicted as cyan spheres, nickel-ligand bonds as grey lines.

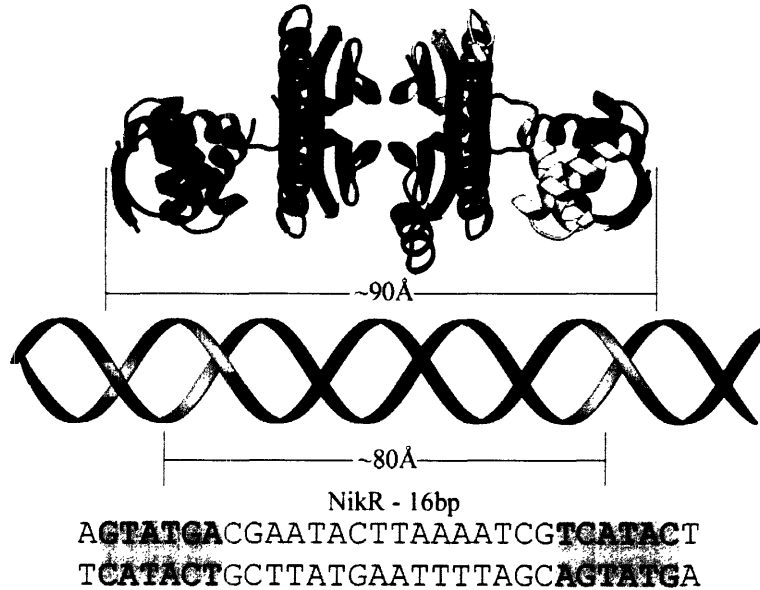


*Figure II.3.b.* Hydrogen-bond network between two nickel-binding sites in the nickel-bound MBD structure. Amino acid side chains of two molecules of the NikR regulatory domain are shown in ball-and-stick with grey or green carbon atoms, respectively. Dashed lines indicate hydrogen bonds.

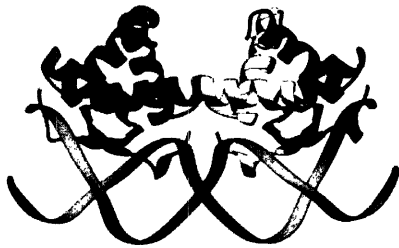
**Figure II.4.** NikR vs. other ribbon-helix-helix transcription factors.



*Figure II.4.a.* Hypothetical model of full-length NikR with nickel ions bound assuming no conformational change of the RHH domains relative to the MBD upon nickel binding. This model was generated by modeling all disordered portions of the Apo-NikR tetramer after complete symmetry-related molecules and positioning of nickel ions after superposition with the nickel-bound MBD tetramer. Inter-domain linkers are denoted with arrowheads, curved arrows show the proposed directions of movement of the RHH domains for operator binding.



*Figure II.4.b.* Apo-NikR structure and DNA operator sequence. Ribbon DNA is idealized B-form with sub-sites highlighted in orange. Below the apo-NikR tetramer is the *nik* operator sequence with sub-sites in orange and the inter-site distance in base-pairs. Distance bars illustrate the approximate distances between the two RHH domains and operator sub-sites. This apo-NikR tetramer conformation would not be able to bind DNA since the RHH domains are too far apart.



Arc - 5bp  
**TACTAG**GTGCT**TOTATCA**  
**ATCATCTCACGAAGATAG**T

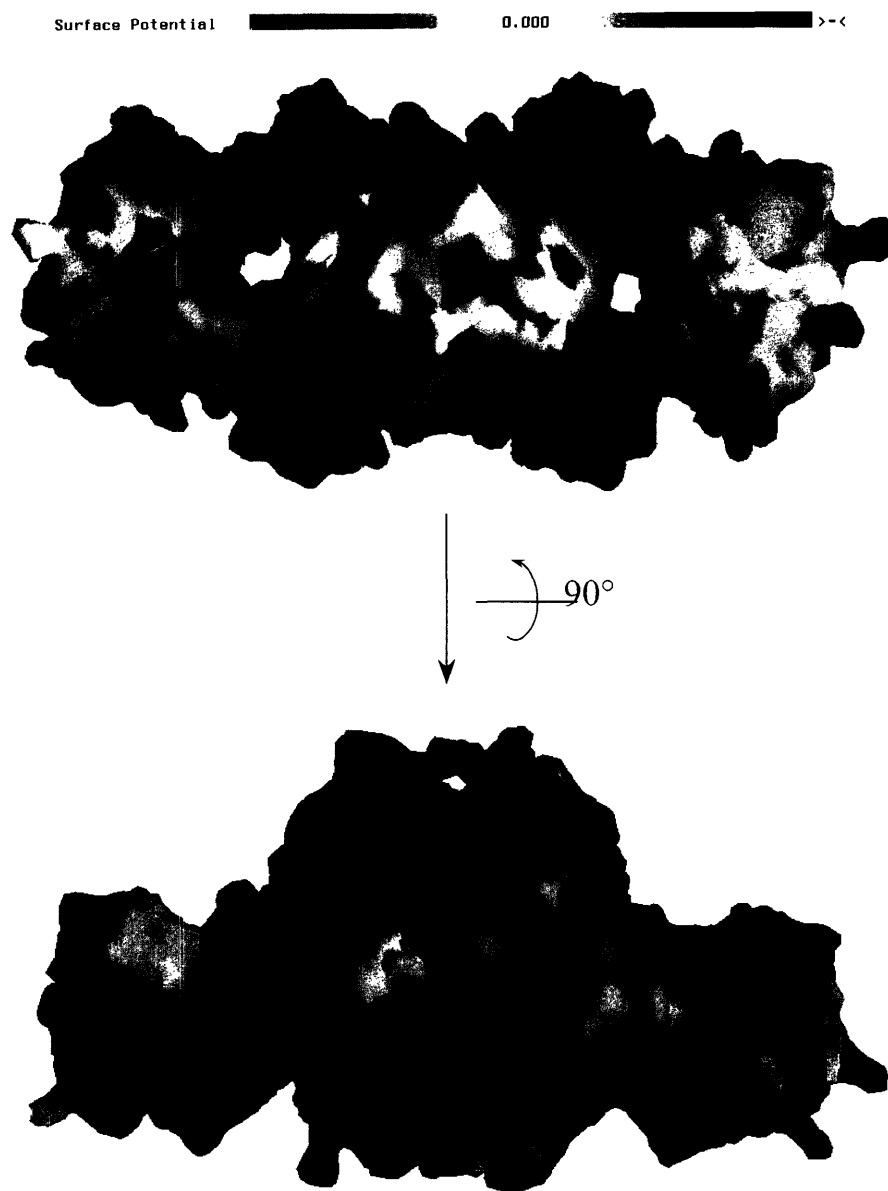
Figure II.4.c. Ribbon diagram of Arc-DNA structure (1PAR)<sup>31</sup> and sequence of operator with binding sites in orange.

CopG - 5bp  
 A**TTCC**ATTGAG**TCCACG**  
 T**AACG**AACTC**ACGTC**

Mnt - 1bp  
 A**GGTCCAC**C**CTGG**ACCT  
 TCC**GGTGGCACCT**GG

Figure II.4.d. Operator sequences of other RHH transcription factors with binding sites highlighted. Note the relatively large binding-site spacing of the NikR operator relative to those of other RHH proteins.

**Figure II.5.** Electrostatic potential surface of a NikR model.



*Figure II.5.* Electrostatic potential surface of a NikR model. The model was produced by building a complete backbone for the NikR tetramer (modeling in missing portions of the apo-NikR model) and building missing side chains by using common rotamers and avoiding steric clashes. An electrostatic potential surface, colored as indicated in the scale bar above in units of  $kT e^{-1}$ , was produced using GRASP. The top view is in the same orientation as Figure II.4.a,b. The bottom view is rotated  $90^\circ$  to show the face of NikR proposed to interact with operator DNA.

## Chapter III: Crystal Structure of the NikR-operator DNA complex and the mechanism of repressor activation by metal ions

### III.A. Summary

Metal ion homeostasis is critical to the survival of all cells. Regulation of nickel concentrations in *Escherichia coli* is achieved via nickel-induced transcriptional repression of the nickel permease, NikABCDE, mediated by the NikR repressor. Here we report two crystal structures of full-length nickel-activated *E. coli* NikR, the isolated repressor at 2.1 Å resolution and in a complex with its operator DNA sequence from the *nik* promoter at 3.1 Å resolution. Binding stoichiometric nickel ions activates NikR by ordering one face of the metal-binding domain to permit complementary interaction with DNA rather than directly altering the conformation of the DNA-binding domains. The conformational changes associated with DNA binding create a second functionally relevant metal-binding site at the interface between the DNA-binding and metal-binding domains.



### III.B Introduction

Metal ions are essential nutrients for all cells, but their intracellular concentrations and distribution must be tightly regulated to avoid toxicity. Nickel ions are particularly important to the physiology of microorganisms such as the model prokaryote *Escherichia coli* (*E. coli*) and the human gastric pathogen *Helicobacter pylori* (*H. pylori*). In these organisms, incorporation of nickel into enzymes such as [NiFe]-hydrogenase and urease is necessary for metabolic adaptation to changing environmental conditions<sup>1</sup>. In *E. coli* nickel is acquired via an ABC-type membrane transporter, NikABCDE<sup>2</sup>. Transcription of the operon encoding this nickel importer, *nikABCDE*, is repressed in the presence of nickel by NikR<sup>3,4</sup>. NikR therefore serves as a cytoplasmic nickel sensor, stopping production of the nickel importer when intracellular levels of nickel are sufficient. In *H. pylori*, which require the enzyme urease to survive and colonize the acidic gastric mucus of the human stomach, NikR plays a more complex regulatory role. Transcriptome analysis of a deletion mutant revealed that *H. pylori* NikR regulates transcription of not only the nickel importer and urease enzyme, but also other regulatory networks, through repression of the gene encoding the ferric uptake regulator (Fur) protein<sup>5,6</sup>.

The NikR protein that has been characterized most extensively at the molecular level is the homolog from *E. coli*. *E. coli* NikR is a homotetramer that binds one nickel(II) ion per subunit with picomolar affinity<sup>7,8</sup>. Stoichiometric nickel ions activate NikR to bind a 28-bp palindromic operator sequence (GTATGA-N<sub>16</sub>-TCATAC) within the promoter of *nikABCDE* with low nanomolar affinity<sup>7,9</sup>. Recently, it was demonstrated that NikR can bind a variety of divalent transition metal ions with high affinity, and that stoichiometric amounts of these other ions can activate NikR to bind

operator DNA with potency comparable to that of nickel<sup>8,9</sup>. In the presence of excess nickel, the affinity of NikR for its operator is enhanced at least 250-fold, and biochemical evidence supports the existence of a second nickel-binding site in the presence of operator DNA with nanomolar to micromolar affinity for nickel<sup>7,9</sup>. The formation of this tighter protein-DNA complex is a nickel-selective response.

The previously reported crystal structure of apo-NikR revealed a modular domain organization, with two dimeric ribbon-helix-helix (RHH) domains flanking a tetrameric metal-binding domain (MBD) (Fig. III.1.a)<sup>10</sup>. This spatial separation of the two DNA-binding domains explained how NikR could recognize two operator sub-sites distantly separated by two turns of DNA. A crystal structure of the nickel-bound form of the isolated MBD showed that the high-affinity metal binding site of NikR is located at the subunit interfaces, with nickel ions each bridging two subunits of the tetramer. Nickel ions bind with a square-planar coordination geometry, ligated by the protein sidechains His87, His89, and Cys95 from one NikR subunit, as well as the sidechain of His76' from an adjacent subunit. In the absence of bound nickel ions, the  $\alpha$ -helix contributing His76' (helix  $\alpha$ 3) to the metal binding site is structurally flexible and was disordered in the apo-NikR crystal structure. The disordered-to-ordered transition of helix  $\alpha$ 3 within the MBD upon nickel binding gave preliminary insight into the structural changes that take place upon activation of NikR by metal ions.

The initial crystal structures of the *E. coli* NikR transcription factor lead to the proposal of several possible mechanisms for activation of repressor activity by nickel ions. These hypotheses included an allosteric effect of nickel binding leading to specific conformational changes in NikR that favor DNA binding; stabilization of the active

tetrameric form of NikR by nickel ions binding at the subunit interface; and creation of a more electropositive DNA-binding surface of the repressor upon binding of positively charged nickel ions. It was further proposed that these mechanisms would not be mutually exclusive and would likely act in concert to increase operator affinity in the presence of nickel ions. We now present two new crystal structures of full-length NikR activated by stoichiometric nickel ions, both alone and bound to a 30 base-pair (bp) oligonucleotide containing its operator DNA sequence from the *nik* promoter. These structures allow us to evaluate functional proposals of NikR repressor activation by metal ions, delineate the drastic conformational changes required for operator recognition, and reveal the formation of a second metal-binding site in the presence of DNA. They further provide a rare set of structural views of a ligand-responsive transcription factor in the unbound, ligand induced, and DNA bound states, adding a new model system to the understanding of this important class of transcription factors.

### **III.C Results and Discussion**

#### *III.C.1 Full-length nickel-activated NikR*

The crystal structure of full-length NikR with stoichiometric nickel reveals domains that are comparable to the two previously determined structures of NikR, with the MBD closely matching the structure of the isolated nickel-bound MBD tetramer (rmsd 0.89-0.90 for 304 C $\alpha$  atoms), and the RHH DNA-binding domains similar to those of the apo-NikR structure (rmsd 0.82-1.21 for 84 C $\alpha$  atoms). The most significant difference in this structure is the relative orientation of the RHH domains with respect to the MBD (Fig. III.1). As shown in Fig. III.1.b, the RHH domains extend almost

completely outward from the MBD, pointing the DNA major-groove binding  $\beta$ -sheets in opposite directions. This rigid-body domain movement occurs about flexible linkers connecting the domains, represented by residues 41-52 as described previously<sup>10</sup>. This crystal form was grown in the presence of a detergent (cyclohexyl-propyl- $\beta$ -D-maltoside), which is seen bound at the interface between the RHH domain and MBD (Fig. III.2). The detergent could be helping to stabilize the extended conformation of Ni-NikR. Crystals of Ni-NikR grown under the same conditions in the absence of detergent that diffracted X-rays poorly showed a similar orientation of the domains, although the RHH domain was partially disordered in this structure (unpublished data).

### *III.C.2 The NikR-operator complex*

In the crystal structure of the nickel-activated NikR tetramer in complex with its operator DNA, the individual RHH and MBDs are again very similar to the previously determined structures of NikR. The RHH domains align well structurally with both apo-NikR (rmsd 1.16-1.35 for 84 C $\alpha$  atoms) and nickel-bound NikR (rmsd 1.06-1.16 for 84 C $\alpha$  atoms), while the MBD closely resembles the nickel-bound forms of both the isolated MBD (rmsd 0.95-1.02 for 304 C $\alpha$  atoms) and full-length NikR (rmsd 0.81-0.84 for 304 C $\alpha$  atoms). The orientations of the RHH DNA-binding domains of the repressor relative to the MBD, however, are dramatically different than in the structures of NikR alone (Fig. III.1). Specifically, the RHH domains rotate about the flexible interdomain linkers to orient their antiparallel  $\beta$ -strands towards the same face of the repressor, allowing each to occupy the DNA major groove of an operator palindrome half-site. Each individual RHH-DNA interaction closely resembles the binding mode of other RHH family

members such as Arc<sup>11</sup>, MetJ<sup>12</sup>, and CopG<sup>13</sup>. To allow for this dramatic rotation, the linker of one subunit of each RHH domain is extended by the unwinding of the final turn of helix  $\alpha 2$  (Fig. III.3). This conformation of NikR results in a new protein-protein interface between the RHH and MBD and creates a second metal-binding site at this interface, apparently occupied by potassium ions from the crystallization solution, in this crystal structure. The DNA in the operator complex is B-form, with no dramatic bends at the interface with protein as observed in some other protein-DNA complexes. The DNA curves smoothly around the face of the NikR tetramer, bending by  $\sim 22^\circ$  over its length and contacting both RHH DNA-binding domains as well as two of the four MBD subunits.

The high-affinity nickel binding site within the MBD of NikR has the same square-planar coordination (within coordinate error) in the NikR-DNA complex structure as in the structure of nickel-activated NikR and the isolated MBD (Fig. III.3.c). This observation is in contrast to X-ray absorption spectroscopy (XAS) results describing a coordination change from square planar, as observed in our structures, to a six coordinate site with loss of thiolate ligation upon specific DNA binding<sup>14</sup>. It is possible that this difference could be attributed to differences in sample preparation. Our crystal structure of the NikR-DNA complex shows an occupied second metal-binding site whereas the XAS samples were prepared with no excess ions in the sample.

### *III.C.3 NikR-operator interactions*

An extensive set of polar interactions (hydrogen bonds and salt bridges of  $< 3.4 \text{ \AA}$  distance) are made between NikR and its operator (Fig. III.4.a). These interactions can

be described in three sets for clarity: specific interactions from the RHH domain, non-specific interactions from the RHH domain, and non-specific interactions from the MBD. First, specific hydrogen bonding interactions are made by the sidechains of Arg3 and Thr5 from the  $\beta$ -sheets of the RHH domains to nucleotide bases in the operator major groove, contacting the five base-pairs of each operator half-site that abrogate NikR binding when mutated (Fig. III.4)<sup>4</sup>. This observation confirms that the crystal structure represents a specific NikR-operator complex. Second, ~14 nonspecific polar interactions are contributed by each RHH domain to contact the phosphate backbone of DNA (Fig. III.4.a). These include hydrogen bonds between two consecutive backbone amide nitrogens at the N-terminus of helix  $\alpha$ 2 and DNA phosphates on either side of the major groove. Analogous positive helix dipole-negative phosphate backbone interactions have been described previously for other RHH family members and serve to anchor the domain to DNA and properly orient the specificity-determinant  $\beta$ -sheets within the major groove<sup>11-13</sup>. Other nonspecific interactions with the DNA phosphate backbone are made by the conserved sidechains of Thr7, Asn27, Arg28, Ser29, and Arg33 from the RHH domain. The final set of interactions, also nonspecific, is from the MBDs of two NikR subunits. The sidechain of Arg119, in a loop connecting the final  $\alpha$ -helix and  $\beta$ -strand of NikR, extends between the phosphates of the DNA minor groove (Fig. III.5). Interactions are also made between a loop leading into helix  $\alpha$ 3 of the MBD (asterisk in Fig. III.3.a) and the center of the NikR operator, including the backbone amide nitrogen of Arg65 and the sidechains of Lys64 and Arg65, which make polar interactions with DNA backbone phosphates (Fig. III.5). This portion of NikR is disordered in the absence

of bound nickel (Fig. III.1.a), implying that positioning of these residues upon metal binding is of functional importance.

#### *III.C.4 From zero to nanomolar affinity for DNA*

The DNA-binding activity of many transcription factors is regulated by the binding of small molecule effectors. In very few cases, however, is the mechanism of activation understood at the molecular level<sup>15-18</sup>. Our structures of apo-NikR, nickel-activated NikR and the NikR-DNA complex allow us to propose and evaluate detailed mechanistic hypotheses regarding the metal ion-activated repressor function of *E. coli* NikR. In the absence of nickel, NikR has no measurable affinity for an oligonucleotide containing its operator sequence<sup>7,9</sup>. Addition of stoichiometric nickel ions leads to a half-maximal protection of the *nik* operator from DNase at a concentration of 5-30 nM NikR. We previously proposed several plausible mechanisms by which stoichiometric amounts of metal ions could activate NikR to bind a specific DNA sequence<sup>10</sup>. First, the location of nickel ions bound at the subunit interface (Fig. III.3), coordinated by amino acid sidechains from both halves of the NikR tetramer, led us to propose that nickel binding could serve to stabilize the active tetrameric form of the repressor. The NikR-DNA complex structure shows that the nickel coordination is unchanged when NikR is bound to DNA, and supports the idea that nickel could stabilize this interface in the presence of DNA, holding the tetramer together in the face of stress induced by conformational change upon operator binding. This type of stress is hinted at by the unwinding of the final turn of helix  $\alpha 2$  of the RHH domain that is required to adopt the DNA-bound conformation. The possible occurrence of dimeric NikR at intracellular

concentrations and the nature and magnitude of strain induced at this interface upon operator binding have yet to be demonstrated.

Second, binding of nickel cations to NikR could increase its affinity for polyanionic DNA by simple electrostatic arguments. Such an electrostatic mechanism has been demonstrated for the RHH family member MetJ, in which binding of the positively charged corepressor S-adenosylmethionine has no effect on protein conformation but increases the calculated surface potential at the interface with DNA<sup>19</sup>. In contrast, the positive charges introduced by nickel binding to the high-affinity site of NikR would be largely neutralized by deprotonation of the cysteine and histidine ligands and are not likely to contribute significantly to the charge at the protein surface. Binding of excess nickel ions to additional sites such as the potassium site observed in our structure could have such an effect, however.

It is possible that metal binding exerts an allosteric effect on the conformation of NikR, thereby increasing the affinity for its operator. Numerous structures of ligand-activated transcription factors support the model that ligand binding alters the conformation of the DNA-binding domain, leading to an increase or decrease in affinity for DNA<sup>17,18,20,21</sup>. For many prokaryotic transcription factors that are dimeric, this allosteric transition results in a change in the spacing of two DNA-binding domains, leaving them better suited to interact with multiple operator sub-sites along the DNA molecule. Our crystal structures of NikR in the unliganded (apo), nickel-bound (activated), and operator-bound forms provide insight into allosteric transitions upon ligand binding (Fig. III.1), and do not support such a mechanism. The apo-NikR structure revealed a conformation that was not suitable for interaction with operator



DNA<sup>10</sup>. The nickel-activated NikR structure now demonstrates that stoichiometric nickel binding alone does not preorganize a conformation capable of operator binding. Furthermore, the RHH domains of NikR are connected to the MBD by flexible linkers that vary significantly in length and amino acid sequence as a function of species, making it unlikely that a signal could be propagated via this route.

Nickel binding does, however, have a very direct short-range effect on the conformation of NikR. As mentioned previously, residues 62-80 of NikR are disordered in the absence of nickel, but form a stable  $\alpha$ -helix (helix  $\alpha$ 3) when nickel is bound due to the direct involvement of His76 from helix  $\alpha$ 3 in nickel coordination. In addition, the NikR-DNA complex structure shows that 6-8 polar interactions are made between the loop containing positively charged residues Lys64 and Arg65 and the negatively charged phosphate backbone of operator DNA (Fig. III.5). Since this loop and the following helix are disordered in the absence of nickel, binding of the nickel corepressor creates a surface of the MBD suitable for interacting with DNA by stabilization of these secondary structure elements. Complementary interaction of the MBD with DNA would localize NikR to the DNA helix and allow cooperative interaction of both RHH domains. Stable binding to the operator then occurs when the specific operator base sequence is recognized and additional specific base contacts can be made.

The hypotheses presented here regarding activation of the NikR repressor activity are not mutually exclusive and could act in concert to create the dramatic increase in affinity of NikR for its operator that has been observed upon stoichiometric nickel binding.

### *III.C.5 From nanomolar to picomolar affinity for DNA*

Addition of excess divalent metal ions to 1:1 nickel-NikR increases the binding affinity of NikR for its operator DNA from nanomolar to as much as 10-20 pM<sup>7,9</sup>, implying the existence of a second metal-binding site. The effect is most dramatic for excess Ni<sup>2+</sup> compared to Cu<sup>2+</sup>, Mn<sup>2+</sup>, Zn<sup>2+</sup>, Cd<sup>2+</sup>, and Co<sup>2+</sup><sup>9</sup>. Although we crystallized the NikR-operator complex with only stoichiometric nickel ions, a second metal-binding site was observed at the interface generated between the RHH domain and MBD. In our crystal structure, the metal in this site is best modeled as a potassium ion from the crystallization buffer. The potassium is coordinated by two strictly conserved amino acid sidechains, Glu30 and Asp34, from the RHH DNA-binding domain, and by the three backbone carbonyl oxygens of Ile116, Gln118, and Val121 from a loop of the MBD (Fig. III.5). The average K<sup>+</sup>-O metal-ligand bond distance at the site is 2.81Å, which agrees well with an average bond length of 2.84Å for similar coordination in the Cambridge Structural Database (CSD)<sup>22</sup>. Conserved Arg33 interacts with the potassium ligand Glu30 and links this metal-binding site to residues responsible for both nonspecific and specific DNA binding (Fig. III.5). Glu30, Arg33, and Asp34 all originate from the second helix of the RHH domain, explaining the high degree of conservation along this helix among NikR sequences that is not seen for other RHH repressors. Thus this metal site appears to stabilize the MBD-RHH domain interface and lock NikR into a conformation that is suitable for interacting with DNA.

To examine whether this second metal-binding site is functionally relevant to the formation of the NikR-DNA complex, a D34A mutation was introduced into NikR and mobility shift assays (MSA) were performed in the presence of excess nickel. D34A-

NikR has the same secondary structure content and high-affinity nickel-binding properties as WT-NikR (data not shown). However, this mutation results in very weak DNA binding and no shift indicative of a 4:1 NikR-DNA complex (Fig. III.6), demonstrating that this residue is critical for the response of NikR to excess metal ions. Preliminary experiments with mutations of the other potassium ligand, E30, suggest that this residue is also required (data not shown).

Since the concentration of  $K^+$  in the NikR-DNA complex crystallization solution is less than that within an *E. coli* cell<sup>23</sup>, we expect that under physiological conditions this site would be occupied by  $K^+$  if no metal with a higher affinity for this site is available. Previous experiments to measure the affinity of nickel for the “low-affinity” nickel-binding site demonstrated that nanomolar concentrations of excess nickel are sufficient to increase the affinity of NikR for its operator, even in the presence of 100 mM potassium<sup>9</sup>. It is possible that excess nickel ions displace potassium at the second metal binding site when present and thereby increase operator affinity, although preliminary experiments suggest that excess  $Ni^{2+}$  ions bind elsewhere on NikR and may increase affinity via increased electrostatic attraction (chapter IV). NikR would therefore be capable of multiple levels of transcriptional repression, perhaps reducing *nikABCDE* transcription with stoichiometric nickel and potassium bound through a “localize and lock” mechanism, and completely arresting transcription when excess nickel ions are available by increasing electrostatic interactions.

Metal-responsive transcription factors play key roles in intracellular physiology, and NikR is especially important for microorganisms that rely on nickel ions for specialized metabolic processes. The crystal structures described here demonstrate that

NikR appears to employ a different mechanism of ligand induction than has been hypothesized for most transcription factors. Instead of having a direct effect on the orientation of the RHH DNA-binding domains, the biggest role of high affinity nickel binding seems to be localization of the repressor to the DNA helix via a short-range allosteric effect that creates a DNA-complementary surface on the MBD. Identification of a functionally relevant second metal binding site suggests a mechanism for stabilizing NikR in the DNA-bound conformation.

### **III.D Materials and Methods**

Nickel-bound selenomethionine (SeMet) labeled NikR was produced and purified as described previously<sup>10</sup>. All crystallization described here is using SeMet NikR.

#### *III.D.1 NikR-DNA complex formation*

Deprotected and desalted deoxyoligonucleotides containing the NikR operator sequence were obtained from Integrated DNA Technologies. The oligos were purified by reverse-phase HPLC on a C8 column using a triethylamine buffered acetonitrile-water gradient. Purified complimentary oligos were annealed by mixing equimolar quantities in water and cooling from 85 to 4 °C over ~12 h and lyophilized. NikR-DNA complex was produced by adding a solution of 10 mg ml<sup>-1</sup> Ni-NikR in 20 mM tris(hydroxymethyl)aminomethane (Tris) pH 8.5, 300 mM NaCl directly to the lyophilized oligos with a ratio of 4 mol NikR/1.2 mol double-stranded oligo. The mixture was then exchanged into 10 mM Tris pH 8.5, 100 mM NaCl using a Micro Bio-Spin P-6 column (BioRad), and incubated on ice for 30 min prior to crystallization.

### III.D.2 Crystallization and data collection

Crystals of full-length NikR with stoichiometric nickel ions bound were grown using sitting-drop vapor diffusion by mixing 2  $\mu\text{L}$  of Ni-NikR (10 mg  $\text{mL}^{-1}$  in 20 mM Tris pH 8.5, 300 mM NaCl) with 2.5  $\mu\text{L}$  of precipitant solution (200 mM  $\text{MgCl}_2$ , 100 mM N-[2-Hydroxyethyl]piperazine-N'-[4-butanefulfonic acid] (HEPES), 30% v/v polyethylene glycol (PEG) 400) and 0.5  $\mu\text{L}$  of 345 mM cyclohexyl-propyl- $\beta$ -D-maltoside (CYMAL-3) at room temperature. Arrowhead-shaped crystals appeared after one day at 20  $^\circ\text{C}$  and were mounted in fiber loops and cooled to 100 K in a gaseous  $\text{N}_2$  stream for data collection without additional cryoprotection. Data were collected at beamline 5.0.2 of the Advanced Light Source on an ADSC q315 CCD detector. Diffraction data were reduced and scaled in DENZO and SCALEPACK<sup>24</sup>, respectively. Crystals of Ni-NikR belonged to the space group  $\text{P3}_121$  with unit cell dimensions  $a=b=50.4 \text{ \AA}$ ,  $c=182.3 \text{ \AA}$ .

Diffraction-quality crystals of the NikR-operator complex were obtained with a 30 bp blunt-ended oligonucleotide encompassing the wild-type *E. coli* NikR operator sequence within the promoter of the *nik* operon, as shown in Fig. III.3.b. Plate-shaped crystals measuring 400x200x50  $\mu\text{m}$  were grown at 20  $^\circ\text{C}$  using the hanging drop vapor diffusion method in a drop consisting of 2  $\mu\text{L}$  of NikR-DNA complex mixed with 2  $\mu\text{L}$  of precipitant solution (200 mM KCl, 50 mM  $\text{MgCl}_2$ , 50 mM Tris pH 7.5, and 10% w/v PEG 4,000), appearing after 2-3 days. These crystals were cryoprotected by sequential 30 second soaks in the precipitant solution supplemented with 10%, 20%, and 30% ethylene glycol prior to cryo-cooling in a gaseous  $\text{N}_2$  stream at 100 K. Data collection and reduction were as described above except that the wavelength was chosen to

maximize the anomalous signal from selenium. NikR-DNA complex crystals belonged to the space group C2 with unit cell dimensions  $a=197 \text{ \AA}$ ,  $b=76 \text{ \AA}$ ,  $c=132 \text{ \AA}$ ,  $\beta=110^\circ$ . NikR is known to bind its operator as a tetramer, and two such NikR tetramer-DNA operator complexes were present in the asymmetric unit (ASU).

### *III.D.3 Structure solution and refinement*

The structure of full-length Ni-NikR was solved by molecular replacement (MR) using the program EPMR<sup>25</sup> and a model of a dimer of the MBD (pdb code 1Q5Y) with data between 15 and 4  $\text{\AA}$  resolution. The top solution from this search had reasonable packing in the space group P3<sub>1</sub>21, and this model was held fixed while searching for the RHH domain using a RHH dimer model from the apo-NikR structure (1Q5V). This combined model was refined by rigid body refinement followed by cycles of positional and B-factor refinement in CNS<sup>26</sup> alternated with manual inspection against simulated annealing composite omit (SA-omit) electron density maps and rebuilding of the protein model in Xfit<sup>27</sup>. This led to a model with good geometry (see Table III.1) comprising two molecules of NikR, two Ni<sup>2+</sup> ions, twenty-seven water molecules, and one molecule of the CYMAL-3 detergent. CNS topology and parameter files for the CYMAL-3 detergent molecule were derived from slight alteration of the MA4 (cyclohexyl-hexyl-beta-d-maltoside) entry from the HIC-Up server<sup>28</sup>. Residues 44-48 within the flexible linker of one NikR molecule and the two C-terminal residues of each NikR molecule were disordered and not included in the final model.

To solve the NikR-operator DNA complex, two MBD tetramers and two RHH DNA-binding domain dimers were located by MR with the program Phaser<sup>29</sup>, using data

between 50 and 3.1 Å resolution. MR models were derived from the Ni-bound MBD (1Q5Y) and apo-NikR (1Q5V) structures, respectively. Electron density maps calculated from this partial model allowed placement of two additional DNA-binding domain dimers. Electron density resembling a DNA double helix was evident in maps calculated from this model, but was of poor quality. An anomalous difference fourier map was calculated in CNS using phases from this molecular replacement model and data collected at the selenium absorption peak, allowing accurate positioning of 16 selenium atoms and 6 nickel ions. These sites were refined and used to calculate SAD phases in the program SOLVE<sup>30</sup> with the Se peak data to 3.4 Å resolution. The electron density map obtained from SAD phasing was modified using the program DM<sup>31</sup> from the CCP4 suite<sup>32</sup> by phase combination with a polyalanine version of the molecular replacement solution, solvent flattening, histogram matching, non-crystallographic symmetry averaging, and phase extension to 3.1 Å. The resulting electron density maps were of excellent quality, clearly showing the DNA double helix with individual base pairs resolved (Fig. III.7). A model of the two crystallographically unique oligonucleotides was built into this density modified map, each comprising all thirty base-pairs of the synthesized sequence. The pseudo-palindromic oligonucleotides in the crystal exhibit chain-orientational disorder, equally populating the two possible directional orientations. Specific nucleotide base contacts are made by NikR to only those portions of the oligo that are strictly palindromic and therefore not affected by the orientational disorder. Crystallographic refinement of the NikR-DNA complex model against a MLHL target was carried out using CNS. Highly weighted NCS restraints were applied during the early cycles of refinement and lowered during later cycles. Sugar pucker and Watson-

Crick base-pairing restraints were applied to the DNA throughout. Several refinement cycles consisting of simulated annealing, positional, and B-factor refinement followed by manual inspection against SA-omit maps and rebuilding using Xfit produced the crystallographic model reported in Table III.1. The final NikR-DNA complex model comprises eight NikR molecules (residues 1-131), two double stranded 30 base-pair deoxyoligonucleotides, eight Ni<sup>2+</sup> ions, and eleven K<sup>+</sup> ions. Ni<sup>2+</sup> ions were identified from an anomalous difference fourier map calculate using data collected at 1.4845 Å (data not shown). K<sup>+</sup> ions were identified based on the contents of the crystallization solution, the size of the difference electron density peak, and the bond lengths and B-factors of the sites after refinement. All structure figures were prepared using PyMOL<sup>33</sup>.

#### *III.D.4 NikR mutants and mobility shift assays*

Mutants were constructed from the pNIK103 parent vector<sup>4</sup> by using the QuickChange PCR mutagenesis method (Stratagene). The D34A mutation was prepared by using the primers 5'-GTTCCGAAGCTATCCGCGcCATTCTGCGTAGCGC-3' and 5'-GCGCTACGCAGAATGgCGCGGATAGCTTCGGAAC-3'. The E30A mutation was prepared by using 5'-GTTATAACAACCGTTCCGcAGCTATCCGCGACATTCTG-3' and 5'-CAGAATGTCGCGGATAGCTgCGGAACGGTTGTTATAAC-3'. The fidelity of the mutagenesis was confirmed by DNA sequencing (ACGT, Toronto). Wild-type NikR and the D34A mutant were purified on a Ni(II)-NTA column (Qiagen), followed by anion exchange chromatography, as previously described<sup>4,8</sup>. The molecular mass of the D34A mutant was confirmed by electrospray mass spectrometry ( $MW_{\text{calc}} = 15049.8$  Da,  $MW_{\text{obs}} =$



15049.2 Da). An assay with Ellman's reagent was used to show that both wild-type NikR and the D34A mutant were fully reduced, and an HPLC assay was used to confirm that both proteins were apo<sup>34</sup>. Nickel binding experiments were performed as described. Electrophoretic mobility-shift assays were performed with a 100-bp DNA fragment containing the *nik* promoter as previously described<sup>9</sup>, and a binding buffer containing 20 mM Tris (pH 7.5), 100 mM KCl, 3 mM MgCl<sub>2</sub>, 0.1% IGEPAL, 5% glycerol, and 0.1 mg/ml sonicated herring sperm DNA (Promega). Nickel was added to the binding buffer, running buffer, and the gel before polymerization, to a final concentration of 35  $\mu$ M.

### III.E Acknowledgements

I would like to thank Sheila Wang and Dr. Deborah Zamble from the University of Toronto for contributing the mutagenesis and MSA data described in this chapter and for many helpful discussions. This research is supported in part by the National Institutes of Health and the Lester Wolfe Predoctoral Fellowship and NSERC (Canada). ALS beamline 5.0.2 is funded by the U.S. Department of Energy.

### III.F References

1. Mulrooney, S.B. & Hausinger, R.P. Nickel uptake and utilization by microorganisms. *FEMS Microbiology Reviews* **27**, 239-261 (2003).
2. 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. *Molecular Microbiology* **9**, 1181-91 (1993).
3. De Pina, K., Desjardin, V., Mandrand-Berthelot, M.A., Giordano, G. & Wu, L.F. Isolation and characterization of the *nikR* gene encoding a nickel-responsive regulator in *Escherichia coli*. *Journal of Bacteriology* **181**, 670-4 (1999).

4. Chivers, P.T. & Sauer, R.T. Regulation of high affinity nickel uptake in bacteria. Ni<sup>2+</sup>-Dependent interaction of NikR with wild-type and mutant operator sites. *Journal of Biological Chemistry* **275**, 19735-41 (2000).
5. Contreras, M., Thiberge, J.M., Mandrand-Berthelot, M.A. & Labigne, A. Characterization of the roles of NikR, a nickel-responsive pleiotropic autoregulator of *Helicobacter pylori*. *Mol Microbiol* **49**, 947-63 (2003).
6. van Vliet, A.H.M., Ernst, F.D. & Kusters, J.G. NikR-mediated regulation of *Helicobacter pylori* acid adaptation. *Trends in Microbiology* **12**, 489-494 (2004).
7. Chivers, P.T. & Sauer, R.T. NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA. *Chemistry and Biology* **9**, 1141-8 (2002).
8. Wang, S.C., Dias, A.V., Bloom, S.L. & Zamble, D.B. Selectivity of metal binding and metal-induced stability of *Escherichia coli* NikR. *Biochemistry* **43**, 10018-28 (2004).
9. Bloom, S.L. & Zamble, D.B. Metal-selective DNA-binding response of *Escherichia coli* NikR. *Biochemistry* **43**, 10029-38 (2004).
10. Schreiter, E.R. et al. Crystal structure of the nickel-responsive transcription factor NikR. *Nat Struct Biol* **10**, 794-9 (2003).
11. Raumann, B.E., Rould, M.A., Pabo, C.O. & Sauer, R.T. DNA recognition by beta-sheets in the Arc repressor-operator crystal structure. *Nature* **367**, 754-7 (1994).
12. Somers, W.S. & Phillips, S.E. Crystal structure of the met repressor-operator complex at 2.8 Å resolution reveals DNA recognition by beta-strands. *Nature* **359**, 387-93 (1992).
13. Gomis-Ruth, F.X. et al. The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO Journal* **17**, 7404-15 (1998).
14. Carrington, P.E., Chivers, P.T., Al-Mjeni, F., Sauer, R.T. & Maroney, M.J. Nickel coordination is regulated by the DNA-bound state of NikR. *Nature Structural Biology* **10**, 126-130 (2003).
15. Raumann, B.E., Brown, B.M. & Sauer, R.T. Major groove DNA recognition by β-sheets: the ribbon-helix-helix family of gene regulatory proteins. *Current Opinion in Structural Biology* **4**, 36-43 (1994).
16. White, A., Ding, X., vanderSpek, J.C., Murphy, J.R. & Ringe, D. Structure of the metal-ion-activated diphtheria toxin repressor/tox operator complex. *Nature* **394**, 502-6 (1998).

17. Orth, P., Schnappinger, D., Hillen, W., Saenger, W. & Hinrichs, W. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat Struct Biol* **7**, 215-9 (2000).
18. van Aalten, D.M., DiRusso, C.C. & Knudsen, J. The structural basis of acyl coenzyme A-dependent regulation of the transcription factor FadR. *Embo J* **20**, 2041-50 (2001).
19. Phillips, K. & Phillips, S.E. Electrostatic activation of Escherichia coli methionine repressor. *Structure* **2**, 309-16 (1994).
20. Zhang, R.G. et al. The crystal structure of trp aporepressor at 1.8 Å shows how binding tryptophan enhances DNA affinity. *Nature* **327**, 591-7 (1987).
21. Otwinowski, Z. et al. Crystal structure of trp repressor/operator complex at atomic resolution. *Nature* **335**, 321-9 (1988).
22. Harding, M.M. Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium. *Acta Crystallogr D Biol Crystallogr* **58**, 872-4 (2002).
23. Epstein, W. & Schultz, S.G. Cation transport in Escherichia coli. VI. K exchange. *J Gen Physiol* **49**, 469-81 (1966).
24. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326 (1997).
25. Kissinger, C.R., Gehlhaar, D.K. & Fogel, D.B. Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr D Biol Crystallogr* **55** ( Pt 2), 484-91 (1999).
26. Brünger, A.T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica. Section D: Biological Crystallography* **54**, 905-21 (1998).
27. McRee, D.E. XtalView/Xfit - A Versatile Program for Manipulating Atomic Coordinates and Electron Density. *Journal of Structural Biology* **125**, 156-165 (1999).
28. Kleywegt, G.J. & Jones, T.A. Databases in protein crystallography. *Acta Crystallogr D Biol Crystallogr* **54**, 1119-31 (1998).
29. Storoni, L.C., McCoy, A.J. & Read, R.J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* **60**, 432-8 (2004).
30. Terwilliger, T.C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallographica. Section D: Biological Crystallography* **55**, 1174-1178 (1999).

31. Cowtan, K.D. & Main, P. Phase combination and cross validation in iterated density-modification calculations. *Acta Crystallographica. Section D: Biological Crystallography* **52**, 43-48 (1996).
32. Collaborative Computational Project, N. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallographica. Section D: Biological Crystallography* **50**, 760-763 (1994).
33. DeLano, W.L. The PyMOL Molecular Graphics System. (DeLano Scientific, San Carlos, 2002).
34. Atanassova, A., Lam, R. & Zamble, D.B. A high-performance liquid chromatography method for determining transition metal content of proteins. *Analytical Biochemistry* **335**, 103-111 (2004).

### III.G Tables and Figures

**Table III.1** Data collection, phase determination and refinement statistics.

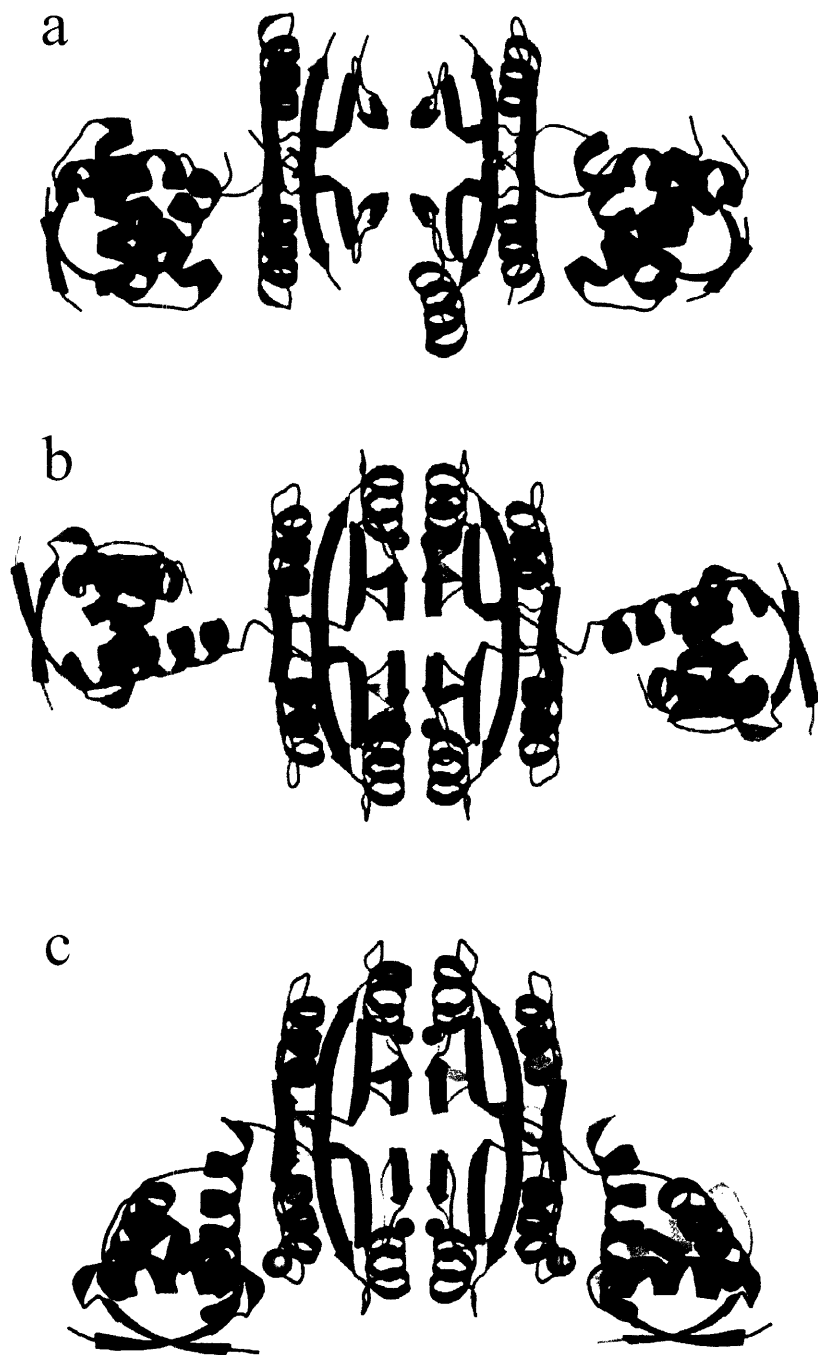
<b>Data Collection</b>			
<b>Crystal</b>	Ni-NikR	NikR+DNA 1	NikR+DNA 2
<b>Data Set</b>	Native	Native	Se-Peak
<b>Space Group</b>	P3 <sub>1</sub> 21	C2	C2
<b>Cell Dimensions</b>			
<b>a (Å)</b>	49.84	197.20	195.97
<b>b (Å)</b>	49.84	76.13	76.02
<b>c (Å)</b>	181.69	132.09	132.28
<b>β (°)</b>	90	110.27	110.37
<b>Wavelength (Å)</b>	1.0000	1.0000	0.9793
<b>Temperature (K)</b>	100	100	100
<b>Resolution Range (Å)</b>	50-2.1	50-3.1	50-3.4
<b>Observations</b>	184200	146795	156641
<b>Unique Reflections</b>	15884	31966	24360
<b>Completeness (%)<sup>1</sup></b>	98.2(90.0)	94.1(70.3)	95.1(74.8)
<b><i>I</i> / σ(<i>I</i>)<sup>1</sup></b>	20.8(3.7)	15.1(4.5)	15.0(6.7)
<b><i>R</i><sub>sym</sub> (%)<sup>1,2</sup></b>	8.0(30.5)	6.1(17.5)	7.0(17.9)
<b>Phasing of NikR+DNA</b>			
<b>Number of sites/asu</b>			22
<b>Resolution Range (Å)</b>			50-3.5
<b>Figure of Merit</b>			0.33
<b>Refinement</b>			
<b><i>R</i><sub>cryst</sub> (<i>R</i><sub>free</sub>) (%)<sup>3</sup></b>	23.9(27.5)	26.1(30.3)	
<b>Resolution Range (Å)</b>	50-2.1	50-3.1	
<b>Number of atoms (Average <i>B</i>-factor (Å<sup>2</sup>))</b>			
<b>Protein</b>	1946(58.6)	8034(101.7)	
<b>DNA</b>	-(-)	2448(92.6)	
<b>Nickel/Potassium ions</b>	2(48.8)	19(104.6)	
<b>Detergent atoms</b>	32(83.5)	-(-)	
<b>Water molecules</b>	25(54.4)	-(-)	
<b>R.m.s. deviations</b>			
<b>Bond lengths (Å)</b>	0.007	0.008	
<b>Bond angles (°)</b>	1.28	1.30	
<b>Ramachandran (%)</b>			
<b>Most favored</b>	88.1	79.4	
<b>Additionally Allowed</b>	11.0	19.2	
<b>Generously Allowed</b>	0.8	1.3	
<b>Disallowed</b>	0.0	0.0	

<sup>1</sup>The number in parentheses is for the highest resolution shell.

<sup>2</sup> $R_{\text{sym}} = \sum_{hkl} |I_{(hkl)}^i - \langle I_{(hkl)} \rangle| / \sum_{hkl} \langle I_{(hkl)} \rangle$ , where  $I_{(hkl)}^i$  is the  $i^{\text{th}}$  measured diffraction intensity and  $\langle I_{(hkl)} \rangle$  is the mean of the intensity for the miller index ( $hkl$ ).

<sup>3</sup> $R_{\text{cryst}} = \sum_{hkl} ||F_o(hkl)| - |F_c(hkl)|| / \sum_{hkl} |F_o(hkl)|$ .  $R_{\text{free}} = R_{\text{cryst}}$  for a test set of reflections (6% for Ni-NikR, 7% for NikR-DNA) not included in refinement.

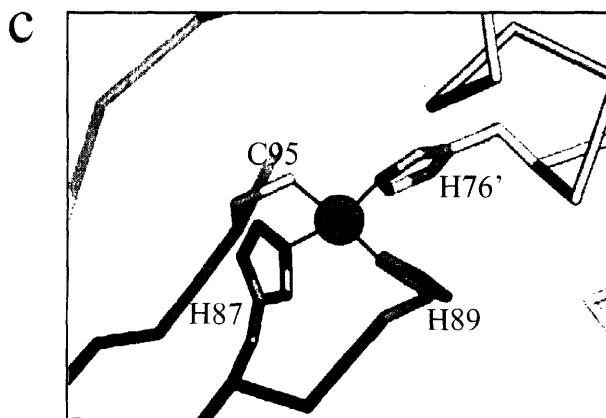
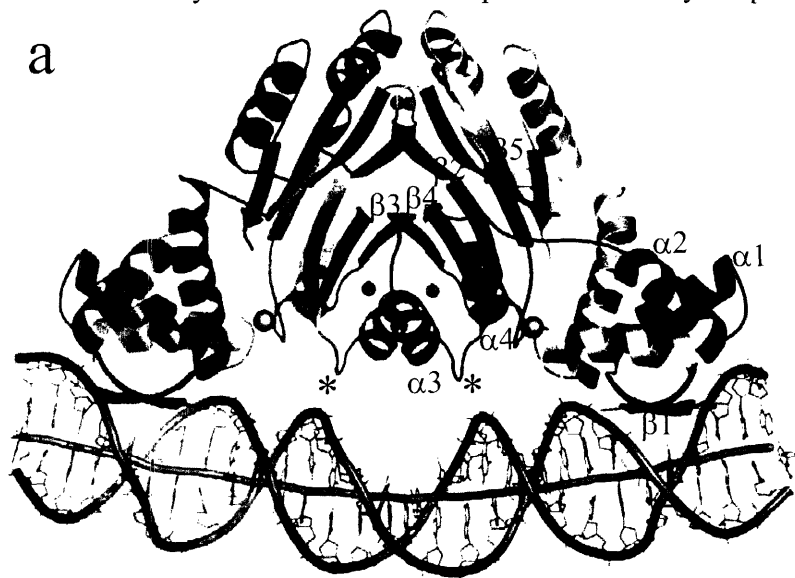
**Figure III.1** Conformational flexibility of NikR. **a:** Apo-NikR tetramer displayed as a ribbon with the MBD colored grey and the RHH domains colored red. **b:** Nickel-activated NikR tetramer displayed as in **a** except the RHH domains are colored green and nickel ions are shown as cyan spheres. **c:** Operator-bound NikR tetramer displayed as in **a** and **b**, except the RHH domains are colored blue and potassium ions are shown as pink spheres.



**Figure III.2** The cyclohexyl-propyl- $\beta$ -D-maltoside detergent bound to Ni-NikR. NikR is represented as a cartoon colored by protein subunit. Bound nickel ions are cyan spheres. The detergent molecule is shown as sticks colored by atom type.

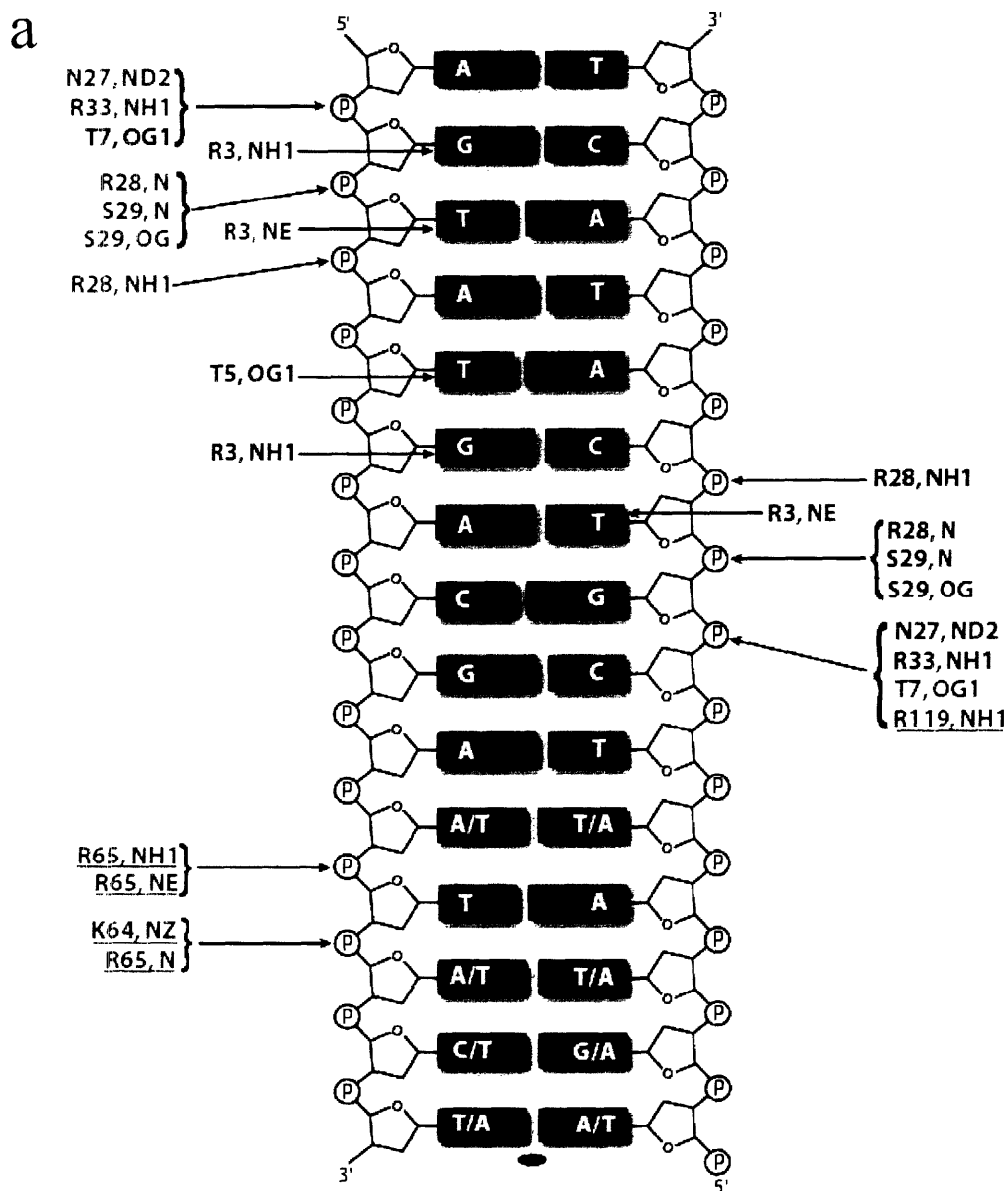


**Figure III.3** The NikR-operator DNA complex. **a:** Ribbon representation of the NikR-operator DNA complex structure. The NikR tetramer is colored by subunit, and secondary structure elements are labeled. Asterisks indicate the loop of the MBD that contacts the operator DNA backbone. DNA is displayed as sticks with a cartoon tube through the phosphorus positions to represent the backbone. The dyad-symmetric operator half-sites are colored orange. The DNA helical axis as calculated in CURVES is shown as a purple tube. Nickel and potassium ions are shown as cyan and pink spheres, respectively. **b:** The dsDNA used for co-crystallization with NikR, which includes the wild-type NikR operator sequence. Dyad-symmetric operator half-sites are highlighted in orange, the -10 region of the *nik* promoter is underscored with red. **c:** The high-affinity nickel-binding site of NikR. Protein is shown as a  $C_{\alpha}$  trace with sidechains as sticks colored by subunit. Nickel is represented as a cyan sphere.

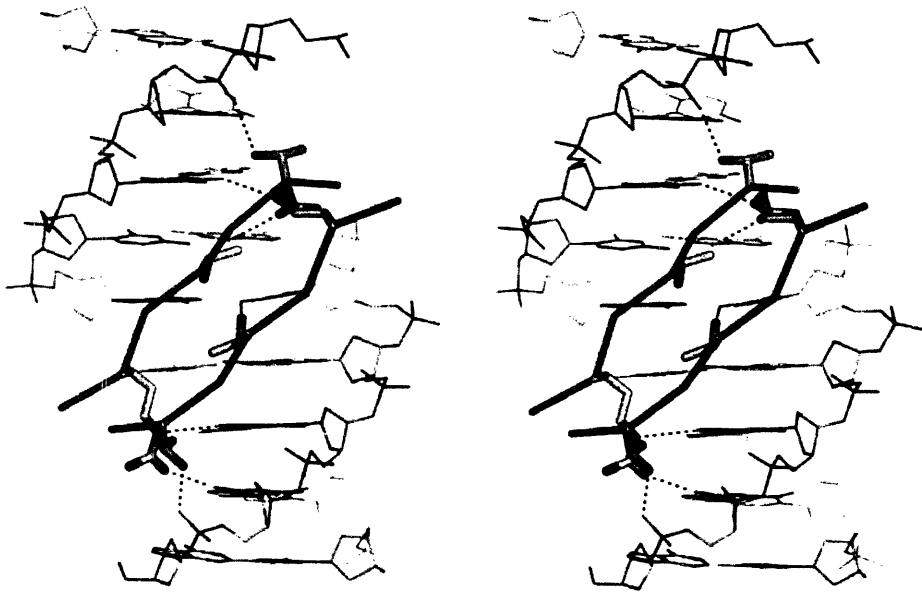




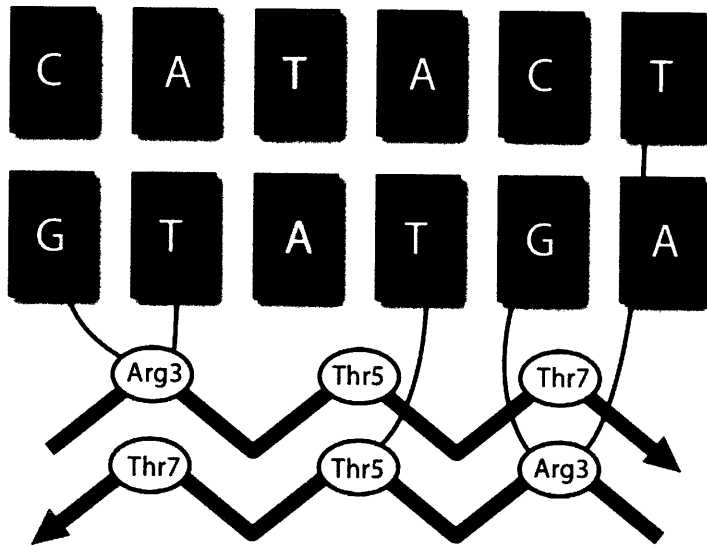
**Figure III.4** NikR-operator DNA interactions. **a:** Schematic representation of polar interactions between NikR and operator DNA. Only half of the operator DNA is shown because the interactions made with the other half are symmetric and equivalent. Base-pairs colored blue were shown to abrogate NikR binding when mutated. The protein atoms making interactions are colored by protein subunit, and correspond to the color scheme in Fig. III.2.a. Interactions contributed by the MBD are underlined. **b:** Stereo view of a NikR RHH domain  $\beta$ -sheet bound in the DNA major groove. DNA is represented as sticks colored by atom type. The RHH  $\beta$ -sheet is represented as a  $C_{\alpha}$  trace colored as in Fig III.2, with sidechains shown as sticks. Hydrogen bonds are shown as dashed lines. **c:** Schematic representation of the specific NikR-operator interactions. Base pairs are colored as in **a**, protein is colored by subunit as in Fig. III.2.a. **d,e,f:** Specific DNA contacts made by Arg3 and Thr5. Stick representation colored by atom type with dashed black lines representing hydrogen bonds.



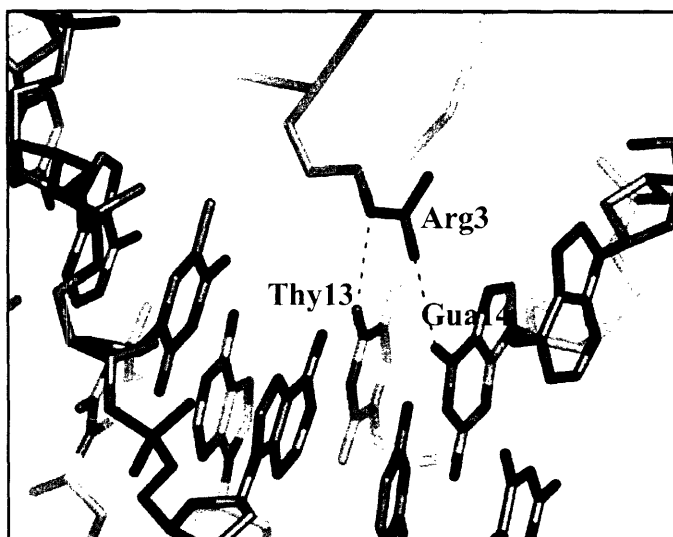
b



c

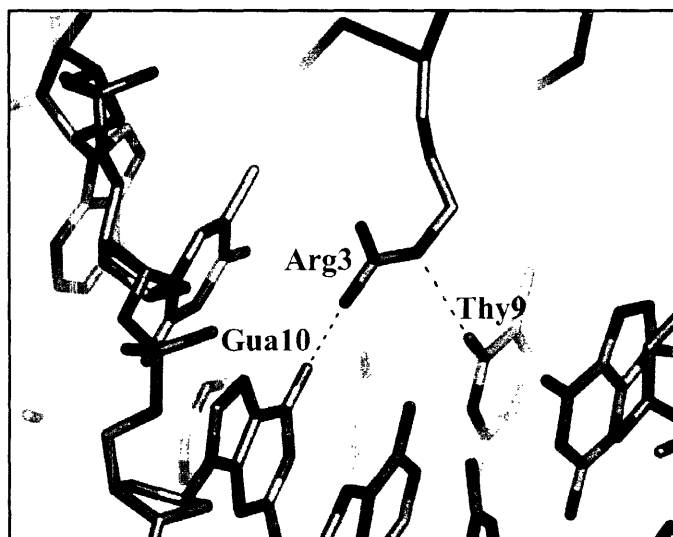


d



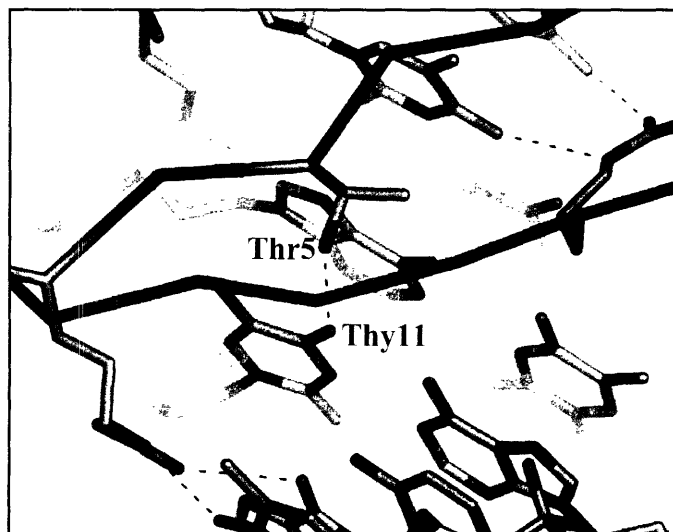
Arg3(Nε) → Thy13(O4)  
Arg3(Nη) → Gua14(O6)

e



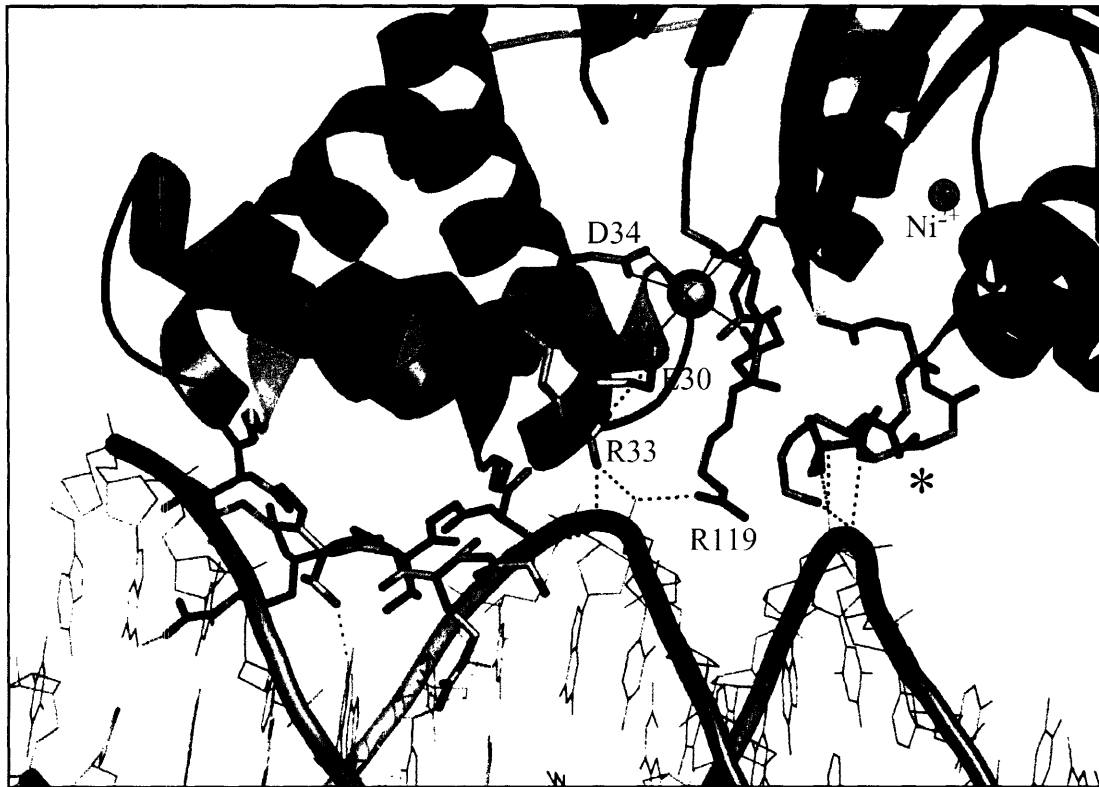
Arg3(Nε) → Thy9(O4)  
Arg3(Nη) → Gua10(O6)

f

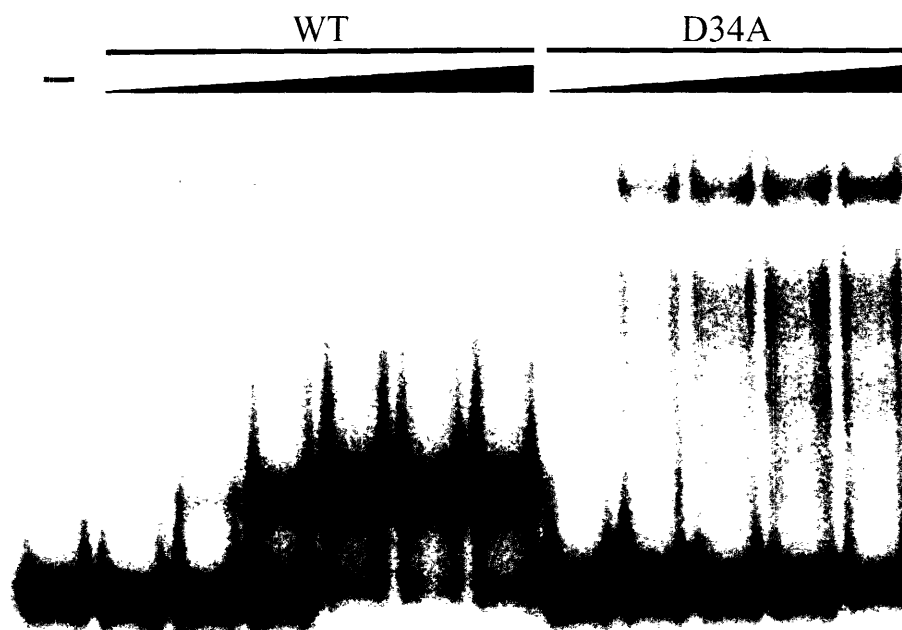


Thr5(Oγ) → Thy11(O4)

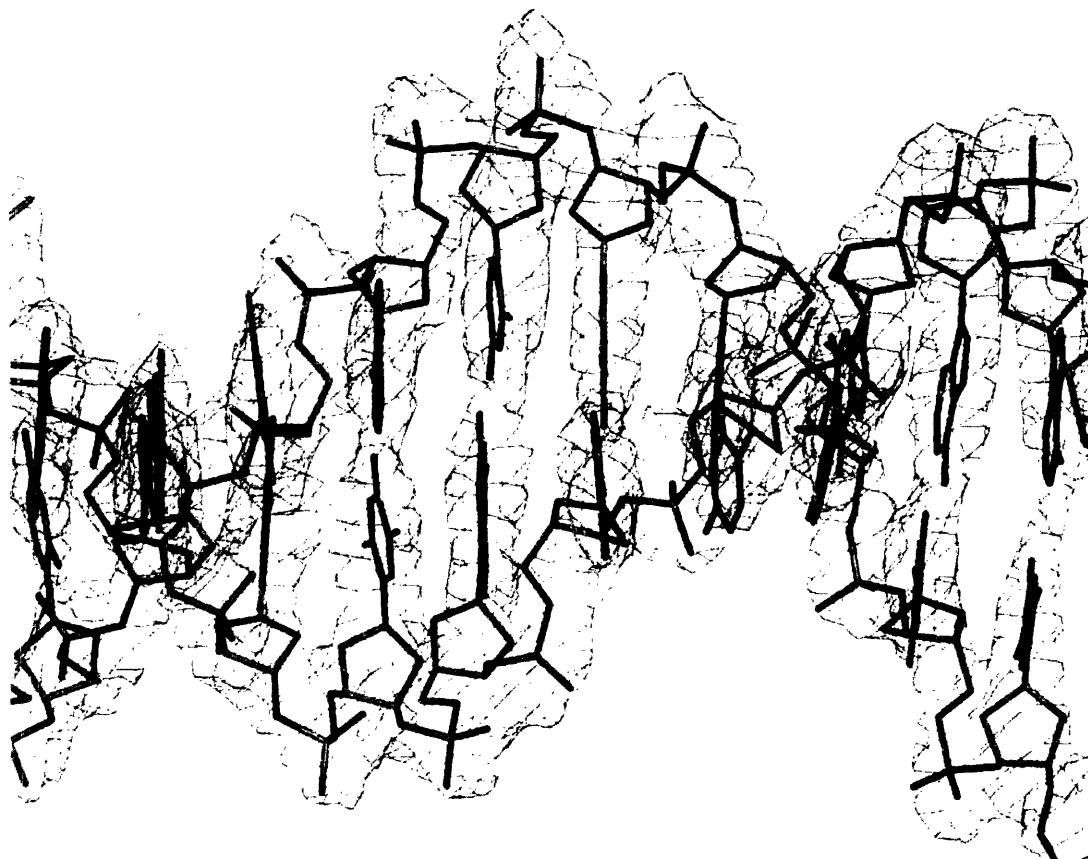
**Figure III.5** The second metal binding site. Protein, DNA, and metal ions are displayed as in Fig. III.2, with sticks for protein sidechains and portions of the backbone. Bonds to the potassium ion are shown as solid black lines and hydrogen bonds as dashed black lines. E30, R33, D34 are conserved among NikR sequences. An asterisk indicates the loop of the MBD that contacts the operator DNA backbone, as described in Fig. III.2.a.



**Figure III.6** DNA binding in the presence of excess metal. WT NikR (1 pM – 100 nM) or the D34A mutant (10 pM – 1  $\mu$ M) containing stoichiometric nickel were incubated with 100-bp *nik* DNA in the presence of 35  $\mu$ M NiSO<sub>4</sub>. The reactions were analyzed on a 7% native gel with 35  $\mu$ M NiSO<sub>4</sub> in the gel and running buffer. Note that although there is 100 mM KCl in the binding buffer, only nickel was added to the gel and running buffer.



**Figure III.7** Averaged experimental electron density map. A portion of the solvent-flattened, averaged experimental electron density map contoured at  $2\sigma$  is shown superimposed on the final refined model of DNA. DNA is shown as sticks colored by atom type.



## Chapter IV: Divalent Transition Metal Ion Binding to NikR

### IV.A. Summary

*Escherichia coli* NikR regulates the transcription of a nickel uptake protein complex, but is capable of binding a range of divalent transition metal ions in addition to  $\text{Ni}^{2+}$ , including  $\text{Co}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ , and  $\text{Cd}^{2+}$ . To understand how the high-affinity nickel binding site of NikR is able to accommodate these other metal ions, we have determined the structure of the metal binding domain (MBD) of NikR in complex with  $\text{Cu}^{2+}$  ions and in the presence of  $\text{Zn}^{2+}$ . We observe that  $\text{Cu}^{2+}$  ions bind in a very similar manner to  $\text{Ni}^{2+}$ , but with longer bond lengths and a geometry distorted from square planar. Crystals grown in the presence of  $\text{Zn}^{2+}$  reveal a protein structure similar to apo NikR, but with an electron density peak near the  $\text{Ni}^{2+}/\text{Cu}^{2+}$  binding site tentatively assigned as a  $\text{Zn}^{2+}$  ion with a pyramidal coordination by His87, Cys95, and Glu97, two of which are directly involved in binding  $\text{Ni}^{2+}$  and  $\text{Cu}^{2+}$ . In addition, to understand how excess  $\text{Ni}^{2+}$  ions affect the affinity of NikR for DNA, we soaked crystals of both Ni-NikR alone and the NikR-DNA complex in excess  $\text{Ni}^{2+}$  ions and collected x-ray diffraction data at the nickel absorption peak wavelength to check for additional binding sites. We observe four  $\text{Ni}^{2+}$  binding sites on the surface of the protein that are common to both crystal forms, as well as two sites unique to the NikR-DNA complex, that delineate potential secondary metal-binding sites.

## IV.B. Introduction

The NikR protein from *Escherichia coli* is responsible for regulating intracellular nickel ion concentrations in this organism. This regulation is accomplished by repressing transcription of the nickel permease complex *nikABCDE* in response to increasing nickel (see chapter I for more details). Previous structural and biophysical work has revealed the quaternary organization of NikR, illustrated conformational changes associated with high-affinity nickel binding, shown how NikR binds its operator DNA sequence, and uncovered a second metal binding site on NikR in the presence of DNA<sup>1-6</sup> (see also chapter 3 of this thesis). Recently, work examining the selectivity of metal binding and activation by NikR was published<sup>5,6</sup>. The authors first found that in addition to Ni<sup>2+</sup>, NikR is capable of binding a number of divalent transition metal ions such as Cu<sup>2+</sup>, Co<sup>2+</sup>, Zn<sup>2+</sup> and Cd<sup>2+</sup> with high affinity<sup>6</sup>. Their binding constants approximately follow the Irving-Williams series of divalent transition metal-complex stabilities: Mn < Co < Ni < Cu > Zn<sup>7</sup>. The binding constants relevant to the work described here are  $9 \times 10^{-13}$ ,  $1 \times 10^{-17}$ , and  $< 1 \times 10^{-12}$  M, for Ni<sup>2+</sup>, Cu<sup>2+</sup>, and Zn<sup>2+</sup>, respectively<sup>6</sup>. Very similar values were obtained for the isolated MBD relative to the full-length protein. Although Cu<sup>2+</sup> binds to NikR with a considerably higher affinity than Ni<sup>2+</sup>, thermal and chemical denaturation studies showed that stoichiometric Ni<sup>2+</sup> was able to stabilize the folded structure of NikR by at least 20°C and 0.7 M guanidine HCl relative other metal ions, suggesting perhaps that it has a unique allosteric effect on the protein. Circular dichroism spectra demonstrated that NikR has a greater  $\alpha$ -helical content with Ni<sup>2+</sup> or Cu<sup>2+</sup> bound, relative to the apo or Zn<sup>2+</sup>-bound forms.



A parallel study investigated the metal ion selectivity of the specific DNA-binding response of NikR<sup>5</sup>. Stoichiometric amounts of Cu<sup>2+</sup>, Co<sup>2+</sup>, and Cd<sup>2+</sup> produced binding affinities to the *nik* operator that were similar to Ni<sup>2+</sup> ( $3 \times 10^{-9}$ ,  $1 \times 10^{-8}$ ,  $2 \times 10^{-8}$ , and  $5 \times 10^{-9}$  M for Cu<sup>2+</sup>, Co<sup>2+</sup>, Cd<sup>2+</sup> and Ni<sup>2+</sup>, respectively). Stoichiometric Zn<sup>2+</sup> also allowed operator binding, but with lower affinity. With excess Ni<sup>2+</sup> present, NikR binds its operator sequence with much higher affinity. This higher affinity binding was somewhat specific for nickel, as an excess of other metal ions produced an affinity 30-100 fold weaker than with Ni<sup>2+</sup>. The results support the hypothesis of a second metal binding site on NikR in the presence of DNA, and suggest that nickel may have some selective effect on the conformation or stability of NikR. In chapter III of this thesis, a structure of the NikR-DNA complex revealed a second metal binding site, but was occupied by potassium from the crystallization buffer rather than nickel.

Here we describe crystallographic experiments aimed at understanding how NikR is capable of binding a range of divalent transition metal ions and examining the effect of excess nickel ions on the protein. We have used the isolated MBD of NikR as a model to examine the binding of metal ions to the high-affinity binding site because the isolated domain crystallizes more readily and the crystals diffract X-rays to higher resolution than the full-length protein. Our previous crystallographic work shows that the structures of the MBD and high-affinity nickel-binding site are extremely similar in the context of full-length NikR (chapter III) and the isolated MBD<sup>4</sup> (RMSD of 0.88 Å for alignment of 316 common C<sub>α</sub> atoms). In addition, metal ion affinities and protein stability in the presence of metal ions are similar for full-length NikR and the isolate MBD, suggesting that the isolated MBD is a good model for studying high-affinity metal ion binding to NikR.

## IV.C. Results and Discussion

### IV.C.1 Crystal structure of $\text{Cu}^{2+}$ -MBD NikR

The structures of the  $\text{Cu}^{2+}$ -MBD and  $\text{Ni}^{2+}$ -MBD complexes are extremely similar, with a RMSD = 0.22 Å for 304  $C_\alpha$  atoms of the MBD tetramer.  $\text{Cu}^{2+}$  ions bind in the same position in the high-affinity site of NikR as  $\text{Ni}^{2+}$  ions, and with similar coordination geometry (Fig. IV.2). The  $\text{Cu}^{2+}$  site is distorted square planar with longer protein-metal ion bond lengths relative to the  $\text{Ni}^{2+}$ -bound structure, particularly for the trans ligands H89 and C95 (Table IV.1). Visible absorption spectra of both  $\text{Ni}^{2+}$  and  $\text{Cu}^{2+}$  bound to NikR or the isolated MBD show features that are suggestive of square-planar or distorted square-planar geometry and at least one thiolate ligand<sup>6</sup>.

Examples of square-planar complexes of both  $\text{Ni}^{2+}$  and  $\text{Cu}^{2+}$  with proteins (from the PDB) and small molecules (from CSD) are relatively common<sup>8</sup>. The distances between the bound  $\text{Cu}^{2+}$  ion and H89 and C95 of MBD NikR (2.44 and 2.43 Å, respectively) are considerably longer, however, than the mean  $\text{Cu}^{2+}$ -ligand bond distances observed for equivalent interactions found in the PDB (2.02 and 2.15 Å). It is possible that the difference in metal-ligand bond lengths observed for  $\text{Cu}^{2+}$  could explain the differential effect on NikR stability relative to  $\text{Ni}^{2+}$  that was observed previously<sup>6</sup>. We proposed previously that ordering of helix  $\alpha 3$  of NikR via coordination of H76 from this helix to a bound metal ion is important for the increase in affinity for *nik* operator observed in the presence of stoichiometric metal ions (chapter III). This proposal is supported by the similarity in operator affinity of the  $\text{Ni}^{2+}$  and  $\text{Cu}^{2+}$  bound proteins, since H76 is involved in coordination of both ions. Although it seems strange that a

transcription factor regulating a nickel uptake operon would respond as efficiently to copper ions, it has been demonstrated that there is likely no copper available in the cytoplasm of *E. coli* since the transcription factor, CueR, regulating production of the copper efflux pump exhibits zeptomolar sensitivity to copper ions<sup>9</sup>.

In comparison to the functionally diverse family of blue-copper proteins, Cu<sup>2+</sup>-NikR and Cu<sup>2+</sup>-MBD are deep orange in color (Fig. IV.1), owing to a broad, intense absorption band centered at 388 nm with an extinction coefficient of 7600 M<sup>-1</sup>cm<sup>-1</sup>.<sup>6</sup> Although both NikR and the blue-copper proteins have coordination of the Cu<sup>2+</sup> ion by histidines and a cysteine thiolate, the coordination geometries are significantly different. Blue-copper proteins generally have a trigonal planar or pyramidal coordination by two histidines and a cysteine, with one or two additional longer interactions<sup>10</sup>. The square-planar Cu<sup>2+</sup>-binding site of NikR is shown in Fig. IV.2.

#### *IV.C.2 Crystal structure of MBD NikR crystallized in the presence of Zn<sup>2+</sup> ions*

MBD NikR was crystallized in the presence of Zn<sup>2+</sup> ions in a new crystal form (Fig. IV.3) and the structure was solved by molecular replacement. Following several cycles of refinement and refitting of the MBD structure determined in the presence of Zn<sup>2+</sup>, a +4.2 $\sigma$  peak in the F<sub>o</sub>-F<sub>c</sub> electron density map was evident near the high-affinity Ni<sup>2+</sup>/Cu<sup>2+</sup> binding site between the sidechains of H87, C95 and E97 (Fig. IV.4). The rotamers of these sidechains more closely resemble those in the apo-NikR structure than the Ni<sup>2+</sup>-bound structure (Fig. II.3.a). The location of the peak is intriguing because it is near the high-affinity nickel binding site and is located between three sidechains that are important for binding nickel<sup>11</sup>. Modeling of this peak as Zn<sup>2+</sup> ion results in a high B-

factor for the  $\text{Zn}^{2+}$  ( $74.6 \text{ \AA}^2$ ) relative to the average B-factors for the protein ( $47.0 \text{ \AA}^2$ ) or for the surrounding sidechains ( $40\text{-}55 \text{ \AA}^2$ ). Modeling the peak as a water molecule, however, results in a relatively low B-factor for the water ( $31.8 \text{ \AA}^2$ ) compared to surrounding residues. Since the B-factor and occupancy of an atom are correlated, the B-factor of the unknown atom was fixed at  $47.0 \text{ \AA}^2$ , the average B-factor for the protein atoms in the structure, and the occupancy of either  $\text{Zn}^{2+}$  ion or a water molecule was refined. Refinement of  $\text{Zn}^{2+}$  at this position gives an occupancy of 0.57, and refinement as water gives an occupancy of 2.18. This result suggests either a low-occupancy  $\text{Zn}^{2+}$  ions or a very well ordered water molecule. The geometry of the site is approximately tetrahedral with an open coordination site, so assignment as a bound metal ion seems possible (Fig. IV.4). The solution from which the crystals grew contained the metal ions  $\text{Na}^+$ ,  $\text{Fe}^{3+}$ , and  $\text{Zn}^{2+}$ . The coordination is not reasonable for  $\text{Na}^+$ , which strongly prefers six-coordinate sites of all oxygen ligands<sup>12</sup>. To evaluate  $\text{Fe}^{3+}$  as a possibility, inverse beam data were collected on the same crystal and the data was scaled anomalously, keeping Friedel pairs separate. These data and phases from the partially refined model were used to calculate an anomalous difference Fourier map, which should exhibit a peak at the position of any iron atoms in the structure since the data was collected at a wavelength ( $1.5418 \text{ \AA}$ ) where iron has a significant anomalous signal ( $\sim 3.2 e^-$ ). The anomalous difference Fourier map showed no significant peak at the position of the unknown atom, ruling out  $\text{Fe}^{3+}$  as a possibility. The peak was therefore tentatively assigned as  $\text{Zn}^{2+}$  with an occupancy of 0.57 (Fig. IV.5). Future multi-wavelength data collection around the Zn absorption edge could confirm the assignment. Although the protein sidechain to  $\text{Zn}^{2+}$  distances are somewhat long ( $\sim 2.5 \text{ \AA}$ ) for a coordinated metal

ion (2.14-2.36 Å for equivalent interactions in the PDB)<sup>8</sup>, it is possible that the model reflects the average of the Zn<sup>2+</sup>-bound state, where the ligands are closer, and the unbound state, where the protein sidechains are in an alternate conformation. The low resolution of the data (2.5 Å), however, prevents the evaluation of this possibility. Superposition of the MBD+Zn<sup>2+</sup> structure with one subunit of the Ni<sup>2+</sup>-MBD structure shows that the Ni<sup>2+</sup> site and putative Zn<sup>2+</sup> site are separated by ~4.5 Å (Fig. IV.6).

One concern regarding the relevance of this structure is the relatively low pH (5.6) at which MBD NikR was crystallized. A pH significantly below seven could affect the protonation state of the histidine and cysteine sidechain ligands and consequently the affinity of Zn<sup>2+</sup> ions for the site. More importantly, the pH difference could alter the binding mode of Zn<sup>2+</sup> ions to the protein if multiple, related sites are accessible.

Square-planar coordination of Zn<sup>2+</sup> in proteins or small molecule complexes is rare<sup>8</sup>, so it could be expected that zinc ions bind to NikR via a different mode than either Ni<sup>2+</sup> or Cu<sup>2+</sup>. The largest difference between the sites with respect to the function of NikR is the loss of ligation by His76' across the tetramer interface. This explains the disorder of helix  $\alpha$ 3 in the structure (His76 is located on helix  $\alpha$ 3) and lower  $\alpha$ -helical content for the Zn<sup>2+</sup>-bound form relative to Ni<sup>2+</sup> or Cu<sup>2+</sup> as measured by circular dichroism<sup>6</sup>. Perhaps it also explains the 100-fold lower affinity of Zn<sup>2+</sup>-NikR for the *nik* operator<sup>5</sup>, since ordering of helix  $\alpha$ 3 upon metal binding could be a requirement for tight operator binding (see chapter III). Finally, it is unlikely that Zn<sup>2+</sup> will activate NikR *in vivo*, since zinc metalloregulatory proteins exist with femtomolar sensitivity for the metal, making the intracellular “free” zinc concentration essentially zero<sup>13</sup>.

### IV.C.3 Additional $Ni^{2+}$ -binding sites on NikR

By soaking crystals of NikR and the NikR-DNA complex (each containing a  $Ni^{2+}$ :NikR ratio of 1:1) in solutions containing millimolar concentrations of  $NiCl_2$ , we were able to identify secondary nickel-binding sites on NikR. Seven unique sites were observed in addition to the high-affinity binding site (Table IV.2) in anomalous difference Fourier maps calculated from data collected at the nickel absorption peak wavelength (Figs. IV.7 and IV.8). Sites 1-4 were observed on multiple NikR subunits and in both crystal forms, while two sites (6 and 7) were observed only in the context of the NikR-DNA complex. Any of these sites could be considered as possibilities for a secondary nickel binding site. Nickel ions at site 1 are coordinated by the free N-terminal amine and carbonyl oxygen of Met1, and are closest in proximity to the operator in the DNA-bound form of NikR (Fig. IV.8). Sites 2-7 are all located on the surface of the MBD of NikR and all include coordination by histidine sidechains except site 7. None of sites 2-7 is closer than  $\sim 15$  Å to the DNA operator in the complex structure, but sites 3, 4, and 7 are located nearby helix  $\alpha 3$  and could stabilize its ordered structure, which has been proposed to be important for operator binding (chapter III). In addition, sites 2 and 6 are located nearby the inter-domain linker and could help to stabilize a conformation of the RHH domains that is suitable for DNA binding.

The goal of these experiments was to identify candidates for a secondary nickel binding site on NikR that was identified biochemically and shown to increase the affinity of NikR for the *nik* operator by  $\sim 1000$ -fold relative to only stoichiometric nickel. The structure of NikR in complex with operator DNA (chapter III) identified a functionally significant second metal binding site on NikR at the interface between the DNA-binding

Ribbon-Helix-Helix (RHH) domains and the MBD, which was modeled as a potassium ion from the crystallization buffer. It was therefore tempting to speculate that although potassium is present at ~250mM in a typical *E. coli* cell<sup>14</sup>, nickel could have a significantly higher affinity for this site and displace potassium. Soaking a crystal of NikR-DNA in a solution containing 200mM K<sup>+</sup> and 5mM NiCl<sub>2</sub><sup>2+</sup> revealed no nickel bound at the potassium site, however (Fig. IV.8). There are two possible explanations for this result. The first is that nickel can bind at this site, but binding was somehow prevented under the experimental conditions. It is also possible that this site is normally occupied by potassium or another abundant metal ion and that excess nickel ions bind to one or more of the unique sites identified in this study. The increase in the affinity of NikR for its operator could then be due to a simple charge increase on NikR, promoting affinity for negatively charged DNA as has been proposed in the case of the RHH family member MetJ<sup>15</sup>, or a consequence of stabilization of either helix  $\alpha$ 3 or the RHH domain conformation by the nickel ions.

The secondary nickel site characterized biochemically showed specificity for Ni<sup>2+</sup> relative to other divalent transition metal ions<sup>5</sup>. Most of the sites identified here involve coordination by only one or two protein sidechains, and the resolution of data is too low to learn about the details of the coordination geometry of the sites. Although most of the sites involve coordination by a histidine sidechain, this is a common ligand for all of the metal ions that have been shown to bind NikR and does not explain the observed specificity for Ni<sup>2+</sup>. The relevance of the newly identified sites to the function of NikR could be tested in a number of ways. First, mutagenesis could be carried out on the amino acids seen here to participate in coordination of additional nickel ions, followed by

DNA-binding assays of the mutants to determine which sites have an effect. Furthermore, soaking crystals in lower concentrations of excess  $\text{NiCl}_2$  to observe which sites on NikR are bound first could lend support to the relevance of one site over another.

Although none of the  $\text{Ni}^{2+}$  sites identified in the soaking experiments described here involves coordination by conserved sidechain positions, it has not been demonstrated that NikR proteins other than that from *E. coli* are responsive to nickel ions above a 1:1 stoichiometry. The physiological relevance of this secondary site has also not been conclusively established. Transcription assays at varied concentrations of nickel will need to be carried out to determine its importance.

We have demonstrated here that  $\text{Cu}^{2+}$  ions bind in a similar manner to the MBD of NikR as  $\text{Ni}^{2+}$  ions, and that  $\text{Zn}^{2+}$  ions potentially bind a related site without ordering helix  $\alpha 3$ . Furthermore, we identified candidates for a secondary nickel binding site on NikR that can be tested by mutagenesis experiments. Studies such as this are key to providing a molecular description of metal ion-dependent transcriptional repression by this interesting metalloregulator.

#### **IV.D. Materials and Methods**

##### *IV.D.1 Crystal structure of $\text{Cu}^{2+}$ -MBD NikR*

MBD NikR was overexpressed and purified with nickel bound as described previously<sup>4</sup>. Nickel was removed from the sample by incubation with 10 mM EDTA (representing a ~10-fold excess over MBD) at 4 °C for 48 hours. The sample changed color from light yellow to colorless. To remove Ni-EDTA and excess EDTA, the sample was passed twice through Micro Bio-Spin P-6 columns (Bio-Rad).  $\text{Cu}^{2+}$  was added by



slow addition with vortex mixing of a 50 mM CuCl<sub>2</sub> stock solution until a stoichiometry of 1.2:1 CuCl<sub>2</sub>:MBD was reached. Upon addition of CuCl<sub>2</sub>, the sample became a dark yellow-orange color. Cu<sup>2+</sup>-MBD was crystallized in the same form and under the same conditions as Ni<sup>2+</sup>-MBD, the structure of which was published previously<sup>4</sup>. Briefly, 1.5 μL of ~10 mg/mL Cu<sup>2+</sup>-MBD in 300 mM NaCl, 20 mM Tris, pH 8.0 was mixed with 1.5 μL of 0.2 M disodium tartrate dehydrate, 20% w/v PEG 3350 in a hanging-drop vapor diffusion experiment. Crystals appeared within 2-4 days and were deep yellow orange in color (Fig. IV.1).

Crystals were cryoprotected using a solution that contained 20% ethylene glycol in the precipitant. A dataset was collected at APS beamline 8BM at 100 K to 1.5 Å resolution. Data was reduced using DENZO/SCALEPACK<sup>16</sup>. Statistics are given in Table IV.3.

Crystals of Ni<sup>2+</sup>-MBD and Cu<sup>2+</sup>-MBD were isomorphous, so the protein component of the Ni<sup>2+</sup>-MBD structure (1Q5Y)<sup>4</sup> was used directly as a starting model for refinement. Refinement was carried out in Refmac 5 from the CCP4 program suite<sup>17</sup>. Rigid body refinement of the MBD tetramer and individual subunits was followed by rounds of positional and B-factor refinement alternated with manual refitting of the model in Xfit<sup>18</sup>. Currently, R<sub>work</sub> = 18.2% and R<sub>free</sub> = 21.6% after refinement of anisotropic thermal displacement parameters and addition of hydrogens in their riding positions (Table IV.4). The final model contains residues 50-132 of NikR. One molecule of the tetramer is missing residues 64-66 from a loop.

All figures of protein structures and electron density maps were produced using PyMOL<sup>19</sup>.

#### *IV.D.2 Crystal structure of MBD NikR + Zn<sup>2+</sup>*

Apo MBD-NikR was produced as described above for the Cu<sup>2+</sup>-MBD NikR structure. Zn<sup>2+</sup> ions were added to Apo MBD-NikR as a stock of 20mM Zn-NTA complex to avoid MBD precipitation. The final Zn:MBD ratio was 3:1. Hexagonal-shaped crystals measuring ~100µm per side (Fig. IV.3) were obtained by mixing 1.5 µL Zn<sup>2+</sup>-MBD solution (~9mg MBD/mL in 300 mM NaCl, 20 mM Tris, pH 8.0) with 1.5 µL of precipitant (10% Jeffamine M-600, 0.1 M Na Citrate, pH 5.6, 0.01 M FeCl<sub>3</sub>) in a hanging-drop vapor diffusion experiment at room temperature.

A dataset was collected at 100 K on a single crystal after cryoprotection with 30% ethylene glycol in the precipitant condition using a Rigaku rotating Cu anode source. Data was reduced using DENZO/SCALEPACK<sup>16</sup>. Statistics are given in Table IV.3.

Since Matthews coefficient analysis suggested only one molecule of MBD-NikR per asymmetric unit (asu), the protein component of one MBD subunit from the Ni<sup>2+</sup>-MBD structure (1Q5Y) was used as a search model for molecular replacement using PHASER<sup>20</sup>. A good solution was found in P6<sub>2</sub>22 and was refined in CNS<sup>21</sup> by rigid body refinement followed by simulated annealing. Analysis of 2F<sub>o</sub>-F<sub>c</sub> and F<sub>o</sub>-F<sub>c</sub> maps calculated using the resulting model indicated that helix α3 of the MBD was mostly disordered, as in the apo-NikR structure (1Q5V)<sup>4</sup>. This helix was removed from the model and cycles of positional and B-factor refinement alternated with manual refitting in Xfit<sup>18</sup> were carried out. The current model includes one subunit of MBD NikR (residues 50-61 and 80-130), three water molecules, and a zinc ion (see Results and Discussion for details of the assignment as zinc ion), and has good geometry (Table IV.4).

#### *IV.D.3 Excess Ni<sup>2+</sup> soak of Ni-NikR crystals*

Crystals of Ni<sup>2+</sup>-NikR were grown as described in chapter 3 of this thesis. One such crystal was soaked in precipitant solution supplemented with 8 mM NiCl<sub>2</sub> at room temperature for 70 min. before cryo-cooling in a gaseous N<sub>2</sub> stream at 100 K. A dataset was collected on this crystal at the nickel peak wavelength (1.4845 Å) at beamline 5.0.2 of the ALS and processed anomalously using DENZO/SCALEPACK<sup>16</sup>. Statistics are given in Table IV.3.

The Ni<sup>2+</sup>-NikR model described in chapter III was refined against this dataset using CNS<sup>21</sup>. Rigid body refinement of the whole structure and individual domains was followed by simulated annealing, positional, and individual B-factor refinement. Phases from the refined model were used to calculate an anomalous difference Fourier map to locate the positions of any nickel ions in the structure. A 4  $\sigma$  contour of this map superimposed on a C <sub>$\alpha$</sub>  trace of the two crystallographically unique NikR subunits is shown in Fig. IV.7.

#### *IV.D.4 Excess Ni<sup>2+</sup> soak of NikR-DNA crystals*

Crystals of NikR-DNA were grown as described in chapter III. One such crystal was soaked in precipitant solution supplemented with 5 mM NiCl<sub>2</sub> at room temperature for 60 min. before cryoprotection with 25% ethylene glycol and cryo-cooling in a gaseous N<sub>2</sub> stream at 100 K. A dataset was collected on this crystal at the nickel peak wavelength (1.4862 Å) at beamline X29A of the NSLS and processed anomalously using DENZO/SCALEPACK<sup>16</sup>. Statistics are given in Table IV.3. Soaking of crystals in

solutions with NaCl substituted for KCl resulted in crystal cracking and loss of diffraction.

Although the dataset from the crystal soaked in excess NiCl<sub>2</sub> scaled well as C2, the unit cell dimensions differed significantly from non-soaked crystals of the NikR-DNA complex (a=199.6 Å, b=82.3 Å, c=125.8 Å, β=104.8° vs. a=197.2 Å, b=76.1 Å, c=132.1 Å, β=110.3°), and simple rigid body refinement using the previous model was unsuccessful. Molecular replacement was therefore carried out using PHASER<sup>20</sup>, with the NikR tetramer and duplex DNA as search models. Two copies of each were found in the space group C2, and were used for subsequent refinement in CNS<sup>21</sup>. Rigid body refinement of the whole structure and individual domains was followed by restrained simulated annealing, positional, and grouped B-factor refinement. Phases from the refined model were used to calculate an anomalous difference Fourier map to locate the positions of nickel ions in the structure. A 4σ contour of this map superimposed on a C<sub>α</sub> trace of one NikR-DNA complex is shown in Fig. IV.8.

#### IV.E. Acknowledgements

I would like to thank Yayi Guo, an undergraduate biology major at MIT, who helped considerably with the purification and crystallization of MBD+Zn<sup>2+</sup>.

#### IV.F. References

1. Chivers, P.T. & Sauer, R.T. NikR is a ribbon-helix-helix DNA-binding protein. *Protein Sci* **8**, 2494-500 (1999).
2. Chivers, P.T. & Sauer, R.T. Regulation of high affinity nickel uptake in bacteria. Ni<sup>2+</sup>-Dependent interaction of NikR with wild-type and mutant operator sites. *Journal of Biological Chemistry* **275**, 19735-41 (2000).

3. Chivers, P.T. & Sauer, R.T. NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA. *Chemistry and Biology* **9**, 1141-8 (2002).
4. Schreiter, E.R. et al. Crystal structure of the nickel-responsive transcription factor NikR. *Nat Struct Biol* **10**, 794-9 (2003).
5. Bloom, S.L. & Zamble, D.B. Metal-selective DNA-binding response of Escherichia coli NikR. *Biochemistry* **43**, 10029-38 (2004).
6. Wang, S.C., Dias, A.V., Bloom, S.L. & Zamble, D.B. Selectivity of metal binding and metal-induced stability of Escherichia coli NikR. *Biochemistry* **43**, 10018-28 (2004).
7. Irving, H. & Williams, R.J.P. Order of Stability of Metal Complexes. *Nature* **162**, 746-47 (1948).
8. Rulisek, L. & Vondrasek, J. Coordination geometries of selected transition metal ions ( $\text{Co}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Cd}^{2+}$ , and  $\text{Hg}^{2+}$ ) in metalloproteins. *J Inorg Biochem* **71**, 115-27 (1998).
9. Changela, A. et al. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* **301**, 1383-7 (2003).
10. Gray, H.B., Malmstrom, B.G. & Williams, R.J. Copper coordination in blue proteins. *J Biol Inorg Chem* **5**, 551-9 (2000).
11. Carrington, P.E., Chivers, P.T., Al-Mjeni, F., Sauer, R.T. & Maroney, M.J. Nickel coordination is regulated by the DNA-bound state of NikR. *Nature Structural Biology* **10**, 126-130 (2003).
12. Harding, M.M. Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium. *Acta Crystallogr D Biol Crystallogr* **58**, 872-4 (2002).
13. Outten, C.E. & O'Halloran, T.V. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488-92 (2001).
14. Epstein, W. & Schultz, S.G. Cation transport in Escherichia coli. VI. K exchange. *J Gen Physiol* **49**, 469-81 (1966).
15. Phillips, K. & Phillips, S.E. Electrostatic activation of Escherichia coli methionine repressor. *Structure* **2**, 309-16 (1994).
16. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326 (1997).

17. Collaborative Computational Project, N. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallographica. Section D: Biological Crystallography* **50**, 760-763 (1994).
18. McRee, D.E. XtalView/Xfit - A Versatile Program for Manipulating Atomic Coordinates and Electron Density. *Journal of Structural Biology* **125**, 156-165 (1999).
19. DeLano, W.L. The PyMOL Molecular Graphics System. (DeLano Scientific, San Carlos, 2002).
20. Storoni, L.C., McCoy, A.J. & Read, R.J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* **60**, 432-8 (2004).
21. Brünger, A.T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica. Section D: Biological Crystallography* **54**, 905-21 (1998).

#### IV.G. Tables and Figures

**Table IV.1.** Comparison of metal binding geometry for the Ni<sup>2+</sup>-MBD and Cu<sup>2+</sup>-MBD structures. Each value is an average of measurements from the four crystallographically independent molecules of MBD-NikR in the ASU. The average estimated standard uncertainty (ESU) in the atomic coordinates for each structure is 0.05 - 0.10 Å, depending upon the method of estimation. See Fig. IV.2 for pictures of the binding sites.

<b>Bond Lengths (Å)</b>	<b>Ni<sup>2+</sup></b>	<b>Cu<sup>2+</sup></b>
<b>H76'(Nε)</b>	1.85	2.00
<b>H87(Nε)</b>	1.94	2.09
<b>H89(Nδ)</b>	2.01	2.44
<b>C95(Sγ)</b>	2.22	2.43
<b>Angles (°)</b>	<b>Ni<sup>2+</sup></b>	<b>Cu<sup>2+</sup></b>
<b>C95 - M<sup>2+</sup> - H87</b>	93.0	99.2
<b>H87 - M<sup>2+</sup> - H89</b>	91.0	83.7
<b>H89 - M<sup>2+</sup> - H76'</b>	87.6	84.4
<b>H76' - M<sup>2+</sup> - C95</b>	88.3	94.1

**Table IV.2.** Ni<sup>2+</sup> binding sites on NikR identified in soaks with excess NiCl<sub>2</sub>. Nickel coordination is by amino acid sidechains unless otherwise noted. A prime denotes an amino acid from another subunit of the NikR tetramer. Sites are displayed relative to the Ni-NikR and NikR-DNA structures in Figs. IV.7 and IV.8. Sites 1-4 are observed in both structures. Site 5 is only in the Ni-NikR soak, and sites 6 and 7 are only observed in the NikR-DNA complex.

Site #	Site description
0	Previously characterized high-affinity metal-binding site in NikR (His87, His89, Cys95, His76')
1	Met1 (free N-terminal amine and backbone carbonyl)
2	His125, His125'
3	His79, His92'
4	His78, His110, D114
5	His123
6	His48
7	Asp104, Asp107



**Table IV.3.** Data collection statistics.

<b>Crystal</b>	<b>Cu<sup>2+</sup>-MBD</b>	<b>MBD+ Zn<sup>2+</sup></b>	<b>Ni-NiKR excess Ni<sup>2+</sup> soak<sup>1</sup></b>	<b>NiKR-DNA excess Ni<sup>2+</sup> soak<sup>1</sup></b>
<b>Space Group</b>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P6 <sub>3</sub> 22	P3 <sub>1</sub> 21	C2
<b>Cell Dimensions</b>				
<b>a (Å)</b>	45.95	46.33	50.73	199.63
<b>b (Å)</b>	78.44	-	-	82.31
<b>c (Å)</b>	81.46	125.27	183.07	125.74
<b>β (°)</b>	-	-	-	104.75
<b>Wavelength (Å)</b>	0.9791	1.5418	1.4845	1.4862
<b>Temperature (K)</b>	100	100	100	100
<b>Resolution Range (Å)</b>	50-1.5	20-2.5	50-2.7	50-3.8
<b>Average Redundancy</b>	8.1	6.4	9.3	3.4
<b>Completeness (%)<sup>2</sup></b>	97.9 (85.2)	98.0 (98.3)	98.3 (89.3)	94.1 (63.4)
<b><i>I</i> / σ(<i>I</i>)<sup>2</sup></b>	16.9 (5.3)	12.9 (3.5)	16.3 (4.5)	10.5 (3.0)
<b><i>R</i><sub>sym</sub> (%)<sup>2,3</sup></b>	5.7 (25.4)	7.5 (46.9)	8.9 (28.8)	7.9 (30.4)

<sup>1</sup>Datasets were scaled anomalously, keeping Friedel pairs separate.

<sup>2</sup>The number in parentheses is for the highest resolution shell.

<sup>3</sup> $R_{\text{sym}} = \sum_{hkl} |I_{(hkl)}^i - \langle I_{(hkl)} \rangle| / \sum_{hkl} \langle I_{(hkl)} \rangle$ , where  $I_{(hkl)}^i$  is the  $i^{\text{th}}$  measured diffraction intensity and  $\langle I_{(hkl)} \rangle$  is the mean of the intensity for the miller index ( $hkl$ ).

**Table IV.4.** Model refinement statistics.

Crystal	Cu <sup>2+</sup> -MBD	MBD+ Zn <sup>2+</sup>	Ni-NikR excess Ni <sup>2+</sup> soak <sup>2</sup>	NikR-DNA excess Ni <sup>2+</sup> soak <sup>2</sup>
<b>R<sub>cryst</sub> (R<sub>free</sub>) (%)<sup>1</sup></b>	18.2 (21.6)	27.0 (32.6)	27.2 (30.4)	43.7 (46.9)
<b>Resolution Range (Å)</b>	15-1.5	20-2.5	50-2.6	50-3.8
<b>Number of protein atoms</b>	2554	476	1959	-
<b>Number of metal ions</b>	4	1	10	-
<b>Number of water molecules</b>	243	3	0	-
<b>R.m.s. deviations</b>				
<b>Bond lengths (Å)</b>	0.017	0.007	0.009	-
<b>Bond angles (°)</b>	1.62	1.29	1.33	-
<b>Ramachandran (%)</b>				
<b>Most favored</b>	91.8	88.9	84.7	-
<b>Additionally Allowed</b>	8.2	11.1	14.0	-
<b>Generously Allowed</b>	0.0	0.0	0.8	-
<b>Disallowed</b>	0.0	0.0	0.4	-
<b>Average B-factor (Å<sup>2</sup>)</b>	22.7	47.0	73.7	-

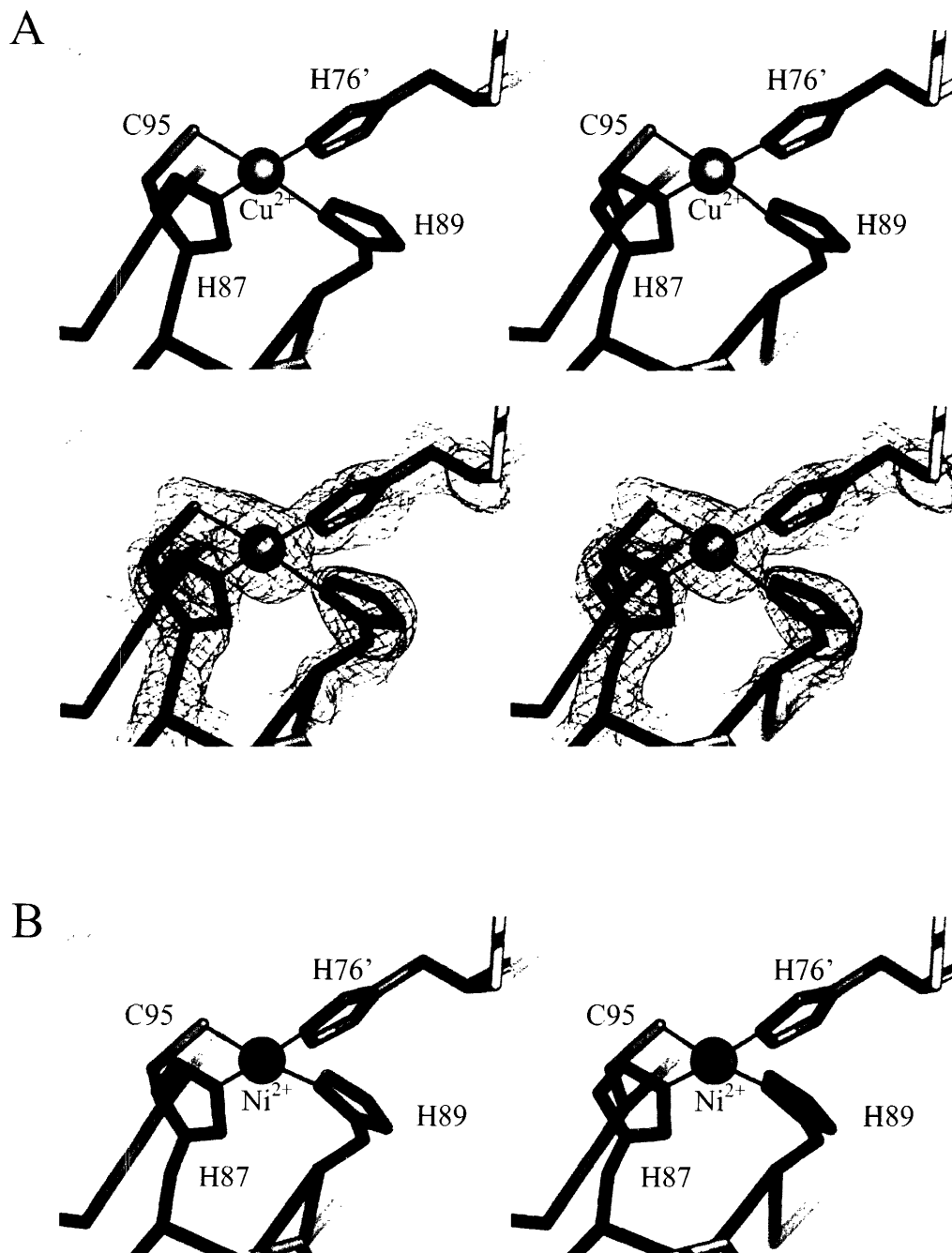
<sup>1</sup> $R_{\text{cryst}} = \frac{\sum_{hkl} \left| |F_o(hkl)| - |F_c(hkl)| \right|}{\sum_{hkl} |F_o(hkl)|}$ .  $R_{\text{free}} = R_{\text{cryst}}$  for a test set of reflections (5%) not included in refinement.

<sup>2</sup>Protein models for the Ni-NikR and NikR-DNA excess Ni<sup>2+</sup> soaks are not fully refined. Each was used after partial refinement to calculate anomalous difference fourier maps. Only rigid-body refinement of the NikR-DNA excess Ni<sup>2+</sup> soak was done, so no model statistics are given.

**Figure IV.1.** Crystals of  $\text{Cu}^{2+}$ -MBD NikR. Pictures taken using a microscope at  $\sim 50\times$  magnification.



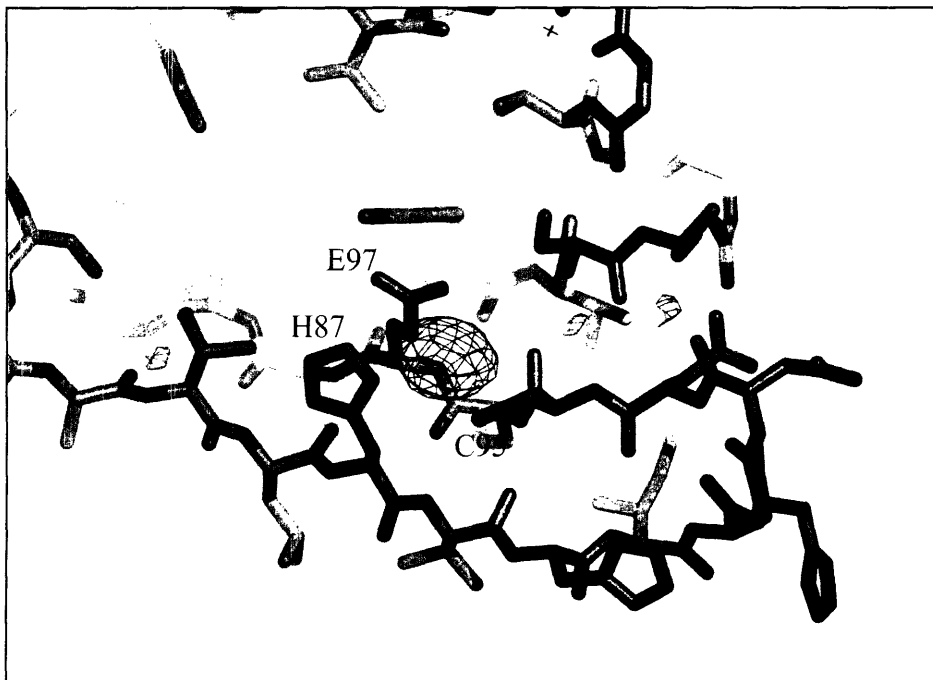
**Figure IV.2.** Stereo view comparison of the high-affinity metal binding site of MBD NikR with  $\text{Cu}^{2+}$  (part A) and  $\text{Ni}^{2+}$  (part B) ions bound. See Table IV.1 for bond length and angle measurements. Protein is represented as a  $\text{C}_\alpha$  trace colored by molecule, with sticks colored by atom type for the metal ligands. Metal ions are represented as colored spheres. A simulated-annealing composite omit map for the  $\text{Cu}^{2+}$ -MBD structure, contoured at  $1\sigma$ , is shown in blue mesh superimposed on the model in the second frame of part A.



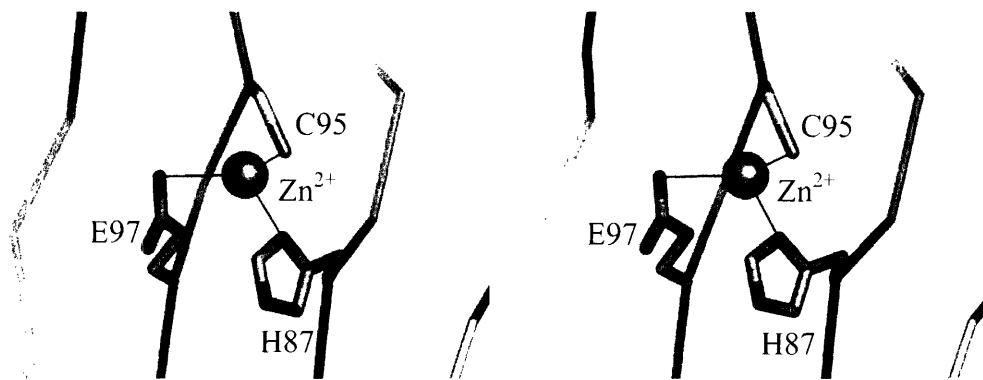
**Figure IV.3.** Crystals of MBD+Zn<sup>2+</sup>. Pictures taken using a microscope at ~50x magnification.



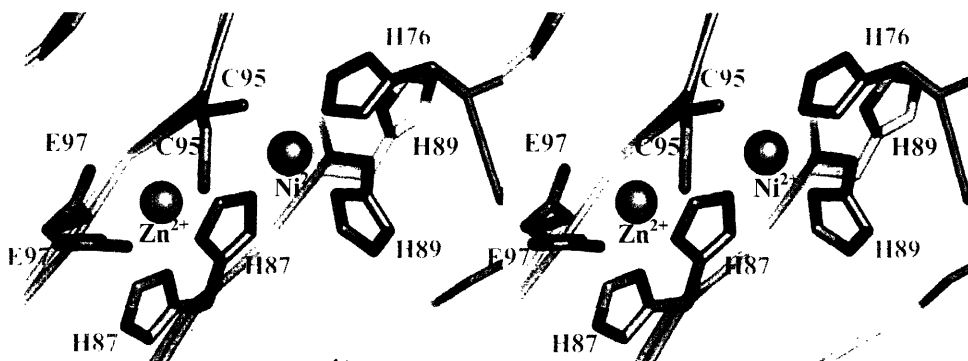
**Figure IV.4.** Positive difference density peak in the MBD+Zn<sup>2+</sup> structure. Protein is represented as sticks, colored by atom type. Helix  $\alpha$ 3, which would normally occlude the view of this site is disordered and absent from the model. A  $F_o-F_c$  electron density map contoured at  $+3\sigma$  is shown in red mesh.



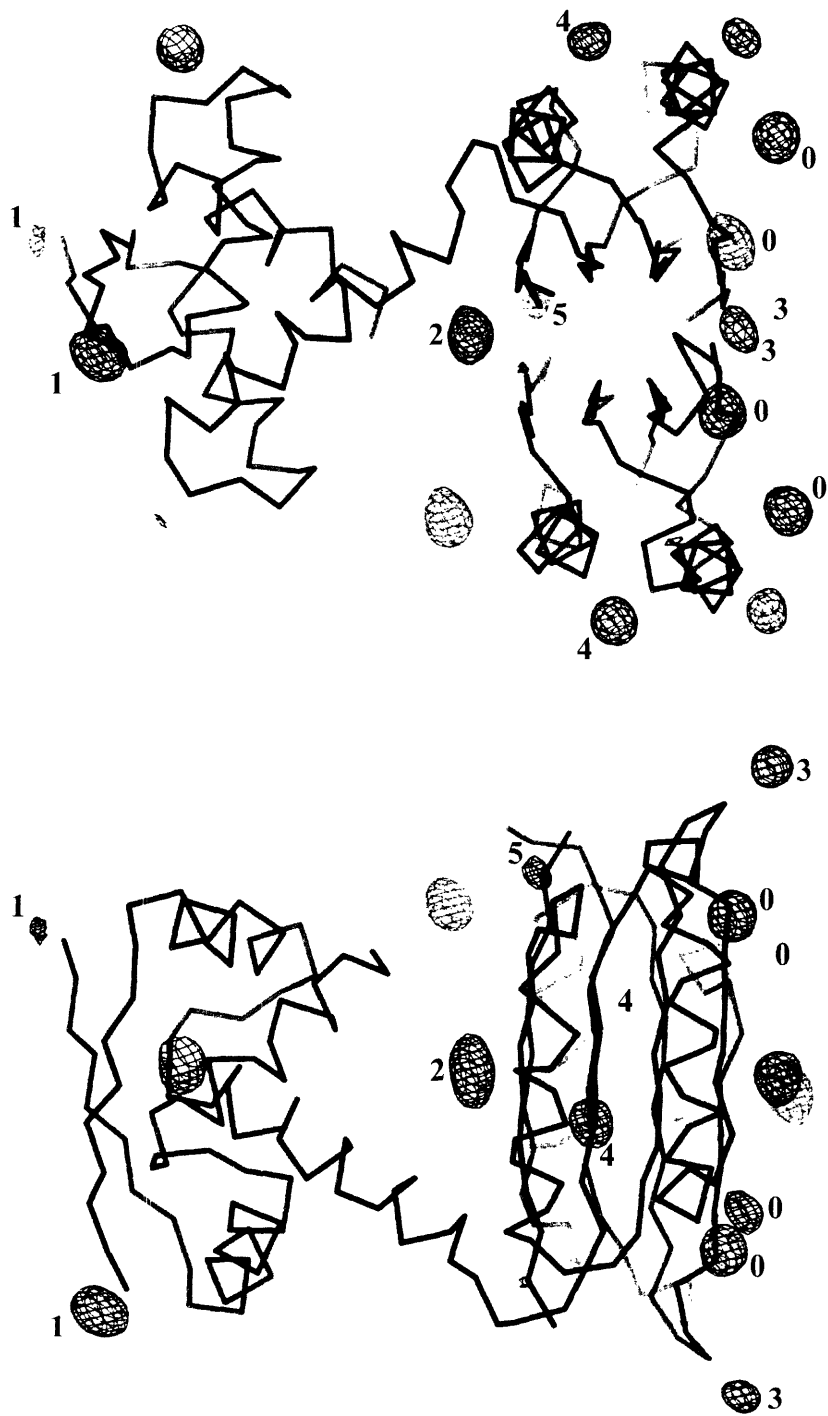
**Figure IV.5.** Stereo view of the putative  $Zn^{2+}$  binding site. Protein is represented as a  $C_{\alpha}$  trace colored by molecule, with sticks colored by atom type for the metal ligands. Zinc ion is represented as a grey sphere. The zinc ion is slightly out of the plan of the three protein sidechains.



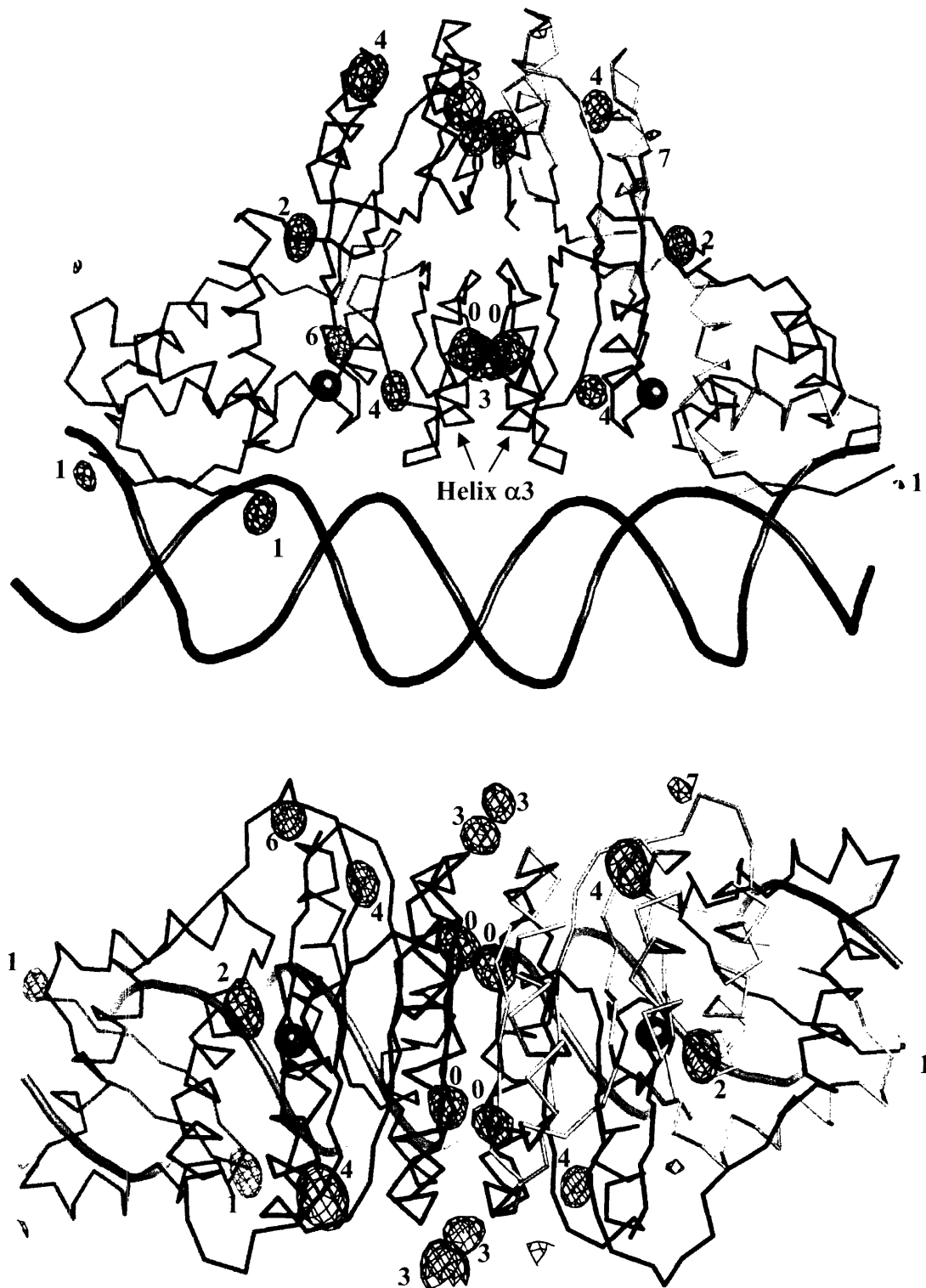
**Figure IV.6.** Stereo view superimposition of  $Ni^{2+}$ -MBD and  $MBD+Zn^{2+}$  and comparison of  $Ni^{2+}/Cu^{2+}$  and  $Zn^{2+}$  binding sites. One subunit each of the  $Ni^{2+}$ -MBD (pink) and  $MBD+Zn^{2+}$  (green) structures were superimposed. Protein is displayed as a  $C_{\alpha}$  trace, metal ligand sidechains as sticks. Metal ions are shown as spheres.



**Figure IV.7.** Nickel binding sites from Ni-NikR crystals (without DNA) soaked in 8 mM NiCl<sub>2</sub>. The two NikR molecules in the asymmetric unit (ASU) of this crystal form are shown as a C<sub>α</sub> trace colored by subunit. An anomalous difference Fourier map calculated using model phases and contoured at 4σ is shown in blue mesh. Numbered sites are described in Table IV.2. Sites without numbers are related to a numbered site by crystallographic symmetry. The two views are related by a 90° rotation.



**Figure IV.8.** Nickel binding sites from NikR-DNA crystals soaked in 5mM NiCl<sub>2</sub>. Representation is the same as Fig. IV.7, except that DNA is represented as an orange cartoon tube passing through the phosphate positions. Bound potassium ions are shown as pink spheres. Numbered sites are described in Table IV.2.





## Chapter V: NikR and the Ribbon-Helix-Helix Family

### V.A. Summary

The amino-terminal DNA-binding domain of NikR belongs to the Ribbon-Helix-Helix (RHH) superfamily of transcription factors. Visual inspection of the structures of RHH family members currently available in the PDB permitted alignment of sequences proposed to belong to this family. Despite low sequence identity among members of this family, several motifs are apparent in the sequence alignment, which primarily contribute to the correct folding of this small, dimeric domain. The alignment also allowed identification of a 36 residue core (72 residues for both molecules of a dimer) that could be used to structurally superimpose the family members. Analysis of root mean square deviation (RMSD) values for pairwise structural alignments of the family members reveals that the DNA-binding domain of NikR is structurally most similar to the Arc repressor from bacteriophage P22 and the ParG repressor from *Salmonella newport*, and least similar to the omega repressor from *Streptococcus pyogenes*. Comparison of the available RHH-DNA complex structures shows that the family shares a generally similar mode of DNA binding, but the details of the specific interactions vary among the members.

## V.B. Introduction

When the crystal structure of the methionine repressor protein, MetJ, was determined in 1989, a model for how MetJ would bind its operator DNA was proposed where  $\alpha$ -helices at the C-terminus of the protein inserted into the DNA major groove to make sequence-specific interactions<sup>1</sup>. The model was based on the structures of transcription factor-DNA complexes available at the time, including zinc fingers, homeodomains, and the bacteriophage repressor and Cro proteins, which showed DNA major groove binding by  $\alpha$ -helices. When the crystal structure of MetJ in complex with operator DNA was reported three years later, it was exciting to see that the antiparallel  $\beta$ -sheet at the N-terminus of the protein, not an  $\alpha$ -helix, was situated in the DNA major groove, making specific nucleotide base contacts<sup>2</sup>. A model for this mode of DNA binding had been proposed two years earlier for the Arc and Mnt repressors from bacteriophage P22 based on an NMR structure of Arc and biochemical data showing where the DNA binding specificity determinants of both proteins was located<sup>3-5</sup>.

According to the Structural Classification of Proteins (SCOP) database<sup>6</sup>, the Ribbon-Helix-Helix (RHH) superfamily is currently represented by seven members whose three-dimensional structures have been determined: NikR from *Escherichia coli*<sup>7</sup>, a nickel-dependent repressor of nickel uptake genes; MetJ from *Escherichia coli*<sup>1,2</sup>, which regulates transcription of the methionine biosynthetic genes in a S-adenosylmethionine (AdoMet)-dependent manner; the Arc<sup>3,8</sup> and Mnt<sup>9</sup> repressors from bacteriophage P22, responsible for regulation of the phage lytic cycle; the streptococcal repressor proteins CopG<sup>10,11</sup> and  $\omega$ <sup>12</sup>, which control plasmid copy number; and ParG from *Salmonella newport*<sup>13</sup>, involved in partition of plasmids during cell division. Structures of each are

depicted in Fig. V.1, shown in the DNA-bound form when available. Recently, a number of other proteins have been proposed to be members of the RHH family on the basis of sequence alignments, nuclear magnetic resonance (NMR) secondary structure assignment, circular dichroism (CD), analytical ultracentrifugation (AUC), gel filtration chromatography, and/or mutagenesis. These include the TraY<sup>14,15</sup> and TrwA<sup>16</sup> proteins from *E. coli*, which regulate bacterial conjugation; AlgZ<sup>17</sup>, responsible for control of alginate biosynthesis in *Pseudomonas aeruginosa*; ParD<sup>18</sup>, an anti-toxin that also controls transcription of a toxin/anti-toxin pair; PutA<sup>19</sup>, a large, multifunctional enzyme from *E. coli* involved in proline catabolism that also regulates transcription of proline utilization genes; and a DNA-binding protein from *Methanopyrus kandleri* of unknown function, nicknamed 7kMk<sup>20</sup>.

The structural aspects of the RHH fold and DNA-binding properties were reviewed a decade ago<sup>21</sup>, when only structures of the Arc repressor and MetJ were available. A wealth of structural information about these and several other family members has become available since then, allowing a more detailed analysis than was previously possible. The RHH domain is dimeric (Figs. V.1, V.2), comprising tightly intertwined identical monomers, except in the case of TraY, which contains two repeats of the RHH sequence (Fig. V.3) and folds as a monomer<sup>14,15</sup>. Short  $\beta$ -strands at the N-termini of each monomer pair to form an antiparallel  $\beta$ -sheet. Binding of the RHH domain to DNA places this  $\beta$ -sheet in the DNA major groove, where two or three amino acid sidechains from each strand point into the groove and make the critical sequence-specific nucleotide base contacts (Fig. V.2.a). A conserved set of non-specific DNA backbone phosphate contacts are made by the two protein backbone amide nitrogens at

the N-terminus of the second  $\alpha$ -helix (helix  $\alpha_2$ ) of the domain on either side of the major groove (Fig. V.2.b). This interaction is electrostatically favorable, since the positive dipole at the helix N-terminus is oriented directly towards a negatively charged phosphate of the DNA backbone. These contacts anchor the domain to DNA and properly orient the base-contacting  $\beta$ -sheet. Other non-conserved DNA phosphate contacts are made by sidechains and backbone of flexible N-terminal extensions (if present), or by sidechains of helix  $\alpha_1$  or helix  $\alpha_2$ . All RHH proteins whose DNA-binding properties have been studied bind as higher-order oligomers to multiple sub-sites within their operators, arranged as inverted or tandem repeats.

## **V.C. Results and Discussion**

### *V.C.1. Sequence alignment of Ribbon-Helix-Helix proteins*

Visual inspection of the available RHH structures was used to create a multiple sequence alignment, to which the sequences of predicted RHH proteins were aligned using secondary structure prediction and profile matching (Fig. V.3). Although the pairwise identity of the aligned sequences is quite low (average pairwise identity of 15.3% for sequences in Fig. V.3), a number of motifs emerge upon inspection of the alignment. First, a clear pattern of hydrophilic-hydrophobic repeats is present along the N-terminal  $\beta$ -strand between positions 2-7, with the hydrophobic sidechains (3,5,7) pointing toward the domain core and the hydrophilic positions (2,4,6) pointing away from the domain to make nucleotide base contacts upon DNA binding. In addition, Arg or Lys is present in position 2 or 6 in each sequence. The  $\omega$  repressor has a lysine in position 3, which alters the conformation of the  $\beta$ -sheet in this protein. Since no structures of the

operator complex or mutagenesis data are available, it remains to be seen how this altered conformation affects the DNA-binding mode of  $\omega$ . Several of the predicted RHH sequences contain a hydrophobic sidechain at position 2, 4, or 6. As NikR, CopG, and MetJ only use two of these three positions to make specific base contacts, it is possible that these substitutions may not impair DNA binding. A hydrophobic amino acid may provide some specificity by allowing van der Waals interactions with the methyl group of thymine bases, for example. A second feature of the sequence alignment is a consensus G-X-S/T/N in the turn between helix  $\alpha 1$  and helix  $\alpha 2$ . A clear exception is MetJ, which has six additional amino acids between the helices, has an extra turn of helix at the C-terminus of helix  $\alpha 1$ , and necessarily adopts a different conformation of this loop. NikR and several predicted RHH sequences are also one amino acid longer in this region. The length and conformation of this turn/loop do not, therefore, appear to be defining characteristics of the RHH domain. Finally, four positions along helix  $\alpha 1$  and helix  $\alpha 2$  (positions 15, 27, 31, and 35) are conserved as hydrophobic, primarily branched-chain amino acids. Taken together with positions 3, 5, and 7 from the N-terminal  $\beta$ -strand mentioned earlier, these sidechains comprise the majority of the hydrophobic core of the RHH domain.

#### *V.C.2. Structural alignment of Ribbon-Helix-Helix proteins*

In order to superimpose and structurally compare the available RHH structures, a core of 36 amino acid positions was chosen, which includes only the N-terminal  $\beta$ -strand, helix  $\alpha 1$ , and helix  $\alpha 2$  (Fig. V.3). Portions of the proteins prior to the  $\beta$ -strand, after helix  $\alpha 2$ , or in the loop connecting helix  $\alpha 1$  and helix  $\alpha 2$  were excluded from the

alignments. Seventy-two  $C_{\alpha}$  positions (36 from each molecule of the dimer) were used to make a best-fit superposition of the RHH structures (Fig. V.4) and to calculate root mean squared deviation (RMSD) values for each pairwise comparison (Table V.1). Visual inspection of both the superposition (Fig. V.4) and RMSD values (Table V.1) leads to the conclusion that, despite low sequence similarity among the RHH family members, they share a very similar fold. NikR is structurally most similar to ParG, followed by Arc and CopG. If the  $\omega$  repressor is removed from Table V.1, nearly all of the remaining pairwise RMSD values are below 2Å, indicating that the structures are highly similar and that perhaps  $\omega$  is an outlier. The  $\beta$ -strand and helix  $\alpha 1$  of the  $\omega$  repressor adopt different orientations relative to helix  $\alpha 2$  than the other RHH proteins, possibly a consequence of substitution of Lys for a hydrophobic amino acid at position 3, which does not allow insertion of this sidechain into the hydrophobic core of the domain. The register of the hydrogen bond pairing between strands of the  $\beta$ -sheet is therefore shifted relative to the other RHH family members, which all show a juxtaposition of the sidechains at position 4 and 4' in the  $\beta$ -sheet. A structure of the complex between  $\omega$  and its operator DNA sequence would confirm that it shares the same binding mode with other RHH family members. In addition, mutagenesis experiments could be carried out to determine which amino acid positions are responsible for sequence-specific DNA contacts in this repressor.

### *V.C.3. DNA binding by Ribbon-Helix-Helix proteins*

Sequence-specific DNA-binding proteins such as transcription factors use both direct readout, involving intermolecular contacts between the protein and nucleotide

bases, and indirect readout, where the conformation of a particular DNA segment is recognized by the protein, to bind DNA in a sequence-specific manner<sup>22</sup>. As mentioned previously, RHH proteins bind as multiple dimers to operators containing repeats of a sub-site, making protein-protein contacts between dimers when bound. Each of the available RHH-DNA co-crystal structures available shows some bending of the DNA operator. Structures of two dimers each of Arc, CopG, or MetJ bound to their respective operator sequences all display an overall bending of the DNA by 50-60° over its length. The molecular details of the cause of this bend vary, however. In the MetJ-DNA complex structure (Fig. V.1.b), the largest local bend is at the center of each sub-site, where the specific contacts from MetJ are located. For Arc and CopG, the majority of the bend is localized in the center of the operator, between the two dimer-bound sub-sites (Fig. V.1.c,d). This difference is not surprising, given that the quaternary interactions between bound RHH dimers are similar for Arc and CopG, and are significantly different for MetJ (Fig. V.1.b-d). NikR represents a unique case since there is no direct contact between the two RHH dimers in the co-crystal structure (Fig. V.1.a), although it also binds its operator DNA sequence as a tetramer. The large, tetrameric, C-terminal metal-binding domain (MBD) of NikR separates the two RHH domains in space, allowing NikR to bind an operator sequence with sub-sites separated by two full turns of DNA. The DNA in the NikR-operator complex curves slightly and uniformly over its length, bending by a total of only ~22°. The curvature of the NikR operator could be due to its interaction with NikR or because of the A-tract sequence (TTAAAAT) comprising the -10 region of the *nik* promoter between the operator sub-sites<sup>23</sup>. The major groove widths around the  $\beta$ -strands of the bound RHH proteins vary as well, although a general

correlation seems to exist between the distance separating the N-termini of helices  $\alpha 2$  of each unbound RHH dimer and the DNA major groove width around the protein when bound (Table V.2 and Fig. V.5). The local DNA distortion in the complex structures could be a result of conforming to a rigid protein framework, or may be a consequence of indirect readout of a sequence-specifically distorted local DNA conformation. This analysis is complicated further by the interesting observation that the phosphates contacted by the N-termini of helices  $\alpha 2$  on either side of the major groove are separated by five base-pairs in the NikR-DNA complex and six base-pairs in the Arc, CopG, and MetJ co-crystal structures. The difference in the binding mode of NikR may be due to the limited range of orientations available to the RHH domains of NikR, since they are each tethered to the large, central metal-binding domain.

The details of the specific DNA base contacts do not appear to be conserved across the RHH superfamily. Each characterized family member binds a unique operator sequence (or sequences). RHH proteins use 2-3 amino acid sidechains from each N-terminal  $\beta$ -strand (positions 2, 4, and 6) to make direct nucleotide base contacts, and the identities of these three sidechains vary across the family (Fig. V.3). Although MetJ binds a symmetric operator sequence by making contacts that follow the two-fold symmetry of the dimer, other members recognize asymmetric sub-site sequences and correspondingly make asymmetric contacts (Fig. V.6). In general, the nucleotide base interactions made by the RHH family are common relative to the known protein-DNA structures<sup>24</sup>, with Lys or Arg at position 2 or 6 of the  $\beta$ -strands making the most specific contacts. Perhaps the most interesting comparison with regards to operator sequence specificity is NikR and CopG. Both of these proteins have Arg-X-Thr-X-Thr at positions



2-6 of their  $\beta$ -strands, but naturally recognize unique operator sub-site sequences (Fig. V.6). Indeed, their structures in complex with operator DNA show different conformations of the base-contacting sidechains. It would be interesting to learn whether substitution of the CopG sub-site sequence into the NikR operator (with the correct spacing) could allow tight binding, or whether other factors such as the large, C-terminal metal-binding domain of NikR play an important role in specificity. The recently identified 7kMk protein also shares this Arg-X-Thr-X-Thr motif along its putative  $\beta$ -strand, but no operator sequence has yet been identified for this protein. Perhaps learning more about sequence specificity in the NikR and CopG systems could allow prediction of an operator sequence for 7kMk.

RHH proteins constitute a functionally diverse structural superfamily that appears to be steadily growing in size. Although the RHH family members use a common set of themes for interacting with DNA, they recognize different operator sequences, make unique protein-protein interactions to stabilize their complexes with DNA, and can be ligand regulated or simply concentration dependent in their ability to bind operator. It is our hope that this detailed structural analysis will aid the identification of new members of this expanding family.

## **V.D. Materials and Methods**

### *V.D.1. Sequence alignment of Ribbon-Helix-Helix proteins*

Sequence alignment of structurally characterized RHH proteins was accomplished by visual inspection of the structures and comparison of structural markers such as

secondary structure elements. Alignment of additional sequences began with limited sequence alignments reported previously<sup>14-20</sup>, followed by manual adjustments.

#### *V.D.2. Structural alignments of Ribbon-Helix-Helix proteins*

Structural alignments and calculation of RMSD values were carried out in LSQKAB, part of the CCP4 program suite<sup>25</sup>. All structural figures were prepared using PyMOL<sup>26</sup>.

#### **V.E. References**

1. Rafferty, J.B., Somers, W.S., Saint-Girons, I. & Phillips, S.E. Three-dimensional crystal structures of Escherichia coli met repressor with and without corepressor. *Nature* **341**, 705-10 (1989).
2. Somers, W.S. & Phillips, S.E. Crystal structure of the met repressor-operator complex at 2.8 Å resolution reveals DNA recognition by beta-strands. *Nature* **359**, 387-93 (1992).
3. Breg, J.N., van Opheusden, J.H., Burgering, M.J., Boelens, R. & Kaptein, R. Structure of Arc repressor in solution: evidence for a family of beta-sheet DNA-binding proteins. *Nature* **346**, 586-9 (1990).
4. Knight, K.L., Bowie, J.U., Vershon, A.K., Kelley, R.D. & Sauer, R.T. The Arc and Mnt repressors. A new class of sequence-specific DNA-binding protein. *Journal of Biological Chemistry* **264**, 3639-42 (1989).
5. Knight, K.L. & Sauer, R.T. DNA binding specificity of the Arc and Mnt repressors is determined by a short region of N-terminal residues. *Proc Natl Acad Sci U S A* **86**, 797-801 (1989).
6. Murzin, A.G., Brenner, S.E., Hubbard, T. & Chothia, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**, 536-40 (1995).
7. Schreiter, E.R. et al. Crystal structure of the nickel-responsive transcription factor NikR. *Nat Struct Biol* **10**, 794-9 (2003).

8. Raumann, B.E., Rould, M.A., Pabo, C.O. & Sauer, R.T. DNA recognition by beta-sheets in the Arc repressor-operator crystal structure. *Nature* **367**, 754-7 (1994).
9. Burgering, M.J. et al. Solution structure of dimeric Mnt repressor (1-76). *Biochemistry* **33**, 15036-45 (1994).
10. Costa, M. et al. Plasmid transcriptional repressor CopG oligomerises to render helical superstructures unbound and in complexes with oligonucleotides. *Journal of Molecular Biology* **310**, 403-17 (2001).
11. Gomis-Ruth, F.X. et al. The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *EMBO Journal* **17**, 7404-15 (1998).
12. Murayama, K., Orth, P., de la Hoz, A.B., Alonso, J.C. & Saenger, W. Crystal structure of omega transcriptional repressor encoded by *Streptococcus pyogenes* plasmid pSM19035 at 1.5 Å resolution. *Journal of Molecular Biology* **314**, 789-96 (2001).
13. Golovanov, A.P., Barilla, D., Golovanova, M., Hayes, F. & Lian, L.Y. ParG, a protein required for active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. *Mol Microbiol* **50**, 1141-53 (2003).
14. Bowie, J.U. & Sauer, R.T. TraY proteins of F and related episomes are members of the Arc and Mnt repressor family. *Journal of Molecular Biology* **211**, 5-6 (1990).
15. Lum, P.L. & Schildbach, J.F. Specific DNA recognition by F Factor TraY involves beta-sheet residues. *J Biol Chem* **274**, 19644-8 (1999).
16. Moncalian, G. & de la Cruz, F. DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochim Biophys Acta* **1701**, 15-23 (2004).
17. Baynham, P.J., Brown, A.L., Hall, L.L. & Wozniak, D.J. *Pseudomonas aeruginosa* AlgZ, a ribbon-helix-helix DNA-binding protein, is essential for alginate synthesis and algD transcriptional activation. *Mol Microbiol* **33**, 1069-80 (1999).
18. Oberer, M., Zangger, K., Prytulla, S. & Keller, W. The anti-toxin ParD of plasmid RK2 consists of two structurally distinct moieties and belongs to the ribbon-helix-helix family of DNA-binding proteins. *Biochem J* **361**, 41-7 (2002).
19. Gu, D. et al. Identification and characterization of the DNA-binding domain of the multifunctional PutA flavoenzyme. *J Biol Chem* **279**, 31171-6 (2004).

20. Pavlov, N.A., Cherny, D.I., Nazimov, I.V., Slesarev, A.I. & Subramaniam, V. Identification, cloning and characterization of a new DNA-binding protein from the hyperthermophilic methanogen *Methanopyrus kandleri*. *Nucleic Acids Res* **30**, 685-94 (2002).
21. Raumann, B.E., Brown, B.M. & Sauer, R.T. Major groove DNA recognition by  $\beta$ -sheets: the ribbon-helix-helix family of gene regulatory proteins. *Current Opinion in Structural Biology* **4**, 36-43 (1994).
22. Michael Gromiha, M., Siebers, J.G., Selvaraj, S., Kono, H. & Sarai, A. Intermolecular and intramolecular readout mechanisms in protein-DNA recognition. *J Mol Biol* **337**, 285-94 (2004).
23. Allemann, R.K. & Egli, M. DNA recognition and bending. *Chem Biol* **4**, 643-50 (1997).
24. Luscombe, N.M., Laskowski, R.A. & Thornton, J.M. Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res* **29**, 2860-74 (2001).
25. Collaborative Computational Project, N. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallographica. Section D: Biological Crystallography* **50**, 760-763 (1994).
26. DeLano, W.L. The PyMOL Molecular Graphics System. (DeLano Scientific, San Carlos, 2002).

## V.F. Tables and Figures

**Table V.1.** Pairwise RMSD values, in Å, for structural alignment of the RHH structures currently available in the PDB. 72 common  $C_{\alpha}$  atoms of each dimeric domain were used for the superposition. See Fig. V.3 for the  $C_{\alpha}$  positions used.

	Apo-NikR	Ni-NikR	NikR-DNA	Arc	Arc-DNA	Apo-MetJ	SAM-MetJ	MetJ-DNA	CopG	CopG-DNA	Mnt	$\omega$	ParG
Apo-NikR													
Ni-NikR	1.04												
NikR-DNA	1.29	1.04											
Arc	1.62	1.43	1.05										
Arc-DNA	1.54	1.37	1.09	0.71									
Apo-MetJ	1.96	1.74	1.44	1.58	1.46								
SAM-MetJ	2.07	1.79	1.47	1.46	1.40	0.40							
MetJ-DNA	2.07	1.83	1.45	1.42	1.36	0.57	0.41						
CopG	1.88	1.83	1.33	1.24	1.29	1.28	1.19	1.17					
CopG-DNA	1.73	1.73	1.29	1.21	1.27	1.31	1.27	1.25	0.39				
Mnt	1.81	1.80	1.66	2.18	1.96	2.11	2.24	2.27	2.13	2.08			
$\omega$	3.20	2.82	2.66	2.70	2.67	2.29	2.27	2.17	2.77	2.80	3.24		
ParG	1.19	0.97	1.26	1.57	1.59	1.71	1.78	1.85	1.88	1.79	1.97	2.77	

**Table V.2.** Major groove width vs. helix  $\alpha 2$  N-termini separation. The major groove width where the operator is contacted by the RHH  $\beta$ -strands is compared with the separation of the backbone amide nitrogen atom of residue 24 (N-24, see Fig. V.3 for numbering, Fig. V.2 for contacts) in the two molecules of the unbound RHH dimer.

<b>RHH Protein</b>	<b>Major groove width (Å)</b>	<b>N-24 separation (Å)</b>
Arc	14.3	15.4
CopG	8.6	13.6
MetJ	9.4	13.2
NikR	12.8	14.6

**Figure V.1.** Structures of RHH proteins.



*Figure V.1.a.* Two views of a ribbon representation of the NikR-DNA complex. The RHH portion of each molecule is colored, other portions of the protein are grey. DNA is represented as orange tubes passing through the phosphates of the DNA backbone. Nickel ions bound to NikR are represented as cyan spheres. The second view is shown from the top relative to the first.



*Figure V.1.b.* Two views of a ribbon representation of the MetJ-DNA complex. The RHH portion of each molecule is colored, other portions of the protein are grey. DNA is represented as orange tubes passing through the phosphates of the DNA backbone. SAM molecules bound to MetJ are represented as cyan sticks.



*Figure V.1.c.* Two views of a ribbon representation of the Arc-DNA complex. The RHH portion of each molecule is colored, other portions of the protein are grey. DNA is represented as orange tubes passing through the phosphates of the DNA backbone.



*Figure V.1.d.* Two views of a ribbon representation of the CopG-DNA complex. The RHH portion of each molecule is colored, other portions of the protein are grey. DNA is represented as orange tubes passing through the phosphates of the DNA backbone.





*Figure V.1.e.* Two views of a ribbon representation of Mnt. The RHH portion of each molecule is colored, other portions of the protein are grey.

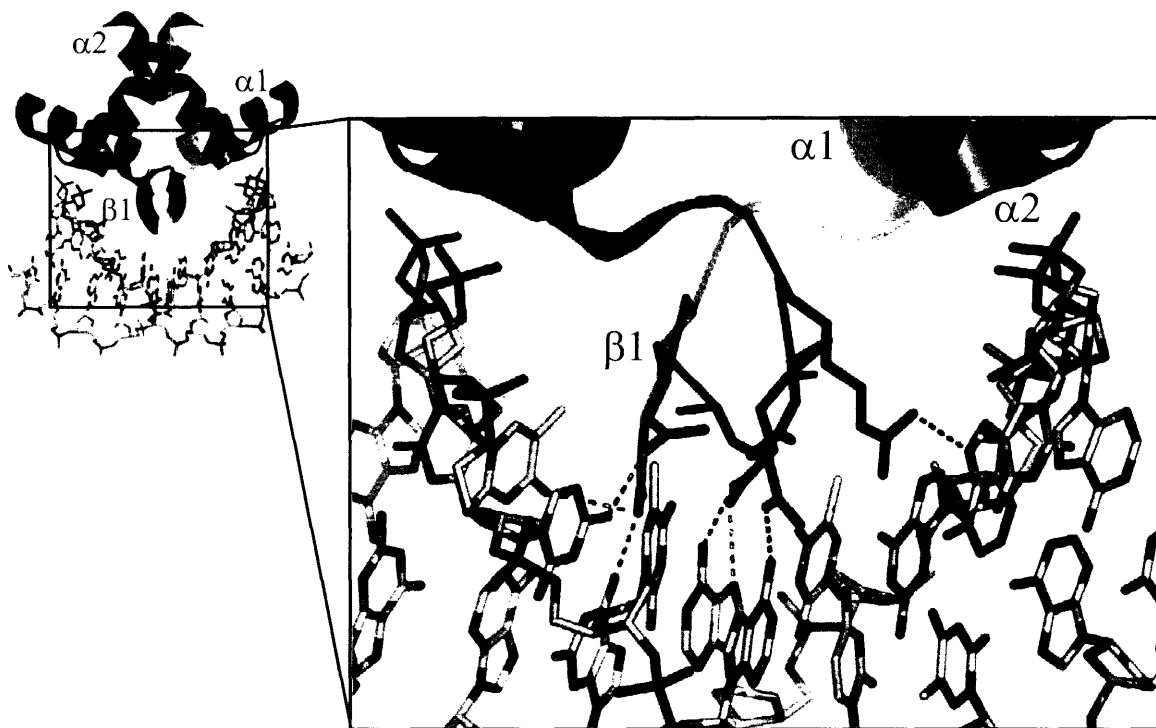


*Figure V.1.f.* Two views of a ribbon representation of ParG. The RHH portion of each molecule is colored, other portions of the protein are grey.

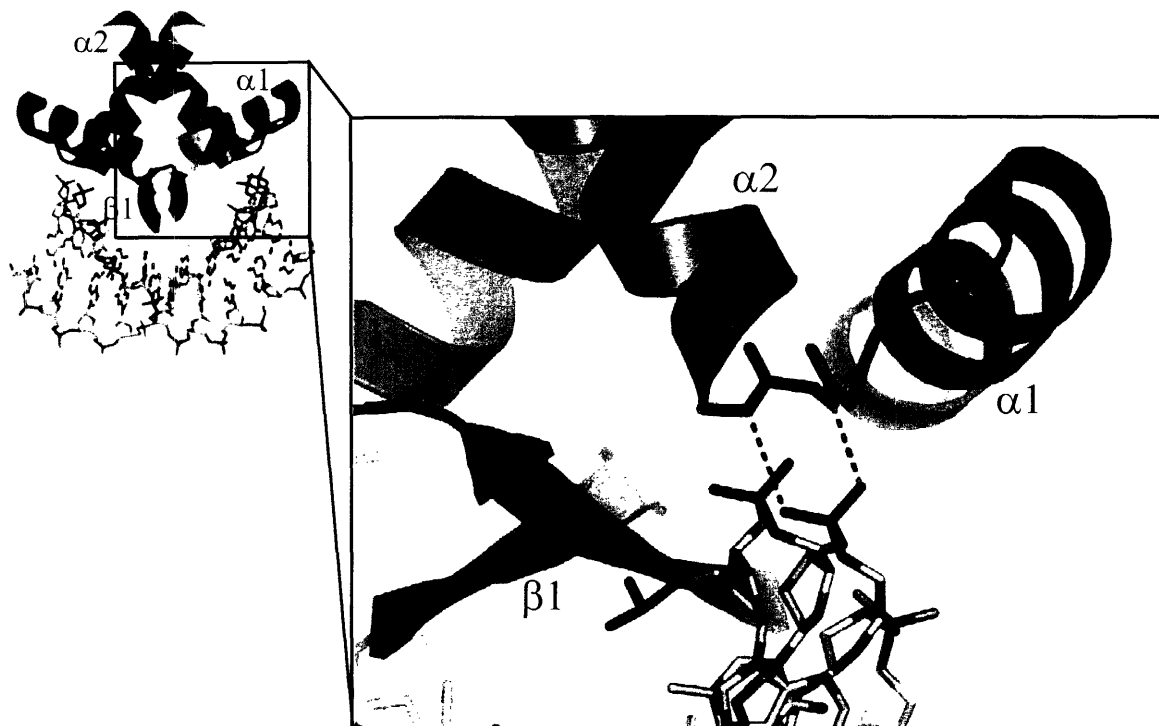


*Figure V.1.g.* Two views of a ribbon representation of  $\omega$ . The RHH portion of each molecule is colored, other portions of the protein are grey.

**Figure V.2.** Common features of DNA binding by RHH proteins.

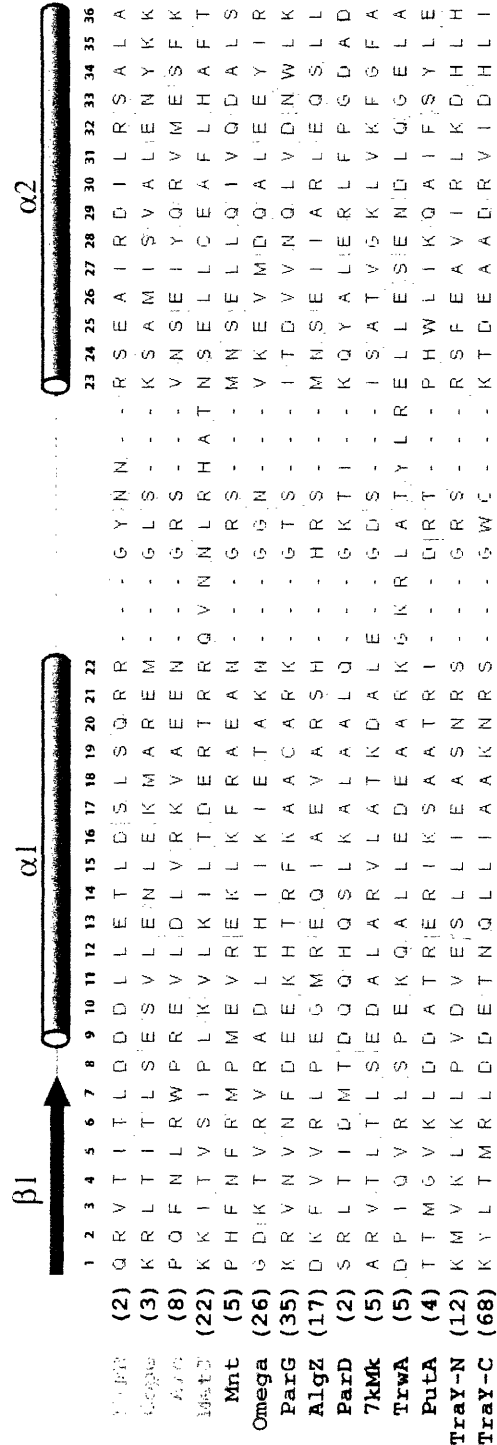


*Figure V.2.a.* Sequence-specific nucleotide base contacts in the DNA operator major groove by RHH proteins. The RHH protein is shown as a cartoon colored by molecule, the backbone of the  $\beta$ -strands is shown as a  $C_{\alpha}$  trace, the three amino acid sidechains from each  $\beta$ -strand and the DNA are represented as sticks, and hydrogen bonds between the protein and DNA bases are shown as dashed lines. The structure depicted here is that of the Arc-DNA complex, but is meant to represent the general RHH protein. The nucleotide base sequence and three amino acid sidechains vary among RHH family members. Secondary structure elements are labeled.

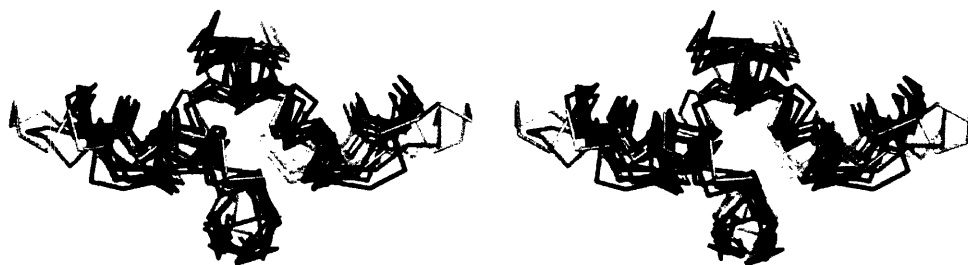


*Figure V.2.b.* RHH backbone-DNA phosphate backbone anchoring contacts. Representations are similar to Fig. V.2.a.

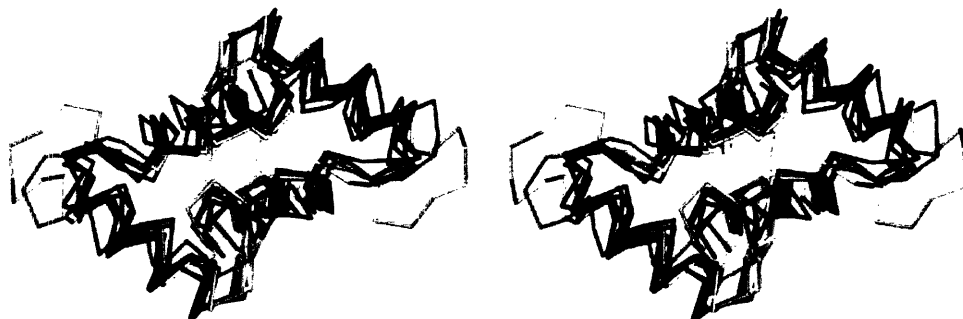
**Figure V.3.** Sequence alignment of characterized RHH proteins. Numbers indicate C $\alpha$  positions used for structural alignment, and are referred to throughout the text. Secondary structure elements are shown above the alignment and labeled. Sequences with red names have structures determined alone and bound to operator DNA. Sequences with blue names have structures determined alone only. Black names have no structures determined. Amino acids are shaded by property: blue – positive charge, red – negative charge, white – hydrophobic, green – neutral hydrophilic, purple – aromatic, orange – Gly or Pro, Yellow – Cys.



**Figure V.4.** Superimposition of RHH domains.

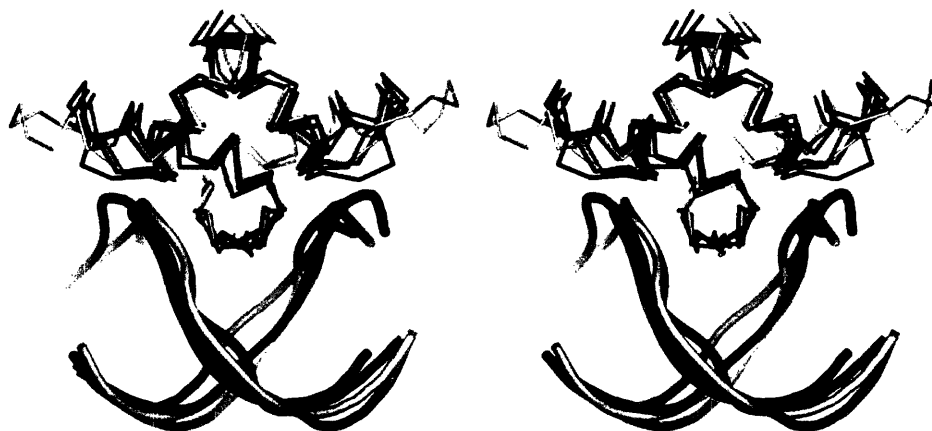


*Figure V.4.a.* Stereo view of a  $C_{\alpha}$  trace of superimposed RHH domains. Colors are as follows: NikR – blue, MetJ – yellow, Arc – red, CopG – green, Mnt – orange, ParG – purple,  $\omega$  – cyan.



*Figure V.4.b.* Stereo view of a  $C_{\alpha}$  trace of superimposed RHH domains. View is from the top relative to Fig. V.4.a.

**Figure V.5.** Superimposition of DNA-bound RHH domains.

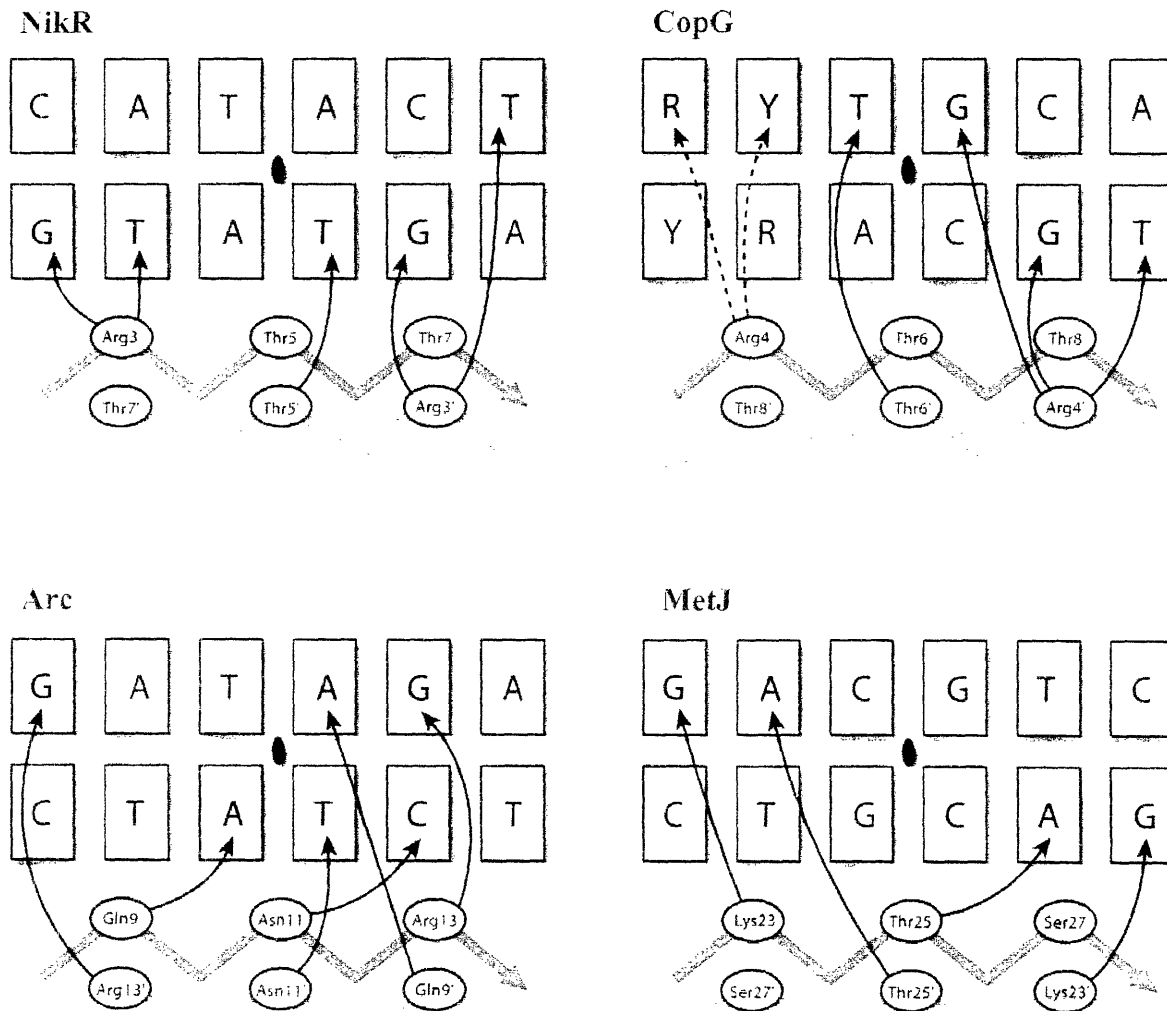


*Figure V.5.a.* Stereo view of superimposed DNA-bound RHH domains. RHH proteins are shown as a C $\alpha$  trace, DNA as a smooth tube passing through the backbone phosphate positions. Colors are as follows: NikR-DNA – blue, MetJ-DNA – yellow, Arc-DNA – red, CopG-DNA – green.



*Figure V.5.b.* Stereo view of superimposed DNA-bound RHH domains. View is from the top relative to Fig. V.5.a.

**Figure V.6.** Specific polar contacts between RHH  $\beta$ -strands and DNA operator sub-sites. Schematic representation of polar interactions between the three amino acid sidechains of each  $\beta$ -strand of a RHH dimer and the nucleotide bases within a DNA operator sub-site. R = purine, Y = pyrimidine. Arrows indicate polar interactions. Dashed arrows indicate interactions that only occur in some structures of CopG but not others. Contacted bases are shaded. The approximate location of the RHH dimer symmetry axis is indicated by an oval.





## Chapter VI: Afterword

### VI.A. Implications of this work

X-ray crystal structures of the NikR protein from *Escherichia coli* provided the first three-dimensional views of this protein, allowed identification of metal ion binding sites, and showed in detail how NikR interacts with its operator DNA sequence. The work described in this thesis has contributed significantly to our knowledge of the structure and mechanism of function of NikR. In a more general sense, it adds to the rapidly expanding field of intracellular metalloregulation, and to the study of the structural mechanisms for ligand induction in transcription factors (see chapter I).

### VI.A. Future experiments

Our results led us to propose hypotheses regarding the function of NikR that can be tested experimentally in the future. In chapter III, we propose that ordering of helix  $\alpha_3$  upon metal ion binding to the high-affinity site of NikR is important for the increase in affinity for operator DNA. This hypothesis could be tested by attempting to stabilize the folded structure of this helix in the absence of metal ion. For example, a disulfide crosslink between this helix and a less flexible part of the protein could be introduced<sup>1</sup>. Mutation of the amino acid sequence of the helix to a sequence with a higher propensity to form a stably folded helix may also be effective. Chapter III and chapter IV described our effort to identify secondary metal ion binding sites on NikR to explain the increased affinity of NikR for operator DNA in the presence of excess metal ions. A collaboration is currently underway with Dr. Deborah Zamble's lab at the University of Toronto to study the importance of the putative potassium binding site we observed at the RHH-

MBD interface by using DNA-binding assays of mutant proteins. A similar experimental approach could be used to study the relevance of additional nickel binding sites at the surface of NikR identified by soaking crystals of the DNA complex in excess NiCl<sub>2</sub>.

The *in vivo* relevance of additional binding sites could be measured by observing *nik* transcription levels within cells growing with various concentrations of nickel in the media. Orthologs of *E. coli* NikR should also be studied in more detail to examine whether secondary metal ion binding sites play a functional role for the NikR family as a whole, or whether the effect is specific to the *E. coli* protein. Transcriptome analysis of a *Helicobacter pylori* NikR knockout revealed that it plays a role in the regulation of a large number of genes in that organism<sup>2</sup>. Similar studies could be done with *E. coli* NikR to see if it regulates transcription of anything besides the characterized *nik* operon.

Additional studies to examine intracellular trafficking of nickel ions will likely contribute in the near future to our understanding of why and how the somewhat relaxed metal ion specificity of NikR is tolerated by *E. coli*. What type of kinetic controls are in place in this organism to deliver nickel ions to the enzyme active sites where they are required when NikR binds nickel so avidly? Furthermore, are there no “free” nickel ions in a normally growing cell, as has been proposed recently for both copper and zinc<sup>3-5</sup>? Answering these questions will move us several steps closer to an overall description of how intracellular metal ion concentrations are regulated.

## VI.C. References

1. Al-Rabee, R., Lee, E.J. & Grant, G.A. The mechanism of velocity modulated allosteric regulation in D-3-phosphoglycerate dehydrogenase. Cross-linking adjacent regulatory domains with engineered disulfides mimics effector binding. *J Biol Chem* **271**, 13013-7 (1996).
2. Contreras, M., Thiberge, J.M., Mandrand-Berthelot, M.A. & Labigne, A. Characterization of the roles of NikR, a nickel-responsive pleiotropic autoregulator of *Helicobacter pylori*. *Mol Microbiol* **49**, 947-63 (2003).
3. Finney, L.A. & O'Halloran, T.V. Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* **300**, 931-6 (2003).
4. Changela, A. et al. Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* **301**, 1383-7 (2003).
5. Outten, C.E. & O'Halloran, T.V. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292**, 2488-92 (2001).

## Chapter VII: Appendices

### VII.A. Protocols

#### *VII.A.1. NikR plasmids*

Expression of NikR or the MBD of NikR (residues 49-133) was from plasmids pNIK103<sup>1</sup> or pNIK102<sup>2</sup>, obtained from Peter T. Chivers (Washington University School of Medicine). Each was created by ligation of the appropriate coding sequence into pET-22b using the restriction enzymes NdeI and XhoI.

#### *VII.A.2. Preparation of competent Escherichia coli cells and plasmid transformation*

*E. coli* strain DL41 was made chemically competent by a variation of the TSS method<sup>3</sup>, and was suitable for storage at -80°C for several years.

1. Grow bacteria in Luria-Bertani (LB) media at 37°C until the culture reaches an optical density at 600 nm (OD<sub>600</sub>) of 0.5.
2. Incubate the culture on ice for 10 min.
3. Pellet the culture by centrifugation at 1000 x g at 4°C for 10 min.
4. Resuspend the bacterial pellet in 1/10 the original culture volume of ice-cold transformation and storage solution (TSS) (10 % polyethylene glycol (PEG) 8,000, 5 % dimethyl sulfoxide (DMSO), 30mM MgCl<sub>2</sub> in LB media) by gentle aspiration with a pipette.
5. Aliquot the 1/10 resuspension in 100 µL volumes into 1.5 mL microcentrifuge tubes and label for storage at -80°C

Generally, *E. coli* were transformed with plasmid stock before each overexpression of NikR. *E. coli* made chemically competent via the procedure described above were transformed with expression plasmids derived from standard plasmid miniprep kits as follows:

1. Thaw a 100  $\mu$ L aliquot of competent cells on ice.
2. Mix 1  $\mu$ L of supercoiled plasmid DNA (from miniprep) into the competent cells and incubate the mixture on ice for 20 min.
3. Place mixture in a 42°C water bath for 60 sec.
4. Incubate the mixture with shaking for 60 min. at 37°C.
5. Plate 10-100  $\mu$ L (depending on the original DNA concentration) of cells onto LB agar containing the appropriate antibiotic selection.
6. Incubate plates overnight at 37°C.

#### *VII.A.3. Overexpression of NikR or MBD in E. coli strain DL41*

1. Pick 4-5 colonies from a recent transformation to inoculate 10 mL starter cultures of LB supplemented with the appropriate antibiotic (usually 100  $\mu$ g/mL ampicillin).
2. Grow starter cultures overnight at 37°C with shaking (or rotating).
3. Pre-equilibrate 1 L aliquots of autoclaved LB in 2 L baffled shaker flasks at 37°C.

4. Mix multiple starter cultures together to ensure equivalent inoculation of larger cultures.
5. Add antibiotics to larger cultures immediately before inoculation.
6. Inoculate 1 L cultures with 4 mL of starter culture.
7. Incubate with shaking at 220 rpm with periodic monitoring of OD<sub>600</sub> to estimate culture doubling time.
8. Grow cultures until OD<sub>600</sub> reaches 1.0-1.2 and remove a 1 mL aliquot (“pre-induction” sample) to analyze later by SDS-PAGE. Pellet 1 mL aliquot by microcentrifugation at full speed for 1 min., discard supernatant, and store cell pellet at -20°C for analysis by SDS-PAGE.
9. Induce NikR overexpression by addition of solid isopropyl-β-D-thiogalactopyranoside (IPTG) to each flask to achieve a final concentration of ~1 mM when the culture OD<sub>600</sub> reaches 1.0-1.2. For convenience, a 1 g bottle of IPTG may be divided into four equivalent 250 mg portions and one portion added to each 1L flask.
10. Allow cultures to incubate with shaking at 37°C for four hours after induction.
11. Remove, pellet, and store another 1 mL aliquot (“post-induction” sample) for SDS-PAGE analysis.
12. Pellet cultures in 1 L centrifuge bottles by centrifugation at 4°C for 30 min. at 10,000 x g.
13. Combine pellets and store at -80°C until purification.

#### *VII.A.4. Overexpression of selenomethionine (SeMet) labeled NikR in DL41*

The *E. coli* strain DL41 is a methionine auxotroph, and was designed to be suitable for expression of SeMet labeled proteins. Expression of SeMet labeled NikR was carried out much the same as described above for WT NikR, with the following modifications:

1. Starter cultures and 1 L cultures are in chemically defined LeMaster medium<sup>4</sup>, supplemented with 1x Kao and Michayluk vitamin supplement (Sigma) and 50 mg/L D,L-selenomethionine (Sigma), instead of LB. The vitamin supplement, selenomethionine, and antibiotic are added to the medium immediately before inoculation.
2. The doubling time of DL41 in the LeMaster medium is considerably slower than in LB, generally around 60 min.
3. After induction with IPTG, the cultures are incubated for 10-12 hours (instead of 4 hours) before harvesting. Addition of another 100 µg/mL ampicillin after induction should help plasmid retention and protein yield since the growth times are considerably longer.

#### *VII.A.5. Purification of NikR or MBD for crystallization*

##### *VII.A.5.a. Immobilized nickel-affinity chromatography*

1. Thaw frozen *E. coli* cell pellet on ice. This procedure is described for a pellet from 8 L of cell culture as described above.

2. Resuspend cell pellet in 50 mL buffer 1 (100 mM potassium phosphate, 300 mM NaCl, 10 mM imidazole, pH 8.0) gently using first a spatula and then a manual tissue homogenizer.
3. Add phenylmethylsulfonyl fluoride (PMSF), a serine protease inhibitor, from a freshly-made 200 mM stock in isopropanol to the resuspended cells to a final concentration of 2 mM.
4. Sonicate the resuspended cells intermittently for 5 min. or until mostly clarified. Use a metal beaker packed in ice to keep the suspension cold.
5. Pellet the insoluble fraction after sonication by centrifugation at 35,000 x g for 30 min. at 4°C. Keep a small aliquot of the supernatant for SDS-PAGE analysis (“cell break supernatant”).
6. Decant supernatant into a clean 100 mL bottle and add NiCl<sub>2</sub> from a concentrated stock to a final concentration of 100 μM with mixing. *Note: Failure to add NiCl<sub>2</sub> will result in NikR removing nickel from the Ni-affinity resin.*
7. Add 10 mL (solid volume) of Ni-affinity resin (Novagen) and nutate at 4°C for at least 10 min. (Batch bind)
8. Pour the mixture into a gravity flow chromatography column with a stopcock (1 in. x 9 in. works well) and allow resin to settle to the bottom.
9. Allow the cell break supernatant to flow through the column and take a small sample for SDS-PAGE (“flow-through”).



10. Wash the resin with 50 resin-bed volumes of buffer 1, flowing at a rate of  $\leq 5$  mL/min, keep a sample for SDS-PAGE (“10mM wash”). Washes can be carried out a room temperature with buffer that has been equilibrated at 4°C.
11. Wash the resin with 10 resin-bed volumes of buffer 2 (100 mM potassium phosphate, 1 M NaCl, 35 mM imidazole, pH 8.0), keep a sample for SDS-PAGE (“35mM wash”).
12. Elute NikR from the column with  $\sim 1.5$ -2 resin-bed volumes of buffer 3 (100 mM potassium phosphate, 300 mM NaCl, 250 mM imidazole, pH 8.0), take a sample for SDS-PAGE (“250mM elution”). The resin should change color during this step from light grey or white to its original light blue color. The eluted NikR should be fairly concentrated and therefore have a light yellow-orange color.
13. Run a SDS-PAGE gel of the samples collected from each step of the expression and affinity chromatography to check the overexpression level and sample purity. Such a gel is show in Fig. VII.1. The 250 mM elution sample should be  $> 90\%$  pure as judged by coomassie-stained SDS-PAGE.
14. NikR can be quantitated at this stage by  $A_{276}$  using an extinction coefficient of  $4400 \text{ M}^{-1} \text{ cm}^{-1}$  for full-length NikR or  $2900 \text{ M}^{-1} \text{ cm}^{-1}$  for the MBD. The sample must first be exchanged into a buffer that does not contain imidazole (which absorbs at 276 nm) and then denatured in 8 M guanidine hydrochloride (Pierce) (to remove bound nickel, which also absorbs at 276 nm when bound to NikR). Typical NikR yields after Ni-affinity chromatography are 20-30 mg of NikR per 1 L of *E. coli* culture.

#### VII.A.5.b. Size-exclusion chromatography

The NikR sample should be mostly pure after Ni-affinity chromatography. Size-exclusion chromatography (SEC) is useful to remove the small amount of aggregated NikR that usually exists and which can inhibit crystallization.

1. Concentrate the NikR sample to 10-15 mg/mL so that a small volume can be loaded onto the SEC column. Overconcentration can result in NikR precipitation, which can be reversed by increasing the sample pH to 9 or adding EDTA to remove nickel from the protein.
2. Equilibrate the SEC (generally Superdex 200 for full-length NikR and Superdex 75 for MBD) with buffer 4 (20 mM tris(hydroxymethyl)aminomethane (Tris), 300 mM NaCl, pH 8.5) that has been filtered. The pH of a Tris-buffered solution is especially sensitive to temperature, so pH the solutions at the temperature where they will be used (4°C).
3. Load NikR onto the column and elute at a flow rate of 0.5-1.0 mL/min, monitoring  $A_{280}$ .
4. NikR will elute as a tetramer in primarily a single peak of molecular weight (MW) ~60,000 (15,096 Da for the NikR monomer). See Fig. VII.2 for a sample chromatogram. Generally a smaller peak of aggregated protein will elute shortly before the main peak.
5. Run a SDS-PAGE of the fractions around the main peak and pool the fractions containing the highest purity NikR.

6. Estimate the concentration of NikR in the pooled fractions by measuring  $A_{276}$  of the denatured protein as described previously. Typical yield after SEC is 10-20 mg of NikR per 1 L of *E. coli* culture.
7. Store the purified sample of NikR at 4°C, where it appears to be stable for years.

#### *VII.A.6. Quantitation of $Ni^{2+}$ :NikR ratio using PAR*

The metallochromic indicator 4-(2-pyridylazo)resorcinol (PAR) can be used to rapidly determine concentrations of metal ions in protein samples<sup>5-7</sup>. The  $Ni^{2+}$ -PAR complex exhibits absorption maxima at 502 nm and 528 nm that do not overlap with the absorbance of denatured protein and can therefore be used to determine the concentration of  $Ni^{2+}$  ions in a sample of NikR.

1. Make a 1 mM stock solution of PAR in water using volumetric glassware. Store the solution at 4°C in the dark.
2. Use nickel concentration standards to produce a calibration curve for  $A_{502}$  and  $A_{528}$  vs.  $[Ni^{2+}]$  in 8 M guanidine HCl. Such a calibration curve is shown in Fig. VII.3, and is generally linear only in the range of 0-12  $\mu$ M.
3. Determine the concentration of NikR using  $A_{276}$  of denatured protein as described above.
4. Dilute the protein sample and PAR together with 8 M guanidine HCl until the estimated  $[Ni^{2+}]$  is 5-10  $\mu$ M and the [PAR] is 100  $\mu$ M. Typically, a sample containing 10  $\mu$ L NikR, 20  $\mu$ L PAR, and 170  $\mu$ L of guanidine was used.

5. Allow at least 30 min. for complete denaturation and nickel dissociation from the NikR sample.
6. Read the absorbance at 502 nm and 528 nm after blanking the UV-Vis with a sample containing buffer in place of the protein solution.
7. Estimate the  $[\text{Ni}^{2+}]$  by comparison with the calibration curve.
8. Divide the  $[\text{Ni}^{2+}]$  by  $[\text{NikR}]$  to obtain the fraction of NikR with nickel bound.

#### *VII.A.7. HPLC purification of deoxyoligonucleotides for co-crystallization with NikR*

Single-stranded deoxyoligonucleotides ( $\mu\text{mole}$  scale synthesis) were purchased from Integrated DNA Technologies (IDT-DNA) after standard desalt purification. These oligonucleotides were further purified by reverse-phase HPLC as described below, a procedure obtained from the laboratory of Carl Pabo:

1. Prepare two 1 L stock solutions of 1 M triethyl amine (TEA). The first stock solution should be adjusted to pH 6.5 using glacial acetic acid in a fume hood (TEAA). The second should be adjusted to pH  $\sim 7$  by bubbling  $\text{CO}_2$  through the solution (TEAB).
2. Make several liters each of 50 mM TEAA in acetonitrile (buffer A) and  $\text{H}_2\text{O}$  (buffer B).
3. Dissolve each oligonucleotide in 10 mM TEAB in  $\text{H}_2\text{O}$  and filter through a  $0.2 \mu\text{m}$  centrifuge filter.
4. Inject sequentially onto a 2.2 cm x 15 cm Vydac C4 column (product #214TP102215) maintained at  $50^\circ\text{C}$  using a column heater and separate using the gradient described in Table VII.1 with a flow rate of 8 mL/min.

5. Collect the peak above  $A_{260} = 0.8$ . A sample chromatogram is shown in Fig. VII.4.
6. Dialyze each oligo three times against 4 L of 10 mM TEAB.
7. Lyophilize oligos to dryness.
8. Resuspend oligos in 1 mL and quantitate using  $A_{260}$ .
9. Anneal complimentary oligos by mixing equimolar amounts, placing them in an insulated waterbath at  $90^{\circ}\text{C}$  and allowing them to cool slowly to  $4^{\circ}\text{C}$  overnight.
10. Aliquot double-stranded oligos into 1.5 mL microcentrifuge tubes in portions suitable for crystallographic experiments and concentrate to dryness using a speed vac.
11. Store double-stranded oligos at  $-20^{\circ}\text{C}$  or  $-80^{\circ}\text{C}$ .

## **VII.B. Crystal structure of the R33M NikR mutant**

### *VII.B.1 Results and Discussion*

The crystal structure of the nickel-bound R33M NikR mutant was solved by molecular replacement and is shown in Fig. VII.5. The MBD of the R33M NikR tetramer is very similar to the structures of the isolated  $\text{Ni}^{2+}$ -MBD and full-length  $\text{Ni}^{2+}$ -NikR (chapter III). The RHH DNA-binding domain (DBD), however is slightly altered in its structure due to the mutation. The largest difference is a twist of the  $\beta$ -strands by  $\sim 13^{\circ}$  in the R33M structure relative to WT NikR (Fig. VII.6). Comparison to the previously determined apo, nickel-bound, and DNA-bound structures of NikR gives RMSD values of 1.06 Å, 1.44 Å, and 1.89 Å, respectively when comparing 80  $C_{\alpha}$  atoms

of the dimeric RHH domain. In the WT structures, Arg33 makes a hydrogen bond to the backbone carbonyl oxygen of Thr7' (Fig. VII.7.a), an interaction that is disrupted by mutation to methionine (Fig. VII.7.b). The orientation of the RHH domains relative to the MBD is also different in this structure than has been observed previously (Fig. VII.8). It seems unlikely that this is a consequence of the mutation, but rather reflects the inherent structural flexibility of NikR.

### *VII.B.2 Materials and Methods*

Ni<sup>2+</sup>-R33M NikR was expressed and purified exactly as the wild-type NikR (see chapter 2). Crystals of nickel-bound R33M NikR were grown by mixing 1.5  $\mu$ L of protein solution (10 mg/mL R33M NikR in 20 mM Tris, 300 mM NaCl, pH 8.0) with 1.5  $\mu$ L of precipitant (0.2M MgCl<sub>2</sub>, 0.1M NaHEPES pH 7.5, 30% PEG 400) in a sitting drop vapor diffusion experiment at room temperature. A crystal from this condition that grew after one week was cooled to 100K in a gaseous N<sub>2</sub> stream directly from the drop without cryoprotection. A dataset was collected at beamline X12-B of the NSLS to 2.3 Å resolution. The data were reduced using DENZO/SCALEPACK<sup>8</sup> in the point group P2, statistics are given in Table VII.2. Matthews coefficient analysis indicated that there were likely four molecules of R33M NikR per ASU. The Ni<sup>2+</sup>-R33M NikR structure was solved by molecular replacement in the space group P2<sub>1</sub> first using AMoRe<sup>9</sup> with the Ni<sup>2+</sup>-MBD tetramer (1Q5Y)<sup>10</sup> as a search model. Four obvious solutions were found corresponding to the four possible orientations of the Ni<sup>2+</sup>-MBD tetramer. Subsequently, EPMR<sup>11</sup> was used to locate the two RHH dimers using a portion of the apo NikR structure (1Q5V)<sup>10</sup> while holding the MBD solution fixed. Inspection of 2F<sub>o</sub>-F<sub>c</sub> maps

following rigid body refinement in CNS<sup>12</sup> revealed that the MBD of Ni<sup>2+</sup>-R33M NikR was nearly identical to the published Ni<sup>2+</sup>-MBD structure. In the RHH domains, however, there was a clear twist of the beta-strands relative to the alpha-helices compared with the apo-NikR model (Fig. VII.6), and the strands were removed to be rebuilt manually. At this stage there was also clear positive  $F_o - F_c$  density for two of the linkers connecting the RHH domain and MBD, corresponding to one RHH dimer. After several cycles of rebuilding in Xfit<sup>13</sup> and positional and individual B-factor refinement in CNS, the current model consists of the entire MBD tetramer (except for a small loop, residues 64-66 in each molecule) as well as a complete RHH dimer and the linkers connecting it (Fig. VII.5). The other RHH dimer, however, shows a considerable amount of disorder and ~30% of this domain has not been modeled owing to uninterpretable density. Neither of the linkers to this RHH dimer have been modeled either because of the disorder. The current model therefore contains 473 out of 532 amino acids, 4 Ni<sup>2+</sup> ions, and 36 water molecules with an  $R_{\text{work}} = 25.7\%$  and  $R_{\text{free}} = 29.5\%$ . Model refinement statistics are given in Table VII.3. Figures were produced using PyMOL<sup>14</sup>.

### **VII.C. Two additional crystal forms of apo NikR**

An interesting question regarding the function of NikR is how flexible the ribbon-helix-helix (RHH) domains are relative to the metal-binding domain (MBD). In our original structures of apo-NikR (1Q5V) and Ni<sup>2+</sup>-NikR, neither NikR tetramer was in a conformation suitable for interaction with operator DNA. This seems to contradict a mechanism where high-affinity nickel binding enforces a conformation of NikR that can bind DNA, but does not rule out that in the absence of nickel the repressor is locked into

a non-productive conformation. Since crystallography provides only a single snapshot of a molecule per crystal (unless multiple copies of the molecule are present in the ASU), multiple structures in unique crystal forms are required to provide information about dynamic processes such as conformational change. To learn more about the range of conformations accessible by apo-NikR, I crystallized apo-NikR and a single point mutant in two crystal forms that are distinct from the original apo-NikR structure (1Q5V).

#### *VII.C.1 Apo-NikR crystal form #2*

The first crystal form of wild-type (WT) apo-NikR was grown by mixing 1.5  $\mu\text{L}$  of  $\text{Cu}^{2+}$ -NikR (10 mg/mL) in 20 mM Tris, pH 8.0, 300 mM NaCl with 1.5  $\mu\text{L}$  of precipitant solution, consisting of 0.1 M NaCl, 0.1 M sodium acetate trihydrate, pH 4.6, 12% v/v 2-methyl-2,4-pentanediol (MPD) at room temperature in a sitting drop vapor diffusion experiment. Diamond-shaped colorless crystals measuring  $\sim 300 \times 250 \times 100 \mu\text{m}$  grew after 24 hours and were cryoprotected by increasing the precipitant MPD concentration to 25 % before cryo-cooling in liquid nitrogen ( $\text{LN}_2$ ). Since the crystals grew at pH 4.6 and were colorless,  $\text{Cu}^{2+}$  ions were no longer bound to the protein. A dataset was collected to 2.8 Å resolution at APS beamline 8BM. Data were reduced using DENZO/SCALEPACK<sup>8</sup> in the point group P222. Statistics are given in Table VII.1.

Matthews coefficient analysis suggests that this crystal form contains 5-9 molecules of NikR in the ASU, with predicted solvent contents of 63-33%, respectively. Systematic absences suggest that the space group is likely  $\text{P}2_12_12_1$  (Table VII.4). Molecular replacement trials were carried out with one, two, or four subunits of the MBD



or a RHH domain dimer from the original apo-NikR structure (1Q5V)<sup>10</sup> as search models, using the programs AMoRe<sup>9</sup>, EPMR<sup>11</sup>, and Phaser<sup>15</sup> in all four of the primitive orthorhombic space groups. No searches yielded reasonable solutions.

### *VII.C.2 Apo-NikR crystal form #3 (E63A point mutant)*

The second new apo-NikR crystal form was grown by mixing 1.5  $\mu\text{L}$  of  $\text{Ni}^{2+}$ -E63A NikR (10 mg/mL in 300 mM NaCl, 20 mM Tris, pH 8.0) with 1.5  $\mu\text{L}$  of precipitant, containing 0.1 M NaCl, 0.1 M sodium acetate, pH 4.6, and 12% w/v PEG 6000 at room temperature in a sitting drop vapor diffusion setup. Irregularly shaped colorless crystals measuring  $\sim 150 \times 100 \times 100 \mu\text{m}$  grew after 2-4 days. Again, because the crystals were grown at low pH and were colorless, they likely contained the apo form of the protein. These crystals were cryoprotected by soaking in the precipitant solution supplemented with 20% v/v ethylene glycol before cryo-cooling in  $\text{LN}_2$ . A dataset was collected to 3.4  $\text{\AA}$  at ALS beamline 8BM. Data were reduced using DENZO/SCALEPACK<sup>8</sup> in the point group P2. Statistics are given in Table VII.1.

Matthews coefficient analysis implies either three or four molecules of NikR per ASU with 54% or 39% solvent, respectively. Systematic absences strongly imply that the space group is P2<sub>1</sub>. A reasonable molecular replacement solution was found by searching first with a tetramer of the MBD followed by a dimer of the RHH domain from the original apo-NikR structure (1Q5V)<sup>10</sup> in P2<sub>1</sub> using the program Phaser<sup>15</sup>. Only one RHH domain was located, searching for a second gave no reasonable solutions. Although the crystal packing was reasonable for the molecular replacement solution and 2Fo-Fc maps calculated from that partial model revealed continuous density for the two

inter-domain linkers, the maps were overall quite noisy and rigid body refinement of the solution in CNS did not significantly reduce the R-factors. A cavity was present in the crystal lattice where the second RHH domain could be expected, but the electron density present there was not continuous and could not be interpreted as a RHH domain. Although the molecular replacement solution for the MBD and one RHH domain seems correct, any positional refinement resulted in a decrease in  $R_{\text{work}}$  but not  $R_{\text{free}}$ , which could not be reduced below 48%.

#### VII.D. References

1. Chivers, P.T. & Sauer, R.T. Regulation of high affinity nickel uptake in bacteria.  $\text{Ni}^{2+}$ -Dependent interaction of NikR with wild-type and mutant operator sites. *Journal of Biological Chemistry* **275**, 19735-41 (2000).
2. Chivers, P.T. & Sauer, R.T. NikR repressor: high-affinity nickel binding to the C-terminal domain regulates binding to operator DNA. *Chemistry and Biology* **9**, 1141-8 (2002).
3. 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**, 2172-5 (1989).
4. LeMaster, D.M. & Richards, F.M.  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear NMR studies of *Escherichia coli* thioredoxin in samples isotopically labeled by residue type. *Biochemistry* **24**, 7263-8 (1985).
5. Jezorek, J.R. & Freiser, H. 4-(Pyridylazo)resorcinol-based continuous detection system for trace levels of metal ions. *Analytical Chemistry* **51**, 373-376 (1979).
6. 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**, 150-7 (1985).
7. Atanassova, A., Lam, R. & Zamble, D.B. A high-performance liquid chromatography method for determining transition metal content in proteins. *Anal Biochem* **335**, 103-11 (2004).
8. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods in Enzymology* **276**, 307-326 (1997).

9. Navaza, J. *AMoRe*: an automated package for molecular replacement. *Acta Crystallographica, section A* **A50**, 157-163 (1994).
10. Schreiter, E.R. et al. Crystal structure of the nickel-responsive transcription factor NikR. *Nat Struct Biol* **10**, 794-9 (2003).
11. Kissinger, C.R., Gehlhaar, D.K. & Fogel, D.B. Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr D Biol Crystallogr* **55 ( Pt 2)**, 484-91 (1999).
12. Brünger, A.T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica. Section D: Biological Crystallography* **54**, 905-21 (1998).
13. McRee, D.E. XtalView/Xfit - A Versatile Program for Manipulating Atomic Coordinates and Electron Density. *Journal of Structural Biology* **125**, 156-165 (1999).
14. DeLano, W.L. The PyMOL Molecular Graphics System. (DeLano Scientific, San Carlos, 2002).
15. Storoni, L.C., McCoy, A.J. & Read, R.J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr* **60**, 432-8 (2004).

## VII.E. Tables and Figures

**Table VII.1.** Gradient programs for reverse-phase HPLC purification of deoxyoligonucleotides. Solvent A is 50 mM TEAA in acetonitrile, solvent B is 50 mM TEAA in H<sub>2</sub>O.

<b>Program</b>	<b>Time (min.)</b>	<b>Flow Rate (mL/min.)</b>	<b>Solvent %B</b>
<b>Start Up</b>	0	8	0
	16	8	0
	24	8	95
<b>Separation</b>	0	8	95
	32	8	83
	37	8	70
	40	8	70
	45	8	95
	53	8	95
<b>Shut Down</b>	0	8	95
	8	8	0
	24	8	0

**Table VII.2.** Dataset statistics.

<b>Crystal</b>	<b>Ni<sup>2+</sup>-R33M NikR</b>	<b>Apo NikR #2</b>	<b>Apo NikR #3</b>
<b>Space Group</b>	P2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub>
<b>Cell Dimensions</b>			
<b>a (Å)</b>	47.41	51.50	51.11
<b>b (Å)</b>	123.25	132.11	50.99
<b>c (Å)</b>	51.62	147.43	93.73
<b>β (°)</b>	116.5	-	93.2
<b>Wavelength (Å)</b>	1.0000	0.9793	0.9793
<b>Temperature (K)</b>	100	100	100
<b>Resolution Range (Å)</b>	50-2.3	50-2.8	50-3.4
<b>Unique Reflections</b>	23,274	25,123	6,590
<b>Average Redundancy</b>	4.5	8.6	3.6
<b>Completeness (%)<sup>1</sup></b>	98.4 (88.7)	98.1 (85.1)	95.8 (72.7)
<b><i>I</i> / σ(<i>I</i>)<sup>1</sup></b>	23.4 (3.7)	25.4 (3.2)	22.4 (4.2)
<b><i>R</i><sub>sym</sub> (%)<sup>1,2</sup></b>	6.3 (21.7)	6.7 (38.6)	6.6 (18.5)

<sup>1</sup>The number in parentheses is for the highest resolution shell.

<sup>2</sup> $R_{\text{sym}} = \sum_{hkl} |I_{(hkl)}^i - \langle I_{(hkl)} \rangle| / \sum_{hkl} \langle I_{(hkl)} \rangle$ , where  $I_{(hkl)}^i$  is the  $i^{\text{th}}$  measured diffraction intensity and  $\langle I_{(hkl)} \rangle$  is the mean of the intensity for the miller index  $(hkl)$ .

**Table VII.3.** Model refinement statistics.

<b>Crystal</b>	<b>Ni<sup>2+</sup>-R33M NikR</b>
<b><math>R_{\text{cryst}} (R_{\text{free}}) (\%)</math><sup>1</sup></b>	25.7 (29.5)
<b>Resolution Range (Å)</b>	50-2.3
<b>Number of protein atoms</b>	3604
<b>Number of nickel ions</b>	4
<b>Number of water molecules</b>	36
<b>R.m.s. deviations</b>	
<b>Bond lengths (Å)</b>	0.008
<b>Bond angles (°)</b>	1.29
<b>Ramachandran (%)</b>	
<b>Most favored</b>	92.6
<b>Additionally Allowed</b>	7.4
<b>Generously Allowed</b>	0.0
<b>Disallowed</b>	0.0
<b>Average <i>B</i>-factor (Å<sup>2</sup>)</b>	69.1

<sup>1</sup> $R_{\text{cryst}} = \sum_{hkl} ||F_o(hkl)| - |F_c(hkl)|| / \sum_{hkl} |F_o(hkl)|$ .  $R_{\text{free}} = R_{\text{cryst}}$  for a test set of reflections (5%) not included in refinement.

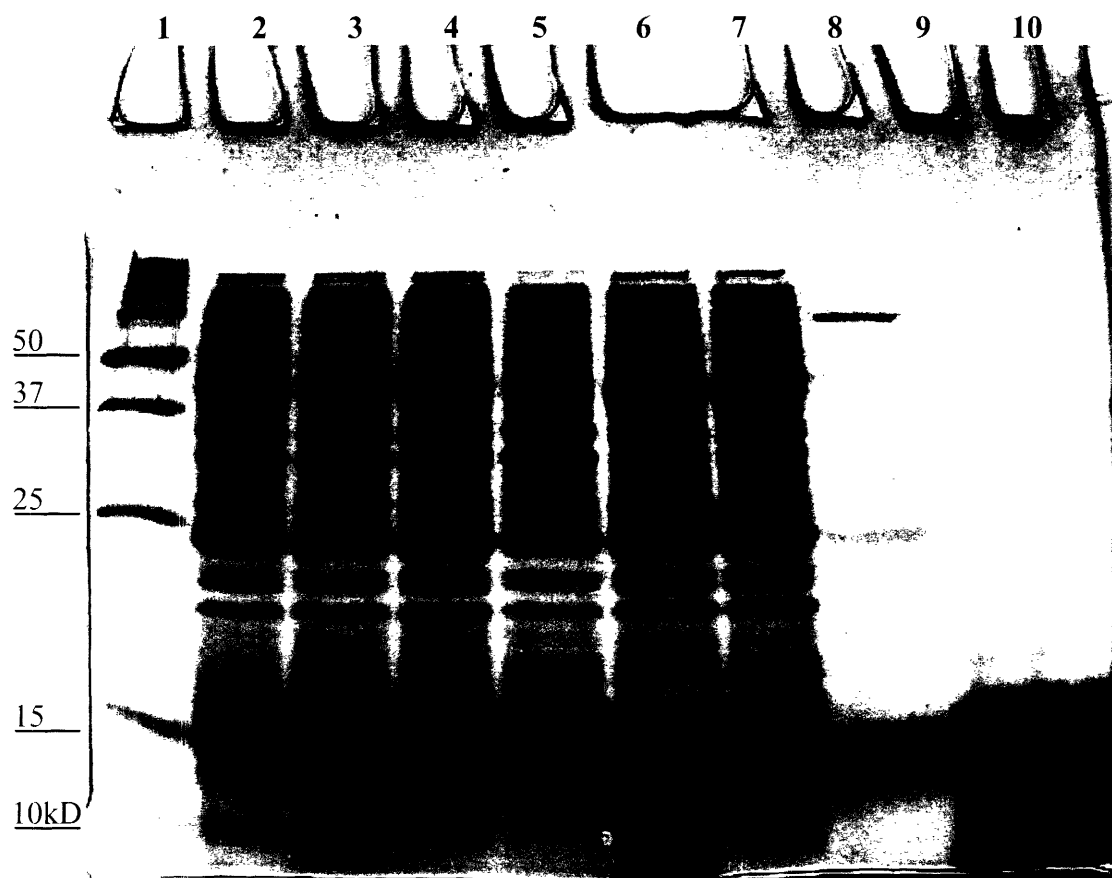
**Table VII.4.** Systematic absences. Intensities of the 00l, 0k0, and h00 reflections for apo-NikR crystal form #2 are shown after processing as P222. The pattern of intensities is consistent with the space group  $P2_12_12_1$ .

0	0	5	17.7	5.2
0	0	6	72203.7	5143.3
0	0	7	17.8	11.4
0	0	8	178445	12661
0	0	10	2334.8	172.5
0	0	11	28.6	26.4
0	0	12	216048	12507
0	0	13	64.1	17.8
0	0	14	193538	13702
0	0	15	-0.6	37.6
0	0	16	68555.0	3977.3
0	0	18	1271.9	81.7
0	0	19	12.4	29.9
0	0	20	8118.4	477.0
0	0	21	70.6	41.5
0	0	22	4886.9	292.0
0	0	23	164.5	45.5
0	0	24	3208.9	194.1
0	0	25	-13.1	68.2
0	0	26	3214.1	192.0
0	0	27	-25.0	41.9
0	0	28	1404.0	98.2
0	0	29	-36.4	59.1
0	0	30	10840.1	633.6
0	0	31	-40.4	63.3
0	0	32	549.3	67.8
0	0	33	50.4	60.3
0	0	34	1631.9	109.3
0	0	35	128.0	118.8
0	0	36	358.4	59.9
0	0	38	224.4	76.6
0	0	39	63.9	76.1
0	0	40	437.3	81.8
0	0	41	53.8	108.7
0	0	42	1601.9	117.7
0	0	43	285.1	129.5
0	0	44	1365.6	118.5
0	0	46	794.5	101.1
0	0	47	-6.3	106.2
0	0	48	937.3	94.8
0	0	49	-33.2	108.8
0	0	50	256.7	87.4
0	0	51	-18.3	93.0

0	0	52	268.8	139.6
0	5	0	1766.8	130.7
0	6	0	71911.2	5108.9
0	7	0	598.4	46.9
0	8	0	229887	16335
0	9	0	101.6	20.6
0	10	0	98029.2	5694.5
0	11	0	1989.9	121.9
0	12	0	289481	14526
0	13	0	152.3	16.1
0	14	0	255114	12800
0	15	0	2722.8	144.5
0	16	0	937.1	54.2
0	18	0	79299.3	3987.3
0	19	0	1095.3	64.3
0	20	0	75165.2	3773.2
0	21	0	228.5	34.1
0	22	0	7937.5	405.2
0	23	0	3120.5	167.7
0	24	0	48525.8	2818.3
0	25	0	1222.7	92.4
0	26	0	19812.4	1003.7
0	27	0	71.6	42.6
0	28	0	36890.9	1679.2
0	29	0	2205.3	117.1
0	30	0	45367.3	2060.3
0	31	0	590.3	67.3
0	32	0	6650.5	347.4
0	34	0	14516.3	1518.5
0	35	0	37.3	49.6
0	36	0	1553.1	104.8
0	39	0	29.6	56.2
2	0	0	290197	20605
3	0	0	53.0	17.8
5	0	0	3162.6	189.0
6	0	0	1879.9	138.3
7	0	0	109.6	18.7
8	0	0	7286.1	429.9
9	0	0	411.0	37.8
10	0	0	21566.3	1536.2
11	0	0	882.6	65.4
12	0	0	9767.6	537.9
13	0	0	463.9	81.2
14	0	0	12502.6	740.6
15	0	0	646.4	114.4

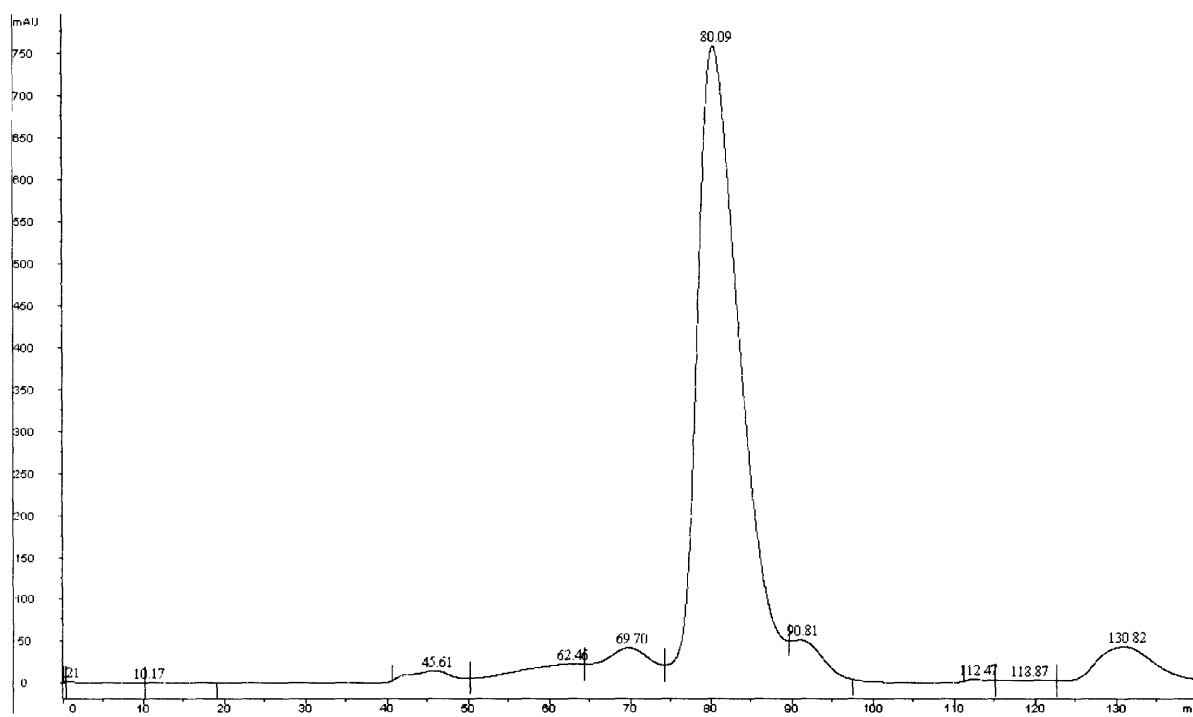


**Figure VII.1.** SDS-PAGE illustrating overexpression and immobilized nickel affinity purification of NikR.

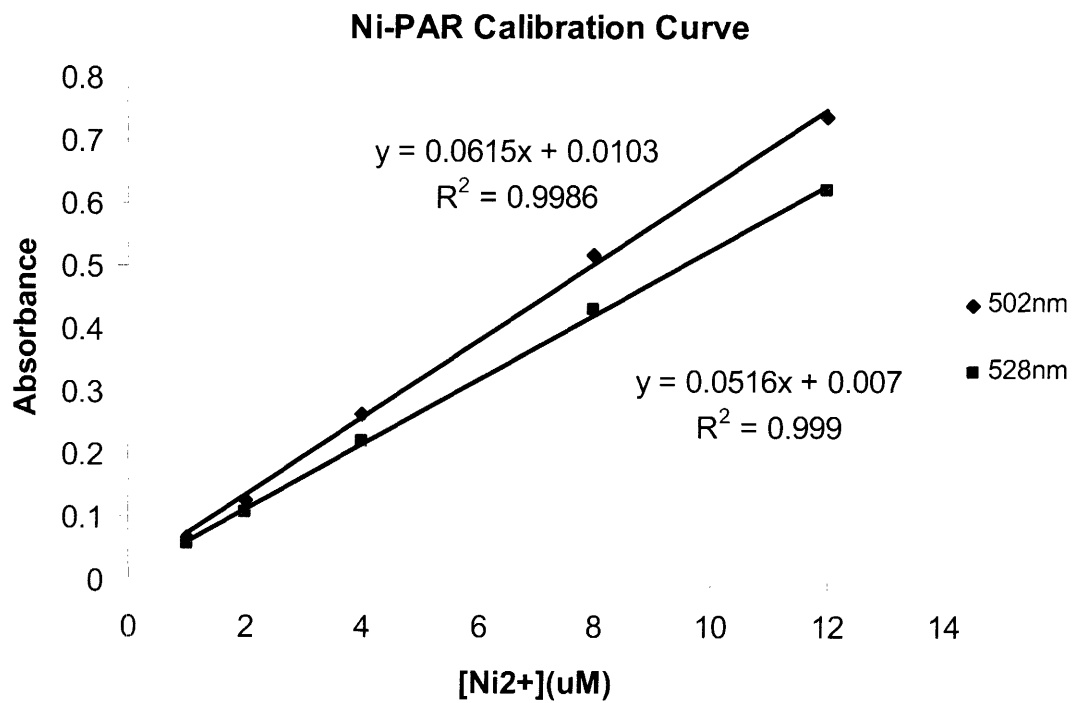


- 1 – Molecular weight markers
- 2 – Before induction
- 3 – After induction
- 4 – Before induction
- 5 – After induction
- 6 – Cell break supernatant
- 7 – Column flow-through
- 8 – 35 mM imidazole wash
- 9 – 250 mM imidazole elution
- 10 – Purified NikR standard (5 µg)

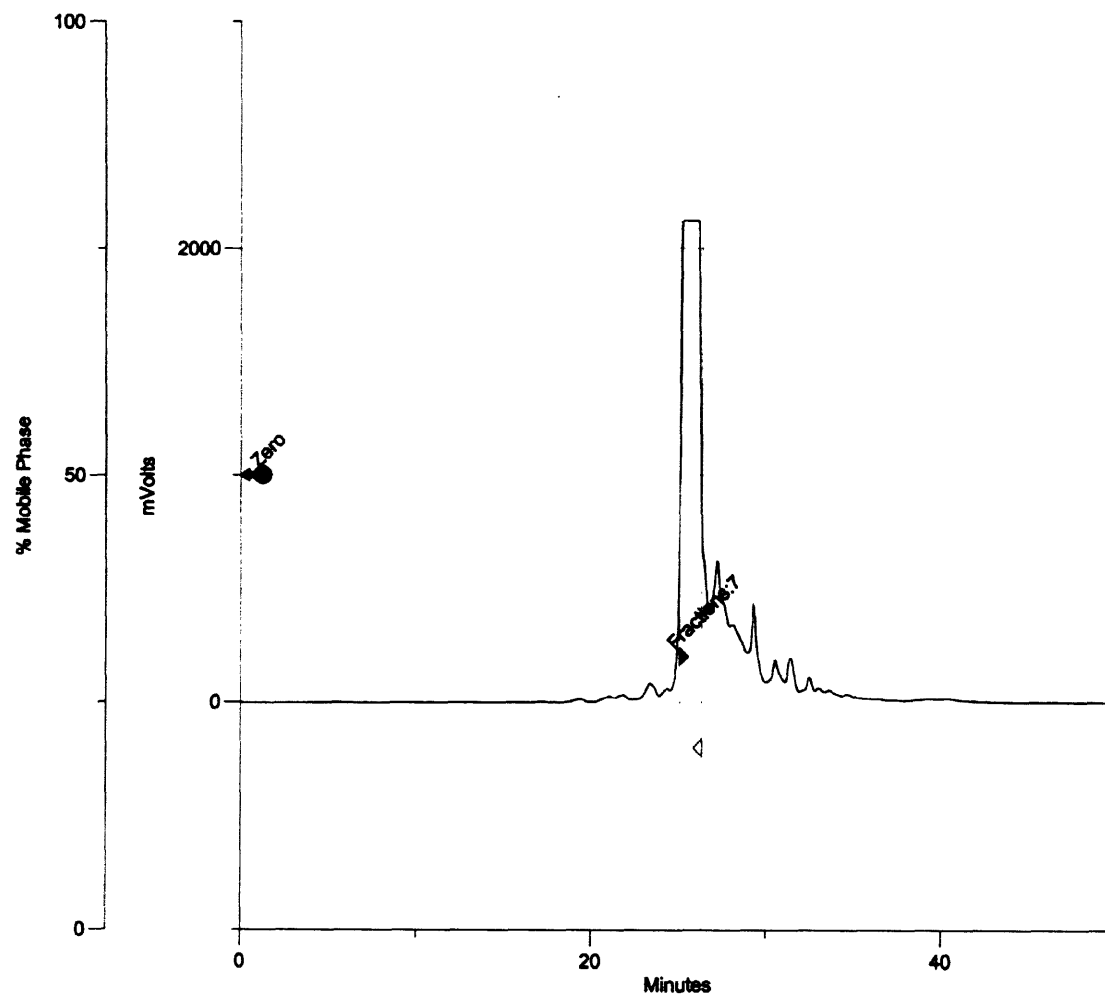
**Figure VII.2.** Size exclusion chromatogram. Typical SEC trace of NikR on Superdex 200 column (Amersham Biosciences).  $A_{280}$  is plotted versus retention volume in milliliters. NikR eluted as a tetramer at 80 mL.



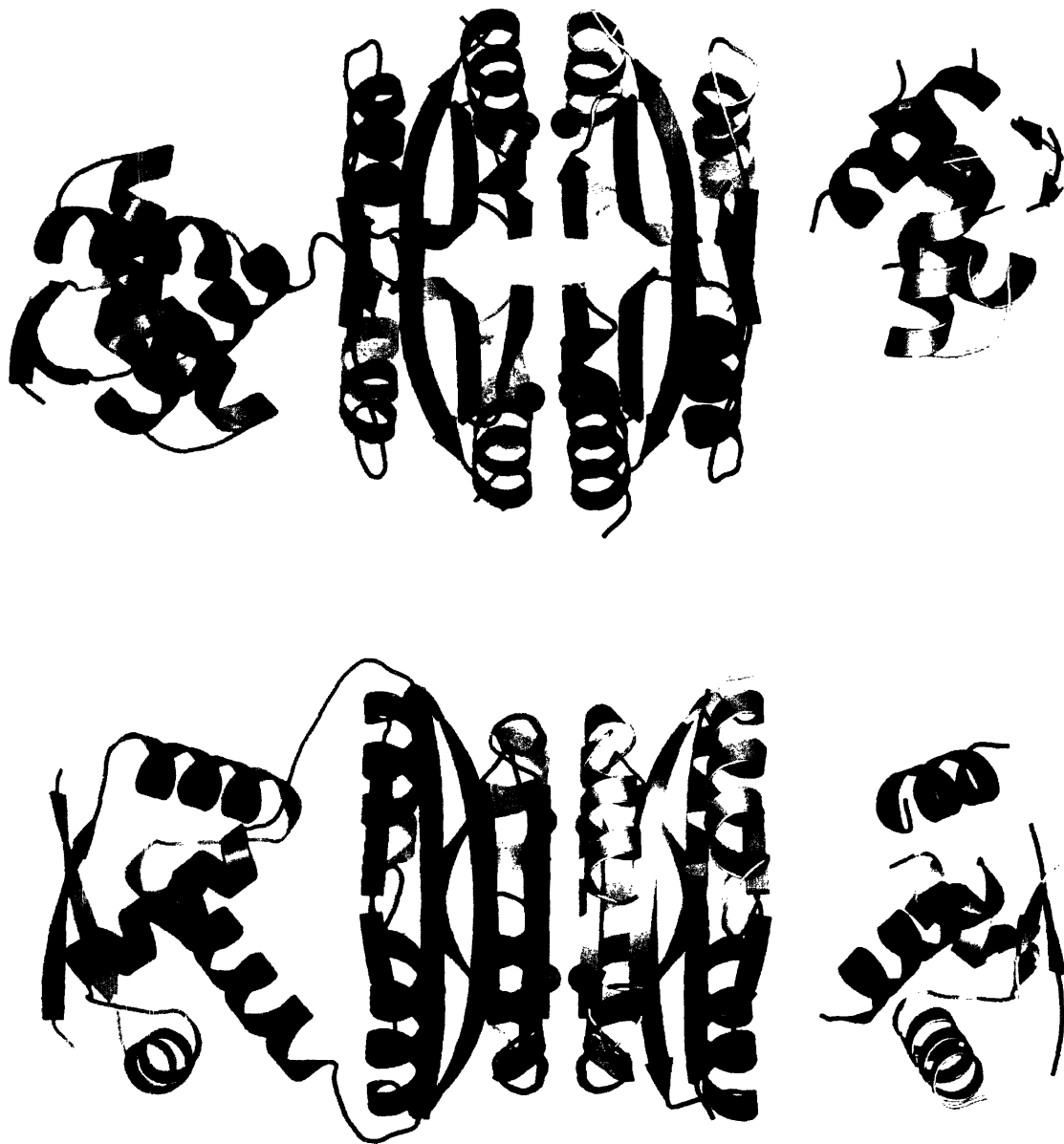
**Figure VII.3.** Ni<sup>2+</sup>-PAR calibration curve. The curve allows estimation of Ni<sup>2+</sup> concentrations in a sample by addition of 4-(2-pyridylazo)resorcinol (PAR) and measurement of A<sub>502</sub> or A<sub>528</sub>.



**Figure VII.4.** Chromatogram showing reverse-phase HPLC purification of a deoxyoligonucleotide.  $A_{260}$  is plotted versus retention time in minutes. The solvent gradient and flow rate are described in Table VII.1.



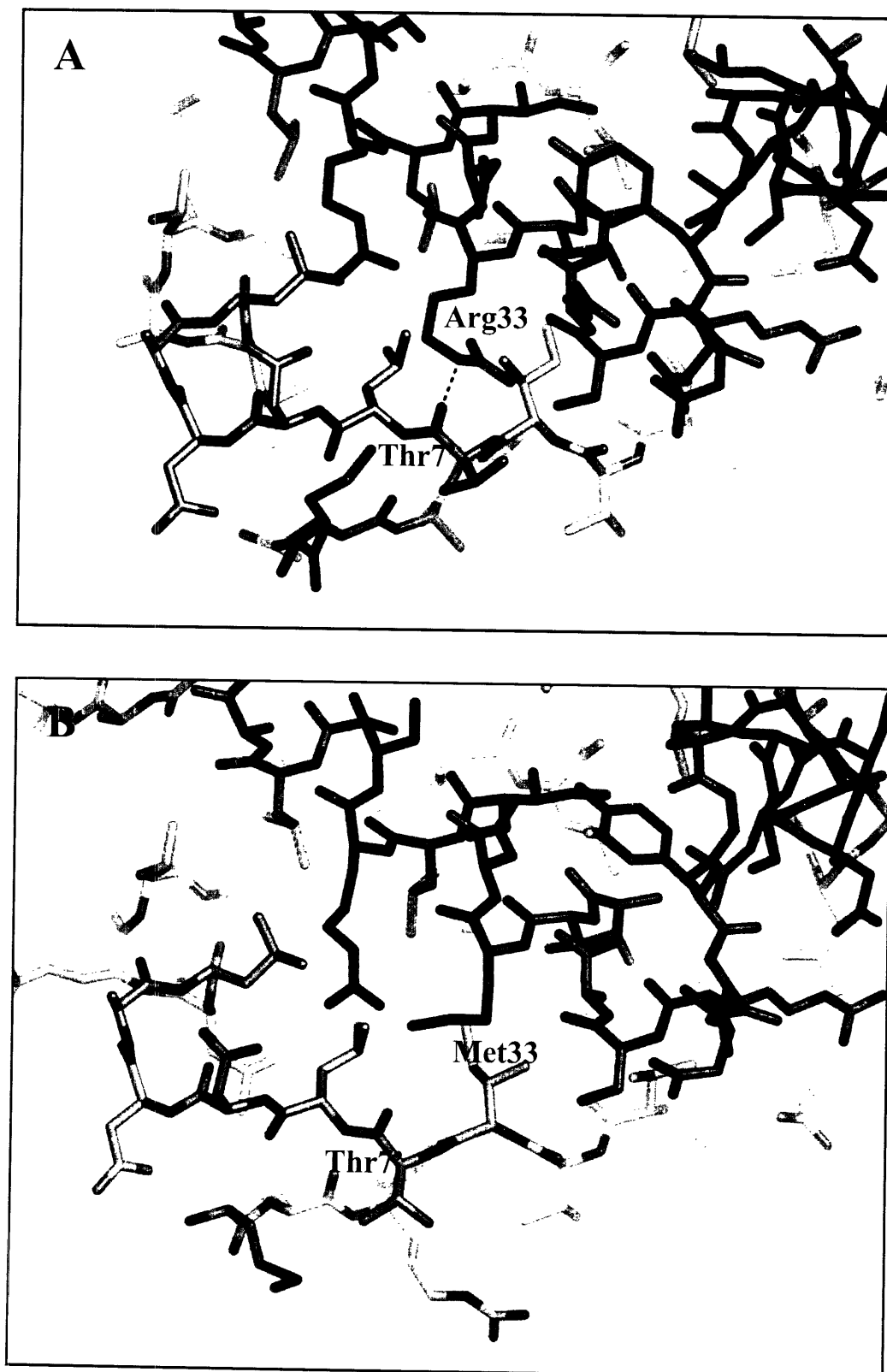
**Figure VII.5.** Ribbon representation of the  $\text{Ni}^{2+}$ -R33M NikR mutant structure, colored by protein subunit.  $\text{Ni}^{2+}$  ions bound in the high-affinity site are represented as cyan spheres. The two views are related by a  $90^\circ$  rotation.



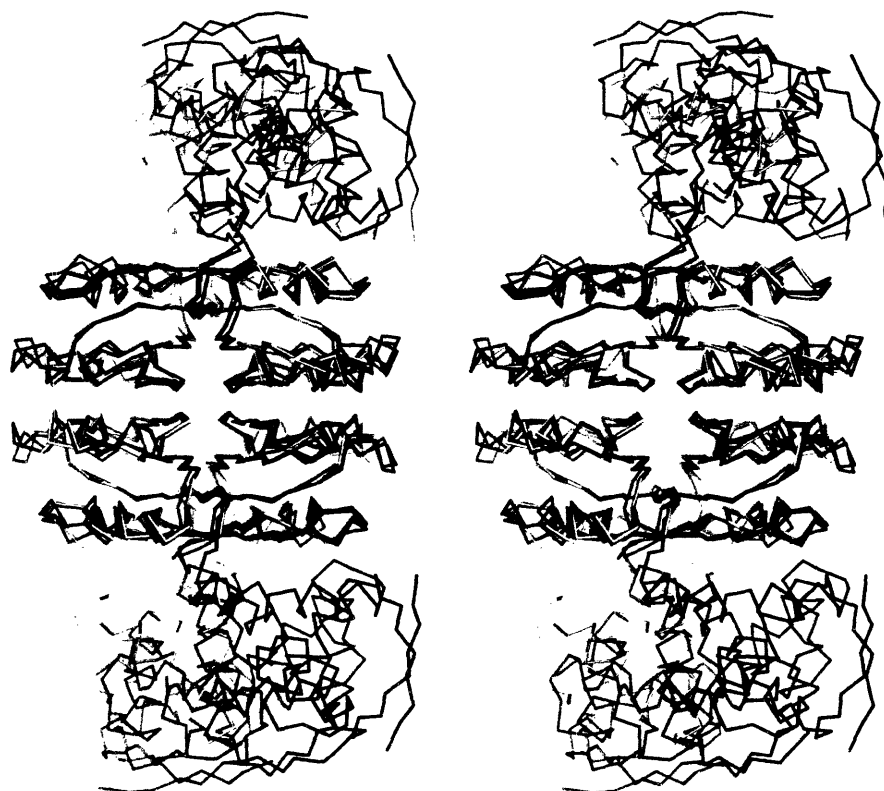
**Figure VII.6.** Superimposition of RHH DBDs. Stereo view of superimposed  $C_{\alpha}$  traces of the RHH domains of R33M (yellow), apo (red),  $Ni^{2+}$ -bound (green), and DNA-bound (blue) NikR. Note the twist of the  $\beta$ -strands in the R33M structure relative to the others.



**Figure VII.7.** Comparison of the protein environment around position 33 of Ni<sup>2+</sup>-WT NikR (A) and R33M (B). Protein is shown as sticks with carbon atoms colored by subunit. Hydrogen bond represented as a dashed black line.



**Figure VII.8.** Superposition of NikR MBDs to illustrate inter-domain flexibility. Stereo view of the superimposed  $C_{\alpha}$  traces of R33M (yellow), apo (red),  $Ni^{2+}$ -bound (green), and DNA-bound (blue) NikR. 312  $C_{\alpha}$  atoms of the MBD tetramer were used for the structural alignment.





## Eric R. Schreiter

---

### Education

- Massachusetts Institute of Technology (MIT), Ph.D. in Biological Chemistry. June, 2005.
- Truman State University, B.S. *summa cum laude* in Chemistry, May 2000.

### Research Experience

- **Graduate Research Assistant**, October 2000 – present, MIT

The laboratory of Dr. Catherine L. Drennan uses protein crystallography as a tool to investigate metalloproteins of medicinal or environmental significance. Specifically, I studied a transcription factor called NikR that regulates nickel uptake in bacteria and archaea. I have solved several crystal structures of this protein with the goal of understanding the mechanism of nickel-dependent DNA binding.

- **Undergraduate Research Assistant**, January 1998 – May 2000, TSU

In the laboratory of Dr. Richard Freeman I synthesized novel room-temperature ionic liquids and characterized these compounds using NMR, Raman, Fluorescence, and UV-Visible spectroscopies.

### Awards Received

- **Sigma Xi Grant-in-Aid of Research**, Fall 1999
- **Truman State University Undergraduate Research Fellowship**, Summer 1999
- **Honors in Chemistry**, received on diploma from TSU, Spring 2000
- **Summa Cum Laude**, received on diploma from TSU, Spring 2000

### Publications

- Schreiter, E. R.; Wang, S.; Zamble, D.; Drennan, C. L. "Crystal Structure of the NikR-operator DNA complex and the mechanism of repressor activation by metal ions" *Manuscript in preparation*.
- Schreiter, E. R.; Sintchak, M. D.; Guo, Y.; Chivers, P. T.; Sauer, R. T.; Drennan, C. L. "Crystal structure of the nickel-responsive transcription factor NikR" *Nature Structural Biology* **2003**, 10, 794.
- Drennan, C. D.; Heo, J.; Sintchak, M. D.; Schreiter, E.; Ludden, P. W. "Life on carbon monoxide: X-ray Structure of *Rhodospirillum rubrum* Ni-Fe-S carbon monoxide dehydrogenase" *PNAS* **2001**, 98, 11973.
- Sitze, M. S.; Schreiter, E. R.; Patterson, E. V.; Freeman, R. G. "Ionic Liquids Based on FeCl<sub>3</sub> and FeCl<sub>2</sub>. Raman Scattering and *ab Initio* Calculations" *Inorganic Chemistry* **2001**, 40, 2298.
- Schreiter, E. R.; Stevens, J. E.; Ortwerth, M. F.; Freeman, R. G. "A room-temperature molten salt prepared from AuCl<sub>3</sub> and 1-ethyl-3-methylimidazolium chloride" *Inorganic Chemistry* **1999**, 38, 3935.