2005

Uric acid-mediated modulation of the transcriptional regulator HucR from Deinococcus radiodurans

Steven P. Wilkinson
Louisiana State University and Agricultural and Mechanical College, swilki2@lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Recommended Citation
https://digitalcommons.lsu.edu/gradschool_dissertations/1748

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
URIC ACID-MEDIATED MODULATION OF THE TRANSCRIPTIONAL REGULATOR HUCR FROM *DEINOCOCCUS RADIODURANS*

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

The Department of Biological Sciences

by

Steven P. Wilkinson
B.S., The College of William & Mary, 1995
May 2005
DEDICATION

My honor in being granted the title Doctor of Philosophy is the consequence of my mom and dad’s lifetime of selfless hard work, perseverance, and love.
ACKNOWLEDGMENTS

I wish to thank my advisor, Dr. Anne Grove, for her unwavering guidance, patience, and support through my graduate research. I feel incredibly fortunate and privileged to have had the opportunity to learn from one whom I hold in the highest regard as a scientist.

Secondly, I wish to express my sincerest gratitude to my committee: Dr. Sue Bartlett, Dr. Vince LiCata, Dr. Grover Waldrop, and Dr. Jill Johnson for their insightful discussions and perceptive suggestions regarding my research and for their encouragement. I owe a special acknowledgment to my fellow lab mates for their friendship and enthusiasm.

I wish to especially thank my family who has always been there for me. Thanks Cindy, David, and John for being my role models and voices of wisdom. Thanks Lauren for reminding me of what is important in life. Thanks Mema and Papa for all that you taught me. Thanks Mindy, April, Sugar, Christy, Toby, Alexis, Corrie, Scottie, Ridgie, Stella, Katie, Luke, and Molly for your companionship.

And finally, I wish to thank my wife, Emily, who has been my constant partner and inspiration. Thank you so much for your understanding and belief in me.
# TABLE OF CONTENTS

DEDICATION .............................................................................................................. ii

ACKNOWLEDGMENTS ............................................................................................. iii

LIST OF FIGURES ........................................................................................................ v

ABSTRACT ................................................................................................................... vii

CHAPTER 1. INTRODUCTION ...................................................................................... 1

Helix-Turn-Helix DNA-Binding Proteins and Discovery of the Winged Helix Motif .... 1
Winged Helix DNA-Binding Motif ........................................................................... 6
DNA Recognition by Winged Helix Proteins (Role of the Wings) ......................... 7
Discovery of the mar Regulon and the Identification of MarR .............................. 10
Structural Analysis of MarR Homologs ................................................................. 11
Structural Evidence for Mechanisms of DNA-Binding by MarR Homologs .......... 14
Structural Data on Phenolic Recognition in the MarR Subfamily ....................... 18
Proposed Mechanisms of Allosteric Regulation of MarR Proteins .................... 20
Other Members of the MarR Family ........................................................................ 22
Deinococcus radiodurans ....................................................................................... 30
References ............................................................................................................... 32

CHAPTER 2. HUCR, A NOVEL URIC ACID RESPONSIVE MEMBER 
of the MARR FAMILY OF TRANSCRIPTIONAL REGULATORS FROM DEINOCOCCUS RADIODURANS ......................................................... 38

Introduction ............................................................................................................. 38
Experimental Procedures ...................................................................................... 40
Results ..................................................................................................................... 47
Discussion .............................................................................................................. 61
References ............................................................................................................. 65

CHAPTER 3. NEGATIVE COOPERATIVITY OF URIC ACID BINDING 
in the DNA-BINDING DOMAIN OF HUCR ............................................................. 69

Introduction ............................................................................................................. 69
Experimental Procedures ...................................................................................... 71
Results ..................................................................................................................... 76
Discussion .............................................................................................................. 96
References ............................................................................................................. 105

CHAPTER 4. SUMMARY AND CONCLUSIONS ......................................................... 108

DNA-Binding by HucR ......................................................................................... 108
Negative Cooperativity in Uric Acid Binding ...................................................... 110
Physiological Considerations .............................................................................. 111
References ............................................................................................................. 113

VITA ......................................................................................................................... 115
LIST OF FIGURES

Fig. 1.1 Prototypical helix-turn-helix DNA-binding motif...........................................3
Fig. 1.2 DNA-binding by a winged helix motif.................................................................9
Fig. 1.3 MarR-salicylate co-crystal structure.................................................................13
Fig. 1.4 Comparison of the predicted MexR “closed” and “open” conformations.......15
Fig. 1.5 MarR salicylate binding sites SAL-A and SAL-B...........................................19
Fig. 2.1 The genetic organization of $dr1159$ and adjacent ORFs...............................48
Fig. 2.2 Multiple sequence alignment of HucR and representative MarR family members.................................................................50
Fig. 2.3 Purified HucR.......................................................................................................51
Fig. 2.4 CD spectral analysis of HucR.............................................................................53
Fig. 2.5 HucR binding to its promoter/operator region, $hucO$.................................54
Fig. 2.6 Footprinting analysis of HucR:$hucO$ complex..............................................56
Fig. 2.7 HucR effector binding assays..........................................................................57
Fig. 2.8 Analysis of $hucR$ and $dr1160$ (uricase) gene expression............................60
Fig. 3.1 Binding-induced conformational changes in DNA and protein......................77
Fig. 3.2 Model of HucR monomer..................................................................................79
Fig. 3.3 Equilibrium sedimentation profile of HucR......................................................81
Fig. 3.4 HucR-ligand interactions, at low ligand concentrations...............................82
Fig. 3.5 HucR fluorescence spectra, normalized to the fluorescence intensity maximum.................................................................84
Fig. 3.6 Purified HucR mutants.......................................................................................87
Fig. 3.7 Affinity measurements of HucR mutants for the HucR binding sequence ($hucO$).................................................................89
Fig. 3.8 Competition assays..........................................................................................90
<table>
<thead>
<tr>
<th>Fig. 3.9</th>
<th>Fluorescence spectra of free and DNA-bound HucR-variants</th>
<th>91</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 3.10</td>
<td>DNA-binding antagonism by uric acid</td>
<td>93</td>
</tr>
<tr>
<td>Fig. 3.11</td>
<td>Fluorescence detection of uric acid binding to HucR-mutants</td>
<td>95</td>
</tr>
</tbody>
</table>
ABSTRACT

The MarR family of transcriptional regulators comprises a subset of winged helix DNA-binding proteins and includes numerous members that function in environmental surveillance of aromatic compounds. This study presents a biochemical characterization of a novel MarR homolog, HucR (hypothetical uricase regulator), from the DNA damage-resistant eubacterium, *Deinococcus radiodurans*.

Circular dichroism spectroscopy suggests that HucR has \( \sim 47\% \) \( \alpha \)-helix and \( 10\% \) \( \beta \)-strand conformation at 25°C, and undergoes a transition to a disordered state with \( T_m = 51.1 \pm 0.0^\circ C \). HucR binds as a homodimer with high sequence-specificity to a single site in its promoter region (*hucO*) with an apparent \( K_d = 0.29 \pm 0.02 \text{ nM} \). DNaseI and hydroxyl-radical footprinting indicate HucR binding site sizes of \( \sim 24 \text{ bp} \) and 21 bp, respectively. The binding site contains a pseudopalindromic sequence comprised of 8 bp inverted repeats separated by 2 bp that overlaps predicted promoter elements for *hucR* and a putative uricase (*dr1160*). Specific phenolic weak acids, notably uric acid, antagonize HucR-*hucO* complex formation. *In vivo*, uric acid increases transcript levels of *hucR* and *dr1160*, \( \sim 1.6\text{-fold} \), and stimulates uricase activity 1.5-fold.

HucR-*hucO* complex formation involves protein conformational changes and a decrease in the helical twist of the DNA duplex. Intrinsic fluorescence measurements show that uric acid induces HucR conformational changes, and its apparent \( K_d = 11.6 \pm 3.7 \text{ \mu M} \) and Hill coefficient of \( 0.7 \pm 0.1 \) suggest negative cooperativity. An amino acid substitution in the predicted HucR wing (HucR-R118A) reduces DNA-binding affinity \( \sim 5\text{-fold} (K_d = 1.60 \pm 0.14 \text{ nM}) \), whereas a substitution in the predicted recognition helix (HucR-S104A) does not significantly alter DNA-binding affinity (\( K_d = 0.23 \pm 0.03 \text{ nM} \)). Each mutation decreases complex stability on the gel, but does not affect sequence-specificity. Intrinsic fluorescence spectra suggest altered conformations...
of the HucR-variants and altered mechanisms of DNA association. The mutations at HucR
positions 118 and 104 also alter a predicted weak ligand-binding site, as indicated by minor
changes in uric acid affinities for HucR-R118A and HucR-S104A ($K_d = 9.7 \pm 3.2 \mu M$ and $7.4 \pm
0.5 \mu M$, respectively) and modest attenuations of protein-$hucO$ complex formation in response to
uric acid.
CHAPTER 1
INTRODUCTION

It has been suggested that approximately 6 – 7% of a eukaryotic genome, and 2 – 3% of a prokaryotic genome encodes DNA-binding proteins (1). However, up to 10% of the genome of a bacterial species capable of surviving exposure to a variety of environmental conditions may encode regulators of gene expression (2). There are multiple systems of classification for DNA-binding proteins, however, several major families may be identified by sequence and structural similarities in the motifs used for DNA recognition, including the helix-turn-helix (HTH), homeodomain, zinc finger, steroid receptor, leucine zipper, helix-loop-helix, and β-ribbon proteins (1, 3). NMR and crystal structure analysis of prokaryotic transcriptional regulators has revealed three recurrent DNA-binding motifs, namely the β-ribbon, HTH, and winged helix motifs (4). The winged helix DNA-binding proteins comprise a subset of the HTH proteins and the winged helix structure is strikingly the most abundant DNA-binding motif in the transcription factors of bacteria and archaea (2). Based upon the wide distribution of winged helix DNA-binding proteins amongst bacterial and archaeal genomes, this structure has been suggested to be the earliest HTH motif to have evolved (2). The MarR (multiple antibiotic resistance regulator) family of prokaryotic transcriptional regulators is one of nine families of winged helix transcriptional regulators suggested to have existed in the common ancestor that preceded the evolutionary split of archaea and bacteria ∼3.5 billion years ago (1, 2, 5).

Helix-Turn-Helix DNA-Binding Proteins and Discovery of the Winged Helix Motif

Structural determination of the HTH motif was first accomplished for the λ Cro and Escherichia coli CAP proteins (6-8). The structures of numerous HTH proteins, free and complexed with DNA, have since been solved at high resolution, revealing general features of
this DNA-binding motif. As the name suggests, the HTH motif consists of two α-helices separated by a turn, but the HTH motif by itself is not a stable, functional domain, but rather exists in the context of a larger DNA-binding domain. The prototypical HTH motif is roughly 20 – 21 amino acids in length, with the first helix averaging seven residues and the second helix (often called the “DNA-recognition helix”), averaging nine residues. The two α-helices are generally separated by a 3 – 4 amino acid β-turn (Fig. 1.1). There is a strong conservation of residues at certain positions in the HTH motif (9). In the first helix of the motif, a hydrophobic residue generally occupies the fourth position and the fifth position is often a residue with a small side chain, such as glycine or alanine, due to steric considerations with the recognition helix. In the turn region, the first and third amino acids are usually hydrophobic, with a glycine at the second position. The fourth and seventh residues in the DNA recognition helix are often hydrophobic. As prolines disrupt α-helical secondary structure, prolines are invariably absent from the interiors of both helices in the HTH motif (9, 11). The result of the conserved, hydrophobic residues in the helices and turn, and the minimal variability in the length of the turn, is that the axes of the two α-helices of the HTH motif are oriented at an angle of ~120°, with little variability observed between canonical HTH proteins.

Side chains from the DNA recognition helix have been shown to mediate hydrogen bonds with both the nucleobases of the DNA duplex and with the phosphodiester backbone. In addition, the DNA recognition helix has been shown to be stabilized on the double helix by hydrophobic contacts with thymine methyl groups in the DNA major groove (3, 9). Side chains outside the DNA recognition helix often are involved in stabilizing the protein-DNA complex. For example, the first helix of the HTH motif in λ repressor mediates a hydrogen bond with the DNA sugar-phosphate backbone, and the first helix of a HTH motif in the prd protein from Drosophila
Fig. 1.1 **Prototypical helix-turn-helix DNA-binding motif.** The HTH motif of the trp repressor from *E. coli* bound to its sequence-specific DNA-binding site (blue). The first helix of the motif is colored green (with the amino-terminal end designated by “N”) and the turn colored yellow. The DNA-recognition helix (red) is inserted into the major groove (10).
makes extensive contacts with the phosphate backbone (12, 13). Additional contacts also exist between the \( \alpha \)-helices themselves, stabilizing their relative orientations and their respective contacts with the DNA duplex. Furthermore, polypeptide backbone-, cation-, and water-mediated (as seen for the trp repressor) contacts with the DNA are sometimes observed for HTH proteins (1, 3, 10). Residues outside the HTH motif also contact the DNA and help anchor the HTH motif in a specific orientation relative to the DNA duplex. For example, an N-terminal arm of the \( \lambda \) repressor wraps around the DNA double helix and mediates contacts in the major groove of the consensus binding half-site (12).

The DNA recognition helix usually tracks the DNA major groove such that the N-terminal protein side chains are in closest contact with the nucleobases of the double helix. The orientation of the recognition helical axis relative to the direction of the major groove varies, such that the angle between the \( \alpha \)-helical axis and the axis of the DNA duplex varies by at least 15° in each direction. The orientation of the recognition helix, determined by other structural components in the protein, in addition to the actual amino acid sequence of the \( \alpha \)-helix, contributes to DNA sequence recognition. The N-terminus of the first \( \alpha \)-helix of the HTH motif often contacts the phosphodiester backbone, but lies above the DNA recognition helix and the DNA duplex (3, 9).

Structural features of a DNA-binding protein form only part of the picture in analyzing mechanisms of sequence specific DNA-binding: DNA sequence, conformation, and dynamics are equally important. The patterns of hydrogen bond donors and acceptors presented in the major and minor grooves are sequence specific, with each of the four base pairs forming a distinctive pattern in the major groove, with considerably less variability in the minor groove (14). In addition, the conformation of the sugar-phosphate backbone is sequence dependent. As
numerous DNA-binding proteins induce bends in the DNA duplex, flexibility of the DNA is a critical determinant of binding affinities. In this regard, it has been shown that A-T steps generally favor duplex bending toward the minor groove, whereas G-C steps favor bending toward the major groove (9). Also, inter-base pair hydrogen bonding stabilizes high propeller twist for AT base pairs when consecutive adenines occur on the same strand, with the result that AT tracts have reduced conformational flexibility (15). Moreover, the average minor groove width of 6 Å is reduced in sequences rich in AT base pairs (9).

Comparisons of HTH proteins from eukaryotic organisms with those of prokaryotes reveal very little sequence similarity, despite the conservation of the HTH structure in their DNA-binding domains (13, 16, 17). Eukaryotic proteins with HTH DNA-binding motifs are usually classified by the structural domain containing the HTH motif (e.g. POU and homeodomain families). Prokaryotic HTH proteins almost always bind as homodimers to a palindromic, or pseudopalindromic, DNA recognition sequence, such that the DNA recognition helices from each half of the dimer bind the symmetric DNA half-sites within the major groove of the duplex. There are exceptions to this rule: for example, proteins from the AraC family, including MarA and Rob, bind as monomers via two HTH motifs to sequence specific DNA sites (18-20).

Interest in winged helix proteins has increased steadily since the co-crystal structure was provided of the DNA-binding domain of HNF-3γ complexed with its target DNA sequence (21). From this structure, it was determined that members of the eukaryotic HNF-3/fork head family of proteins mediate contacts with DNA through a novel α/β DNA-binding domain. The N-terminal half of this domain formed a HTH-like structure, with the α-helices extended in length, relative to those in the canonical HTH proteins, with the first and recognition helices being 10
and 14 amino acids in length, respectively. In addition, the turn in the HTH region of HNF-3γ was 8 residues, at least twice the length seen in canonical HTH domains, creating an angle between the two helices of 140° (22). The DNA recognition helix was positioned in the major groove of the DNA duplex, inducing a narrowing of this groove, and an overall bend of ∼13° (21). The average helical twist of the DNA was increased slightly, to 35°. The HNF-3γ DNA-binding fold revealed an unusual, three-stranded, antiparallel β-sheet from which two loops, or “wings” extended. Interestingly, residues from each of these wings were shown to contact DNA, prompting the designation of this DNA recognition motif as the “winged helix” motif.

**Winged Helix DNA-Binding Motif**

Numerous DNA-binding proteins with a winged helix DNA recognition motif have since been characterized in eukarya, prokarya, archaea, and viruses. Particular interest, in regards to elucidating archaeal transcriptional regulation, has stemmed from analyses of the known archaeal genomes which suggest that most of the predicted HTH proteins, and most of the putative transcriptional regulators, in these organisms belong to the winged helix subfamily of DNA-binding proteins (2, 16, 23). However, this DNA-binding motif is not restricted to transcriptional activators and repressors; histone H5 also binds DNA via a winged helix motif (22).

The winged helix DNA binding motif, also referred to as the winged helix-turn-helix (wHTH) motif, is defined topologically by secondary structure elements arranged in the following order: H1-S1-H2-H3-S2-W1-S3-W2, where “H” represents α-helix, “S” represents β-strand, and “W” represents a loop (5). The sequence spanning α-helices H2 through H3 constitutes the general HTH motif, with H3 being the DNA recognition helix. The two eponymous “wing” structures are actually formed by β-strands and loops; “wing 1” is a β-hairpin motif, comprised of the S2-W1-S3 secondary structure elements and “wing 2” is formed by the
S3-W2 elements. The three $\beta$-strands interact to form an antiparallel $\beta$-sheet, however, in some winged helix proteins S1 is represented by a single, hydrophobic residue (21, 24-26). Whereas wing 1 is invariably present in winged helix proteins, some members of this family do not contain wing 2, as observed in the crystal structures of the winged helix DNA-binding domains of E2F4, DP2, histone H5 and MarR (22, 26, 27). As observed in HNF-3$\lambda$, the length of the turn in the winged helix motif can vary significantly from the 3-4 residues found in canonical HTH proteins (21). Consequently, greater variation in the angle between the two helices of the HTH motif is observed amongst the proteins in the winged helix family, than in the canonical HTH proteins. For example, the turn in DP2 is approximately 10 residues in length, whereas the turn in BirA is approximately 3 amino acids, allowing angles of 100° and 150°, respectively (5, 25, 27).

**DNA Recognition by Winged Helix Proteins (Role of the Wings)**

Structural analysis of winged helix proteins complexed with their target DNA sites reveals that this family is similar to HTH proteins, in that proteins similar in the tertiary structures of their DNA-binding motifs often differ in the manner in which they contact the DNA (28). The DNA-recognition helix (H3) almost invariably contributes most of the contacts determining sequence-specificity. However, the role of the wing(s) in contributing to DNA-binding affinity and specificity seems to vary widely, with the wings of some proteins being critical for site-specific DNA-binding by mediating numerous base-specific and sugar-phosphate backbone contacts, while in other proteins, the wings mediate few contacts. Moreover, the relative contributions from each wing (when both are present) also varies. The co-crystal structure of HNF-3$\lambda$ with its target DNA-site revealed that of the 17 amino acid-mediated DNA contacts, 6 were contributed from the wing 2 structure, including three hydrogen bonds with
backbone phosphates, one hydrophobic interaction with a backbone ribose, one direct hydrogen bond with a base in the minor groove, and one water-mediated contact with a base in the major groove. In contrast, wing 1 contributed one hydrogen bond to the phosphate backbone from each of the S3 and W1 elements (21). Solution NMR analyses of the eukaryotic protein genesis indicate that wing 1 makes contacts with the DNA minor groove but that this wing is conformationally flexible, even in the DNA-bound state. Wing 2 appears to become less flexible upon complex formation, suggesting its importance in stabilizing the interaction. Evidence indicates that it is unlikely to contribute to sequence specificity (29-31). Conversely, wing 1 appears to be more important than wing 2 in stabilizing the BmrR-DNA complex; wing 1 from BmrR mediates four hydrogen bond and hydrophobic contacts with the sugar-phosphate backbone at its target DNA sequence, and forms a hydrogen bond and a hydrophobic interaction with two bases in the minor groove (Fig. 1.2) (32). NMR structural analysis of the bacteriophage protein MuR in complex with its cognate DNA site reveals that the β-hairpin wing 1 undergoes a transition from a disordered state to a defined conformation upon DNA-binding, and that side chains from the wing hydrogen bond with bases in the minor groove, thus immobilizing the wing (33). Similarly, wing 1 of FadR mediates sequence-specific contact between a histidine with two bases in the minor groove; substitution of this histidine with a glycine completely disrupts DNA binding (34, 35). The wing-mediated DNA contacts in FadR allows an unusual mode of binding in which only the N-terminal end of each recognition helix of the homodimer contacts the DNA, allowing both recognition helices to occupy the same major groove (34). However, the wing structures do not always appear to be important in stabilizing the protein-DNA complex. The co-crystal structure of the E2F4-DP2 heterodimer with its cognate DNA-binding site revealed that the single wing from E2F4 mediates two contacts with backbone phosphates and the single wing
Fig. 1.2  **DNA-binding by a winged helix motif.** The DNA recognition helix (green) of BmrR (from the MerR family of transcriptional regulators) is inserted into the DNA major groove (blue). “N” identifies the amino-terminal end of the helix. The β-hairpin, wing 1, motif is comprised of secondary structure elements topologically arranged (in the N- to C- terminal direction): β-strand (red), loop (yellow), β-strand (orange). Only a single strand of the cognate BmrR DNA-binding site is shown.
from DP2 mediates only one hydrogen bond with a backbone phosphate. For this protein-DNA complex it is apparent that the wings make only slight contributions to the complex stability (27).

The co-crystal structure of the winged helix protein RFX1 with its site-specific DNA-binding site revealed a strikingly different mode of interacting with the DNA duplex (36). The β-hairpin, wing 1, structure makes extensive base-specific hydrogen bonds with the DNA major groove in one half-site of the recognition sequence, narrowing the major groove by 1 Å. Surprisingly, the α-helix corresponding to the DNA recognition helix mediates only a single hydrogen bond between a lysine and a cytosine in the minor groove on the opposite face of the same half-site, causing the minor groove to widen by more than 3 Å.

**Discovery of the mar Regulon and the Identification of MarR**

The identification of the MarR (multiple antibiotic resistance regulator) family of transcriptional regulators began with the identification of a chromosomally encoded mechanism of multidrug resistance in *E. coli* K-12 (37, 38). Genetic selections identified mutants that exhibited resistance to a broad range of antibiotics including tetracycline, chloramphenicol, β-lactams, puromycin, nalidixic acid, penicillins, fluoroquinolones, and organic solvents (37-40). The *mar* (multiple antibiotic resistance) phenotype was shown to be conferred by the *marRAB* operon, specifically by the expression of *marA*, which encodes a transcriptional activator belonging to the AraC family (20, 41-43). MarA is an activator of the *marRAB* operon and induces the expression of a number of genes responsible for the *mar* phenotype, including the expression of the AcrAB-ToIC multidrug efflux system and the gene, *micF*, that downregulates the synthesis of the porin OmpF (39, 44-47). *In vivo* upregulation of *marRAB* expression and the *mar* phenotype was shown to be inducible by a range of antibiotics and anionic lipophilic compounds, including 2, 4-dinitrophenol, menadione, plumbagin, and salicylic acid (48, 49).
The product of the first gene of this locus, MarR (144 amino acids), was shown to be a transcriptional repressor of the \textit{marRAB} operon (48, 50). MarR binds to two separate 21 bp sites in the \textit{marRAB} promoter/operator region (\textit{marO}) (51). An apparent K\textsubscript{d} of \(~1\) nM was calculated for MarR binding to \textit{marO}. DNaseI footprinting indicated that MarR binding site I overlaps the predicted –35 and –10 promoter elements and site II overlaps the putative ribosome binding site and ends immediately before the first codon of \textit{marR} (51). The size of each footprint, and the fact that each binding site is palindromic, with 5 bp half-sites separated by 2 bp, is consistent with MarR binding as a homodimer at each site and is supported by size-exclusion chromatography evidence that uncomplexed MarR exists as a dimer, or higher order oligomers, in solution (51). Interestingly, a number of phenolic compounds that have been shown to increase \textit{marRAB} expression \textit{in vivo} also antagonize MarR-\textit{marO} complex formation \textit{in vitro}, including 2, 4-dinitrophenol, menadione, plumbagin, and salicylic acid (52). MarR was measured to bind salicylic acid with an apparent K\textsubscript{d} of 0.5 mM by equilibrium dialysis (51). These results, \textit{in toto}, revealed a system of intrinsic multidrug resistance in \textit{E. coli} that is under the control of a novel transcriptional repressor, MarR, that responds to cytoplasmic phenolic compounds. Functional \textit{marRAB} operons have since been identified in \textit{Salmonella typhimurium} and \textit{Enterobacter aerogenes} (53, 54).

\textbf{Structural Analysis of MarR Homologs}

A search of the sequenced eukaryotic and prokaryotic genomes reveals numerous predicted MarR homologs throughout the bacterial and archaeal domains. However, to date, structural data has only been provided for two members of the MarR subfamily of winged helix DNA-binding proteins.
The co-crystal structure of a MarR-salicylate complex was determined at 2.3 Å resolution and reveals the protein to exist as a dimer with a pyramidal shape and overall dimensions of 50 x 55 x 45 Å³ (Fig. 1.3) (26). Each monomer consists of 6 α-helices and two β-strands. The N-terminal region encompassing α-helix 1, and the C-terminal region encompassing helices 5 and 6, interdigitate with the corresponding regions of the other subunit to form a dimerization domain with a buried surface area of 3570 Å². The stabilization of MarR as a homodimer is predominantly mediated by hydrophobic contacts involving 10 residues from each subunit. Several inter-subunit hydrogen bonds (H-bonds) in this domain contribute to dimer stability: the Nζ of a lysine in helix 1 H-bonds with the main chain carbonyl of the C-terminal residue, and the Nζ of a lysine in helix 6 H-bonds with the main chain carbonyl in the N-terminal coil region.

Helices 1 and 5 of each subunit connect the dimerization domain to a globular DNA-binding domain, such that the two DNA-binding “lobes” of the dimer are juxtaposed and related by a two-fold rotational symmetry. The topological arrangement of the secondary structure elements in the DNA-binding domain (H1-S1-H2-H3-S2-W1-S3) indicates that MarR binds via a winged helix motif. Each DNA-binding domain is stabilized by hydrophobic interactions involving 14 residues from each of the secondary structural elements, which serve to make the domain compact. The residues spanning MarR helices 3 and 4 (corresponding to H2-H3 above) constitute the HTH motif. The β-hairpin (S2-W1-S3) following the HTH motif forms the eponymous “wing” structure. The β-strands (S1, S2, S3) in each MarR subunit form an antiparallel β-sheet, with S1 being comprised of a single isoleucine residue, similar to the structures of OmpR and BirA (24, 25). The wing extending from S3 in some winged helix proteins is absent from the DNA-binding domain of MarR, resembling the structures of E2F4, DP2, and histone H5 (22, 27). The surface potential of each DNA-binding domain is highly
Fig. 1.3. **MarR-salicylate co-crystal structure.** MarR is shown as a monomer with the N- and C- termini labeled. MarR \( \alpha \)-helices are identified as \( \alpha_1-6 \) and \( \beta \)-strands as \( \beta_1 \) and \( \beta_2 \). The wing is comprised of \( \beta_1 \), \( \beta_2 \), and the intervening loop (yellow). The DNA-recognition helix (green) is flanked by salicylate binding pockets SAL-A and SAL-B. Salicylate molecules are depicted in red.
electropositive, and a 6 Å wide cleft separates the lobes from each subunit. In this salicylate-bound structure, the DNA-binding lobes of MarR interact through two salt bridges: D67 from the turn of the HTH contacts R73’ in the putative recognition helix of the other subunit (and the reciprocal interaction).

MexR is a MarR homolog from *Pseudomonas aeruginosa* that serves as a repressor of the MexAB-OprM multidrug efflux operon (55). The crystal structure of the MexR dimer in the absence of effector compounds was solved at a resolution of 2.1 Å and revealed an overall, pyramidal structure very similar to MarR (Fig. 1.4) (56). Four copies of the MexR dimer were present in the crystallographic asymmetric unit, providing several conformational views of the dimer. The MexR monomer is 147 amino acids in length and is comprised of 6 α-helices and 3 β-strands. The N- and C-terminal regions encompassing α-helices 1, 5, and 6 from each subunit intertwine with the reciprocal regions in the other subunit to create a dimerization interface with a total buried surface area ranging from 4360 – 4930 Å³. Stabilization of the dimer is provided almost entirely by hydrophobic contacts. Helices 1 and 5 of each subunit connect the dimerization domain to a compact, globular DNA-binding domain (residues 36 – 97). The topological arrangement of the secondary structure elements of the DNA-binding domain indicates a winged helix fold, with the sequence spanning helices 3 and 4 forming the HTH. β-strands 2 and 3 and the intervening sequence form the wing. As for MarR, MexR lacks a wing 2 structure. The two salt bridges that connect the DNA-binding lobes of the MarR dimer are absent in the MexR structure.

**Structural Evidence for Mechanisms of DNA-Binding by MarR Homologs**

The biochemical and genetic data, to date, is insufficient to explain the DNA-binding mechanism of MarR. A MarR-*marO* co-crystal structure, and biophysical analysis of MarR are
Fig. 1.4 **Comparison of the predicted MexR “closed” and “open” conformations.** The individual subunits of the MexR dimer are colored red and blue (dimer AB) or yellow and green (dimer CD). MexR dimer AB is in the “closed” conformation, predicted to resemble its ligand-bound state. MexR dimer CD is in the “open” conformation, suggested to resemble its DNA-binding conformation. Indicated below each structure is the distance separating the central residues in the DNA recognition helices from each half of the dimer.
critically needed to solve this mystery. The structure of salicylate-bound MarR does not allow modeling onto B-DNA with the known MarR binding sequences. However, this finding is consistent with evidence that salicylate is a negative effector of MarR-\textit{marO} complex formation. Nevertheless, the combined data from the MarR-salicylate co-crystal structure, MarR footprinting experiments, and MarR mutational studies, are suggestive of possible modes of DNA-recognition. The identification of 5 bp inverted repeats separated by 2 bp, within each 21 bp MarR binding site, would place the half-site sequences on opposite faces of the DNA double helix (51). As prokaryotic transcriptional regulators almost invariably bind DNA as homodimers to palindromic sequences, it is reasonable to suspect that MarR does so as well. Mutational studies of MarR have shown that both the wing and the DNA recognition helix are critical for MarR-DNA complex formation, so it is highly probable that the winged helix motif of each MarR subunit binds at each half-site (11). The DNA-binding lobes of the salicylate-bound MarR structure would have to undergo significant conformational changes to accommodate binding to DNA half-sites on opposite faces of the double helix. Such conformational shifts of the globular DNA-binding domains would require flexibility in the \(\alpha\)-helices that connect the lobes to the dimerization domain. That MarR possesses such intrinsic flexibility is supported by the crystal structure, which indicates poorly ordered loop regions that might allow conformational shifts in the \(\alpha\)-helices that form the dimerization domain (26). Also, the van der Waals interactions that stabilize the dimer do not require specific geometries of the residues involved, suggesting that flexibility would be allowed at the dimer interface. In addition, \textit{in vivo} selections of MarR mutants with enhanced DNA-binding activity indicated that residues in the N- and C-terminal regions of MarR (in addition to the wing) are important in defining DNA-binding affinity (57). \textit{In vivo} selection of mutants with reduced DNA-binding activities primarily identified residues in
the recognition helix and wing, but also indicated that mutations to residues in the C-terminal region reduced DNA-binding (11). Furthermore, MarR crystals grown in the absence of ligand were determined to be highly disordered, consistent with MarR being intrinsically flexible (26). Protein-induced conformational changes in the DNA-binding site could also accommodate binding of a MarR dimer. For example, overwinding or underwinding of the DNA duplex, to increase or decrease the helical twist, respectively, would serve to position the major groove of each half-site in an orientation that would require a less drastic conformational change in the MarR dimer. Biophysical experiments to ascertain the conformational flexibility of MarR and its cognate DNA-binding sequence would be crucial in testing these possibilities.

The crystal structure of MexR in the absence of a ligand revealed a dimer in a conformation that could be docked to a linear, B-form representation of its known DNA-binding sequence with a reasonably good fit (56). The sequence-specific MexR binding site contains 5 bp half-sites separated by 5 bp, so that the center of each half-site would likely be positioned on the same face of the DNA duplex (58). The MexR dimer (denoted “CD”) modeled onto the DNA-binding site was determined to be in an “open” conformation, such that the globular DNA-binding lobes were separated by a distance optimal for the insertion of the recognition helices into consecutive major grooves, and the wings were positioned to mediate contacts with the sugar-phosphate backbone and minor groove (56). The MexR dimer CD structure reveals the centers of the two recognition helices to be separated by 29.2 Å, close to the 34 Å distance that separates the center of each half-site in linear B-form DNA. This “open” conformation is likely maintained by electrostatic repulsions between positively charged residues lining the crevice between the DNA-binding domains. The model suggests that bending of the DNA duplex, or an increase in the helical twist, might occur upon MexR binding to accommodate a tighter fit.
Additionally, the DNA-binding lobes of the MexR dimer might undergo conformational shifts, a possibility suggested by the intrinsic flexibility of the MexR dimer (see below).

**Structural Data on Phenolic Recognition in the MarR Subfamily**

The MarR-salicylate co-crystal structure revealed two ligand binding sites per subunit, labeled SAL-A and SAL-B (Fig. 1.5) (26). Interestingly, the bound salicylates flanked the proposed DNA-recognition helix on either side. Ligand binding site SAL-A is packed in the interior of the globular DNA-binding domain, and is formed by residues from both helices of the HTH motif and from the wing. The side chain hydroxyl of T72 from the recognition helix forms an H-bond with the salicylate hydroxyl group, the guanidinium group of R86 H-bonds with the salicylate carboxylate, and the aliphatic pyrrolidone ring of P57 is positioned within 3.5 Å over the hydrophobic ring of salicylate. Site SAL-B is exposed to the surrounding solvent and most contacts with salicylate are mediated by residues from the recognition helix. The backbone carbonyl of A70 H-bonds to the salicylate hydroxyl, the guanidinium group of R77 H-bonds with the salicylate carboxylate, and the hydrophobic ring of salicylate is within 3.5 Å of the hydrophobic side chain of M74. Q42 from helix 2 may also H-bond with the salicylate carboxylate.

The H-bond forming residues in site SAL-A are strictly conserved in MexR, and the MarR proline mediating hydrophobic contact with the salicylate ring is replaced with a leucine in MexR. Site SAL-B, however, is not conserved in MexR. The natural ligands of MexR are unknown and the structure was determined in the absence of any potential effectors (56). However, the C-terminal polypeptide from an adjacent dimer was inserted in the cleft between the DNA-binding domains of MexR dimer “AB”, resulting in a dimer significantly different in conformation than the dimer CD that was modeled onto the MexR DNA-binding site (see
Fig. 1.5 **MarR salicylate binding sites SAL-A and SAL-B.** Salicylates are colored red. Oxygen, nitrogen, and sulfur atoms from MarR side-chains are colored red, blue, and yellow, respectively. Predicted hydrogen bonds are indicated by dashes with approximate distances.
below). The polypeptide was stabilized between the DNA-binding lobes by electrostatic interactions between the positively charged residues of the dimer lining the crevice and negatively charged glutamate and aspartate residues from the inserted polypeptide. Several van der Waals contacts between proline, leucine, and isoleucine residues in the polypeptide with hydrophobic residues lining the crevice of the dimer also stabilized the interaction.

**Proposed Mechanisms of Allosteric Regulation of MarR Proteins**

As discussed above, the salicylate-bound structure of MarR is in a conformation that is unlikely to bind its sequence specific DNA-binding site (26). Such binding would require significant conformational flexibility in the protein and/or the DNA duplex. The location of the two salicylate binding sites on either side of the proposed DNA recognition helix is suggestive of possible mechanisms by which phenolic ligands antagonize MarR-DNA interaction. As both ligand-binding sites SAL-A and SAL-B are composed of residues from the DNA-recognition helix and wing motif, and as both of these regions of the winged helix fold have been shown to be critical for DNA-binding, it is clear that the ligand- and DNA-binding sites are not separate in MarR. It is plausible that ligand binding at one, or both, sites coordinates residues required for direct, or water-mediated, contacts with the cognate DNA-binding site, thus preventing sequence-specific MarR-\textit{marO} complex formation (11). Alternatively, or in combination with the effect posited above, ligand binding might stabilize a MarR dimer conformation that cannot accommodate the insertion of the recognition helices into the major grooves at the binding sequence half-sites. The suggested intrinsic flexibility of MarR at the dimerization interface is consistent with a ligand-mediated conformational shift in the relative positions of the DNA-binding lobes, from a DNA-binding state, to the state observed in the crystal structure.
The four dimers present in the crystallographic asymmetric unit of MexR provided a fortuitous glimpse of a possible mechanism of ligand-mediated allosteric control of MexR-DNA binding (56). A comparison of the “open” MexR dimer CD conformation, that was readily modeled onto its linear B-form DNA binding site, with the MexR dimer AB-polypeptide conformation, revealed the latter to be incompatible with binding to its recognition sequence. Notably, the distance separating the centers of the recognition helices had been reduced to 22.6 Å in the MexR dimer AB, compared to the distance of 29.2 Å observed in the “open” conformation (Fig. 1.4). Comparisons of all four dimer representations revealed that the basis for this conformational shift resided in the intrinsic flexibility of the dimerization domain. Whereas the winged helix DNA-binding domains appeared to shift in orientation as a rigid body, flexible loop regions allow for significant conformational flexibility in α-helices 1, 5, and 6. The helix orientations of α-helices 1, 5, and 6 vary by 17°, 6.8°, and 12.1°, respectively, with their midpoint positions deviating by 4.9 Å, 1.7 Å, and 8.2 Å. This flexibility of the helices that comprise the dimerization domain is consistent with variable geometries being allowed for the van der Waals contacts that stabilize the dimer. As α-helices 1 and 5 connect the dimerization domain to the DNA-binding domain, their flexibility results in concomitant shifts in the DNA-binding lobes. The “closed” conformation observed for MexR dimer AB suggests that ligands may disrupt MexR-DNA complex formation by neutralizing the electrostatic repulsions that otherwise maintain the dimer in an “open”, DNA-binding conformation. Additional hydrophobic contacts between the ligand and residues lining the crevice of the MexR dimer contribute to bring the DNA-binding lobes into closer proximity. As all known ligands of MarR homologs are anionic lipophilic compounds, it is tempting to speculate that members of this family share a similar mechanism of allosteric control.
**Other Members of the MarR Family**

Since the discovery of MarR, a number of MarR homologs have been predicted from the genomes of Gram-positive and Gram-negative bacteria, mycobacteria, and archaea. However, only a small subset of the potential MarR homologs have been characterized biochemically or genetically.

All members of the MarR family possess a winged helix DNA-binding motif. All characterized homologs exist as dimers in both the uncomplexed and DNA-bound states and this may be a defining characteristic of this family (lower proportions of higher order oligomers have been observed for some uncomplexed homologs). Consistent with the observed DNA-binding stoichiometries, MarR homologs invariably bind to palindromic or pseudopalindromic DNA sequences that presumably reflect the two-fold rotational symmetry of the protein dimer. The gene encoding each MarR homolog is generally part of a gene cluster containing the gene(s) under its regulation (with the possible exceptions of FarR and PecS). In some cases, the MarR homolog is encoded in its regulated operon. A large proportion of the characterized family members are adjacent to the divergently transcribed gene(s) they regulate, such that the MarR homolog binding site(s) reside in the intergenic region containing the associated, divergent promoters. More than half of the characterized MarR homologs have been shown to be autoregulatory. Most members of this family serve as repressors of gene transcription, but several activators have been identified. Response to environmental phenolic ligands has been demonstrated for a number of MarR proteins. Specifically, ligand-responsive MarR proteins almost invariably respond to anionic lipophilic compounds in their capacity to bind their cognate DNA sequences. The physiological roles of MarR proteins can be classified into three general categories, with some proteins serving multiple regulatory roles: 1) regulation of response to
environmental stress, 2) regulation of aromatic catabolic pathways, and 3) regulation of virulence factors.

**MarR Regulators of Stress Response**

The MarR homolog, EmrR from *E. coli*, was first demonstrated to be repressor of microcin B and C production and later shown to be the encoded by the first gene of the *emrRAB* operon, which encodes a multidrug resistance pump (59, 60). Analyses *in vivo* using *lacZ* fusions demonstrated that *emrR* expression represses the *emrRAB* locus and that this repression could be relieved by certain antibiotics and protonophores that are the targets of the EmrAB pump (60). In addition, the MarR ligands, salicylate and 2,4-dinitrophenol, also induced expression of this operon. Gel electrophoresis under non-reducing conditions suggested that EmrR exists as a dimer in solution (61). DNaseI footprinting analysis revealed a surprisingly large EmrR binding site of 42 bp that partially overlaps the –35 promoter element and extends past the start site of transcription for the *emrRAB* operon (62). This site contains a pseudopalindromic sequence comprised of 9 bp half-sites separated by 3 bp. Several compounds were shown to negatively effect EmrR-DNA complex formation. Direct binding between potential ligands and EmrR was demonstrated using equilibrium dialysis and the ligand binding affinities of 2,4-dinitrophenol and two protonophores were measured spectrophotometrically, revealing apparent dissociation constants of approximately 2.0, 3.0, and 15.0 µM, respectively (61). Interestingly, fitting the data to the Scatchard equation suggested that each ligand bound to a single site in the EmrR monomer.

The gene encoding the MarR homolog, MexR from *P. aeruginosa*, is adjacent to the oppositely oriented *mexAB-oprM* operon that encodes a non-ATPase, multisubstrate efflux pump that contributes to this organism’s intrinsic multidrug resistance (55). Initial genetic evidence
suggested a regulatory role for MexR. A *mexR* mutant strain was shown to have enhanced resistance to antibiotics, and MexR was shown to reduce the expression of *mexA: lacZ, mexA: phoA, and mexR: lacZ* fusions, suggesting that MexR is an autorepressor and repressor of the *mexAB-oprM* operon (55). Despite suggestions that MexR might possess dual repressor and activator roles *in vivo*, extensive genetic experiments indicate that MexR functions only as a repressor of this operon (63). DNasel footprinting identified two MexR binding sites within the 274 bp intergenic region that separates *mexR* and *mexA* (58). Site I is ~29 bp and overlaps the predicted –10 promoter element for *mexR* and a putative –35 promoter element for the *mexAB-oprM* operon. Site II is ~28 bp and overlaps the predicted –35 promoter element for *mexR* and a putative –10 promoter element for the operon. Each site contains a palindromic sequence comprised of 5 bp half-sites separated by 5 bp. Curiously, the footprints of sites I and II are separated by only 3 bp (58). In addition to the crystal structure evidence, two-hybrid experiments are consistent with MexR existing as a dimer in solution, and it is therefore likely that MexR binds each of its palindromic sites as a dimer (64). It will be interesting to see if cooperativity exists in MexR binding to its closely spaced binding sequences. Surprisingly, the selection of trans-dominant MexR mutants that were defective in DNA-binding, but not dimerization, predominantly identified single-amino acid substitutions of hydrophobic residues in the DNA-binding domain. Only 2 of the 25 mutations were to charged residues, suggesting the importance of hydrophobic contacts in this region in stabilizing the winged helix motif in a conformation that can accommodate DNA-binding (65).

A member of the MarR family has been characterized from *Neisseria gonorrhoeae* that likely mediates the resistance of this organism to antimicrobial hydrophobic agents. The *farAB* operon of *N. gonorrhoeae* encodes an efflux pump that exports out of the cell host-derived
antimicrobial agents such as long-chain fatty acids (FAs). Using lacZ reporter fusions, it was shown that farAB and farR expression was enhanced in strains mutated at farR (66). Also, strains mutated at the farR site were less resistant to FAs. Electrophoretic mobility shift assays demonstrated that FarR binds sequence specifically in the farAB and farR promoter regions. Thus, FarR is autoregulatory and represses the farAB efflux system, but it remains to be determined if its regulatory activity is affected by certain fatty acids or other potential ligands.

MarR Regulators of Aromatic Catabolism

A phenolic sensing protein from Rhodopseudomonas palustris has been characterized as an inducer of the badDEFG operon, which encodes benzoyl-CoA reductase, an enzyme involved in the anaerobic catabolism of benzoate (67). Analysis in vivo using a badE: lacZ fusion construct demonstrated that, in the presence of benzoate or 4-hydroxybenzoate, BadR increases lacZ expression approximately 5-fold. Strains deficient in badR grew slowly under anaerobic conditions with benzoate as a carbon source. Though not shown, it is likely that the stimulatory effect of BadR occurs through direct binding of benzoyl-CoA reductase substrates to this protein, which induces DNA-binding of BadR to the promoter of badDEFG. This proposed mechanism is interesting in light of the fact that the known ligands of other MarR homologs almost invariably antagonize DNA-binding. The badR gene is part of a gene cluster including the badDEFG operon, but is transcribed from a separate promoter. The mechanism by which the proposed DNA-binding by BadR activates gene expression, and whether or not BadR is autoregulatory, remains to be determined.

The MarR homolog, CbaR from Comamonas testosteroni BR60, is a phenolic-sensing modulator of the cbaABC operon that encodes enzymes involved in the oxidation of 3-chlorobenzoate (68). An intergenic region of 667 bp separates cbaR from the oppositely oriented
cbaABC operon. Electrophoretic mobility shift assays, and in vivo analysis monitoring lacZ expression, indicated that CbaR does not regulate its own expression. However, gel shifts and DNaseI footprinting indicated that CbaR forms two distinct complexes, dissimilar in binding affinities, in the promoter region of the cbaABC operon, suggesting that CbaR represses gene expression. The high-affinity \( \sim 22 \) bp binding site is located \( \sim 40 \) bp downstream of the transcriptional start site and contains a palindromic sequence comprised of 4 bp half-sites separated by 6 or 9 base pairs. The low-affinity binding site contains a pseudopalindromic sequence similar to the high-affinity sequence. The substrate of the enzymes encoded by the cbaABC operon, 3-chlorobenzoate, is an efficient antagonist of CbaR complex formation at both sites. Interestingly, the downstream product of 3-chlorobenzoate catabolism, protocatechuate, is also an efficient antagonist, suggesting a mechanism of positive feedback in cbaABC expression. Curiously, the phenolic compounds 3-hydroxybenzoate and 3-carboxybenzoate might promote a slight increase in the affinity of CbaR for each of its binding sites, suggesting that these compounds enhance CbaR-mediated gene repression.

A MarR homolog has been characterized from the ruminal bacterium Butyrivibrio fibrisolvens E14 that regulates the expression of an enzyme involved in the catabolism of polysaccharide derivatives in plant cell walls (69). It was shown that cells overexpressing a 142 residue, 16 kDa protein, CinR, and a cinnamoyl ester hydrolase (CinB) displayed less activity from this enzyme than did cells overexpressing CinB alone. The open reading frames of cinR and cinB are oriented in the opposite direction in the Butyrivibrio genome and are separated by an intergenic region of 170 bp containing predicted, overlapping promoter elements for each gene. Gel shifts demonstrated that CinR binds a DNA fragment containing the intergenic region between cinR and cinB. Sequence analysis identified two identical 16 bp palindromic sites in this
intergenic region, comprised of 8 bp half-sites. One of these potential CinR binding sites overlaps the predicted cinR transcriptional start site and the other lies just downstream of the putative cinB transcriptional start site, suggesting that CinR represses its own expression and that of cinB. Two cinnamic acid sugar esters, potential substrates of CinB, were shown to antagonize CinR-DNA complex formation. Notably, certain cinnamic acids, sugars, and a non-sugar cinnamic acid ester were not negative effectors of CinR DNA-binding activity, indicating that both the sugar and cinnamic groups are essential components of a CinR effector.

The MarR homolog HcaR, from Acinetobacter sp. strain ADP1, regulates the hcaABCDE operon that encodes genes required for the catabolism of plant-derived hydroxycinnamates. Genetic analyses indicated that HcaR represses this operon and that this repression is relieved by hydroxycinnamoyl-CoA thioesters (70). The gene encoding HcaR and the hcaABCDE operon are part of the same gene cluster, but are divergently transcribed. Direct interaction of HcaR with the operon promoter has not been demonstrated and the HcaR recognition sequence remains unknown.

The hpa-meta operon of E. coli encodes a set of genes required for the catabolism of 4-hydroxyphenylacetic acid (4-HPA). An intergenic region of 219 bp separates the first ORF of this operon from a divergently transcribed gene encoding HpaR. In vitro transcription assays demonstrated that HpaR repressed the expression from the hpa-meta operon promoter and that 4-HPA effectively relieved this repression (71). In addition, 4-HPA, and two structurally similar compounds (3-HPA and 3,4-HPA), increased expression from a lacZ fusion with the hpa-meta operon promoter. In vivo assays involving a lacZ translational fusion with the hpaR promoter demonstrated that HpaR is an autorepressor, and the same phenolic compounds listed above relieve that repression. EMSAs and DNaseI footprinting identified two 27 bp HpaR binding sites
in the intergenic region between \textit{hpaR} and the \textit{hpa-meta} operon. The affinity of HpaR for one site (OPR1) was 10-fold higher than for the other site (OPR2). Site OPR1 overlapped the –10 promoter element and transcriptional start site of the \textit{hpa-meta} operon and contained a palindromic sequence comprised of 9 bp half-sites separated by 4 bp. Site OPR2 was located between the transcriptional start site and ribosome binding site of \textit{hpaR}. OPR2 contained a pseudopalindromic sequence, with one of its half-sites differing in 5 bp from the half-sites in OPR1. Interestingly, gel shifts comparing HpaR binding affinities to DNA fragments containing one, or both, sites indicated that HpaR binding is positively cooperative at these sites. Potassium permanganate footprinting demonstrated that HpaR likely represses transcription of the \textit{hpa-meta} operon by blocking promoter escape by RNA polymerase. In the presence of HpaR, the open complex was shifted upstream of the HpaR binding site, and partially overlapped the –35 promoter element (71).

\textbf{MarR Regulators of Virulence Factors}

A MarR homolog from the nitrogen-fixing bacterium, \textit{Sinorhizobium meliloti}, has been shown to activate transcription of three \textit{exp} operons that are involved in the production of galactoglucan, an exopolysaccharide required for this organism to infect plant roots for nodule formation (72, 73). ExpG was shown to mediate this activation by binding to promoter regions in the \textit{exp} gene cluster (74). Size exclusion chromatography suggested that this protein, ExpG, exists as a dimer in solution (75). Analysis of ExpG-DNA binding using atomic force microscopy (AFM) revealed that the DNA curvature at the ExpG binding site changes upon complex dissociation, suggesting that ExpG induces DNA bending upon complex association. Electrophoretic mobility shift assays and AFM experiments demonstrated that ExpG binding sites require a palindromic sequence comprised of 6 bp half-sites separated by 3 bp.
Measurements of the binding kinetics for ExpG at 3 different promoter binding sites indicated that the ExpG dimer binds with a very high on-rate (ranging from $1.0 - 5.0 \times 10^5$ M$^{-1}$s$^{-1}$) and dissociates with a very low off-rate (ranging from $4.3 - 1.3 \times 10^{-4}$ s$^{-1}$) indicating dissociation constants ranging from $0.58 - 1.3$ nM. Interestingly, the DNA-binding affinity of ExpG is enhanced by the presence of short conserved sequences (“boxes 1 and 2”) located on either side of the palindrome. The distance of these boxes from the conserved 21 bp sequence containing the palindrome ranges from 5 – 30 bp. Dynamic force spectroscopy was used to calculate the natural thermal off-rates for the binding kinetics between ExpG and a wild-type binding site, a binding-site mutated in box 1, and a binding site mutated in box 2. The off-rate was increased approximately 10-fold by a mutation in box 1 and approximately 100-fold by a mutation in box 2. Sites analogous to boxes 1 and 2 have not been described for any other MarR family member.

Gel shift experiments determined that the MarR homolog, PecS from *Erwinia chrysanthemi*, binds specifically to the regulatory regions of pectinase and cellulase enzymes, virulence factors required for this organism to infect plant cells (76, 77). PecS binds with a $K_d$ of $\sim 200$ nM to the promoter regions of two pectinase genes and has been shown to be involved in a complex regulatory system of these enzymes, involving CRP and another repressor, KdgR (77). DNaseI footprinting experiments determined that PecS protects an $\sim 50$ bp site in the pectinase promoter region, which contains a high-affinity binding site for the activator, CRP, suggesting a mechanism for PecS-induced transcriptional attenuation (77). A 23 bp consensus DNA-binding sequence for PecS was defined and shown to contain a palindromic sequence comprised of 5 bp half-sites separated by 3 bp (78). In addition, PecS has been shown to regulate the expression of genes involved in the production of flagella, by binding to a regulatory region with an apparent $K_d$ of 20 nM (78). Furthermore, genetic experiments identified PecS as a repressor of a cluster of
genes involved in the synthesis of the blue pigment, indigoidine (79). Gel shifts determined that PecS binds to the promoter regions of two genes in this cluster, with apparent K\(_d\)'s of 5 and 20 nM. Interestingly, increased production of indigoidine, conferred by a pecS mutation in \(E. \) chrysanthemi, resulted in a significant increase in tolerance to oxidative stress, presumably due to the ability of this aromatic compound to scavenge reactive oxygen species (79).

**Deinococcus radiodurans**

*Deinococcus radiodurans* is one of 11 members belonging to the family *Deinococcaceae*, which includes at least 7 members that are intrinsically resistant to extreme levels of radiation and other sources of DNA damage, such as oxidative stress and dessication (80-84). Though *Deinococcus* often stains Gram-positive, sequence comparisons of 16S rDNA suggests that this lineage is most closely related to the thermophilic, Gram-negative genus, *Thermus* (85). This relationship is consistent with the peptidoglycan chemotype shared by these two genera (86, 87).

The *D. radiodurans* genome is comprised of two chromosomes, one megaplasmid, and one plasmid and is predicted to encode 3195 ORFs (88). *D. radiodurans* maintains 8-10 copies of its genome during exponential growth and approximately 4 copies during stationary phase (89).

Interest in *D. radiodurans* has focused primarily on its ability to repair extensive levels of DNA damage. It has been shown that *D. radiodurans* is approximately 50-fold more resistant to ionizing radiation than is *E. coli*, demonstrating 10% survival at 800 kilorads, whereas *E. coli* has 10% survival at 15 kilorads (90-92). This organism can survive high doses of acute gamma radiation (greater than 1500 kilorads) and grow under conditions of prolonged exposure at 6 kilorads, remarkably without accumulating abnormal levels of mutations (93-95). Similar levels of resistance are observed when *D. radiodurans* is exposed to prolonged dessication, UV radiation, or oxidative stress (83, 92, 96). The damage accrued by the *D. radiodurans* genome
under these conditions suggests that this organism’s intrinsic resistance is a consequence of a remarkable DNA repair system (93, 97). Whereas *E. coli* can only sustain 10-15 double strand breaks in its genome, *D. radiodurans* can repair up to 2000 such breaks (92). Efficient homologous recombination is most likely the primary means by which *D. radiodurans* can withstand such high levels of DNA damage (98, 99). However, numerous non-resistant organisms possess the same complement of genes for recombination, leaving unresolved the full mechanism by which *D. radiodurans* maintains resistance to extreme levels of DNA damage (92). Recent evidence demonstrates that the *D. radiodurans* nucleoid adopts an unusual toroidal morphology, which suggests that efficient DNA recombination may be facilitated by the alignment of homologous segments from multiple copies of the genome (100).

It has been noted that *D. radiodurans* undergoes a dose-dependent growth lag after exposure to DNA-damaging events (92). Damaged nucleotides and nucleosides are exported from the cell during this period of growth inhibition (101). In addition, *D. radiodurans* must handle the enhanced levels of reactive oxygen species that are the direct products of ionizing and UV radiation, dessication, and oxidative stress agents. Catalase and superoxide dismutase enzymes have been shown to be critical in this regard, as mutations to these genes reduce *D. radiodurans* resistance to ionizing radiation (96). In addition, the *D. radiodurans* genome encodes two Dps (DNA protection during starvation) homologs. Dps proteins are structurally related to ferritin, and have been shown to protect DNA from damage by physically binding DNA and by chelating iron (102, 103). Iron, in its ferrous (Fe^{2+}) state, can be particularly toxic to aerobically respiring organisms such as *D. radiodurans* by reacting with hydrogen peroxide to generate reactive oxygen species that are damaging to amino acids and lipids, as well as DNA nucleobases. One of these Dps homologs in *D. radiodurans*, Dps-1, has been recently
characterized and shown to possess ferroxidase activity, oxidizing Fe$^{2+}$ to Fe$^{3+}$ (104). The dodecameric form of this protein stores the oxidized iron, while a dimeric form of Dps-1 protects DNA from DNaseI and hydroxyl radical cleavage, suggesting a role for this protein in mediating oxidative stress resistance in *D. radiodurans* (104).

As numerous MarR proteins have been shown to regulate prokaryotic stress responses to toxic compounds and oxidative stress agents, the identification of two predicted MarR homologs within the *D. radiodurans* genome suggested the possibility of other uncharacterized stress response proteins that contribute to the extreme-resistance phenotype of this organism.

References


CHAPTER 2
HUCR, A NOVEL URIC ACID RESPONSIVE MEMBER OF THE MARR FAMILY OF TRANSCRIPTIONAL REGULATORS FROM DEINOCOCCUS RADIODURANS

Introduction

Members of the MarR family of winged helix transcriptional regulators control a variety of biological functions in bacteria and archaea (1). Several transcriptional activators have been identified in this family, but the majority of MarR homologs are transcriptional repressors (2, 3). A number of bacterial MarR proteins regulate environmental stress responses and the expression of pathogenic factors. For example, MexR represses the MexAB-OprM operon of Pseudomonas aeruginosa, an operon that encodes a tripartite efflux system responsible for this organism’s intrinsic resistance to multiple antibiotics (4, 5). PecS from Erwinia chrysanthemi regulates pectinase and cellulase production, the main virulence determinants of this plant pathogen, and the synthesis of indigoidine, an apparent scavenger of reactive oxygen species (6, 7). Other members of the MarR family regulate the catabolism of aromatic compounds, such as HpaR, which mediates the catabolism of 4-hydroxyphenylacetic acid in Escherichia coli, and HpcR, which regulates homoprotocatechuate catabolism (8, 9). Autoregulation is a characteristic of a number of MarR proteins, including the operon repressors EmrR and MarR, and several homologs that are not encoded in the context of a contiguous operon, such as MexR and HpaR (8, 10-12).

Phenolic ligands have been shown to regulate gene expression by negatively affecting interactions of MarR homologs with their cognate promoter/operator regions. Repressor activity

of EmrR on the multidrug resistance operon, *emrRAB* in *E. coli*, is antagonized *in vitro* by sodium salicylate and a variety of structurally unrelated phenolic drugs that are putative ligands of the multidrug pump encoded by this operon (10, 13). The affinity of CinR, a repressor of a cinnamoyl ester hydrolase from *Butyrivibrio fibrisolvens* E14, for its binding region is reduced *in vitro* by cinnamic acid derivatives (14). HpcR repression of homoprotocatechuate metabolic genes, and HpaR repression of the *hpa-meta* operon, are relieved by the respective aromatic substrates of these catabolic pathways (8, 9). Similarly, the regulators of benzoic acid and 3-chlorobenzoate catabolism, BadR and CbaR, respectively, respond to the aromatic substrates of their pathways (2, 15).

The prototypical member of this protein family, MarR from *E. coli*, negatively regulates the *marRAB* operon, the expression of which confers an intrinsic phenotypic resistance to structurally diverse antibiotics, organic solvents, oxidative stress agents, and household disinfectants (11, 16, 17). MarR binds as a homodimer to two sites on the operator/promoter (*marO*) of this operon with an apparent K_d of ~ 1nM, with site I partially overlapping the -35 and -10 regions of the promoter and site II overlapping the putative ribosome binding site (18). A range of phenolic compounds induce *marRAB* transcription *in vivo*, including sodium salicylate, cinnamate, 2,4-dinitrophenol, acetaminophen, sodium benzoate, tetracycline, and chloramphenicol and a subset of these, as well as plumbagin and menadione, antagonize DNA binding activity of MarR *in vitro* (16, 19-21). Sodium salicylate has been shown to bind MarR with an apparent K_d of 0.5 – 1 mM (18). The 2.3 Å crystal structure of MarR bound to sodium salicylate revealed two binding sites per monomer for this inducer, with each site flanking the putative recognition helix of the DNA binding motif (22).
The heterotrophic, mesophilic bacterium *Deinococcus radiodurans* is best known for its resistance to high levels of ionizing and UV radiation, bulky chemical-adducts, and other agents that damage DNA (23-25). Genomic analysis revealed that *D. radiodurans* also encodes orthologs of nearly every known bacterial stress response protein, suggesting mechanisms of resistance to osmotic, temperature, pH, starvation, toxin, phage, dessication, antibiotic, and oxidative stresses (25, 26). The presence of two genes encoding MarR homologs, *dr1159* and *dra0248* (TIGR gene annotation), within the genome suggested an uncharacterized stress response regulatory system or a novel mode of metabolic regulation in *D. radiodurans*.

In this paper, we report the cloning of *dr1159* from *D. radiodurans* and the subsequent purification of a hypothetical uricase regulator (HucR) belonging to the MarR family of transcriptional regulators. Characterization of this novel protein reveals that it binds as a dimer with very high affinity to a promoter region shared between *hucR* and a neighboring uricase. Through electrophoretic mobility shift assays (EMSAs) we demonstrate that this affinity is antagonized by specific phenolic compounds, notably uric acid. These results, in conjunction with *in vivo* analyses, indicate a novel catabolic regulatory system in *D. radiodurans*.

**Experimental Procedures**

**Cloning, Overexpression, and Purification of HucR**

*D. radiodurans* R1 was kindly provided by J. Battista and genomic DNA was isolated as described (27). PCR amplification of the *hucR* ORF (*dr1159*) from the genome was achieved using primers HucR-fwd (5'-GCT CGT GTT CAT ATG TCA GCC CGC-3'), which introduced an NdeI site (bold) overlapping the first codon (underlined), and HucR-rev (5'-CCT TTC CG AAT TC GGG AAT C-3'), which introduced an EcoRI site (bold) downstream of the *hucR* stop codon. The resulting 589 bp PCR product was cloned into pET-5a, generating pHucR.
Plasmid pHucR was transformed into *E. coli* TOP10 (Invitrogen). Fidelity of the construct was verified by DNA sequencing. Plasmid pSPW1 was subsequently transformed into *E. coli* BL21(DE3)pLysS. Cultures were grown in LB containing 100 µg/ml ampicillin at 37°C to $A_{600} = 0.5$ and HucR overexpression was induced with 0.2 mM isopropyl-1-thio-β-D-galactopyranoside for 1 hour. Cells were pelleted at 4°C and stored at -80°C. All subsequent steps were carried out at 0 - 4°C. Cells were resuspended in lysis buffer (50 mM Tris-HCl (pH 8), 25 mM NaCl, 5% glycerol (v/v), 5 mM Na$_2$EDTA, 0.15 mM phenyl methyl sulfonyl fluoride (PMSF), 10 mM 2-mercaptoethanol) and incubated with 200 µg/ml lysozyme for 1 hour. Triton X-100 and NaCl were added to final concentrations of 0.05% (v/v) and 0.5 M, respectively. DNA was removed from the lysate by slow addition of Polymin P to a final concentration of 0.5% (v/v) followed by centrifugation at 11,000g for 20 minutes. The supernatant was dialyzed overnight against 30 volumes of HA buffer (20 mM Tris-HCl (pH 8.7), 50 mM KCl, 4.8% glycerol (v/v), 1 mM Na$_2$EDTA, 4.3 mM 2-mercaptoethanol, 0.2 mM PMSF) and centrifuged at 11,000g for 20 minutes. The supernatant was loaded onto CM-Sepharose and DEAE-Sepharose columns, linked in tandem and equilibrated with HA buffer, pH 8.7. The HucR-containing flow through and wash fractions were combined and concentrated using a Centriprep centrifugal filter device (Millipore). The concentrated retentate was dialyzed for 3 hours against 60 volumes of HAP buffer (20 mM potassium phosphate (pH 7), 50 mM KCl, 4.8% glycerol (v/v), 1 mM Na$_2$EDTA, 4.3 mM 2-mercaptoethanol, 0.2 mM PMSF) and loaded onto a hydroxyapatite column equilibrated with HAP buffer, pH 7. The wash fractions containing HucR were loaded onto a heparin column equilibrated with HAP buffer, pH 7, and HucR was eluted with a linear gradient of 50 mM – 1 M KCl in HAP buffer, pH 7. Peak fractions were pooled, concentrated, and the glycerol concentration was increased to 20%. The purity of HucR was established by
SDS-PAGE and Coomassie staining. HucR concentration was ascertained spectrophotometrically using $\varepsilon_{280} = 13,512 \text{ M}^{-1}\text{cm}^{-1}$ (calculated from amino acid extinction coefficients), and verified by SDS-PAGE using bovine serum albumin (BSA) as a standard.

**Circular Dichroism Spectroscopy**

Circular dichroism (CD) spectroscopy was performed on an Aviv 202 CD spectrophotometer. HucR was diluted to 0.2 mg/ml in CD buffer A (20 mM potassium phosphate (pH 7), 34 $\mu$M EDTA, 0.8% glycerol, 13 mM KCl). Ellipticity measurements were performed in triplicate at 25°C using a quartz cuvette with a 0.1 cm pathlength. Measurements were made at 1 nm steps over the wavelength range 250 – 190 nm, and were corrected for buffer contributions to the signal. HucR secondary structure composition was calculated using the secondary structure algorithm CDSSTR and protein reference set 7 provided by the authors (28-32). The goodness-of-fit was determined from the NRMSD value of 0.013.

For measurement of thermal stability, HucR was diluted to 0.05 mg/ml in CD buffer B (20 mM potassium phosphate (pH 7), 8 $\mu$M EDTA, 0.2% glycerol, 3 mM KCl) and measurements were made using a quartz cuvette with a 1 cm pathlength. Ellipticity readings from 230 – 200 nm (1 nm steps) were taken over the temperature range 19 – 70°C, with steps of 3°C (19 - 37°C and 61 - 70°C) or 1.5°C (37 – 61°C). Each sample also underwent a reverse scan from 67 – 19°C. Three minutes was allowed for thermal equilibration after each step. Wavelength scans from 240 – 200 nm were performed at 19°C and 70°C to verify the native and denatured states of HucR, respectively. CD measurements of HucR melting were performed in triplicate and corrected for buffer contributions to the signal. Calculation of the $T_m$ of HucR was made based on ellipticity measurements from 224 – 220 nm. Temperature dependent ellipticity
values were plotted at each wavelength and fit to a two state model for protein unfolding (33). The Tm of HucR is reported as the average ± standard deviation.

HucR-hucO Stoichiometry and Affinity Determination

Primers were designed according to the *D. radiodurans* genome to amplify a 241 bp segment that included the entire intergenic region between *hucR* and *dr1160* and extended 61 bp and 79 bp into the coding region of each gene, respectively. The resulting PCR product, *hucO*, was gel purified and ^32^P-labeled with T4-polynucleotide kinase (T4-PNK).

For binding assays under stoichiometric conditions, 0.1 µM ^32^P-labeled *hucO* was titrated with HucR up to 0.35 µM, in a total reaction volume of 10 µl in Binding Buffer (20 mM Tris-HCl (pH 8), 0.1 mM Na$_2$EDTA, 0.075% BRIJ58, 50 mM NaCl, 5 mM MgCl$_2$, 50 µg/ml BSA, and 4% (v/v) glycerol). Protein-DNA complexes were equilibrated at 22°C for 1 hour. A non-denaturing 6.5% polyacrylamide gel was pre-run for 30 minutes in 0.5xTBE buffer (45 mM Tris-borate (pH 8.3), 1.25 mM Na$_2$EDTA) and samples were loaded with the power on. After 1.25 hours of electrophoresis the gels were dried and protein-DNA interactions analyzed by phosphorimaging using a STORM 840 phosphorimager and ImageQuant 1.1 software. Fractional complex formation was plotted against ([HucR]/[hucO]) and fit to a spline curve. Tangents were generated from data points in the upward slope and the plateau, and the stoichiometry of HucR:hucO complex formation extrapolated algebraically. Experiments were performed in duplicate.

EMSAs for K$_d$ determination were performed as described above, except that binding reactions involved 0.1 nM ^32^P-labeled *hucO* titrated with HucR up to 30 nM. The binding isotherm for K$_d$ determination was generated by non-linear fit of three data sets to the binding equation: normalized fractional saturation of *hucO* = (n(P)/K$_d$)/(1 + (P/K$_d$)) where n is the
number of HucR binding sites, P is free protein concentration, and $K_d$ is the observed equilibrium dissociation constant. The $K_d$ value is reported as the mean ± standard deviation. In the competition assay, binding conditions were as described above, involving 0.1 nM $^{32}$P-labeled \( hucO \) titrated with up to 10 nM of unlabeled \( hucO \) or pGEM5. HucR was added last to the binding reactions, at a final concentration of 1 nM.

**Methidiumpropyl-EDTA (MPE)-Fe(II) and DNaseI Footprinting**

Complimentary 77-mer oligonucleotides were gel purified and the “top” strand was 5'-end-$^{32}$P-labeled with T4-PNK. Annealing of the oligonucleotides was accomplished by slow cooling from 90°C to 16°C. Binding reactions were in a total volume of 10 µl and included 500 fmol of DNA in modified Binding Buffer with 0.06% BRIJ58, 20 µg/ml BSA, and 1.5% glycerol. Protein-DNA complexes were equilibrated for 1 hour at 22°C. For MPE-Fe(II) footprinting, 1 µl of 10 mM sodium ascorbate was added followed by 2 µl of MPE mix (25 µM MPE, 25 µM FeNH$_4$SO$_4$), and digestion allowed to proceed for 2 minutes at room temperature. DNaseI footprinting samples were incubated with $10^{-3}$ to $10^{-2}$ units of DNaseI (Epicentre) for 30 seconds at room temperature. Digestion was terminated by phenol-chloroform extraction and samples were ethanol precipitated. A/G chemical sequencing ladders were generated according to Sambrook et al. (34). DNA fragments were separated on a 12% polyacrylamide sequencing gel. Density profiles were obtained by phosphorimaging.

**Ligand Binding Assays**

A 180 bp sequence spanning bp -158 to +22 relative to the \( hucR \) translational start site (bp -76 to +104 relative to the predicted \( dr1160 \) transcriptional start site) was amplified, gel purified, and $^{32}$P-labeled with T4-PNK. For each ligand assayed, binding conditions were established to buffer pH effects from the compound. Binding reactions were assembled in
Binding Buffer with 0.1 nM DNA, 0.75 nM HucR, and up to 25 mM sodium salicylate. EMSA were also performed without BSA in the Binding Buffer, with no detectable effect. For acetylsalicylate assays, acetylsalicylate was dissolved in ethanol and binding reactions were assembled in modified Binding Buffer with 400 mM Tris-HCl (pH 8) and 20% ethanol with up to 25 mM of the compound. Uric acid was dissolved in 0.35 M NaOH to a concentration of 125 mM. Binding reactions were assembled in modified Binding Buffer with 500 mM Tris-HCl (pH 8), 0.05% BRIJ58, 7.5 µg/ml BSA, and 0.6% glycerol with up to 20 mM uric acid. Reactions were equilibrated for 1 hour at 22°C and protein-DNA complexes were analyzed by EMSA as described above. All experiments were carried out in triplicate. Data was fit to a single exponential equation: normalized fractional complex formation = e^{-KL}, where k is the exponential decay constant and L is the ligand concentration.

RNA Dot Blot Hybridization

Ten µl of an overnight culture of *D. radiodurans* was added to 5 ml of TGY broth (0.5% tryptone, 0.3% yeast extract, 0.1% glucose) or 5 ml of TGY broth supplemented with 10 mM uric acid. Before inoculation, the pH of broths containing uric acid was adjusted to that of TGY broth. Cultures were grown at 30°C to A_{600} = 0.45 and cells were harvested by centrifugation at 4°C at 12,000g for 10 minutes and stored at -80°C. Pellets were resuspended in 3 ml of 95% ethanol and held at 4°C for 10 minutes and harvested by centrifugation at 4°C at 12,000g for 10 minutes. Total RNA was prepared as described (35). DNA contamination was removed by digestion with 1 µg RNase-free DNaseI (Epicentre) at 37°C for 1 hour.

RNA dot blots were performed essentially as described (35). For each RNA sample, 15 µg was adhered to a (+)-charged nylon membrane using a Bio-Dot microfiltration apparatus (Bio-Rad) and crosslinked using a Stratalinker UV crosslinker (Stratagene) at 120,000 µJ/cm² for
50 seconds. Oligonucleotide probes complementary to the dr1160 transcript (+30 to +49 relative to the predicted start site of translation) and to the hucR transcript (+21 to +40 relative to the start site of translation) were used. As groESL (dr0606-0607) expression in D. radiodurans has been shown to be constitutively expressed under different growth conditions, a probe was designed complementary to residues +310 to +330 of groES to serve as a normalizing factor for RNA samples (36). The probes were 32P-labeled with T4-polynucleotide kinase, separately hybridized to the crosslinked membranes overnight and washed under moderate conditions, as described (35). The densitometric count obtained with the dr1160 or hucR probe for each RNA sample was normalized by multiplying by the ratio of sample 32P-groES counts to control RNA 32P-groES counts. RNA dot blot analysis was performed in duplicate from separate RNA preparations. The levels of hucR and dr1160 transcript from cells grown in 10 mM uric acid are reported as the mean ± standard deviation relative to the transcript level from control cells.

Analysis of Uricase Activity

D. radiodurans was grown in TGY or TGY + 10 mM uric acid as described above. Cell pellets were resuspended in chilled HAP buffer, pH 7.5 without 2-mercaptoethanol and lysed by sonication. Insoluble material was removed by centrifugation at 4°C. Protein concentrations of whole cell lysates were determined from a BSA standard using the modified Lowry protein assay (Pierce).

Uricase activity in the whole cell lysate was analyzed using an Amplex® Red Uric Acid/Uricase Assay Kit (Molecular Probes). Each reaction included 45 µg of whole cell lysate protein. Uricase activity is normalized to the activity recorded for cells grown in absence of added uric acid and is reported as the mean ± standard deviation (n = 5).
Results

The hucR Regulatory Region

Locus dr1159 in chromosome I of D. radiodurans encodes a putative 181 residue, 19.7 kDa MarR homolog. An intergenic region of 101 bp separates dr1159 from the oppositely oriented ORF of a hypothetical uricase (dr1160) (Fig. 2.1). Pairwise alignment of the predicted protein product of dr1160 with a characterized uricase from Bacillus subtilis indicates 29.5% identity and 54% similarity, suggesting that the D. radiodurans homolog is functional (37). Uricase catalyzes the conversion of uric acid into allantoin during purine catabolism, and its activity is correlated with oxidative stress response in mammals (38, 39). A downstream locus, dr1161, encoding a hypothetical, transthyretin-like protein is oriented in the same direction as dr1160, with the ORFs overlapping by one codon. A study of D. radiodurans promoter elements indicated a strong similarity to the E. coli σ70 consensus -10 and -35 regions, amongst the promoters surveyed (40). We identified σ70-like promoter elements for dr1159 and dr1160 in the sequence between these two genes. These promoter regions are partially overlapping, with the putative transcription initiation site of each gene positioned in the Pribnow box of the other. The spacing and orientation of the hypothetical MarR homolog encoded at locus dr1159, relative to the ORFs of dr1160 and dr1161 suggested a potential regulatory design analogous to systems observed with other MarR proteins, such as MexR from P. aeruginosa and HpaR from E. coli (8, 12). We were therefore prompted to designate dr1159 as hucR for hypothetical uricase regulator.

Sequence and Structural Analysis of HucR

Pairwise alignment of the amino acid sequence of HucR with the prototype of this family, MarR, reveals 29% identity and 49% similarity between these two proteins. Pairwise alignment with other characterized MarR regulators shows that HucR shares higher homology with EmrR
Fig. 2.1. The genetic organization of *dr1159* and adjacent ORFs. The relative orientations of the ORFs are represented by open arrows. HucR is encoded by *dr1159*, *dr1160* encodes a hypothetical uricase, and *dr1161* encodes a transthyretin-like protein. An intergenic region of 101 bp separating *hucR* and *dr1160/dr1161* contains putative σ° -10 and -35 promoter elements for each gene, designated by bars above (*dr1160/dr1161*) or below (*dr1159*) the sequence. Predicted transcription and translation start sites are identified by asterisks and arrows, respectively. The shaded box marks the HucR binding site. Each half-site of the pseudopalindromic sequence within the binding site is identified in bold.
and MexR, demonstrating 34% identity with each, and 55% and 49% similarity with these repressors, respectively. Multiple sequence alignment of HucR with eight representative MarR homologs reveals seven identical residues in the C-terminal half of these proteins (Fig. 2.2). Five of these sites (residues 118, 124, 126, 132, and 133) occur within the β-sheet and turn structural elements of MarR that form the “wing 1” motif, and a sixth identical residue lies in an α-helix immediately adjacent to this region (22, 41). The wing 1 structural element of the winged helix motif flanks the DNA recognition helix, and has been shown to make direct contacts with DNA, in either the minor or major groove (41). Of the 14 residues identified from the MarR crystal structure to form the hydrophobic core of the monomeric DNA binding domain, 10 are conserved or identical in the MarR homologs analyzed (22). Direct comparison of HucR and MarR reveals that 5 of these residues are identical and an additional 6 are conserved.

The residues in the MarR monomer that were shown to make contacts with the two salicylate ligands are highly conserved in HucR. In ligand binding site “A” of MarR, Thr104 hydrogen bonds with the salicylate hydroxyl, Arg118 hydrogen bonds with the salicylate carboxylate, and Pro89 is located within 3.5 Å of the unsubstituted side of the salicylate ring. In HucR, Arg118 and Pro89 are conserved, while serine replaces threonine at position 104. Arg109 in MarR, which is conserved in HucR, hydrogen bonds to the salicylate carboxylate group in site “B” of MarR. Val128 of MarR is positioned 3.6 Å above the salicylate ring in site “A” of the adjacent half of the dimer. This valine is replaced by alanine in HucR.

The coding region of hucR was cloned into a pET5a expression vector and the product, HucR, was purified to greater than 95% homogeneity, as revealed by SDS-PAGE (Fig. 2.3). The far-UV circular dichroism spectrum from 250–190 nm was recorded at 25°C to determine the
Fig. 2.2. Multiple sequence alignment of HucR and representative MarR family members. The alignment was generated using ClustalX and conserved residues were colored using MacBoxshade. Numbering is based upon the HucR sequence. Residues that are identical in all nine homologs are shaded red. Residues that are ≥ 80% conserved are shaded blue/green, where green indicates non-identity to the HucR residue at that position. Plus signs above the sequence designate residues identified from the MarR crystal structure to form the hydrophobic core of the monomeric DNA binding domain (22). The helix-turn-helix DNA-binding motif, and flanking β-sheet and “wing 1” region identified from the MarR crystal structure, are designated below the alignment. Numbering of the secondary structural elements is as reported for MarR, with α4 being the DNA recognition helix. Proteins are HucR from *D. radiodurans* R1, a homolog from *Sinorhizobium meliloti* (NP_384406), two homologs from *Agrobacterium tumefaciens* (NP_530978 labeled as “1” and NP_353303 as “2”), EmrR from *E. coli* (P24201), MexR from *P. aeruginosa* (C83593), PecS from *Erwinia chrysanthemi* (P42195), MarR from *E. coli* (P27245), and HpaR from *E. coli* (Q07095).
Fig. 2.3. **Purified HucR.** HucR was purified to > 95% homogeneity. *Lane 1,* molecular mass marker, in kDa; *lane 2,* 1 µg of purified HucR. Monomeric HucR migrates at approximately 19 kDa, close to its predicted molecular mass of 19.7 kDa.
secondary structure composition of HucR (Fig. 2.4). The $\alpha$-helical and $\beta$-sheet content for HucR was calculated to be approximately 47% and 10%, respectively, with 17% turns and 25% random coil (28-32). In comparison, the crystal structure of MarR revealed its composition to be ~58% helical and 10% $\beta$-sheet (22). The MexR crystal structure also revealed a large $\alpha$-helical content and a relatively small $\beta$-sheet contribution (42). Ellipticity measurements at five wavelengths spanning the negative ellipticity maximum characteristic of $\alpha$-helices (220–224 nm) were recorded over the temperature range of 19 - 70°C and plotted to measure protein denaturation (Fig. 2.4B). From the CD melting curve, the $T_m$ of HucR was calculated to be 51.1 ± 0.0 °C. HucR did not refold during the reverse temperature scan, so $\Delta H^\circ$ values for folding transitions could not be determined.

### Binding of HucR to its Promoter/Operator Region

Given the prevalence of autoregulatory MarR transcription factors, we biochemically assessed HucR for such a capacity. To test the affinity of HucR for its promoter/operator region ($hucO$), we amplified a 241 bp region of the *D. radiodurans* genome, extending from 180 bp upstream of the putative $hucR$ translation start site to 61 bp within the coding region. EMSA revealed HucR to have high affinity for $hucO$, with an apparent $K_d$ of 0.29 ± 0.02 nM (Fig. 2.5) and a concomitant $\Delta G_{assoc}$ of -12.9 kcal/mol. The single complex observed, and the goodness of fit to a single-site binding polynomial, suggested that HucR binds $hucO$ at a single site. The specificity of this interaction was verified by the inability of pGEM5 to compete for HucR binding (Fig. 2.5B). Complex formation between 1 nM HucR and 0.1 nM $^{32}$P-labeled $hucO$ was reduced to half saturation at approximately 1 nM unlabeled $hucO$, whereas only a minor reduction in complex formation was elicited by 10 nM pGEM5. As pGEM5 contributes an approximately twelve-fold greater molar excess of base pairs than $hucO$, the competition assay,
Fig. 2.4. **CD spectral analysis of HucR.**  

**A,** Far-UV CD spectrum of HucR at 25°C. Ellipticity measurements are expressed in machine units (millidegrees).  

**B,** Thermal unfolding transition of HucR. Ellipticity measurements were collected over the temperature range of 19 - 70°C at five wavelengths spanning the negative ellipticity maximum characteristic of α-helices:  

- (●) = 220 nm,  
- (■) = 221 nm,  
- (▲) = 222 nm,  
- (▲) = 223 nm, and  
- (▼) = 224 nm.
Fig. 2.5. HucR binding to its promoter/operator region, hucO.  A, EMSA.  0.1 nM hucO was titrated with HucR.  Uncomplexed hucO and the single hucO:HucR complex are identified by arrows.  Protein concentrations are indicated above the corresponding lanes.  B, Competition assay.  0.1 nM of labeled hucO and 1 nM HucR was titrated with up to 10 nM of either unlabeled hucO or pGEM5 and complexes resolved by EMSA.  The first lane contains only labeled hucO.  Concentrations of unlabeled competitor DNA are indicated above the corresponding lanes.  Complexed and uncomplexed hucO are identified by arrows.  C, Binding isotherm depicting normalized hucO fractional saturation as a function of uncomplexed HucR.  Data were collected in triplicate.  D, Stoichiometry of HucR:hucO complex formation.
in conjunction with footprinting analyses (see below), unequivocally demonstrates site specific binding of HucR and \textit{hucO}. From EMSA performed under stoichiometric conditions, HucR was shown to bind \textit{hucO} at a ratio of 1.85:1, suggesting that HucR binds its promoter/operator region as a dimer (Fig. 2.5D). This finding is consistent with crystallographic and biochemical analyses of other MarR homologs which have also been shown to form homodimers (13, 22, 42).

Methidiumpropyl-EDTA-Fe(II) (MPE-Fe(II)) and DNaseI footprinting was performed using the top strand of a 77 bp \textit{hucO} fragment. MPE-Fe(II) footprinting revealed the HucR dimer to protect 21 bp of \textit{hucO} spanning the region from -19 to +3 relative to the putative uricase (\textit{dr1160}) transcriptional start site (Fig. 2.6). The footprint was extended by partial protection of bp -20 and -21. The footprint generated by DNaseI (Fig. 2.6) shows protection from -18 to +6, relative to the uricase transcriptional start site. Analysis of the DNaseI footprint also revealed sites of hypersensitive cleavage flanking each end of the HucR dimer binding site (+8, +7, and -19), suggesting that HucR distorts \textit{hucO} upon binding. Winged helix proteins from the MarR family characteristically bind as dimers to inverted repeat sequences in their cognate recognition sites (12). Footprinting analysis revealed the HucR dimer binding site within \textit{hucO} to contain an imperfect 8 bp inverted repeat, with two bp separating each half of the palindrome (Fig. 2.1).

\textbf{Ligand Binding Assays}

Members of the MarR family of transcriptional regulators are natural phenolic sensors, and hence play critical roles in environmental surveillance. Given the significant conservation in HucR of residues that are involved in binding salicylate in MarR, we tested this anionic lipophilic compound as a potential ligand of HucR. Normalized fractional HucR:\textit{hucO} complex formation was analyzed as a function of sodium salicylate concentration (Fig. 2.7). Sodium salicylate antagonized the binding of HucR to \textit{hucO}, with 5.2 mM of the compound reducing the
Fig. 2.6. **Footprinting analysis of HucR:hucO complex.** *A,* DNaseI (*lanes 1 – 3*) and MPE-Fe(II) (*lanes 4 – 9*) footprinting of top strand of hucO. *Lanes 1 and 4,* A/G ladder; *lane 2,* DNaseI treatment of hucO in the absence of HucR; *lane 3,* DNaseI treatment of hucO incubated with 200 nM HucR. *Lanes 5 -9,* MPE-Fe(II) treatment of hucO after incubation with 0 nM, 25 nM, 100 nM, 200 nM, or 400 nM HucR, respectively. The predicted positions of the *dr1160*-10 promoter element and transcriptional start site are indicated at the right. *B,* Densitometric profile of the HucR:hucO complex, as determined by MPE-Fe(II) footprinting. MPE-Fe(II) treated hucO in the absence of HucR is represented by the gray trace. The dark line represents DNA incubated with 200 nM HucR (*lane 8 in panel A*). *C,* Densitometric profile of the HucR:hucO complex, as determined by DNaseI footprinting. The gray line shows the densitometric trace of DNaseI treated hucO, in the absence of HucR. DNaseI cleavage of hucO incubated with 200 nM HucR is represented by the dark trace. Sequence numbering in panels *A – C* is relative to the predicted +1 transcriptional start site of *dr1160*, as in Fig. 2.1.
Fig. 2.7. **HucR effector binding assays.**  

A, Structures of the compounds tested: salicylic acid, acetylsalicylic acid, and uric acid (shown in its more stable keto tautomeric conformation).  

B, EMSA demonstrating the capacity of uric acid as a negative effector of HucR binding to its cognate DNA-binding site. Uric acid concentrations (in mM) are shown above the corresponding lanes. Bands corresponding to free *hucO* and *hucO*:HucR complex are marked by arrows.  

Lane 1, *hucO* incubated with 15 mM uric acid in the absence of HucR.  

C, Normalized *hucO*:HucR complex formation as a function of ligand concentration. Experiments were performed in triplicate.  

(▲) = sodium salicylate; (■) = acetylsalicylate; (◊) = uric acid.
normalized fractional saturation to 0.5. At 25 mM sodium salicylate, the fraction of complexed \( hucO \) approached zero. As BSA is known to bind salicylate, it was possible that, due to the BSA in the binding reaction, salicylate’s antagonistic effect on HucR: \( hucO \) interaction was greater than we observed. However, removal of BSA from the binding reactions resulted in no observable change in salicylate’s role as a negative effector of HucR (data not shown).

Acetylsalicylate was suggested to induce transcription of the \( marRAB \) operon \textit{in vivo}, yet \textit{in vitro} analysis suggested that acetylsalicylate does not bind MarR (16, 18). Acetylsalicylate caused a gradual decrease in complexed \( hucO \) with increasing drug concentration, however, even at 25 mM acetylsalicylate, the normalized fractional saturation of \( hucO \) remained at \(~0.6\). Approximately 46 mM acetylsalicylate is required to reduce the fractional saturation to 0.5. Therefore, an acetyl group esterified to the 2-hydroxyl of salicylate decreases by 9-fold the negative effector capacity of the ligand.

The apparent phenolic sensing capability of HucR, and its high affinity for a site within the promoter/operator region of a putative uricase (Fig. 2.1), suggested a potential catabolic regulatory system in which HucR regulation of uricase expression is responsive to levels of uric acid, the substrate for this enzyme. Like salicylic acid, uric acid is a planar, aromatic compound, but consists of two conjugated ring systems. As seen in Fig. 2.7, uric acid is a potent regulator of HucR; the normalized fraction of complexed \( hucO \) approached zero at \(~1\) mM uric acid, with 0.26 mM uric acid being sufficient to reduce the normalized fractional complex to 50%. Uric acid is therefore an approximately 20-fold stronger antagonist of HucR: \( hucO \) interaction than salicylate, and over 175-fold stronger than acetylsalicylate. This pronounced attenuation of complex formation elicited by uric acid, relative to the effects observed from the other aromatic compounds, suggests that this compound is the natural ligand of HucR.
In vivo Analysis of Gene Regulation by HucR

The in vitro investigations described above are consistent with a model in which transcription of \textit{hucR} and \textit{dr1160} (putative uricase) is regulated by the high-affinity interaction of HucR in the intergenic region. The dramatic reduction in HucR:\textit{hucO} affinity educed by uric acid suggests that this compound would weaken HucR-mediated repression. Furthermore, the homology between the hypothetical enzyme encoded by \textit{dr1160}, the only putative uricase in the \textit{D. radiodurans} genome, and the characterized uricase from \textit{B. subtilis} suggests that this gene encodes a functional enzyme (37). We investigated this model through a combination of RNA dot blot hybridization and uricase activity experiments.

Transcript levels of \textit{hucR} and \textit{dr1160} were compared from \textit{D. radiodurans} grown in the presence versus absence of 10 mM uric acid (Fig. 2.8A). RNA dot blot hybridization of a $^{32}$P-labeled probe complementary to the sense strand of \textit{hucR} revealed a higher level of transcript in cells grown in 10 mM uric acid, at a ratio of 1.6 ± 0.1 relative to control cells. Hybridization of a probe complementary to the sense strand of \textit{dr1160} revealed upregulation of the putative uricase in the cells grown in 10 mM uric acid, at a ratio of 1.6 ± 0.0 relative to control cells.

A coupled enzymatic assay was performed to assess uricase activity in protein from whole cell lysate of cells grown in the presence versus absence of 10 mM uric acid (Fig. 2.8B). The assay links the activity of uricase from the samples to the activity of horseradish peroxidase, producing resorufin, which has an absorption maximum near 560 nm. Uricase activity was detected in cells supplemented with or without uric acid, suggesting that \textit{dr1160} does indeed encode a functional uricase. Uricase activity was 1.5 ± 0.0 times higher in cells grown in the presence of 10 mM uric acid compared to its absence. The enhanced levels of \textit{hucR} and \textit{dr1160} transcript and the elevated uricase activity in cells grown in excess uric acid is consistent with
Fig. 2.8. **Analysis of hucR and dr1160 (uricase) gene expression.** *A*, Expression of hucR and dr1160 in log-phase cells grown in the absence (black bars) or presence (cross-hatched bars) of 10 mM uric acid, as determined by RNA dot blot analysis. Transcript levels of hucR and dr1160 were calculated using radiolabeled probes complementary to the sense strand of the respective gene, as indicated on the x-axis. The transcript levels of uric acid supplemented cells are reported relative to those of control cells. The error bars indicate the standard deviations from two experiments. *B*, Analysis of uricase activity in log-phase cells grown in the absence (black bars) or presence (cross-hatched bars) of 10 mM uric acid. Uricase activity is measured by the absorbance of resorufin at 560 nm. Activity in uric acid supplemented cells is reported relative to that in control cells. The error bar indicates the standard deviation from five experiments.
the *in vitro* ligand binding studies with HucR, and strongly support a model in which transcription of these two divergent genes is de-repressed by the effector uric acid.

**Discussion**

The identification of MarR homologs within the *D. radiodurans* genome suggested the presence of uncharacterized regulatory systems responsive to phenolic compounds within this stress-resistant microorganism. In this study, we demonstrate that HucR binds with very high affinity (*K_d* = 0.29 ± 0.02 nM) at a single site in its regulatory region (Fig. 2.5 and 2.6). In comparison, *E. coli* MarR binds its cognate site with an apparent *K_d* of 1 nM and the winged helix protein, Ptr1, binds its cognate sites with apparent *K_d* values of 1-2 nM (18, 43). Winged helix proteins from the MarR family characteristically bind to sites containing palindromic, or pseudopalindromic sequences. HpaR, for example, binds to sequences possessing two 9 bp half-sites separated by 2 bp and CinR binds to a 16 bp palindrome composed of two adjacent 8 bp half-sites (8, 14). The binding site of EmrR contains an imperfect 9 bp inverted repeat with each half-site separated by 3 bp (10). The binding site for HucR contains an imperfect 8 bp inverted repeat, with each half-site separated by 2 bp (Figs. 2.1 and 2.6). The center of each half-site of the palindrome is therefore separated by 10 bp, thus positioning the binding site for each half of the HucR homodimer on the same face of the double-helix. This contrasts with the binding site of MarR, which contains two inverted 5 bp sequences separated by 2 bp, thus positioning the half-sites on different faces of the double helix (18). In the case of MexR, the 5 bp inverted repeat sequences are separated by 5 bp, which would orient the major groove of each half-site on the same face of the DNA helix (12). There is variation within the winged helix family in the mode of DNA binding, but the recognition helix typically makes most of the sequence-specific contacts within the major groove (38, 44). It is therefore likely that HucR shares a similar mode
of DNA binding with its two closest, characterized homologs, EmrR and MexR, in which the recognition helix of each half of the homodimer binds in the major groove on the same face of the double-helix.

A number of autoregulatory members of the MarR family have been described, including repressors that, like HucR, are not encoded in a contiguous operon. The position of the HucR binding site in \( hucO \) suggests a mechanism of simultaneous transcriptional repression of the divergent \( hucR \) and uricase genes, involving steric inhibition of RNA polymerase recruitment to the promoter or elongation. The orientation of \( hucR \) and the uricase gene, with an intervening regulatory region, is similar to the genetic organization of other MarR homologs. For example, MexR binds to two sites in the \( mexR-mexA \) intergenic region, with each binding site containing promoter elements of each gene, explaining the mechanism of MexR repression (12). HpaR regulates its own expression, and that of an oppositely oriented gene cluster, by binding in the central operator/promoter region (8).

Our finding that salicylate is a negative effector of HucR:\( hucO \) interaction is consistent with HucR’s conservation of residues that form the salicylate binding site in MarR (Fig. 2.7) (22). The decreased affinity of HucR (and MarR) for acetylsalicylate is likely due to the extra acetyl group and the concomitant loss of hydrogen bonding capacity with Ser104 (Thr104 in MarR) (18). An explanation of HucR’s apparent higher affinity for uric acid compared to salicylate awaits structural details of this protein. The efficient antagonism of HucR:\( hucO \) interaction by uric acid suggests a regulatory mechanism of uric acid catabolism (Fig. 2.7). Several repressors of aromatic catabolic pathways from the MarR family have been characterized (45). For instance, HpaR represses the transcription of genes involved in the catabolism of 4-hydroxyphenylacetic acid, and this repression is relieved by the binding of the substrate, 4-
hydroxyphenylacetic acid, to HpaR (8). The repression of a cinnamoyl ester hydrolase by CinR is relieved by aromatic substrates of this enzyme (14). The catabolism of uric acid in *D. radiodurans* is likely regulated in an analogous manner, with uric acid directly relieving HucR mediated repression of uricase levels in the cell. This model of co-repression of *hucR* and *dr1160* (predicted uricase) is supported by RNA dot blot analysis which reveals uric acid-induced upregulation of both genes to similar transcript levels in *D. radiodurans* (Fig. 2.8A) and by the upregulation of uricase activity in the presence of excess uric acid (Fig. 2.8B). As *dr1160* is the only uricase homolog in the *D. radiodurans* genome, we interpret these results as support of a model of HucR mediated regulation.

The physiological significance of uricase regulation in *D. radiodurans* is unclear. This enzyme acts as part of the purine degradation pathway in prokaryotes and in eukaryotes. In primates, birds, terrestrial reptiles, and many insects, uric acid is the final product of purine catabolism, and is excreted into the environment. Selected soil bacteria are capable of catabolizing uric acid as a carbon and energy source, including strains of *Bacillus fastidiosus*, which require uric acid (or its degradation products allantoin and allantoic acid) as a substrate for growth (46). *B. subtilis* can utilize uric acid as its sole nitrogen source (47). As *D. radiodurans* is also a soil dwelling microorganism, it is tempting to speculate on the possibility that it, too, has acquired the capacity to use uric acid as a source of carbon, energy, or nitrogen.

Analysis of 16S ribosomal DNA sequences suggests that *Deinococcus* forms a separate bacterial phylum with the thermophilic *Thermus* genus (48). However, the T_m of 51.1 ± 0.0°C for HucR, is similar to the T_m values (at similar salt concentrations) of other helix-turn-helix transcriptional regulators from mesophilic organisms, such as the LacI DNA-binding domain, which has a T_m of 45.6 ± 0.2°C and the DNA-binding domain of MATα2, which has a T_m of
56.5°C (49, 50). Despite the close relationship of Deinococcus to a thermophilic genus, HucR does not appear to be an unusually thermostable protein. HucR shares the highest sequence similarity to uncharacterized MarR homologs from Sinorhizobium meliloti and Agrobacterium tumefaciens (Fig. 2.2), both of which are soil dwelling, mesophilic, plant symbionts from the family Rhizobiaceae. This family is grouped with the Proteobacteria, a eubacterial lineage that is distantly related to Deinococcus. Pairwise alignment of HucR and its homolog from Sinorhizobium shows 38% sequence identity and 55% similarity. Alignment with each of the Agrobacterium homologs reveals 36.5% identity and 55.7% similarity. As D. radiodurans is a soil dwelling mesophile, it is plausible that HucR’s similarity to these Rhizobial proteins is a result of horizontal gene transfer (25). Multiple genes in D. radiodurans appear to have been acquired via horizontal transfer, including at least seven stress response genes that were likely acquired from species belonging to the family Rhizobiaceae (25). HucR’s potential role in regulating uric acid levels, suggests that it, too, is involved in D. radiodurans stress response.

Uric acid is an efficient scavenger of reactive oxygen species, including hydroxyl radicals, superoxide anion, and singlet oxygen and is thought to be a critical antioxidant in mammals (51, 52). Indeed, D. radiodurans demonstrates extreme resistance to oxidative damage (53-55). HucR mediated regulation of uric acid levels in D. radiodurans could therefore contribute to this organism’s observed resistance to high levels of oxidative stress. It has been demonstrated that oxygen increases the lethality of ionizing radiation in D. radiodurans presumably by generating reactive oxygen species (56). By regulating levels of uric acid, HucR could therefore also participate in the response of D. radiodurans to ionizing radiation. It is likely that the sensitivity of HucR activity to levels of uric acid serves to maintain an optimum level of this scavenger of peroxynitrite and other reactive oxygen species in the cytoplasm. The low solubility of uric acid...
is well documented and is manifested in precipitate-related health problems such as gout in humans. Other proteins in this family have been shown to regulate resistance to oxidative stress, including MarR, SlyA, and OhrR (57-59). PecS control of indigoidine, an apparent radical scavenger, regulates *E. chrysanthemi* resistance to reactive oxygen species (7). It is thus tempting to speculate that HucR shares with these MarR homologs a common functional role of mediating oxidative stress response.

**References**


CHAPTER 3
NEGATIVE COOPERATIVITY OF URIC ACID BINDING IN THE DNA-BINDING DOMAIN OF HUCR

Introduction

The identification of the MarR (multiple antibiotic resistance regulator) family of transcriptional regulators was initiated by genetic experiments that identified strains of *Escherichia coli* K-12 with enhanced resistance to structurally diverse antibiotics, oxidative stress agents, and organic solvents (1-5). The locus responsible for the resistance phenotype was identified as the *marRAB* (multiple antibiotic resistance) operon (6, 7). Transcriptional repression was shown to be conferred by the product of the first gene of this operon, MarR, which binds with high affinity ($K_d \sim 1 \text{ nM}$) as a homodimer to two sites in the *marRAB* promoter/operator region (4, 8). MarR-mediated repression was shown to be relieved, both *in vitro* and *in vivo*, by structurally diverse phenolic compounds, including some of the antimicrobial agents to which the operon confers resistance (4, 9-11). Structural determination identified MarR as a winged helix DNA-binding protein (12).

Numerous prokaryotic MarR homologs have since been identified with physiological regulatory roles in antibiotic and oxidative stress response (MarR, MexR, EmrR, PecS, HucR), the production of virulence factors (PecS), and the catabolism of aromatic compounds (HpaR, HpcR, CinR, BadR, CbaR, HucR) (9, 13-17). The DNA-binding capacities of a number of these MarR family members have been shown to be altered by anionic lipophilic compounds, often by the suggested target of the regulated genes.

The structure of the MarR dimer bound to the effector salicylate has been determined at 2.3 Å resolution, revealing two ligand binding sites, designated as SAL-A and SAL-B, on either side of the proposed DNA recognition helix (12). The proximity of the ligand-binding sites in
MarR to residues shown to be critical for MarR binding activity suggested a potential mechanism of regulation (18). In addition, the crystal structure of a MarR family member, MexR from *Pseudomonas aeruginosa*, has been solved at a resolution of 2.1 Å (19). Several conformational states of the MexR dimer were observed in the asymmetric unit, suggesting that the mechanism of ligand-induced effects on the DNA binding activity of this protein involves alterations to the relative orientations of the DNA binding domains from each half of the dimer. However, without combined structural data for each protein in its DNA-bound, ligand-bound, and apo-conformations, the DNA-binding and ligand-response mechanisms of these MarR homologs remain unknown.

We previously reported the biochemical characterization of HucR, a novel MarR homolog from *Deinococcus radiodurans* (17). HucR binds as a homodimer with very high affinity (K\textsubscript{d} = 0.29 nM) to a single 21 bp site containing *E. coli* σ\textsuperscript{70}-like promoter elements driving its own expression and that of an adjacent, putative uricase (*dr1160*). The affinity of HucR for its cognate binding sequence is antagonized by certain anionic lipophilic compounds, most notably uric acid. Analyses *in vivo* indicate that uric acid upregulates transcription of *hucR* and *dr1160* in *D. radiodurans* and suggest that uricase activity in *D. radiodurans* total protein extracts is increased in cells grown in the presence of uric acid.

In this study, we provide the first biochemical investigation into the mechanisms of DNA interaction and ligand response for a member of the MarR family. Spectroscopic evidence reveals that the mechanism by which HucR binds its sequence-specific DNA binding site involves conformational flexibility in both the protein and in the DNA helix. We show that HucR possesses at least two uric acid binding sites with dissimilar ligand-binding affinities. Moreover, our data suggest that the highly conserved ligand-binding site corresponding to SAL-A in MarR
is a weak ligand-binding site in HucR, indicating that uric acid’s antagonistic effect on HucR-DNA interaction is not mediated primarily by site SAL-A.

**Experimental Procedures**

**Circular Dichroism (CD) Spectroscopy**

Complimentary synthetic oligonucleotides were gel purified and annealed to generate a 31 bp fragment (*hucO-31*) containing the HucR binding site, spanning the sequence -23 to +8 relative to the putative transcriptional start site of *dr1160*. CD spectroscopy was performed on an AVIV 202 CD spectrophotometer using a quartz cuvette with a 1 cm pathlength. DNA was diluted to a final concentration of 2 µM in CD buffer (20 mM potassium phosphate, pH 7, 50 mM KCl) and titrated with concentrated HucR to a final concentration of 4 µM monomer. After each titration, samples were equilibrated at 25°C for 15 minutes before data was collected. Ellipticity measurements were collected at 25°C from 340 - 190 nm in steps of 1 nm. All data was corrected for buffer contributions to the signal and samples containing HucR were corrected for additional ellipticity contributions from the protein. Experiments were performed in duplicate (with representative data reported from one experiment).

**Intrinsic Fluorescence Measurements**

Fluorescence emission spectra from 280 - 440 nm were recorded on a Jasco FP-6300 spectrofluorometer with an excitation wavelength of 280 nm at 25°C using a 0.5 cm pathlength cuvette. All experiments were performed with 0.03 mg/ml protein (1.52 µM wild type or mutant HucR) in 40 mM Tris-Cl, pH 8.0, 0.2 mM EDTA, 0.1% BRIJ58, 100 mM NaCl, and 10 mM MgCl₂, unless stated otherwise. Parallel absorbance spectra were recorded for each sample from 190 - 450 nm. For measurements of protein-*hucO-31* complexes, HucR (wt or mutant) was titrated with concentrated *hucO-31*, mixed, and equilibrated for 15 minutes before scanning. For
the protein-ligand binding assays, stock solutions of uric acid, salicylic acid, and acetylsalicylic acid were prepared as described (17). Serial ligand dilutions were prepared such that the addition of 0.5 µl to the protein sample would attain the desired ligand concentration. Samples were mixed and incubated for 60 seconds before each scan.

The corrected protein fluorescence intensity at each wavelength \( (F_{\text{corr}}(\lambda)) \) was obtained from the observed fluorescence by first correcting for dilution from titration and background fluorescence to obtain \( F_c(\lambda) \). Inner filter effects were then resolved by the following correction factor,

\[
F_{\text{corr}}(\lambda) = F_c(\lambda) \times 10^{(A_{\text{ex}}/2 + A_{\text{em}}/2)}
\]

where \( A_{\text{ex}} \) and \( A_{\text{em}} \) are the absorbances at the excitation and emission wavelengths, respectively (20). \( F_{\text{corr}}(\lambda) \) is only reported for samples where both \( A_{\text{ex}} \) and \( A_{\text{em}} \) are less than 0.2. For ligand binding assays where \( A_{\text{ex}} \) exceeded 0.2, normalized fluorescence intensities were obtained by partial correction of the inner filter effect \( (F_{\text{pcorr}}(\lambda)) \),

\[
F_{\text{pcorr}}(\lambda) = F_c(\lambda) \times 10^{(A_{\text{em}}/2)}
\]

before being normalized to the maximal \( F_{\text{pcorr}} \). For fluorescence measurements involving acetylsalicylic acid, \( F_c(\lambda) \) was calculated by incorporating an additional, experimentally determined factor to correct for quenching resulting from the solvent, ethanol.

Percent quenching induced by DNA binding was calculated by:

\[
\%Q = 100(1 - (F_{\text{corr}}(338 \text{ nm}) \ {\text{Protein + DNA}} / F_{\text{corr}}(338 \text{ nm}) \ {\text{Protein}}))
\]

For ligand binding at low ligand concentrations, quenching at 338 nm was calculated by:

\[
\text{quenching (338 nm)} = 1 - (F_{\text{corr}}[X] / F_{\text{corr}}[0])
\]

where \( F_{\text{corr}}[X] \) and \( F_{\text{corr}}[0] \) are the corrected fluorescence intensities at 338 nm for ligand concentrations \( X \mu\text{M} \) and \( 0 \mu\text{M} \), respectively. Uric acid binding isotherms were generated by
nonlinear fits to the Hill equation: quenching (338 nm) = \{n(1/K_d)(U)^{n_H}\} / (1 + (1/K_d)(U)^{n_H}),

where n represents the quenching plateau, U is uric acid concentration, K_d is the observed
dissociation constant, and n_H is the Hill coefficient. Data are reported as the mean ± standard
deviation for two separate experiments.

**Sedimentation Equilibrium**

A 150 µl HucR sample at A_{280} = 0.65 was dialyzed overnight at 4°C against 350 ml of
AU buffer (20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 5 mM MgCl_2). The reference and solution
sectors of an analytical cell with a double-sector centerpiece were loaded with 125 µl AU buffer
and 110 µl HucR sample, respectively. Equilibrium analysis was performed at 20°C at 10,000
rpm using a Beckman Optima XL-A analytical ultracentrifuge equipped with a An-60 Ti rotor.
The cell was scanned at 5-hr intervals at 294 nm with a step size of 0.004 cm until the system
reached equilibrium. The partial specific volume of HucR at 20°C was determined to be 0.7387
cm³/g from the primary amino acid sequence using the program SEDNTERP, and the solution
density calculated to be 1.00120 g/cm³. Equilibrium sedimentation data were analyzed using
Origin Equilibrium software and fit to an equation describing a single ideal protein species.

**Cloning, Overexpression, and Purification of HucR Mutants**

Primers were designed to introduce the desired amino acid substitutions into the hucR
coding sequence via whole plasmid PCR using the recombinant plasmid pHucR, which contains
the hucR coding region as a template (17). Arginine 118 was converted to alanine using primers
R118A-FWD (5’- CTG ATC GAG GCC CGC GAG GAC -3’; nucleotide substitutions yielding
R118 → A in bold) and R118A-REV (5’- GCC CTT TTC GAG CAG CCG CAC- 3’). Serine
104 was converted to alanine using primers S104A-FWD (5’- CCT TCG ACG GCC CGC ACG
ATC -3’; nucleotide substitutions yielding S104 → A in bold) and S104A-REV (5’- CCC GGA
AAT GGC GGC CAG G -3'). PCR products were treated with DpnI and transformed into *E. coli* TOP10 (Invitrogen). The fidelity of the recovered plasmids, identified as pR118A and pS104A, was verified by DNA sequencing. Plasmids pR118A and pS104A were individually transformed into *E. coli* BL21(DE3)pLysS. HucR mutants were expressed, and cells lysed essentially as described for wild type HucR, with the exception that protein expression was induced with 0.8 mM isopropyl-1-thio-β-D-galactopyranoside for 75 minutes (17). Cell lysates were precipitated on ice with ammonium sulfate to 60% saturation and the resulting pellets resuspended in 100 ml of the appropriate buffer: for HucR-R118A purification, HA buffer (20 mM Tris-HCl (pH 7.5), 50 mM KCl, 4.8% glycerol (v/v), 1 mM Na$_2$EDTA, 4.3 mM 2-mercaptoethanol, 0.2 mM phenylmethylsulfonyl fluoride) and for HucR-S104A purification, HAPT buffer (20 mM potassium phosphate (pH 6.0), 50 mM KCl, 10% glycerol (v/v), 1 mM Na$_2$EDTA, 0.02% Tween-20, 4.3 mM 2-mercaptoethanol, 0.2 mM phenylmethylsulfonyl fluoride) and subsequently dialyzed overnight against 40 volumes of the respective buffer (this and all subsequent purification steps were performed at 4°C). Each dialysate was subsequently centrifuged at 11,000 x g for 20 min and the supernatants processed as described below.

For HucR-R118A purification, the supernatant was loaded onto CM-Cellulose and DEAE-Sepharose columns linked in tandem and equilibrated with HA buffer, pH 7.5. The HucR-R118A-containing flow-through and wash fractions were dialysed against HAP buffer, pH 6.0 (same as HA buffer, except that 20 mM potassium phosphate (pH 6.0) replaces Tris-HCl) and loaded onto a CM-Cellulose column equilibrated with HAP buffer, pH 6.0. The HucR-R118A-containing flow-through was loaded onto a hydroxyapatite column equilibrated with HAP buffer, pH 6.0. The HucR-R118A-containing flow-through and wash fractions were loaded onto a heparin column equilibrated with HAP buffer, pH 6.0. HucR-R118A was eluted with a
linear gradient of 50 mM to 1 M KCl in HAP buffer, pH 6.0. Peak fractions were pooled, dialyzed against HAP buffer, pH 6.0, and loaded onto a Cibacron Blue 3GA column equilibrated with HAP buffer, pH 6.0. The column was eluted with a linear gradient of 50 mM to 1 M KCl and peak fractions pooled. HucR-R118A was concentrated, and the KCl concentration reduced to ∼50 mM, using a Centriprep centrifugal filter device. The glycerol concentration was subsequently increased to 20%.

For HucR-S104A purification, the supernatant was loaded onto a CM-Cellulose column equilibrated with HAPT buffer, pH 6.0. The HucR-S104A-containing flow-through and wash fractions were loaded onto a heparin column equilibrated with HAPT buffer, pH 6.0. The HucR-S104A-containing wash was loaded onto a Cibacron Blue 3GA column equilibrated with HAPT buffer, pH 6.0. The column was eluted with a linear gradient of 50 mM to 1 M KCl and peak fractions pooled. HucR-S104A was concentrated and the KCl concentration reduced to ∼50 mM. The glycerol concentration was subsequently increased to 20%.

HucR-R118A and HucR-S104A concentrations were determined spectrophotometrically using $\varepsilon_{280} = 13,512 \text{ M}^{-1} \text{ cm}^{-1}$ (calculated from amino acid extinction coefficients) and verified by SDS-PAGE using bovine serum albumin (BSA) as a standard.

Electrophoretic Mobility Shift Assays (EMSAs)

The DNA used for protein-DNA affinity measurements was prepared by PCR amplification from the *D. radiodurans* genome, as described, and spans a 180 bp sequence containing a single HucR binding site (17). The resulting PCR product, *hucO*, was gel purified and $^{32}$P-labeled with T4-polynucleotide kinase. EMSAs were performed as described for the wild type protein; 0.1 nM $^{32}$P-labeled *hucO* was titrated with up to 15 nM of protein dimer (17). Each binding isotherm was generated by a nonlinear fit of three data sets to the equation: fractional
saturation of $hucO = \frac{n(P)/K_d}{(1 + (P/K_d))}$, where $n$ represents maximal fractional saturation, $P$ is the concentration of free protein dimer, and $K_d$ is the apparent equilibrium dissociation constant. Each $K_d$ is reported as the mean ± S.D. Binding conditions for the competition assays were as described above, except that a slightly larger, gel purified PCR product of 241 bp was used (containing a single HucR binding site). Each binding reaction contained 0.1 nM $^{32}$P-labeled $hucO$ titrated with up to 15 nM unlabeled $hucO$ or pGEM5. Protein was added last to the binding reactions, at a final concentration of 2.5 nM HucR-S104A dimer or 10 nM HucR-R118A dimer. Protein-$hucO$ binding conditions for HucR-S104A and HucR-R118A, analyzing uric acid as an effector, were as described previously for wild type HucR, except that protein dimer concentrations were 3.75 nM and 12.5 nM for HucR-S104A and HucR-R118A, respectively.

Data were fit to an equation for two single-exponential decays: normalized fractional complex formation $= (n_1 e^{-jL}) + (n_2 e^{-kL})$, where $L$ is uric acid concentration, $n_1$ and $n_2$ are fractional complex amplitudes, and $j$ and $k$ are decay constants.

Results

Complex Formation is Accompanied by Conformational Changes in Both HucR and hucO

Previous footprinting analyses of the binding interaction between HucR and its sequence-specific dsDNA binding site suggested that the process involved deformation of the DNA (17). To assess the role of conformational changes in the DNA double helix in the HucR binding mechanism, we used CD spectroscopy to analyze HucR-mediated changes in the secondary structure of a 31 bp DNA fragment ($hucO-31$) containing the 21 bp HucR binding site. Protein contributions to the CD signal in the near UV are minimal; above 240 nm, ellipticity is dominated by contributions from the purine and pyrimidine bases of the DNA duplex. The CD spectrum of $hucO-31$ is characteristic of B-form DNA (Fig. 3.1a). Under stoichiometric binding
Fig. 3.1. **Binding-induced conformational changes in DNA and protein.** (a) Near-UV CD spectra of 2 µM *huCO* in the absence of protein (closed circles) or in complex with 4 µM HucR (open triangles). Data were corrected for protein and solvent contributions to the signal. HucR:*huCO-31* complex formation was verified by an electrophoretic mobility shift assay (data not shown). (b) DNA induced changes in the intrinsic fluorescence spectrum of HucR. The intrinsic fluorescence spectrum of 1.52 µM HucR was measured in the absence or presence of increasing concentrations of *huCO-31*: 0:1 molar ratio of *huCO-31* to HucR (black, closed circle); 0.25:1 (red, closed square); 0.5:1 (green, diamond); 0.75:1 (blue, triangle); 1:1 (orange, inverted triangle).
conditions, HucR induced a significant increase in the magnitude of the positive CD band at 275 nm. Changes in the ellipticity band at 275 nm have been correlated with changes in the DNA winding angle (21-24). The HucR mediated effect on the hucO-31 CD band at 275 nm is consistent with an increase in the number of base pairs per turn of the DNA, or a decrease in the helical twist of the DNA duplex.

Secondary structure prediction algorithms suggest that the HucR monomer is composed of 7 α-helices and two β-strands (25, 26) with helices 4 and 5 comprising the two α-helices of the helix-turn-helix motif. HucR contains two tryptophans located at positions 20 and 72 in the primary amino acid sequence, and two tyrosines at positions 62 and 79. Secondary structure prediction places W20 in helix 1, Y62 in helix 2, and W72 and Y79 in helix 3. Structural modeling of HucR, and structural analysis of the homologs MexR and MarR, places helix 3 in the globular interior of the monomer and helices 1 and 2 at the dimer interface (Fig. 3.2). None of these residues are expected to be in the wing or DNA recognition helix and hence changes in the intrinsic fluorescence spectrum of HucR would indicate conformational changes in the dimerization domains and hydrophobic cores of the homodimer. Using an excitation wavelength of 280 nm, the measured emission signal is dominated by the tryptophan residues due to this amino acid’s much higher quantum yield relative to tyrosine. The intrinsic fluorescence of HucR was measured as a function of hucO-31 concentration to monitor protein conformational changes upon complex formation (Fig. 3.1b). The intrinsic fluorescence spectrum of HucR is characterized by an emission maximum at 338 nm and a “shoulder” at 328 nm. Titration of HucR with hucO resulted in a significant quenching of the entire fluorescence spectrum, with no change in the wavelength of maximal fluorescence or in the relative intensity of the shoulder.
Fig. 3.2. **Model of HucR monomer.** The known structure of the closely related MarR family member, MexR, was used as a template to model HucR via SWISS-MODEL. Residues 178-181 of HucR are not shown.
This quenching effect was saturated at a *hucO*-HucR-monomer molar ratio of 0.5:1, consistent with previous work demonstrating that HucR binds its cognate DNA site as a homodimer.

The equilibrium sedimentation profile of HucR was attained to determine the oligomeric state of HucR in the absence of DNA (Figure 3.3). Data were best fit to a model describing a single, non-associating protein species, yielding a molecular weight average of 43,016 ± 4,374 for HucR. As the calculated molecular weight of a HucR monomer from its primary amino acid sequence is 19,711.56, the equilibrium sedimentation data suggests that HucR exists predominantly as a homodimer in the absence of DNA, consistent with crystallographic and biochemical data for other MarR homologs (12, 19). This supports the accuracy of the $K_d$ value of 0.29 nM that we reported for the HucR dimer- *hucO* association, as the binding mechanism is uncomplicated by monomer-dimer equilibrium.

**Uric Acid-HucR Interactions**

Uric acid was shown to be an efficient negative effector of HucR’s capacity to bind its cognate DNA, with a concentration of 0.26 mM being required to reduce HucR-DNA complex formation to 50%, compared to 5.2 mM for salicylic acid, and 46 mM for acetylsalicylic acid (17). To monitor binding of HucR to its ligands, we measured the intrinsic fluorescence of HucR as a function of ligand concentration. Absorption limitations at the excitation wavelength of 280 nm determined the maximum concentration of ligand for which fluorescence intensities could be fully corrected for inner filter effects. At concentrations up to 50 µM, uric acid induced a strong quenching effect on the emission spectra of HucR. The quenching of the fluorescence emission maximum at 338 nm as a function of uric acid concentration was fit to the Hill equation, yielding an apparent dissociation constant for HucR-uric acid interaction of 11.62 ± 3.71 µM (Fig. 3.4a). The fit provided a Hill coefficient ($n_H$) of 0.73 ± 0.08, suggesting negative cooperativity.
Fig. 3.3. **Equilibrium sedimentation profile of HucR.** The absorbance of 40 μM HucR (monomer equivalents) is shown as a function of the radial cell position (lower panel). Data were fit to a model describing a single, non-associating, species. Residuals to the fit were randomly distributed (upper panel).
Fig. 3.4. **HucR-ligand interactions, at low ligand concentrations.** (a) Quenching of HucR fluorescence at 338 nm is plotted against ligand concentration: uric acid (closed circles, solid line); salicylic acid (closed squares, dashed line); acetylsalicylic acid (closed triangles, dotted line). (b) Scatchard plot: quenching (338 nm)/uric acid concentration is plotted against quenching (338 nm). Double-reciprocal plot (inset).
Analysis of the data using equations for non-equivalent, non-interacting sites resulted in fits with high error, suggesting that measurements of fluorescence quenching at uric acid concentrations below 50 µM predominantly monitors binding at higher affinity binding site(s), and that the observed negative cooperativity is either between equivalent sites in a protein homodimer, or between the higher affinity site and one, or more, lower affinity ligand binding sites within the same monomer. Confirmation of negative cooperativity is provided by linear transforms of the data (Fig. 3.4b). As described by Hensley, Scatchard and double-reciprocal representations of data provide a diagnostic method for qualitative confirmation of cooperativity (27). A Scatchard plot of uric acid-induced quenching of HucR fluorescence at 338 nm reveals a clear deviation from linearity, and the direction of the curvature is indicative of negative cooperativity. The double-reciprocal plot is hyperbolic, also indicating negative cooperativity (Fig. 3.4b) (27, 28). Separate titrations of HucR with salicylic acid and acetylsalicylic acid up to final concentrations of 100 µM resulted in essentially no effect on the intrinsic fluorescence of HucR (Fig. 3.4a), consistent with previous data showing little effect on DNA-HucR complex formation at these compound concentrations (17).

To monitor the effects of higher ligand concentrations on HucR fluorescence, the emission spectra were corrected for inner filter effects of ligand absorbance at emission wavelengths, and absorbance at the excitation wavelength was removed by normalizing the emission spectrum to the maximal fluorescence intensity. In this way, alterations to the HucR emission spectrum could be observed, in addition to the fluorescence quenching observed at low uric acid concentrations. Titration of HucR with uric acid resulted in a gradual red-shift in the wavelength of maximal fluorescence, from 338 nm for unbound HucR to 341 nm at 200 µM uric acid (Fig. 3.5a). This red-shift saturated at 200 µM uric acid, as the fluorescence maximum
Fig. 3.5. **HucR fluorescence spectra, normalized to the fluorescence intensity maximum.** (a) HucR + uric acid; (b) HucR + salicylic acid; (c) HucR + acetylsalicylic acid. Representative spectra from one of two separate experiments, showing HucR in the absence of ligand (black, closed circle) or in the presence of 30 µM ligand (red, closed square), 200 µM ligand (blue, closed diamond), and 1 mM ligand (green, closed triangle). The divergence of the spectrum of HucR + 1mM salicylic acid from the other spectra in panel (b) is likely due to difficulties in correcting for the inherent fluorescence of salicylic acid (wavelength of maximum fluorescence ~410 nm).
remained at 341 nm at uric acid concentrations up to 1 mM. The fluorescence shoulder, characteristic of the HucR spectrum, is sensitive to uric acid concentration. The shoulder intensity, relative to the fluorescence maximum, underwent a gradual quenching with increasing uric acid concentration. These effects on the HucR emission spectrum were not induced by the low affinity ligands, salicylic acid and acetylsalicylic acid (Fig. 3.5b and 3.5c). In comparison to the red-shift observed with uric acid, separate titrations up to 1 mM of each compound did not shift the wavelength of maximum fluorescence from 338 nm. In addition, neither salicylic acid nor acetylsalicylic acid induced quenching of the fluorescence shoulder. From this data, we conclude that three characteristic alterations to the fluorescence spectrum of HucR are indicative of uric acid binding: (1) quenching, observable at low ligand concentrations, (2) red-shift in the wavelength of maximum fluorescence intensity, and (3) quenching of the intensity of the shoulder at an emission wavelength 10 nm below the wavelength of maximum intensity.

**HucR Mutants**

The observation of negative cooperativity in uric acid binding to HucR (Fig. 3.4) is consistent with three possible models: (1) there are at least two separate uric acid binding sites per HucR monomer and the negative cooperativity exists between these sites, (2) there is only one uric acid binding site per HucR monomer and the negative cooperativity is between homologous sites in a homodimeric assembly, or (3) there are multiple uric acid binding sites per monomer and the negative cooperativity is between homologous sites in a homodimer. The co-crystal structure of MarR with salicylate reveals two binding pockets for salicylate within each half of a homodimer, designated as sites SAL-A and SAL-B (12). However, the physiological relevance of these binding pockets remains to be determined, as the protein crystals were saturated at the exceedingly high salicylate concentration of 250 mM.
Primary sequence alignment of HucR with its most closely related homologs from the MarR family revealed a high degree of conservation of residues that form the ligand binding pockets of MarR (12, 17). Most striking in the alignment is the conservation of residues that form ligand binding pocket SAL-A in MarR. The guanidinium group of R118 (numbering according to HucR primary sequence) forms a hydrogen bond with the carboxylate group of salicylate at this site (~3 Å); R118 is strictly conserved in the alignment. The threonine residue that hydrogen bonds to the hydroxyl group of salicylate in site SAL-A (~3 Å) is strictly conserved amongst the HucR homologs in the alignment, but is replaced by a homologous serine in HucR (S104). The hydrogen bonding residues in site SAL-B are less well conserved in the alignment, but HucR retains R109 which coordinates with the carboxylate of salicylate.

To determine whether or not the mechanism of ligand recognition suggested by the MarR co-crystal structure is retained by HucR, we generated two HucR mutants, each containing a single amino acid substitution in ligand binding site SAL-A. HucR-R118A and HucR-S104A were overexpressed and purified to greater than 95% homogeneity, as shown by SDS-PAGE (Fig. 3.6). Secondary structure prediction, sequence alignment against MarR, and structural modeling of HucR (Fig. 3.2) positions R118 in the β-hairpin motif that forms the characteristic “wing” structure of winged helix proteins and positions S104 in the fifth alpha-helix of HucR that forms the DNA recognition helix of the winged helix motif. Therefore, changes in both the ligand- and DNA-binding properties of HucR were expected from each mutation.

**DNA Binding Properties of HucR Mutants**

We measured the affinities of HucR-R118A and HucR-S104A for the HucR binding site, *hucO*. EMSA revealed an apparent K_d of 0.23 ± 0.03 nM (ΔG_assoc = -13.0 kcal/mol) for the interaction of HucR-S104A with *hucO*, nearly identical to the apparent K_d of 0.29 ± 0.02 nM.
Fig. 3.6. **Purified HucR mutants.** Lane 1, molecular mass marker in kDa; lane 2, 1 μg of purified HucR-R118A; Lane 3, 1 μg of purified HucR-S104A.
previously measured for wild type HucR under identical conditions (Fig. 3.7). The affinity of HucR-R118A for *hucO* was reduced approximately five-fold, revealing an apparent *K_d* of 1.60 ± 0.14 nM (*ΔG_{assoc} = -11.9 kcal/mol*). The complex of each HucR mutant with *hucO* dissociates during electrophoresis, as reflected in a fractional saturation less than 100%. In addition, whereas only a single protein-DNA complex is observed in each HucR and HucR-R118A titration, a second complex may be distinguished at high HucR-S104A concentrations.

The protein-DNA complex instability observed for both HucR-S104A and HucR-R118A suggested the possibility that the DNA binding specificities of the HucR mutants were reduced, relative to the wild type. A measure of the sequence specificity in binding by HucR-S104A and HucR-R118A was obtained by a direct comparison of the abilities of unlabeled *hucO* and pGEM5 to compete for protein binding with 0.1 nM ³²P-labeled *hucO* (Fig. 3.8). Each HucR mutant is highly sequence specific in its DNA binding capacity, as reported for the wild type protein. Titration of unlabeled *hucO* to a concentration of 5 nM elicited a clear reduction in labeled complex formation with HucR-S104A and HucR-R118A. However, when using pGEM5 as a competitor, only a slight reduction in the labeled *hucO*-protein complex is observed at a concentration of 15 nM pGEM5, for each mutant. Since pGEM5 contributes a greater than twelve-fold molar excess of base pairs compared to *hucO*, the competition assays demonstrate the retention of a definite sequence specific binding preference for both HucR mutants.

Consistent with the apparent DNA binding constants for the HucR variants (Fig. 3.7), the DNA-responsive intrinsic fluorescence spectra of the proteins suggest altered mechanisms of DNA association (Fig. 3.9). A comparison of the intrinsic fluorescence spectra of HucR, HucR-S104A, and HucR-R118A reveals significant differences. Whereas the measured fluorescence maximum is at 338 nm for each protein, the peak intensity is 86.2 for HucR, 73.2 for HucR-
Fig. 3.7. **Affinity measurements of HucR mutants for the HucR binding sequence (hucO).** Panels a-c depict EMSAs in which 0.1 nM hucO was titrated with protein. Uncomplexed hucO and protein-DNA complexes are identified by arrows. (a) Wild type HucR: 0, 0.05, 0.125, 0.25, 0.5, 1.0, 2.5, 5.0, 15.0 nM (b) HucR-S104A; protein concentrations as for panel a (c) HucR-R118A: 0, 0.25, 0.5, 2.5, 5.0, 10.0, 15.0, 20.0 nM (d) Binding isotherms for HucR-S104A (solid line, closed circles) and HucR-R118A (dashed line, closed squares). Data were collected in triplicate. Fractional saturation is reported as the mean ± S.D.
Fig. 3.8. **Competition assays.** Binding reactions involved 0.1 nM $^{32}$P-labeled $hucO$ and titration of up to 15 nM unlabeled $hucO$ or pGEM5, followed by the addition of protein. Complexes were resolved by EMSAs. The first lane in each gel contains only labeled DNA. Concentrations of unlabeled competitor DNA are indicated above the corresponding lanes. Complexed and uncomplexed $hucO$ are identified by arrows. (a) HucR-S104A (b) HucR-R118A.
Fig. 3.9. **Fluorescence spectra of free and DNA-bound HucR-variants.** The intrinsic fluorescence spectrum is shown for 1.52 µM of each HucR variant before (solid line, closed circles) or after (dashed line, open triangles) incubation with 0.76 µM hucO-31. (a) Wild type HucR. (b) HucR-S104A. (c) HucR-R118A.
S104A, and 111.6 for HucR-R118A, suggesting that the environments of either one, or both, of the tryptophans differ in these HucR variants. When complexed with a stoichiometric quantity of hucO, such that all protein is expected to be in complex with DNA, fluorescence quenching is observed for each protein, indicating altered environments for one, or both, tryptophans. However, the degree of quenching induced by complex formation differs significantly between the proteins. HucR interaction with its cognate DNA binding site leads to a quenching of intrinsic fluorescence by 10.0%, compared to 5.4% for HucR-S104A, and 22.2% for HucR-R118A. From these two observations, we conclude that protein conformational changes upon DNA binding by of HucR, HucR-S104A, and HucR-R118A also differ.

Uric Acid Binding by HucR Mutants

As both R118 and S104 mediate hydrogen bond contacts with a ligand in the crystal structure of MarR, and as both of these residues are highly conserved in the MarR family, we also analyzed the roles of each of these side chains in uric acid recognition by HucR. For each HucR mutant, increasing concentrations of uric acid reduced protein-hucO complex formation (Fig. 3.10a,b). However, in contrast to the case of wild type HucR, where DNA complex formation was abolished in the presence of approximately 1 mM uric acid, we observed faint smears extending above the band of free hucO at uric acid concentrations up to 10 mM. This effect was shown not to be a result of uric acid or its solvent affecting the DNA migration (data not shown). Additional controls confirmed that the uric acid solvent did not affect protein-hucO complex formation (data not shown). We therefore interpreted the smearing above the band of free DNA to be residual DNA complex formation with HucR-S104A and HucR-R118A. At uric acid concentrations below 1 mM, a steep decrease in normalized DNA complex formation is seen for all three variants (Fig. 3.10c). At uric acid concentrations above 1 mM, the antagonistic
Fig. 3.10. **DNA-binding antagonism by uric acid.** Protein-*hucO* association was challenged with increasing concentrations of uric acid, and complexes resolved by EMSA. Reactions contained 0.1 nM $^{32}$P-*hucO*. Uric acid concentrations are 0, 0.1, 0.5, 1, 5, 10 mM (a) HucR-S104A. (b) HucR-R118A. (c) normalized protein-*hucO* complex formation as a function of uric acid concentration for wild type HucR (solid line, closed circles), HucR-R118A (dashed line, closed squares), and HucR-S104A (dotted line, closed triangles). Experiments were performed in triplicate and data reported as mean ± standard error.
effect of uric acid approaches saturation; for wild type HucR, fractional complex saturation approaches zero while fractional saturation for the HucR-mutants near saturation at \( \sim 30\% \). At 10 mM uric acid, approximately 20\% of the HucR-S104A complex remains and approximately 10\% of the HucR-R118A complex remains.

In addition to affects on DNA affinity, the S104A and R118A amino acid substitutions in HucR, therefore, each appear to reduce the protein’s response to high concentrations of uric acid. This data is consistent with there being at least two separate uric acid binding sites in HucR with dissimilar ligand affinities. Our data suggests that the mutations at HucR positions 104 and 118, corresponding to ligand binding pocket SAL-A in MarR, each affect a lower affinity uric acid binding site.

The interaction of uric acid with HucR-S104A and HucR-R118A was measured by intrinsic fluorescence (Fig. 3.11a). For HucR-S104A, the apparent \( K_d \) for interaction with uric acid was \( 7.4 \pm 0.5 \mu M \), compared to \( 9.7 \pm 3.2 \mu M \) for HucR-R118A, and \( 11.6 \pm 3.7 \mu M \) for HucR. Interestingly, the negative cooperativity in uric acid binding by HucR was reduced, or lost, through the S104A and R118A mutations. Whereas HucR displayed an \( n_H \) of 0.73 \( \pm 0.08 \), fits for HucR-S104A and HucR-R118A revealed \( n_H \) values of 1.10 \( \pm 0.06 \) and 0.94 \( \pm 0.16 \), respectively. As for the wild type protein, fits of the data for the mutants to an equation for non-equivalent, independent sites resulted in high error. As measurements of fluorescence quenching at low concentrations of uric acid predominantly monitors binding at a high affinity site(s), the unchanged, or small increase in, affinity for uric acid observed for both HucR mutants is consistent with each mutation residing in a low affinity uric acid binding site.

We then compared the effects of uric acid on the normalized fluorescence spectra of HucR, HucR-S014A, and HucR-R118A. The uric acid induced red-shift in the wavelength of
Fig. 3.11. **Fluorescence detection of uric acid binding to HucR-mutants.** (a) Binding isotherms for protein-uric acid interactions. Quenching of protein fluorescence at 338 nm is plotted against uric acid concentration for wild type HucR (solid line, closed circles), HucR-R118A (dashed line, closed squares), and HucR-S104A (dotted line, closed triangles). (b and c) Fluorescence spectra, normalized to the fluorescence intensity maximum, for protein in the absence (black, closed circles) or presence of 30 µM (red, closed squares), 200 µM (blue, closed diamond), and 1 mM (green, closed triangles) uric acid. (b) HucR-S104A. (c) HucR-R118A.
maximal fluorescence intensity for HucR was also observed in both HucR mutants (Fig.
3.11b,c). For each protein, the wavelength of maximal intensity shifted from 338 nm, for the
apo-protein, to 341 nm at 200 µM uric acid and remained at 341 nm at a ligand concentration of
1 mM. The reduction of the shoulder fluorescence intensity relative to the peak intensity
observed in HucR, was also retained in both HucR mutants, even at low uric acid concentrations.
We therefore concluded that these effects on the fluorescence spectra were induced by ligand
binding at a site that was unaltered by the separate mutations at S104 and R118. Moreover, as the
red-shift and shoulder-quenching effects were observed at ligand concentrations below 50 µM,
we conclude that these effects correspond to binding at a high affinity site(s) and that the S104A
and R118A mutations perturb a low-affinity ligand binding site(s).

Discussion

HucR-hucO Association – Conformational Changes

The conformational changes observed in HucR and the DNA upon binding indicates that
the interaction is not a static one. Rather, a DNA distortion is observed that may indicate a
decrease in the duplex winding angle (i.e. an increase in the number of base pairs per turn) (Fig.
3.1). This finding is consistent with the mechanism of site specific DNA binding by the winged
helix protein MerR from *E. coli* which has been shown to decrease the helical twist of the DNA
duplex, in addition to inducing a 25° bend, upon complex formation (29). Binding to DNA by a
*Bacillus subtilis* winged helix protein from the MerR family, BmrR (in its activated ligand bound
state), is accommodated by a significant distortion in the cognate DNA binding site, with a
reduction in the helical twist and a bending of the DNA duplex by ~50° (30). Insertion of the
recognition helix of the *E. coli* helix-turn-helix protein TetR into the major groove causes a
widening of the major groove by 2.8 Å and a concomitant reduction in the DNA helical twist,
but does not alter the overall curvature of the DNA binding site (31). Similarly, the *Staphylococcus aureus* protein QacR, from the TetR family, does not bend DNA upon binding, but insertion of the recognition helix into the major groove induces a significant reduction in the helical twist to 32.1° per bp (and 11.2 bp per turn) and a concomitant widening of the major groove (32). As the center of each inverted repeat in the HucR binding site is separated by one full turn of the double helix, it is tempting to speculate that unwinding of the double helix is required to accommodate insertion of the recognition helices of a HucR homodimer into consecutive, widened major grooves. In the absence of a crystal structure of a DNA-bound MarR homolog, this is the first piece of evidence suggesting that untwisting of the DNA duplex and concomitant widening of the major groove is required to accommodate binding by these proteins. Our finding of HucR-induced changes to DNA structure is interesting in comparison with suggestions that MarR-DNA binding is accommodated strictly by protein conformational changes (33). However, these suggestions were based on observations that MarR-binding to DNA did not confer sites of hypersensitivity to DNaseI, which does not preclude the possibility of alterations to the helical twist of the DNA duplex (8, 34). However, it is interesting to note that DNA binding by the archaecal winged helix protein Sac10a from *Sulfolobus acidocaldarius* induces the opposite effect on the DNA duplex; the DNA is distorted such that the number of base pairs per turn decreases (winding angle increases) (35).

Quenching of the intrinsic fluorescence of HucR upon complex formation with *hucO* suggests conformational changes in the homodimer that extend beyond the helix-turn-helix and wing motifs that are predicted to make all direct contacts with DNA; the two HucR tryptophans are expected to be in the globular interior of the protein and at the dimer interface (Fig. 3.1b, 2). Quenching of the intrinsic fluorescence of the indole ring of tryptophan has been associated with
conformational shifts to three altered environments: (1) increased solvent exposure, (2) closer contact with amide bonds in the peptide backbone, or (3) closer contact with polar side chains (36, 37). Additional work will be required to elucidate the nature of the conformational changes induced in HucR and its cognate DNA binding site upon complex formation.

**HucR DNA-Binding Mechanism**

Given the predicted locations of S104 and R118 in the DNA recognition helix and wing of HucR, respectively, it is not surprising that alanine substitutions at these positions affected HucR-*hucO* association (Fig. 3.7). The protein-DNA complex stability for each HucR-mutant, relative to that of the wild type protein, is reduced on the gel; we cannot say if this relative instability is also present in solution. However, an increased $k_{off}$ for the complex equilibria suggested from EMSA would explain the increased $K_d$ observed for HucR-R118A, relative to HucR. Also consistent with the increased apparent $K_d$ for HucR-R118A is the intrinsic fluorescence data suggesting an altered DNA-binding mechanism for this mutant: the greater intrinsic fluorescence of HucR-R118A relative to HucR and the proportionally greater extent of quenching induced by DNA-binding is consistent with an altered protein conformation that must undergo a different conformational change to accommodate DNA-binding (Fig. 3.9). The intrinsic fluorescence data also offers an explanation for the nearly identical apparent $K_d$ values for the S104A and wild type HucR-variants (Fig. 3.9). The combination of the lower intrinsic fluorescence of HucR-S104A, relative to HucR, and the proportionally lower degree of quenching induced by DNA-binding suggests that the state of the free protein is closer to the DNA-binding conformation, perhaps leading to a higher $k_{on}$ for complex formation. The combined effects of an increased $k_{off}$, suggested by EMSA, and an increased $k_{on}$ is consistent with the essentially unaltered apparent $K_d$ for HucR-S104A. In addition, the modest alterations to
DNA binding affinities observed in our HucR mutants is consistent with mutational analysis of MarR, in which the selection of trans-dominant mutants that interfered with the activity of wild type MarR did not isolate individual mutations of the arginine or threonine corresponding to HucR positions 118 and 104 (18).

These effects, induced by the individual mutations, upon HucR’s site specific DNA-binding support the prediction that HucR mediates binding via a winged helix fold. In addition, the decreased affinity for the HucR binding site, elicited by the R118A mutation, strongly suggests that the $\beta$-hairpin wing, and not just the recognition helix, makes direct contacts with the DNA duplex. Moreover, the high degree of amino acid conservation in the wing region of MarR homologs suggests its importance in directing association with DNA (17). It will be interesting to determine whether the HucR wing motif contacts the minor groove, as seen with the prokaryotic winged helix proteins MuR and FadR, or with the sugar-phosphate backbone, as observed with BmrR, or within the major groove, as observed with the eukaryotic protein, RFX1 (30, 38, 39).

**HucR-Uric Acid Association and Negative Cooperativity**

The apparent $K_d$ of $11.62 \pm 3.71 \mu M$ for binding of uric acid to HucR indicates a ligand-binding affinity significantly higher than has been observed for MarR. The apparent dissociation constants for MarR binding to the phenolic compounds salicylate, plumbagin, 2,4-dinitrophenol, and menadione have been measured to be 500 $\mu M$, 250 $\mu M$, 250 $\mu M$, and 800 $\mu M$, respectively (8, 11, 33). However, the HucR-uric acid affinity is similar to the drug binding affinities of the MarR homolog, EmrR, which binds phenolic ligands with $K_d$ values ranging from approximately 2.0 – 15.0 $\mu M$ (40). The ligand dissociation constants for the multi-drug binding protein QacR, range from 0.1 – 5 $\mu M$ (33).
Our uric acid binding assays lead us to conclude that the HucR monomer contains at least two separate uric acid binding sites with dissimilar affinities, and that the HucR site corresponding to the MarR ligand-binding pocket SAL-A binds uric acid with low affinity. These conclusions are based upon the following observations: 1) Uric acid-induced quenching of HucR intrinsic fluorescence reveals negative cooperativity in uric acid binding, indicating that at least two uric acid binding sites are involved (Fig. 3.4). Uric acid quenching predominantly monitors binding at a high affinity site(s) as indicated by the poor fit of the data to equations for multiple, non-equivalent, non-interacting sites, suggesting that negative cooperativity is between homologous sites in each monomer. 2) The S104A and R118A mutations result in little to no change in uric acid binding affinity and abolish cooperativity of uric acid binding, consistent with each mutation residing in the same binding site (Fig. 3.11a). 3) The alterations to the normalized fluorescence spectra observed at uric acid concentrations up to 1 mM were nearly identical for all three HucR variants, indicating that these effects were induced by ligand binding at a site unaltered by the separate S104A and R118A mutations (Fig. 3.5; 3.11b,c). Moreover, as these alterations to the emission spectra also occurred at uric acid concentrations below 50 µM, we conclude that these effects correspond to binding at high affinity site(s). 4) Increasing concentrations of uric acid up to 1 mM result in a similar, steep antagonism of HucR-hucO complex formation for HucR, HucR-S104A, and HucR-R118A (Fig. 3.10). However, unlike HucR, residual protein-DNA complex is observed for both HucR mutants up to 10 mM uric acid. These observations are consistent with each mutation residing in a low affinity ligand binding site. Saturation of this lower affinity site in wild type HucR nearly completely abolishes complex formation with DNA at a uric acid concentration of 2 mM, whereas approximately 30% of the initial protein-DNA complex remains at this ligand concentration for both mutants.
The loss of uric-acid binding cooperativity induced by each mutation, combined with the differences in the intrinsic fluorescence spectra of the HucR variants (Fig. 3.9), suggests that each mutation propagated a conformational change in HucR that disrupted the cooperativity between the high affinity site(s) in each half of the homodimer. Alternatively, uric acid binding at the low affinity site (containing HucR residues S104 and R118) might stabilize the recognition helix in a position such that ligand binding at the high affinity site might reduce the affinity of the corresponding site in the adjacent monomer. This latter model is consistent with the expected ligand-binding site SAL-B in HucR. The data does not rule out the possibility that the negative cooperativity in HucR is between the low affinity site comprised of residues S104 and R118, and a separate high affinity site(s). However, ligand binding in the low-affinity, SAL-A site is unlikely to be involved in negative cooperativity, based on our observation that negative cooperativity is seen at uric acid concentrations below 50 µM, where site SAL-A is probably not filled.

It has been proposed that the physiological benefit of negative cooperativity is that it increases the concentration range over which the protein can respond (36, 41). In addition, negative cooperativity can increase the protein’s sensitivity to low ligand concentrations. Given the tendency of uric acid to precipitate and the consequential deleterious effects to the cell, it would likely be advantageous for D. radiodurans to upregulate uricase in response to low uric acid concentrations.

Possible Basis for Changes in the Fluorescence Spectrum of HucR on Ligand or DNA Binding

W20 of HucR is predicted to reside in an N-terminal α-helix that does not exist in either MarR or MexR. As the N-terminal helices of both MarR and MexR form their respective hydrophobic dimer interfaces, it is likely that the extra N-terminal helix of HucR serves in such a
capacity. It is likely that this helix is quite flexible in the uric acid-mediated allosteric mechanism that regulates DNA binding of HucR. Structural determination of MarR was based upon a single homodimer in the asymmetric unit (12). The DNA binding domains of the ligand-bound structure were in a conformation that precluded DNA binding to the known, palindromic binding site (8, 12). Movement of the DNA binding domains to accommodate binding to the sequence-specific sites would require significant distortion of MarR α-helix 1 and the C-terminal α-helices. Structural determination of MexR allowed for multiple views of this protein, as the asymmetric unit was comprised of 4 separate MexR homodimers. There is significant flexibility in the MexR dimerization domain, as observed in a 17° deviation in the helix orientation of α-helix 1, and a 4.9 Å shift in its helix midpoint position. Moreover, as dimerization of both MexR and MarR are mediated primarily by van der Waals contacts involving their respective N- and C-terminal helices, flexibility in this domain is expected as such hydrophobic interactions do not require specific molecular orientations. Except for the “wing” itself, the MexR winged-helix DNA binding domain appears to be considerably less flexible; rather, each domain moves as a rigid body relative to the homologous domain in the dimer. HucR α-helix 3, which contains W72, corresponds to the first helix in this motif. The alterations of the HucR fluorescence spectrum observed upon DNA binding and ligand binding are therefore consistent with conformational changes in the dimerization domain containing HucR tryptophan 20 and/or rigid movements of HucR helix 3 in the winged-helix DNA binding domain.

Proposed Mechanism for Uric Acid-Mediated Allosteric Control of HucR-DNA Complex Formation

Though there is considerable variability in the mechanisms of allosteric control of the DNA-binding properties of homodimeric winged helix proteins, the effector binding domains in
these proteins are generally distal to the winged helix DNA binding domains, and ligand-induced conformational changes are propagated to the winged helix motif (39). Such domain separation is observed in MerR and other MerR family members including BmrR and MtaN as well as in FadR from the GntR winged helix family (33). For the CRP protein, signals are transmitted from the N-terminal cAMP-binding domain to the C-terminal DNA-binding domain (42). Likewise, the phosphorylated “receiver” domain of OmpR is separated by a linker region from its winged helix DNA-binding domain (43). Indeed, all characterized prokaryotic multidrug binding transcriptional regulators other than MarR incorporate a spatial separation of ligand- and DNA-binding domains (33, 44). Even though the MarR family is proposed to share a common evolutionary history with the GntR, MerR, and Crp families, MarR proteins appear to have an interesting structural organization such that the ligand- and winged helix DNA-binding domains almost completely overlap in the residues involved.

The MexR structure reveals an “open” dimer conformation in which basic charge repulsions between the two DNA binding domains maintains the dimer in a conformation that can be docked onto the known MexR DNA binding site. Comparison to another MexR dimer suggested a mechanism of allosteric control in which the binding of a potential effector between the two DNA binding domains perturbs the relative orientation of the monomers to become incompatible with DNA binding. The mechanism involves effector-mediated neutralization of the charge repulsion between the DNA binding domains and hydrophobic interactions between the effector and protein side chains, with the result that the distance between the DNA binding domains is reduced to be no longer optimal for DNA binding. Support for this mechanism comes from the MarR-salicylate co-crystal structure, which reveals the orientation of the DNA-binding lobes to be incompatible with binding to the MarR binding sites. Indeed, the DNA-binding
capacities of a number of homodimeric winged helix proteins are mediated by allosteric changes in the proximal distance between relatively rigid DNA binding lobes (39). In the multidrug-binding protein QacR, ligand binding induces a coil-to-helix transition that propagates a protein conformational change resulting in a $36.7^\circ$ rotation and $9.1\,\AA$ translation of the DNA-binding domain; the altered distance between the recognition helices of the homodimer precludes binding to its DNA site (44).

As for MexR, MarR, and other winged-helix proteins, HucR is strongly electropositive in its proposed DNA binding region. Uric acid could serve to neutralize charge repulsions that maintain the DNA binding lobes of the HucR dimer in a DNA binding conformation. Uric acid’s hydrophobic rings could also potentiate van der Waals interactions with hydrophobic HucR side chains, further distorting the relative conformations of the monomers. It is likely that the uric acid binding site mediating such conformational events corresponds to ligand binding site SAL-B in MarR. Ligand binding site SAL-A is located between the DNA recognition helix and the wing region, whereas site SAL-B is located on the other side of the DNA recognition helix; the two SAL-B sites in the homodimer would be adjacently positioned in the channel between the DNA binding lobes (6 Å wide in MarR structure). Binding of uric acid at such a position could potentially disrupt the relative orientation of the HucR DNA binding domains. Such a perturbation of the spatial separation of the recognition helices in the homodimer would be dependent upon the flexibility of the dimerization domain suggested by the fluorescence data.

Support for our observation that site SAL-A in HucR is a low affinity ligand binding site and our hypothesis that site SAL-B mediates allosteric events comes from the fact that the residues that comprise ligand binding site SAL-A, but not SAL-B, are highly conserved in MexR and MexR does not respond to salicylic acid as an effector (19). The conservation of residues in the
proposed site SAL-B in HucR, relative to site SAL-B in MarR, is greater than for MexR. The observed differences may explain the high affinity of HucR for uric acid.

The significant alterations in the fluorescence spectrum of HucR that are induced upon uric acid and DNA binding likely correspond to movements of both tryptophans (though we know not to which degree each contributes). According to this hypothesis, uric acid would distort the relative positions of the DNA binding domains in the homodimer, thus altering the environment of tryptophan 72. Such movements of the winged-helix domains would require flexibility of the N-terminal α-helices of each monomer, altering the environment of HucR tryptophan 20. Similarly linked movements would be expected for the HucR dimer to adopt a DNA binding conformation.

References


CHAPTER 4

SUMMARY AND CONCLUSION

Interest in the MarR family of transcriptional regulators has increased since the eponym, MarR from *E. coli*, was identified as a repressor of a chromosomally encoded operon that confers intrinsic resistance to structurally diverse drugs, organic solvents, and oxidative stress agents (1-4). Gene regulatory activity by MarR homologs is mediated by DNA-binding via a conserved winged helix motif that likely originated before the divergence of archaea and bacteria ~3.5 billion years ago (5, 6). MarR proteins appear to be especially well adapted to binding anionic lipophilic compounds, particularly aromatic ligands, and this may prove to be a defining feature of all members of this family. The identification of two putative MarR homologs in the *D. radiodurans* genome, *dr1159* from chromosome I and *dra0248* from chromosome II, suggested the presence of novel phenolic-sensitive regulatory systems in this eubacterium. Particularly enticing, given the extreme resistance of *D. radiodurans* to radiation, desiccation, and oxidative stresses, is the involvement of numerous MarR homologs in stress response systems.

DNA-Binding by HucR

The high affinity of HucR for its cognate DNA-binding sequence (*hucO*) ($K_d = 0.29 \pm 0.02$ nM) (Fig. 2.5) is comparable to the binding affinities of ExpG ($K_d = 0.58-1.3$ nM) and MarR ($K_d \sim 1$ nM) for their respective sites (7, 8). The homolog PecS, however, binds site-specifically with apparent $K_d$’s ranging from 5 nM to $\sim 200$ nM (9, 10). Despite the dearth of quantitative DNA-binding characterization for MarR proteins, it is clear that these proteins bind site-specifically to palindromic, or pseudopalindromic, sequences, consistent with the fact that most prokaryotic transcription factors bind DNA as homodimeric assemblies. HucR binds *hucO* as a dimer (Fig. 2.5)
2.5D and Fig. 3.1) and exists as a dimer in solution (Fig. 3.3) consistent with other evidence that MarR proteins are predominantly dimers in the absence of DNA (7, 8, 11, 12).

Sequence alignment of HucR with its most closely related homologs indicates high conservation of residues in the predicted β-hairpin (wing) motif in the DNA-binding domain, suggesting the importance of this structure in gene regulatory activity (Fig. 2.2). Binding site sizes for most characterized MarR homologs range from 21 bp (for HucR) (Fig. 2.6) to 29 bp (for MexR), similar to the regions of DNaseI protection observed for most dimeric, canonical HTH proteins (13-15). However, biophysical, genetic, and structural data clearly implicates the involvement of the wing motif in mediating DNA contacts in most winged helix proteins (16-18). Accordingly, the substitution of an arginine with an alanine in the wing motif of HucR (HucR-R118A) decreased the binding affinity for hucO ∼5-fold (Fig. 3.7). Moreover, HucR-S104A and HucR-R118A demonstrated reduced DNA-complex stability on the gel, but not reduced sequence-specificity (Fig. 3.7 and 3.8). It is likely that the reduced complex stabilities observed on the gel are a consequence of the amino acid substitutions inducing global conformational changes propagated throughout the protein, as suggested by the altered intrinsic fluorescence spectra of the HucR variants (Fig. 3.9). Thus, the high affinity and stability of the HucR-hucO complex appears to depend not only on specific sequence and structural relationships in the winged helix domain, but also on the conformations of the residues that comprise the dimerization domain. Atomic force microscopy data suggests that DNA bending is critical in the binding mechanism of the MarR homolog, ExpG (7). We provide the first evidence that alterations to the helical twist of the DNA duplex may also be part of the DNA-binding mechanisms of MarR homologs (Fig. 3.1).
Negative Cooperativity in Uric Acid Binding

Interestingly, the MarR-salicylate co-crystal structure identified residues from the winged helix DNA-binding domain mediating all of the H-bond and hydrophobic contacts with the two salicylate molecules bound at sites SAL-A and SAL-B (19). This co-localization of ligand- and DNA-binding domains is unusual in transcriptional regulators, and the physiological relevance of the MarR structural data is questionable, as the protein crystals had been soaked in a relatively concentrated solution of salicylate. Nevertheless, residues comprising both sites are conserved in the MarR family, particularly in site SAL-A (Fig. 2.2) (19). Whether such conservation is related to the importance of these residues in ligand- and/or DNA-binding had not been rigorously analyzed.

Our results indicate that in HucR, site SAL-A is a weak uric acid binding site and that a high affinity uric acid binding site(s) exist elsewhere in this protein (Fig. 3.4 and 3.11). This notion is supported by observations that mutations at site SAL-A resulted in only modest alterations to uric acid-responsive HucR-\( hucO \) complex formation at low uric acid concentrations, but allowed residual protein-DNA complex at uric acid concentrations that eliminated wild type complex (Fig. 3.10). It is likely that site SAL-B is a high affinity ligand binding site, and the relatively reduced conservation of residues at this site might be the consequence of adaptation to structural variation of the phenolic ligands. To address this hypothesis, efforts are underway to purify a HucR variant with an arginine to asparagine substitution at position 106, corresponding to M74 in MarR site SAL-B. Additional efforts are underway to generate a HucR variant with an arginine to asparagine substitution at position 109, corresponding to R77 in MarR site SAL-B.

Our data suggesting a high affinity ligand binding site, possibly corresponding to site SAL-B, in conjunction with our observations of intrinsic flexibility in HucR (Fig. 3.1 and 3.5)
supports a model of ligand-responsive allosteric control similar to the one proposed by Lim (11). Uric acid binding at a high affinity site could alter the relative orientations of the DNA-binding lobes of the HucR dimer such that DNA-binding is no longer accommodated. Our data indicating that HucR mutations at positions 118 and 104 attenuate, or remove, the negative cooperativity in uric acid binding, in concert with the fluorescence data suggesting altered intrinsic conformations for the HucR variants, is consistent with uric acid binding with negative cooperativity at a high affinity site in each monomer (Fig. 3.9 and 3.11). Ongoing efforts by Tee Bordelon in the lab of Dr. Marcia Newcomer to obtain a high-resolution crystal structure of HucR will likely enhance our understanding of the allosteric mechanism by which uric acid modulates DNA-binding activity of HucR.

**Physiological Considerations**

HucR is a uric acid-responsive autoregulator and regulator of the expression of a putative uricase enzyme (Fig. 2.7 and 2.8). Uricase converts uric acid to allantoin in the purine degradation pathway in both prokaryotes and eukaryotes. The genetics of purine catabolic pathways have been delineated in relatively few prokaryotic species, including *Bacillus subtilis* and *Methanococcus vannielii*, but indicate that they are diverse, and include both aerobic and anaerobic systems (20-22). However, catabolism beyond uric acid occurs only in the aerobic pathways, as O₂ is a required electron acceptor in uricase activity (22). It is worth noting that *D. radiodurans* genes *dr1160* and *dr1161* encoding the putative uricase and a predicted, transthyretin-like protein share significant homology to the genes *pucL* and *pucM* from *B. subtilis*, respectively, which have been shown to encode uricase activity (21). Moreover, uricase activity was detected in total protein extracts from *D. radiodurans* (Fig. 2.8), and we are in the process of cloning the HucR-regulated gene encoding this putative uricase to confirm that it is a
functional enzyme. It is likely that *D. radiodurans* possesses a full complement of genes for purine catabolism, as a genome search identifies the following putative enzymes of aerobic purine catabolism: adenine deaminase (*dra0270*), xanthine dehydrogenase (N-terminal subunit, *dra0177*; C-terminal subunit, *dra0178*), allantoinase (*dr1153*), and N-carbamyl-L-amino acid amidohydrolase (*dr1154*).

Upregulation of uricase activity in response to increased concentrations of its substrate, uric acid, is logical given the low solubility of this compound. However, given the ability of uric acid to scavenge damaging reactive oxygen species (23, 24), it seems beneficial to the cell for this uric acid-modulated response to be attenuated, such that an optimal concentration of uric acid is maintained in the cell below a critical threshold where precipitation would occur. Though it remains to be determined how many uric acid binding sites must be filled in the HucR dimer to inhibit HucR-*hucO* complex formation, the negative cooperativity observed in uric acid binding in HucR could be critical in maintaining optimal uric acid levels in the cell. Negative cooperativity could extend the uric acid concentration range over which HucR can respond, by effectively increasing the uric acid concentration at which all HucR-mediated uricase repression is relieved, thus allowing a basal level of uric acid to remain in the cell. Simultaneously, negative cooperativity of uric acid binding would effectively decrease *hucR* transcription at low uric acid concentrations, likely to maintain an optimal level of uricase in the cell. RNA dot blot analysis suggests that *hucR* and *dr1160* transcript levels are proportionally upregulated in response to uric acid (Fig. 2.8), but a full understanding of this tight regulatory system will require knowledge of the relative transcriptional and translational efficiencies of these genes, as well as the half-lives of their protein products.
The ability of *D. radiodurans* to survive extreme damage to its DNA requires that it be able to cleanse itself efficiently of the reactive oxygen radicals that result from ionizing and UV radiation, desiccation, and oxidative stress agents (25). In addition, *D. radiodurans* must quickly degrade, export, or recycle damaged DNA components to prevent mutations. As a uric acid responsive transcriptional regulator, HucR is a likely candidate in contributing in both capacities.

**References**


VITA

Steven Wilkinson was born in Chesterfield, Virginia, where he attended Clover Hill High School. In the fall of 1991, he entered The College of William & Mary in Williamsburg, Virginia, and received his Bachelor of Science in biology in December 1995, with a minor in anthropology. He enrolled in the biochemistry doctoral program in the Department of Biological Sciences at Louisiana State University in the fall of 1999. In the spring of 2000 he joined the laboratory of Dr. Anne Grove where his primary research involved the biochemical characterization of a novel transcriptional regulator from the radiation-resistant eubacterium, *Deinococcus radiodurans*. As a graduate student, he taught laboratory courses in microbiology, protein biochemistry, and the biochemistry of nucleic acids. In addition, he served as a graduate assistant in the scope-on-a-rope outreach program, funded by the Howard Hughes Medical Institute. He will complete the requirements for the Doctor of Philosophy degree in biochemistry in May 2005. In June 2005 he will join the laboratory of Dr. E. Peter Geiduschek at the Center for Molecular Genetics at the University of California, San Diego as a postgraduate research scientist.