Exploring the determinants of metal sensing in Salmonella typhimurium using FrmR, a non-metal sensing RcnR/CsoR family member

CECILIA PIERGENTILI

Exploring the determinants of metal sensing in *Salmonella typhimurium* using FrmR, a non-metal sensing RcnR/CsoR family member

CECILIA PIERGENTILI

School of Biological and Biomedical Sciences

Durham University

England

October 2015

Thesis submitted in fulfilment for the degree of Doctor of Philosophy

Exploring the determinants of metal sensing in *Salmonella typhimurium* using FrmR, a non-metal sensing RcnR/CsoR family member

Salmonella FrmR, a member of the RcnR/CsoR family of metalloregulators, has been characterised during the course of this work and was coincidentally confirmed to bind specifically to the *frmRA* operon, which encodes a putative Zn(II)-requiring class III alcohol dehydrogenase. FrmR shares a high degree of similarity with Ni(II)/Co(II)-sensing RcnR, in particular conserving two residues of a so-called WXYZ motif required to detect metals. Metal-binding properties of FrmR were therefore extensively investigated *in vitro* and its ability, or otherwise, to respond to metals explored *in vivo*.

FrmR binds Zn(II), Cu(I), Co(II) and Ni(II), adopting different geometries, and always involving a mercapto group from the only cysteine residue (Cys35). Moreover, $K_{Zn(II)}^{FrmR}$ is only slightly below the range of affinity found for other zinc sensors.

Since FrmR fails to sense metals in cells, where only formaldehyde is detected, questions about which parameters are required in metal regulation in *Salmonella* and, in general, in bacteria were investigated. A single-point mutation (Glu64 → His) allows FrmR to sense cellular zinc and cobalt. FrmR and E64HFrmR have been consequently used as a case of study to test hypotheses about the mechanisms determining which metals are detected by a given sensor in cells.

In addition, the ability of FrmR to detect cellular formaldehyde has been investigated, and a reaction mechanism tested by site-directed mutagenesis *in vitro*. *Salmonella* Ni(II)/Co(II)-sensing RcnR has been characterised, and employed to test the specificity of formaldehyde responsiveness of FrmR. By a single point-mutation (Ser2 → Pro), the Ni(II)/Co(II)-sensing RcnR has been successfully switched to a formaldehyde sensor *in vitro*, further endorsing the proposed mechanism.

Investigation of FrmR structure has been pursued by producing apo- and Zn(II)-bounded FrmR and E64HFrmR crystals, which were then analysed at the Diamond Light Source. The best dataset has been processed to obtain a 3D-structure.

Table of contents

Title page **i**

Abstract **ii**

Table of contents **iii**

List of abbreviations **xii**

List of Figures **xv**

List of Tables **xxi**

List of Scripts **xxii**

Declaration **xxiii**

Statement of copyright **xxiii**

Acknowledgements **xxiv**

Dedication **xxv**

Chapter 1 Introduction 1

1.1. Control of the intracellular metal landscape **2**

1.1.1. *Metal in biological systems* *2*

1.1.2. *Physico-chemical factors determining metal-ligand interaction* *3*

1.1.3. *Metal preferences regulation in proteins* *4*

1.2. Metalloregulatory proteins **5**

1.2.1 *CsoR/RcnR-like de-repressors* *6*

1.2.1.1 Ni(II)- and Co(II)-sensing repressor RcnR 9

1.2.1.2 Cu(I)-sensing repressor CsoR 12

1.2.1.3 Regulated in copper repressor RicR 13

1.2.1.4 Dimer metal efflux repressor DmeR 14

1.2.1.5 Internal nickel-responsive sensor InrS 14

1.2.1.6 CsoR-like-sulfur-transferase repressor CstR 15

1.2.1.7 Formaldehyde repressor FrmR 16

1.2.2 *Other families of transcriptional regulators* *18*

1.2.2.1 MerR family 18

1.2.2.2 ArsR/SmtB family 19

1.2.2.3 Ferric uptake regulator (Fur) family 20

1.2.2.4 ModE-like regulators 20

1.2.2.5 NikR-like regulators 21

1.3 How metal sensors detect the right metal **21**

1.3.1 *Relative metal affinity* *22*

1.3.2 *Allosteric mechanism triggered by effector binding* *22*

1.3.3 *Channeling* *24*

1.4 DNA recognition **25**

1.4.1 *Structure of DNA-binding proteins* *25*

1.4.1.1 Mainly α motifs 25

1.4.1.2 Mainly β motifs 25

1.4.1.3 Mixed α/β motifs 26

1.4.1.4 Multi domains proteins 26

1.4.2 *Sequence-dependent variations of DNA-structure* *26*

1.4.2.1 Global shape variations 26

1.4.2.2 Local shape variations 27

1.4.3 *Mechanisms of protein-DNA recognition* *29*

1.4.3.1 Base readout 29

1.4.3.2 Shape readout 30

1.4.4 *DNA recognition in RcnR and CsoR homologues* *30*

1.5 Formaldehyde: Toxicity, origins and detoxification mechanisms **31**

1.5.1 *Reactivity* *32*

1.5.1.1 Nucleophilic addition of oxygen and sulfur nucleophiles 32

1.5.1.2 Nucleophilic addition of nitrogen nucleophiles 35

1.5.2 *Toxicity* *35*

1.5.2.1 Protein-DNA crosslinking with formaldehyde in vitro 35

1.5.3 *Origins* *39*

1.5.4 *Detoxification* *39*

1.5.4.1 GSH-dependent formaldehyde detoxification pathways 40

1.5.4.2 Other formaldehyde detoxification pathways 40

1.5.5 *Regulation of Formaldehyde detoxification pathways* *41*

1.6 Project aims and objectives **43**

Chapter 2 Materials and methods **45**

2.1 Reagents and chemicals **46**

2.2 Molecular biology methods **46**

2.2.1 *Growth conditions* *46*

2.2.2 *Antibiotics* *46*

2.2.3 *List of oligonucleotides* *46*

2.3 DNA manipulation **50**

2.3.1 *Amplification of DNA by Polymerase Chain Reaction (PCR)* *50*

2.3.2 *3'-A tailing of PCR products* *50*

2.3.3 *Agarose gel electrophoresis* *50*

2.3.4 *Ligation into pGEM-T plasmid* *50*

2.3.5 *Blue/white identification of transformant cells and colony PCR* *51*

2.3.6 *Ligation into pET29a plasmid* *51*

2.3.7 *Isolation of plasmid DNA* *52*

2.3.8 *Restriction digestion* *52*

2.3.9 *Site directed Mutagenesis using the 'Quick Change' method* *53*

2.3.10 *Production and transformation of E. coli chemically competent cells* *53*

2.3.11 *Production and transformation of S typhimurium electrocompetent cells* *54*

2.3.12 *Generation of Salmonella deletion mutants* *55*

2.3.13 *Generation of Promoter-lacZ fusion constructs and β–galactosidase assays* 55

2.4 Protein Production and Purification **55**

2.4.1 *SDS PAGE analysis* *55*

2.4.2 *Native gel analysis* *55*

2.4.3 *Overexpression and solubility testing* *55*

2.4.4 *Overexpression conditions for recombinant FrmR and RcnR in E. coli* *56*

2.4.5 *Overexpression conditions for recombinant ZntR and CueR in E. coli* *56*

2.4.6 *Purification of recombinant wild-type and mutated FrmR proteins* *56*

2.4.7 *Purification of recombinant wild-type and mutated RcnR proteins* *57*

2.4.8 *Purification of recombinant wild-type ZntR and CueR* *57*

2.4.9 *Estimation of protein concentration* *58*

2.4.10 *Detection of Met1 cleavage in FrmR and E64HFrmR by liquid chromatography (LC)* *58*

2.5 Anaerobic manipulation of proteins **58**

2.5.1 *Production of chelex-treated, anaerobic buffers* *58*

2.5.2 *Production of anaerobic, metal- free protein samples* *59*

2.5.3 *Determination of reduced sulphur content of proteins* *59*

2.5.4 *Measurement of metal content of purified proteins* *59*

2.6 Experimental procedures **60**

2.6.1 *Preparation of metal stocks* *60*

2.6.2 *Fractionation of protein-bound metal complexes by size-exclusion chromatography* *60*

2.6.3 *UV-visible absorption spectroscopy* *60*

2.6.4 *Monitoring quenching of protein intrinsic fluorescence upon metal-binding* *61*

2.6.5 *Measurement of the metal binding affinities of recombinant FrmR and RcnR* *61*

2.6.5.1 Measurement of the Zn(II)-binding affinities of FrmR and E64HFrmR by competition with quin-2 and mag-fura-2 61

2.6.5.2 Measurement of the Cu(I)-binding affinity of FrmR and E64HFrmR by competition with BCS 62

2.6.5.3 Measurement of the Cu(I)-binding affinity of FrmR and E64HFrmR by competition with BCA 62

2.6.5.4 Measurement of the Co(II)-binding affinities of FrmR and E64HFrmR by competition with Fura-2 and Bis-Tris 62

2.6.5.5 Estimation of the Ni(II)-binding affinity of FrmR and E64HFrmRs by competition with mag-fura-2 62

2.6.6 *Interprotein metal-exchange* *63*

2.6.6.1 Metal exchange between *Salmonella* ZntR, CueR, RcnR and FrmR 63

2.6.6.2 Metal exchange between *Salmonella* ZntR and E64HFrmR 63

2.6.7 *Analysis of protein-DNA interaction by fluorescence anisotropy (FA)* *63*

2.6.7.1 Production of fluorescently labelled annealed, double stranded DNA oligonucleotides 63

2.6.7.2 Analysis of protein-DNA binding by fluorescently anisotropy 63

2.6.7.3Assessment of FrmR-, E64HFrmR-, and RcnR-DNA stoichiometry by fluorescence anisotropy (FA) 64

2.6.8 *Analysis of DNA binding by Electrophoretic mobility shift assays (EMSA)* *64*

2.6.8.1 Production of probe DNA fragments by PCR 64

2.6.8.2 Binding of RcnR to DNA monitored by EMSA 65

2.7 Crystallography **65**

2.7.1 *Crystal trials* *65*

2.7.2 *Data collection* *67*

2.8 Bioinformatics **67**

Chapter 3 Characterisation of *Salmonella* FrmR **68**

3.1 Aims and objectives **69**

3.2 Identification of recombinant FrmR **69**

3.3 Production and purification of recombinant FrmR **69**

3.4. Analysis of FrmR metal binding properties **74**

3.4.1. *Co(II) and Zn(II) binding properties* *74*

3.4.2. *Cu(I) and Ni(II) binding properties* *80*

3.5. Metal binding affinities of FrmR **86**

3.5.1. *Measurement of the Zn(II) binding affinity of FrmR by competition with mag-fura-2 and quin-2* *86*

3.5.2. *Measurement of the Co(II) binding affinity of FrmR by competition with Fura-2 and BisTris* *88*

3.5.3. *Measurement of the Cu(I) binding affinity of FrmR by competition with BCS and BCA* *88*

3.5.4. *Probing Ni(II) binding affinity of FrmR by competition with mag-fura-2* *90*

3.6. Identification of FrmR DNA binding site and characterisation of the DNA-protein interaction **94**

3.6.1. *FrmR:DNA interactions* *94*

3.6.2. *Fluorescence anisotropy (FA) principles* *94*

3.6.3. *Fluorescence anisotropy oligonucleotides production* *95*

3.6.4. *FrmR:DNA stoichiometry studied by fluorescence anisotropy* *97*

3.6.5. *FrmR:DNA interaction studied by fluorescence anisotropy* *97*

3.6.6. *FrmR:DNA interaction is weakened by Zn(II) and Cu(I)* *101*

3.7. Identification of residues required for DNA binding **106**

3.7.1. *Production and purification of recombinant H60LFRmR and E81AFrmR* *108*

3.7.2. *H60LFrmR and E81AFrmR DNA binding properties studied by fluorescence anisotropy* *108*

3.8. Identification of residues required for metal binding **111**

3.8.1. *Production and purification of recombinant C35AFrmR* *111*

3.8.2. *Zn(II) binding properties of C35AFrmR investigated by size exclusion chromatography* *111*

3.8.3. *Apo and Zn(II)-C35AFrmR : DNA interactions studied by fluorescence anisotropy* *114*

3.8.4. *Zn(II) and Cu(I) binding properties of H60LFrmR investigated by size exclusion chromatography* *114*

3.9. Concluding remarks **118**

Chapter 4 Generation of a metal-sensing transcriptional regulator by gain of function single point mutation **120**

4.1 Aims and objectives **121**

4.2 Wild-type FrmR responds only to formaldehyde *in vivo* **121**

4.2.1 *In vivo expression from pRSfrmR following exposure to metals and formaldehyde* *121*

4.2.2 *In vivo expression from pRSfrmRE64H following exposure to metals and formaldehyde* *121*

4.3 Production and purification of recombinant E64HFrmR **123**

4.4 *In vitro* analysis of E64HFrmR metal-binding properties **127**

4.4.1 *Co(II)-binding properties* *127*

4.4.2 *Zn(II)-binding properties* 127

4.4.3 *Cu(I)- and Ni(II)-binding properties* 130

4.5 Metal binding affinities of E64HFrmR **133**

4.5.1 *Measurement of the Zn(II)-binding affinity of E64HFrmR by competition with mag-fura-2 and quin-2* 133

4.5.2 *Measurement of the Co(II)-binding affinity of E64HFrmR by competition with mag-fura-2* 135

4.5.3 *Measurement of the Cu(I)-binding affinity of E64HFrmR by competition with bincinchoninic acid (BCA) and bathocuproine disulfonate (BCS)* 138

4.6 Relative metal-binding affinities of wild-type FrmR and E64HFrmR **141**

4.6.1 *Metal-binding competition between FrmR and Salmonella Zn(II)-, Co(II)- and Cu(I)-sensors ZntR, RcnR and CueR* 141

4.6.2 *Salmonella Zn(II)-sensor ZntR competition with E64HFrmR for zinc binding* 145

4.7 Contribution of glutathione (GSH) to metal sensing by FrmR and E64HFrmR **145**

4.7.1 *Assessing GSH contribution to zinc-sensing by FrmR and E64HFrmR* 147

4.8 *In vitro* interaction between E64HFrmR and *frmRA* promoter **151**

4.8.1 *E6AHFrmR:DNA stoichiometry determined by fluorescence anisotropy* 151

4.8.2 *E64HFrmR:DNA interaction determined by fluorescence anisotropy* 151

4.8.3 *E64HFrmR:DNA interaction in the presence of Zn(II)* 154

4.9 Concluding remarks **154**

Chapter 5 Formaldehyde sensing by *Salmonella* FrmR and RcnR **157**

5.1 Aims and objectives **158**

5.2 Identification of formaldehyde-sensing fingerprints **158**

5.2.1 *Formaldehyde weakens FrmR K_{DNA}* 158

5.2.2 *In vivo study of FrmR function* 160

5.3 Mechanism of formaldehyde responsiveness by FrmR **160**

5.3.1 *Proposed formaldehyde-binding mechanism* 160

5.4 Probing the formaldehyde-sensing mechanism of FrmR **165**

5.4.1 *C35AFrmR:DNA K_{DNA} is unaffected by formaldehyde* *166*

5.4.2 *Production and purification of P2SFrmR* *166*

5.4.3 *P2SFrmR:DNA interactions in presence of various formaldehyde concentrations by FA* *166*

5.5 *In vitro* analysis of the RcnR:*rcnRA* promoter interaction **170**

5.5.1 *Recombinant Salmonella RcnR production and purification* *170*

5.5.2 *RcnR unlabelled oligonucleotides production and competitive EMSA to study RcnR:DNA interaction* *173*

5.5.3 *RcnR labelled oligonucleotides production* *176*

5.5.4 *RcnR:DNA stoichiometry explored by fluorescence anisotropy* *176*

5.5.5 *RcnR:rcnRAPro interaction explored by fluorescence anisotropy* *179*

5.5.6 *Analysis of the specificity of RcnR:rcnRAPro interaction* *179*

5.5.7 *Effect of Co(II)- and Ni(II)-binding on formation of RcnR:rcnRAPro interaction* *182*

5.6 Assessing *Salmonella* RcnR *in vitro* responsiveness to formaldehyde **182**

5.6.1 *RcnR:DNA interaction explored by fluorescence anisotropy in the presence and absence of formaldehyde* *182*

5.7 RcnR Pro2→Ser: Generation of a formaldehyde sensor *in vitro* by a single point mutation **185**

5.7.1 *Production and purification of recombinant S2PRcnR* *185*

5.7.2 *S2PRcnR:DNA interaction studied by fluorescence anisotropy in presence and in absence of formaldehyde* *188*

5.8 Concluding remarks **188**

Chapter 6 Crystal structure determination of FrmR 192

6.1 Aims and objectives **193**

6.2 Crystallography background **193**

6.2.1 *Crystal assembly* *193*

6.2.2 *Crystallization techniques* *194*

6.2.3 *Diffraction experiment theory and principles* *196*

6.2.4 *Data processing and phase problem* *198*

6.3 Crystallization of apo- and zinc-forms of FrmR and E64HFrmR **200**

6.4 Diffraction data collection on FrmR and E64HFrmR crystals **200**

6.5 Data processing of apo-E64HFrmR diffraction data 201

6.6 Concluding remarks 205

Chapter 7 Discussion and future work 208

7.1 Summary of results 209

7.2 Formaldehyde-sensing Salmonella FrmR is capable of binding metals 213

7.3 *Salmonella* FrmR allosteric regulation of transcription 217

7.3.1 *Salmonella FrmR binds a type-1 operator promoter region 217*

7.3.2 *Zn(II) and Cu(I) negatively regulate in vitro binding of frmA promoter by Salmonella FrmR 217*

7.4 One single point mutation allows Salmonella FrmR to sense Zn(II) and Co(II) *in vivo* 224

7.4.1 *FrmR and E64HFrmR cannot compete for metal-binding with Salmonella cognate metal sensors 226*

7.4.2 *Combination of relative affinity and relative-allostery explains Zn(II) gain of function 228*

7.4.3 *Glutathione enhanced metal-detection by E64HFrmR may explain Co(II)-sensing 231*

7.4.4 *Detection of two distinct E64HFrmR conformations 234*

7.5 Specific formaldehyde-sensing by *Salmonella* FrmR 236

7.5.1 *Cys35 and Pro2 play an essential role in formaldehyde detection 236*

7.5.2 *frmRA expression is not induced by other aldehydes or alcohols 237*

7.5.3 *Glutathione-formaldehyde adducts 239*

7.6 *Salmonella* RcnR regulates expression of *rcnRA* operon 241

7.6.1 *Ni(II) and Co(II), but not formaldehyde, negatively regulate in vitro binding of rcnRAPro by Salmonella RcnR 241*

7.7 Single point mutation enhances the capacity to detect low concentration of formaldehyde *in vitro* by *Salmonella* RcnR 242

7.8 Future work 242

Chapter 8 Appendix 249

Chapter 9 Reference 271

List of Abbreviations

A	adenine
AAA	Amino Acid Analysis
ABC	ATP binding cassette superfamily
ADH	alcohol dehydrogenase
ADP	adenosine diphosphate
Ala	alanine
Arg	arginine
ATP	adenosine triphosphate
BCS	bathocuproine disulfonate
BLAST	basic local alignment search tool
bp	base pair
BSA	bovine serum albumin
BSH	bacillithiol
Bsu	*Bacillus subtilis*
C	cytosine
CCD	charge-coupled devices
CD	circular dichroism
cDNA	complementary DNA
CFSE	Crystal Field Stabilization Energy
CORM-2	tricarbonyldichlororutenium (II)
Cys	cysteine
Da	Dalton
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTNB	5,5´-dithiobis-(2-nitrobenzoic) acid
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
DUF	domain of unknown function
Ec	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
ESI-MS	electrospray ionisation mass spectrometry
EtOH	ethanol
EXAFS	extended X-ray absorption fine structure

FA	fluorescence anisotropy
FRET	fluorescence resonance energy transfer
x G	number of times the gravitational force
GD-FALDH	formaldehyde dehydrogenase
Glu	glutamic acid
GOE	Great Oxidation Event
GSH	reduced glutathione
Gt	*Geobacillus thermodenitrificans*
HCOH	formaldehyde
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEX	hexachlorofluorescein
His	histidine
HSQC NMR	heteronuclear single quantum coherence nuclear magnetic resonance
ICP-MS	inductively coupled plasma mass spectrometry
IPTG	isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
kb	kilobase
LB	Luria-Bertani growth medium
LC-MS	Liquid chromatography-mass spectrometry
Leu	leucine
LMCT	ligand to metal charge transfer
LMW	low molecular weight
Lys	lysine
MALS	multi angle light scattering
MDH	methanol dehydrogenase
mf2	mag-fura-2
MME	monomethylether
MW	molecular weight
Mtb	*Mycobacterium tuberculosis*
N	any amino acid
OD	optical density
ONP	ortho-nitrophenol
ONPG	ortho-nitrophenol-β-D-galactosidase
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycole
Phe	phenylalanine

PMSF	phenylmethylsulfonyl fluoride
RNAp	ribonucleic acid polymerase
r_{obs}	observed anisotropy
rpm	revolutions per minute
RT	room temperature (25 ºC)
RuMP	Ribulose monophosphate methabolism
Sau	*Staphylococcus aureus*
SCOP	Structural Classification of Proteins
SDS	sodium dodecyl sulphate
SEC	size-exclusion chromatography
Sli	*Streptomyces lividans*
SMT	Stroboscopic single molecule tracking
SOD	superoxide dismutase
Sty	*Salmonella typhimurium*
Syn	*Synechocystis PCC 6803*
T	temperature
TMAO	trimethylamine-N-oxide
TMA	trimethylamine
TOF-MS	time of flight-mass spectrometry
TPP	tetraphenylphosphonium
Tth	*Thermus thermophilus*
Tyr	tyrosine
UV	ultraviolet
V	volt
Vis	visible
v/v	volume to volume
WHO	World Health Organization
WT	wild-type
w/v	weight to volume
XANES	X-ray absorption near edge structure
XAS	X-ray absorption spectroscopy
X-gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside

List of Figures

Figure 1. 1	Deduced metal-sensing transcriptional regulators from *S. typhiumurium*	**7**
Figure 1. 2	RcnR/CsoR family members	**8**
Figure 1. 3	Stability order of carbocations derived from carbonyl compounds	**33**
Figure 1. 4	Acetal synthesis reaction mechanism	**34**
Figure 1. 5	Hemiacetal synthesis reaction mechanism	**36**
Figure 1. 6	Enamine synthesis reaction mechanism	**37**
Figure 1. 7	Reactions between formaldehyde and DNA nucleotides	**38**
Figure 3. 1A	Genomic region of *frmRA*	**70**
Figure 3. 1B	DNA-binding sites of RcnR/CsoR family members	**70**
Figure 3. 2	Sequence alignment of RcnR/CsoR family members	**71**
Figure 3. 3	SDS-PAGE analyses following FrmR purification	**72**
Figure 3. 4A	FrmR Co(II)-binding monitored by UV-vis spectroscopy	**75**
Figure 3. 4B	Size-exclusion chromatography analysis of Co(II)-FrmR	**75**
Figure 3. 5	Zn(II)-quenches Co(II)-FrmR absorbance spectral features	**77**
Figure 3. 6A	FrmR Zn(II)-binding monitored by fluorescence spectroscopy	**79**
Figure 3. 6B	Size-exclusion chromatography analysis of Zn(II)-FrmR	**79**
Figure 3. 7	Concentration of Cu(I) stock determined by BCS calibrated assay	**81**
Figure 3. 8A	FrmR Cu(I)-binding monitored by UV-vis spectroscopy	**82**
Figure 3. 8B	Size-exclusion chromatography analysis of Cu(I)-FrmR	**82**
Figure 3. 9A	FrmR Ni(II)-binding monitored by UV-vis spectroscopy	**84**
Figure 3. 9B	Size-exclusion chromatography analysis of Cu(I)-FrmR	**84**
Figure 3. 10	FrmR Ni(II)-binding monitored by fluorescence spectroscopy	**85**
Figure 3. 11A	Zn(II)-binding competition between FrmR and mag-fura-2	**87**
Figure 3. 11B	Zn(II)-binding competition between FrmR and quin-2	**87**
Figure 3. 12A	Co(II)-binding competition between FrmR and fura-2	**89**
Figure 3. 12B	Co(II)-binding competition between FrmR and BisTris	**89**
Figure 3. 13A	Cu(I)-binding competition between FrmR and BCS (single addition)	**91**
Figure 3. 13B	Cu(I)-binding competition between FrmR and BCS	**91**
Figure 3. 14	Cu(I)-binding competition between FrmR and BCA	**92**
Figure 3. 15	Ni(II)-binding competition between FrmR and mag-fura-2	**93**
Figure 3. 16	Production of *frmRA*Pro probe for FA assay	**96**
Figure 3. 17A	*frmRA*Pro:FrmR stoichiometry assessed by FA	**98**
Figure 3. 17B	Scheme of proposed *frmRA*Pro:FrmR interaction	**98**
Figure 3. 18A	*frmRA*Pro-T/A-mod:FrmR stoichiometry assessed by FA	**99**
Figure 3. 18B	Scheme of proposed *frmRA*Pro-T/A-mod:FrmR interaction	**99**
Figure 3. 19A	Titration of *frmRA*Pro with FrmR monitored by FA	**100**
Figure 3. 19B	Titration of *frmRA*Pro-T/A-mod with FrmR monitored by FA	**100**
Figure 3. 20A	Titration of *frmRA*Pro with Zn(II)-FrmR monitored by FA	**102**

Figure 3. 20B Titration of *frmRA*Pro with Cu(I)-FrmR monitored by FA **102**

Figure 3. 20C Titration of *frmRA*Pro with apo-, Cu(I)- and Zn(II)-FrmR (by FA) **102**

Figure 3. 21 Allosteric coupling scheme applied to homotetrameric FrmR **103**

Figure 3. 22A Titration of pre-formed *frmRA*Pro:FrmR complex with Zn(II) (by FA) **105**

Figure 3. 22B Titration of pre-formed *frmRA*Pro:FrmR complex with Cu(I) (by FA) **105**

Figure 3. 23A Interaction between helices α2 and α3 in FrmR **107**

Figure 3. 23B Interaction between helices α2 and α3 in InrS **107**

Figure 3. 24A SDS-PAGE analyses following H60LFrmR purification **109**

Figure 3. 24B SDS-PAGE analyses following E81AFrmR purification **109**

Figure 3. 25A Titration of *frmRA*Pro with H60LFrmR monitored by FA **110**

Figure 3. 25B Titration of *frmRA*Pro with E81AFrmR monitored by FA **110**

Figure 3. 25C Titration of *frmRA*Pro with wt-, H60L-, E81A-FrmR proteins (by FA) **110**

Figure 3. 26 SDS-PAGE analyses following C35AFrmR purification **112**

Figure 3. 27AB Size-exclusion chromatography analyses of Zn(II)-C35AFrmR **113**

Figure 3. 28A Titration of *frmRA*Pro with C35AFrmR monitored by FA **115**

Figure 3. 28B Titration of *frmRA*Pro with Zn(II)-C35AFrmR monitored by FA **115**

Figure 3. 28C Titration of pre-formed *frmRA*Pro:C35AFrmR complex with Zn(II) monitored by FA **115**

Figure 3. 29 Size-exclusion chromatography analysis of Cu(I)-H60LFrmR **116**

Figure 3. 30AB Size-exclusion chromatography analyses of Zn(II)-H60LFrmR **117**

Figure 3. 31 H60LFrmR Ni(II)-binding monitored by fluorescence spectroscopy **119**

Figure 4. 1A *β*-galactosidase activity measured in Δ*frmR* containing P_{frmRA}-*frmR* fused to lacZ upon incubation with no metal, or MNIC of Mn(II), Fe(III), Co(II), Ni(II), Cu(II), Zn(II), or formaldehyde **122**

Figure 4. 1B *β*-galactosidase activity measured in Δ*frmR* containing P_{frmRA}-*frmR* fused to lacZ upon incubation with no metal, or MNIC of Mn(II), Fe(III), Co(II), Ni(II), Cu(II), Zn(II), SeO_3^{2-}, TeO_3^{2-}, or formaldehyde **122**

Figure 4. 2A E64HFrmR amino acid sequence **124**

Figure 4. 2B *β*-galactosidase activity measured in Δ*frmR* containing P_{frmRA}-*frmRE64H* fused to lacZ upon incubation with no metal, or MNIC of Mn(II), Fe(III), Co(II), Ni(II), Cu(II), Zn(II), or formaldehyde **124**

Figure 4. 3 SDS-PAGE analyses following E64HFrmR purification **125**

Figure 4. 4A UV-vis absorbance spectrum of apo-E64HFrmR **126**

Figure 4. 4B UV-vis absorbance spectrum of apo-FrmR **126**

Figure 4. 5 E64HFrmR Co(II)-binding monitored by UV-vis spectroscopy **128**

Figure 4. 6A E64HFrmR Zn(II)-binding monitored by fluorescence spectroscopy **129**

Figure 4. 6B Size-exclusion chromatography analysis of Zn(II)-E64HFrmR **129**

Figure 4. 7A E64HFrmR Cu(I)-binding monitored by UV-vis spectroscopy **131**

Figure 4. 7B Size-exclusion chromatography analysis of Cu(I)-E64HFrmR **131**

Figure 4. 8A E64HFrmR Ni(II)-binding monitored by UV-vis spectroscopy **132**

Figure 4. 8B Size-exclusion chromatography analysis of Ni(II)-E64HFrmR **132**

Figure 4. 9A Zn(II)-binding competition between E64HFrmR and mag-fura-2 **134**

Figure 4. 9B Zn(II)-binding competition between E64HFrmR and quin-2 **134**

Figure 4. 10A Comparison of the data obtained for Zn(II)-binding competition between E64HFrmR or FrmR and mag-fura-2 **136**

Figure 4. 10B Comparison of the data obtained for Zn(II)-binding competition between E64HFrmR or FrmR and quin-2 **136**

Figure 4. 11A Co(II)-binding competition between E64HFrmR and fura-2 **137**

Figure 4. 11B Co(II)-binding competition between E64HFrmR and BisTris **137**

Figure 4. 12 Cu(I)-binding competition between E64HFrmR and BCA **139**

Figure 4. 13 Cu(I)-binding competition between E64HFrmR and BCS **140**

Figure 4. 14 Cu(I)-binding competition between *Salmonella* CueR and FrmR **142**

Figure 4. 15 Zn(II)-binding competition between *Salmonella* ZntR and FrmR **143**

Figure 4. 16 Co(II)-binding competition between *Salmonella* RcnR and FrmR **144**

Figure 4. 17 Zn(II)-binding competition between *Salmonella* ZntR and E64HFrmR **146**

Figure 4. 18A Zn(II) titration of mag-fura-2 in the presence/absence of glutathione **148**

Figure 4. 18BC Absorbance spectra at the end of Zn(II) competition between FrmR or E64HFrmR with mag-fura-2 upon addition of EDTA **148**

Figure 4. 19A Time-course analyses of Zn(II) migration from mag-fura-2 to FrmR in the presence/absence of glutathione monitored by absorbance spectroscopy **149**

Figure 4. 19B Time-course analyses of Zn(II) migration from mag-fura-2 to E64HFrmR in the presence/absence of glutathione monitored by absorbance spectroscopy **149**

Figure 4. 20A Time-course analyses of Zn(II) migration from mag-fura-2 to FrmR or E64HFrmR in the presence of glutathione monitored by absorbance spectroscopy **150**

Figure 4. 20B Time-course analyses of Zn(II) migration from mag-fura-2 to FrmR or E64HFrmR in the absence of glutathione monitored by absorbance spectroscopy **150**

Figure 4. 21A *frmRA*Pro:E64HFrmR stoichiometry assessed by FA **152**

Figure 4. 21B *frmRA*Pro-T/A-mod:E64HFrmR stoichiometry assessed by FA **152**

Figure 4. 22A Titration of *frmRA*Pro with E64HFrmR monitored by FA **153**

Figure 4. 22B Titration of *frmRA*Pro with Zn(II)-E64HFrmR monitored by FA **153**

Figure 4. 22C Comparison of the data obtained for the titration of *frmRA*Pro with apo- and Zn(II)-E64HFrmR monitored by FA **153**

Figure 4. 23 Titration of pre-formed *frmRA*Pro:E64HFrmR complex with Zn(II) monitored by FA **155**

Figure 5. 1AB Titration of *frmRA*Pro with FrmR in the presence of various concentrations of formaldehyde monitored by FA **159**

Figure 5. 1C Comparison of the data obtained for the titration of *frmRA*Pro with FrmR in the absence/ presence of formaldehyde monitored by FA **159**

Figure 5. 2AB Titration of *frmRA*Pro with FrmR in the presence of acetaldehyde or ethanol monitored by FA **161**

Figure 5. 3A *β*-galactosidase activity measured in Δ*frmR* containing either P_{frmRA}-*frmR::lacZ* or P_{frmRA}*::lacZ* in the absence/presence of MNIC of formaldehyde **162**

Figure 5. 3B *β*-galactosidase activity measured in Δ*frmR* containing P_{frmRA}-*frmR::lacZ* in the absence/presence of MNIC of various alcohols and aldehydes **162**

Figure 5. 4 Proposed mechanism of action of formaldehyde-FrmR cross-linking **164**

Figure 5. 5A-C Titration of *frmRA*Pro with C35AFrmR in the absence/presence of various concentrations of formaldehyde monitored by FA **167**

Figure 5. 5D Comparison of the data obtained for the titration of *frmRA*Pro with C35AFrmR in the absence/ presence of formaldehyde monitored by FA **167**

Figure 5. 6 SDS-PAGE analyses following P2SFrmR purification **168**

Figure 5. 7A-D Titration of *frmRA*Pro with P2SFrmR in the absence/presence of various concentrations of formaldehyde monitored by FA **169**

Figure 5. 8 Comparison of the data obtained for the titration of *frmRA*Pro with P2SFrmR in the absence/ presence of formaldehyde monitored by FA **171**

Figure 5. 9 SDS-PAGE analyses following RcnR purification **172**

Figure 5. 10 Sequence alignment of *S. typhimurium* and *E. coli* RcnR proteins **174**

Figure 5. 11A Nucleotide sequence of the start of the coding region and upstream region of *rcnRA*ProEM **175**

Figure 5. 11B Competitive EMSA assessing RcnR binding to *rcnRA*ProEM **175**

Figure 5. 12 Production of *rcnRA*$Pro^{sites1,2}$ and *rcnRA*Pro^{site1} probes for FA assay **177**

Figure 5. 13A *rcnRA*$Pro^{sites1,2}$:RcnR stoichiometry assessed by FA **178**

Figure 5. 13B *rcnRA*Pro^{site1}:RcnR stoichiometry assessed by FA **178**

Figure 5. 14A Titration of *rcnRA*$Pro^{sites1,2}$ with RcnR monitored by FA **180**

Figure 5. 14B Titration of *rcnRA*Pro^{site1} with RcnR monitored by FA **180**

Figure 5. 15 Titration of *frmRA*Pro with RcnR monitored by FA **181**

Figure 5. 16A Titration of *rcnRA*$Pro^{sites1,2}$ with Ni(II)-RcnR monitored by FA **183**

Figure 5. 16B Comparison of the data obtained for the titration of *rcnRA*$Pro^{sites1,2}$ with apo- and Ni(II)-RcnR monitored by FA **183**

Figure 5. 17A Titration of *rcnRA*$Pro^{sites1,2}$ with Co(II)-RcnR monitored by FA **184**

Figure 5. 17B Comparison of the data obtained for the titration of *rcnRA*$Pro^{sites1,2}$ with apo- and Co(II)-RcnR monitored by FA **184**

Figure 5. 18AC Titration of *rcnRA*Pro^{sites1,2} with RcnR in the presence of various concentrations of formaldehyde monitored by FA **186**

Figure 5. 18D Comparison of the data obtained for the titration of *rcnRA*Pro^{sites1,2} with RcnR in the absence/presence of various concentrations of formaldehyde monitored by FA **186**

Figure 5. 19 SDS-PAGE analyses following S2PRcnR purification **187**

Figure 5. 20AD Titration of *rcnRA*Pro^{sites1,2} with S2PRcnR in the absence/presence of various concentrations of formaldehyde monitored by FA **189**

Figure 5. 21 Comparison of the data obtained for the titration of *rcnRA*Pro^{sites1,2} with S2PRcnR in the absence/presence of various concentrations of formaldehyde monitored by FA **190**

Figure 6. 1 Schematic representation of the solubility curve **195**

Figure 6. 2 Schematic representation of the sitting-drop and hanging-drop vapour diffusion techniques **197**

Figure 6. 3 Graphical interpretation of Bragg's Law **199**

Figure 6. 4AD Pictures taken on E64HFrmR crystals obtained by hanging-drop **202**

Figure 6. 5AD Pictures taken on FrmR crystals obtained by hanging-drop and sitting-drop **203**

Figure 6. 6 Diffraction pattern obtained from a plate-like apo-E64HFrmR crystal **204**

Figure 7. 1 Sequence alignment of *S. typhimurium* STM1627 and *E. coli* FrmA **211**

Figure 7. 2AD Time-course dissociation of pre-formed *frmRA*Pro:FrmR complex in the presence/absence of formaldehyde and/or Zn(II) (by FA) **212**

Figure 7. 3 Putative E64HFrmR metal binding site **215**

Figure 7. 4 Different coordination geometries adopted by FrmR depending on the nature of the metal ion **216**

Figure 7. 5AC Proposed scheme of RcnR homologues, *S. lividans* CsoR and *S. typhimurium* FrmR binding to their respective operators **218**

Figure 7. 6 Phylogenetic tree and group-specific amino acid signatures of functionally characterized DUF_{156} (RcnR/CsoR) proteins **220**

Figure 7. 7AB Preliminary model of E64HFrmR crystal structure (chains A,B) **222**

Figure 7. 8AB Least-square superposition of preliminary E64HFrmR crystal structure (chains A,B,C,D) with *Streptomyces lividans* apo-CsoR model **223**

Figure 7. 9AB Least-square superposition of preliminary E64HFrmR crystal structure (chains A,B,C,D) with *Mycobacterium tuberculosis* Cu(I)-CsoR model (chain A) converted to a tetramer by COOT **225**

Figure 7. 10 Scheme showing metal-affinities of *Salmonella* ZntR, ZuR, RcnR, CueR, FrmR, and E64HFrmR **229**

Figure 7. 11 Scheme representing the effects that the weaker K_{DNA} and tighter $K_{Zn(II)}$ of E64HFrmR, compared to FrmR, have in the cell **230**

Figure 7. 12 Six possible metal-coordination sites of glutathione **233**

Figure 7. 13 Putative formaldehyde-binding site consisting of Cys35 and Pro2 **238**

Figure 7. 14A Acetaldehyde reaction with Pro2 yielding an enamine **240**

Figure 7. 14BD Different adducts produced by glutathione reacting with formaldehyde **240**

Figure 7. 15 X-ray crystal structures of DNA-bound form of metal-transcriptional regulators available in the PDB at the time of writing **244**

Figure 8. 1 UV-visible spectra of apo-FrmR samples used for Amino Acid Analysis **263**

Figure 8. 2 UV-visible spectrum of apo-RcnR sample used for Amino Acid Analysis **264**

Figure 8. 3AB Extracted LC-MS chromatograms of specific ion transitions monitored for the FrmR and E64HFrmR PHSPEDK peptide detected in trypsin digested samples **265**

Figure 8. 4A UV-visible spectrum of apo-FrmR recorded anaerobically **266**

Figure 8. 4B UV-visible spectra of E64HFrmR recorded aerobically, or anaerobically upon incubation with glutathione, or formaldehyde **266**

Figure 8. 4C UV-visible spectra of E64HFrmR* recorded aerobically, or anaerobically upon incubation with glutathione, or formaldehyde **266**

Figure 8. 5 Zn(II)-binding by E64HFrmR* monitored by fluorescence spectroscopy **267**

Figure 8. 6A Control titration of BCA with Cu(I) monitored by absorbance spectroscopy **268**

Figure 8. 6B Control titration of BCS with Cu(I) monitored by absorbance spectroscopy **268**

Figure 8. 7A Control titration of mag-fura-2 with Zn(II) monitored by absorbance spectroscopy **269**

Figure 8. 7B Control titration of quin-2 with Zn(II) monitored by absorbance spectroscopy **269**

Figure 8. 8A Titration of *frmRA*Pro with apoE64HFrmR* monitored by FA **270**

Figure 8. 8B Titration of *frmRA*Pro-T/A-mod with apoE64HFrmR* monitored by FA **270**

Figure 8. 9A Titration of *frmRA*Pro with Zn(II)-E64HFrmR* monitored by FA **271**

Figure 8. 9B Comparison of the data obtained for the titration of *frmRA*Pro with apoE64HFrmR* and Zn(II)-E64HFrmR* monitored by FA **271**

Figure 8. 10 Surface of effector-binding site in apo-E64HFrmR crystal structure **272**

List of Tables

Table 1. 1	Structural properties of B-DNA, A-DNA, and Z-DNA conformations	**28**
Table 2. 1	List of primers used in the course of this work	**49**
Table 2. 2	Hampton Research crystallization kits used for a preliminary screening of crystallization conditions of FrmR and E64HFrmR	**66**
Table 2. 3	Crystallization cocktails adopted for the optimization screening of crystallization conditions of FrmR and E64HFrmR	**66**
Table 3. 1	Similarity and identity of amino acid sequences of MtbCsoR, SynInrS, EcRcnR, and StyRcnR compared to StyFrmR	**71**
Table 5. 1	Similarity and identity of amino acid sequences of StyRcnR compared to EcRcnR	**174**
Table 6. 1	List and description of programmes adopted to process E64HFrmR diffraction data-set collected at Diamond Light Source facility	**206**
Table 6. 2	Similarity and identity of amino acid sequences of MtbCsoR, TtCsoR, and SynInrS compared to StyFrmR	**206**
Table 6. 3	Data collection statistics for apo-E64HFrmR high resolution synchrotron data set	**207**
Table 7. 1	Similarity and identity of amino acid sequence of EcFrmA compared to *S. typhimurium* STM1627	**211**
Table 7. 2	List of the approximate shortest distances between the side chain atoms of E64HFrmR putative metal ligands plus Asp63	**215**
Table 7. 3	Metal affinities of ZuR, ZntR, RcnR and CueR	**229**
Table 8. 1	List of buffers used during the course of this work	**250**
Table 8. 2	DNA-binding affinities and allosteric coupling free energies of protein used during the course of this work	**252**
Table 8. 3	Metal-binding affinities of FrmR and E64HFrmR	**253**
Table 8. 4	FrmR Amino Acid Analysis data	**261**
Table 8. 5	RcnR Amino Acid Analysis data	**262**

List of Scripts

- Dynafit script used to determine $K_{Co(II)}$ in competition experiment between FrmR and BisTris **254**
- Dynafit script used to determine $K_{Cu(I)}$ in competition experiment between FrmR and BCA **255**
- Dynafit script used to determine $K_{Cu(I)}$ in competition experiment between E64HFrmR and BCA **256**
- Dynafit script used to determine $K_{Zn(II)}$ in competition experiment between FrmR proteins and mag-fura-2 **257**
- Dynafit script used to determine $K_{Zn(II)}$ in competition experiment between FrmR proteins and quin-2 **258**
- Dynafit script used to determine $K_{DNA}^{apo\text{-}protein}$ of FrmR and RcnR proteins by FA **259**
- Dynafit script used to determine $K_{DNA}^{effector\text{-}protein}$ of FrmR and RcnR proteins by FA **260**

Memorandum

Parts of this work have been published as:

Osman D., Piergentili C., Chen J., Chakrabarti B, Foster A. W., Lurie-Luke E., Huggins T. G. & Robinson N. J., Generating a metal-responsive transcriptional regulator to test what confers metal sensing in cells. *J. Biol. Chem.* 290(32):19806-22 (2015).

Declaration

No portion of this work has been submitted in support of an application for another degree or qualification from this or any other University or institute of learning.

Statement of copyright

Acknowledgements

I would like to express my sincere gratitude and special thanks to my supervisor Prof. Nigel Robinson for his continuous support, motivation, enthusiasm and advice given to me during these years. Thanks for allowing me to work at this project and for the great help in the research and the writing of this thesis.

I would also like to thank for his help and collaboration Dr. Ehmke Pohl, who gave me much advice in scientific and personal matters, and who introduced me to the fascinating subject of protein crystallography.

I thank our collaborators Dr. T. Huggins and Dr. J. Chen (Procter and Gamble Mason Businness Centre, Cincinnati, Ohio) for the mass-spectrometry analyses conducted on *Salmonella* FrmR and RcnR, and Dr. J. S. Cavet (University of Manchester) who kindly provided the *S. enterica* sv. *Typhimurium* SL1344 and LB5010a strains used in this work.

A special thanks goes to all members of the Robinson lab, past and present, for support, guidance and training provided. In particular I would like to thank Dr. Deenah Osman, who also provided access to her *in vivo* data, allowing a deeper interpretation of my *in vitro* results, and Dr. Andrew Foster for the patience shown in teaching me even the most basic molecular biology lab techniques (e.g. how to use an automated pipette!) when I transitioned from Chemistry to Biology.

I would like to thank all my friends here, who made my time in Durham a special period of my life, in particular the Dept. of Chemistry "Italian gang", the "Sons of Durham" and le "Banane in Pigiama", I will miss all of you. Of course, I cannot forget to thank my long-term friends who shared with me the joys and sorrows of a PhD student life: Simona, Anastasia, Nicola, il Miglio and Valentina.

I take this opportunity to thank my parents, my brother Giacomo and my granny, who always supported me during these years and showed me their affection by visiting me so often to almost do not make me feel homesick.

At the end, the most grateful thanks goes to Gabriele, whose love and patience are the most precious things in my life.

To my family and Gabriele

"The important thing in Science is not so much to obtain new facts as to discover new ways of thinking them"

William Bragg

Chapter 1

Introduction

1.1 Control of the intracellular metal landscape

1.1.1 *Metals in biological systems*

Carbon (C), hydrogen (H), nitrogen (N), oxygen (O), phosphorous (P) and sulfur (S) are the constituents of the cell: proteins, nucleic acids, lipids, membrane, sugars, and metabolites; the chemistry of these six elements and their compounds is considered by classical biochemistry. However, additional elements belonging to the *d*-block and the alkaline and alkaline-earth groups play critical roles and are essential for life.

Sodium (Na), potassium (P), magnesium (Mg) and calcium (Ca) have an important role in living organisms, in relatively high proportions and for this reason they are called "bulk-metals". The reason for their high bioavailability lies in the solubility of their salts, together with their natural abundance. The main biological functions of Na^+ are associated with membrane potentials, while its chemistry as enzyme activator is very limited (Alberts *et al.* 2002; Rana *et al.* 2011). Potassium, however, in addition to the important role in stabilizing nucleic acids structures (Marathias & Bolton 2000; Owczarzy *et al*. 2008), is also necessary for the activation of a large variety of enzymes, such as pyruvate kinase, pyruvate phosphatase, DNA polymerase, aldehyde dehydrogenase, etc. (Oria-Hernandez *et al*. 2005; Boyer *et al.* 1942; Di Cera 2006; Kachmar & Boyer 1953, Suelter 1970, Bostian & Betts 1978;). Amongst the alkaline metals, only K^+ has been shown to act as stoichiometric cofactor (in particular, in the pyruvate kinase, in association with Mg^{2+}) (Oria-Hernandez *et al*. 2005). The low charge density of potassium precludes its use as a Lewis acid and so its role consists of orienting the pyruvate in the active site, acting as a bridge between the enzyme and the substrate (Oria-Hernández *et al.* 2005). The Lewis-acid character of Mg(II) and its polarizing power against water molecules, oxygen atoms of phosphate groups, or phosphoric esters, together with its bioavailability, makes it indispensable in the catalysis of many biochemical processes (Cowan 1995, 2002; Black & Cowan 1995). A representative example of a process where Mg(II) plays an essential role is the reversible transfer of the phosphate group from the phosphoenolpyruvate to the ADP, producing pyruvate and ATP (Boyer P. D *et al*. 1942; Garcia-Olalla & Garrido-Pertierra 1986).

The remaining metals belong to the *d*-block and can be further divided in "trace" and "ultra-trace" metals. Iron (Fe), copper (Cu) and zinc (Zn) belongs to the first subset and essential to virtually all characterised organisms, whereas the ultra-trace subset includes manganese (Mn), molybdenum (Mo), cobalt (Co), nickel (Ni), tungsten (W), vanadium (V) and chromium (Cr) (Bertini *et al*. 2000).

The *d*-block metal ions are fundamental constituents of approximately one- quarter to one-third of all proteins and have roles encompassing structural components of biomolecules, signaling molecules, catalytic cofactors in reversible oxidation-reduction and hydrolytic reactions, and in

structural rearrangements of organic molecules and electron transfer chemistry (Bertini *et al.* 2000).

A useful classification is to discern between metal ions possessing structural and functional roles depending on whether the metal ion is required to stabilize the three-dimensional structure of the macromolecule (e.g. Zn, which is redox-inactive), or is involved in the reactivity of the biomolecule. For instance metals that have more than one oxidation state (e.g. Fe, Mn, Cu, Ni, Mo, V, W) are used in redox processes. They can act as conveyors of electrons (in the electron-transfer chain), catalyse oxidation and reduction reactions, or bind to small exogenous molecules (e.g. O_2) (Bertini *et al.* 2000; Valko *et al.* 2015). Metals with good Lewis acid properties (such as Zn) are used to catalyse hydrolytic reactions exploiting their ability to polarize chemical bonds (representative example is its role in alcohol dehydrogenase) (Coleman *et al.* 1972; Dunn & Hutchison 1973).

1.1.2 *Physico-chemical factors determining metal-ligand interaction*

The high assortment of properties of metal ions makes them simple and convenient co-factors available to biomolecules in the cell. The association between metal and ligand depends on physico-chemical properties of both species. Metal ions bind a given ligand according to several parameters such as its valence state (a higher metal's net charge is reflected in a stronger ionic interaction with an anionic ligand) and its ionic radius (the complexation reaction between a given ligand a metal become less favorable as the metal increases its size because this occurs in association with a decrease in charge density) (Dudev & Lim 2014). In addition to these, preferred coordination geometry, coordination number and charge-accepting ability must be considered (e.g. Mg^{2+} and Zn^{2+} possess the same charge, +2, and a similar ionic radius, 86 and 88 Å respectively, but zinc ion is a better Lewis acid therefore the final zinc-ligand complex will be more stabilized than with Mg^{2+}) (Dudev & Lim 2014).

For a given metal, preferences toward different amino acid residues are driven by the ligand net charge (negatively charged residues are more favorable than neutral residues as they can stabilize the positively charge located on the metal ion), charge-donating/accepting ability (better Lewis bases such as oxygen and sulphur, which are more polarizable, are preferable), number of ligands, dipole moment and polarizability (asparagine, glutamine and histidine residues are more favorites as they are polar and uncharged) (Dudev & Lim 2014). In addition, in many amino acid residues there are more than one donor atom: in these cases the ligands are called multidentate and the affinity for the metal species is incremented.

Not only should the properties of first-shell ligands be taken in account but also those of second-shell ligands. In fact these residues can stabilize the metal-binding site (resulting in an increased metal affinity) (Dudev & Lim 2014; Dudev *et al.* 2003), interact with and stabilize ligands belonging to the inner coordination sphere (Dudev & Lim 2006), and assist metal selection by

operating a steric selection (Dudev *et al.* 2003). They can also "protect" the inner coordination shell from unfavorable interactions (Lee & Lim 2011) or aid the proper orientation of the first-coordination sphere (Dudev *et al.* 2003; Lesburg *et al.* 1995; Kiefer *et al.* 1995; El Yazal *et al.* 2000.

Another fundamental parameter is the effects (increase of coordination site rigidity, effects on access of water molecules, repercussions on enzyme catalysis, etc.) that the remaining amino acid residues have on the formation of the metal complex. (Dudev & Lim 2014; Kuppuraj *et al.* 2009)

1.1.3 *Metal preference regulation in proteins*

Although the overview of parameters involved in metal-binding presented in the previous paragraph (Section 1.1.2) could suggest that protein metallation is a finely tuned process, when metals compete with each other mismetallation (protein binding the wrong metal) can occur. There is not much knowledge regarding how many proteins in the cell get mismetallated and if they are recycled or selectively degraded (Foster *et al.* 2014b).

The affinities of proteins for metal have a tendency to follow a universal order of preference, which for essential divalent metals is the Irving-Williams series:

Mg(II), Ca(II),< Mn(II) < Fe(II) < Co(II) < Ni(II) < Cu(II) > Zn(II)

In fact, the Irving-Williams series predicts an order of stability for divalent metal complexes, including those coordinated by protein ligands (Irving & Williams 1948, 1953; Foster *et al.* 2014b). Although not entirely exhaustive, several theoretical bases have been proposed to explain this trend of relative stabilities. These include the evaluation of the Crystal Field Stabilization Energy (CFSE), which refers to the number of electrons of a metal located in the bonding molecular orbitals (low-energy, stabilizing effect) compared to those in the anti-bonding molecular orbitals (high-energy, destabilizing effect). CFSE value of Ni(II), which is a d^8 low spin in the tetrahedral geometry, is the highest in the Irving-Williams series, whereas CFSE values of Zn(II) and Mn(II) are zero (Weller *et al.* 2014). The distortion undergone by Cu(II) complexes (Jahn-Teller effect) helps explain why these complexes are more stable than Ni(II) ones even if $CFSE^{Ni(II)} > CFSE^{Cu(II)}$ (Weller *et al.* 2014). Another factor is the ionic radius length of the metal since a smaller radius allows a higher degree of overlap between metal and ligand orbitals, increasing the stabilization of the resulting complex. Although the ionic radius of an ion is a property depending on several factors, such as the coordination number, it is possible to observe a trend in its variation through the periodic table. For example, Zn(II) has the smallest radius in the set, resulting in a short metal-ligand distance and a stabilization of the complex (Weller *et al.* 2014).

In addition to the Irving-William series, which deals only with divalent metal ions, monovalent copper and trivalent iron are also highly competitive as are cadmium, mercury and silver. Notably Cu(I) and Fe(III) are also sparingly soluble. Based only on these affinity considerations alone it is unclear how many proteins acquire and bind uncompetitive metal ions such as manganese when copper and zinc are also present in the cytosol.

In order to minimize mismetallation, cells must discern the elements to maintain optimal buffered levels of each metal. Ultimately, intracellular bioavailability becomes a major determinant of metal-protein speciation because the sum total effects of the parameters described above are insufficient to give perfect discrimination between the elements. Bacterial and archaeal cells restrict access to competitive metals, keeping these metals out of binding sites for less competitive metals (those lower down in the stability series). Each protein competes with other proteins for a limited pool of metals therefore the relative metal affinities of the different proteins for a particular metal become more important than their absolute metal affinities. This is achieved in part by the action of metalloregulatory proteins (sensors) used to control the expression of genes encoding proteins that manage metal homeostasis and resistance (Waldron *et al.* 2009).

Metal sensors (Section 1.2) are specialized allosteric proteins, in which the direct binding of a specific metal ion(s) drives a conformational change in the regulator that allosterically activates or inhibits operator DNA binding, or alternatively, distorts the promoter structure thereby converting a poor promoter to a strong one.

1.2 Metalloregulatory proteins

Metal sensor proteins bind metal ions and repress, derepress or activate the transcription of operons that encode metal-specific efflux pumps, metal transport proteins, metal-sequestering proteins and often the metal-responsive transcriptional regulator itself. In this way the correct concentration of a particular transition metal in the cell is maintained. Structural (and other) studies reveal distinct families of metal sensor proteins: the MerR, ArsR/SmtB and RcnR/CsoR families regulate the expression of genes required for metal ion detoxification, efflux and sequestration (Pennella & Giedroc 2005). Metal binding leads to activation (MerR) or derepression (ArsR/SmtB, RcnR/CsoR) of the resistance operon. In contrast, the DtxR, Fur, and NikR families classically regulate genes encoding proteins involved in metal ion utilization (Pennella & Giedroc 2005; Giedroc & Arunkumar 2007). Metal binding to these sensors leads to repression. This work focused on the metallo-regulatory network of *Salmonella enterica serovar typhimurium* (herein *Salmonella typhimurium* or *Salmonella*), which was chosen as model organism because is an important foodborne pathogen and a target for anti-microbial products. Additionally, there exists a range of preliminary studies of its metal-sensors and of homologous proteins from the closely related organism *Escherichia coli*.

Salmonella is a facultative anaerobe, Gram-negative bacterium widely distributed in nature. It is the bacterial agent most commonly isolated in cases of foodborne infections, both sporadic and epidemic. This organism is present in nature with more than 2400 known variants (serotypes) but the strains most frequently widespread in humans and animal species, in particular those reared for the food chain, are *S. enteritidis* and *S. typhimurium*. Infections from Salmonella are distinguished between typhoid forms (caused by *S. typhi* and *S. paratyphi*, responsible for typhoid and enteric fevers in general), in which humans are the only reservoir for the organism, and the non-typhoid forms, caused by other serotypes (such as *S. typhimurium* and *S. enteritidis*), responsible for clinical forms with predominantly gastrointestinal event. *Salmonella enterica subspecies I, serovar typhimurium* (*S. typhimurium*) causes serious medical and veterinary problems world-wide causing millions of infections characterized by persistent diarrhoea, abdominal pain, fever, and headache, and many deaths each year (Coburn *et al.* 2007, Osman *et al.* 2010, Majowicz *et al.* 2010). *Salmonella typhimurium* strain LT2, the principal strain studied as a model for cellular and molecular biology in *Salmonella*, was isolated in the 1940s and used in the first studies on phage-mediated transduction. Virulence has been related to deficiency and excess of different transition metal ions (such as copper, cobalt, iron, manganese, nickel, zinc) (Boyer E. *et al.* 2002; Zaharik *et al.* 2004) and the control of transition metal availability in this organism depends on the actions and the interactions of metal sensing regulatory proteins. *Salmonella* possesses sixteen deduced DNA-binding metal-sensing transcriptional regulators (Figure 1. 1) five of which have previously been characterized. In the following sections the families of metalloregulators characterized in various organisms will be presented.

1.2.1 *CsoR/RcnR-like de-repressors*

Escherichia coli RcnR (resistance to cobalt and nickel repressor) and *Mycobacterium tuberculosis* (Mtb) CsoR (copper-sensitive operon repressor) are the founding members of a large family of bacterial metal-responsive DNA-binding proteins (Blaha *et al.* 2011) (Figure 1. 2).

RcnR was first identified as a repressor of *E. coli rcnA* (Ni(II)/Co(II) inducible resistance gene) expression (Iwig *et al.* 2006). The induction of this gene, which is divergently transcribed from *rcnR*, is Ni(II)-dependent therefore questions were raised as to whether its regulation was correlated with that of *nikABCDE* expression (Iwig *et al.* 2006). Iwig and coworkers used *lacZ* reporter gene assay and EMSA analyses to confirm that RcnR is a Ni(II)/Co(II) dependent *rcnA* regulator. It is thought that RcnR prevents premature shutdown of the *nik* import system by preventing NikR sensing spill over nickel ions (Iwig *et al.* 2006). When the copper sensing *Mycobacterium tuberculosis* CsoR

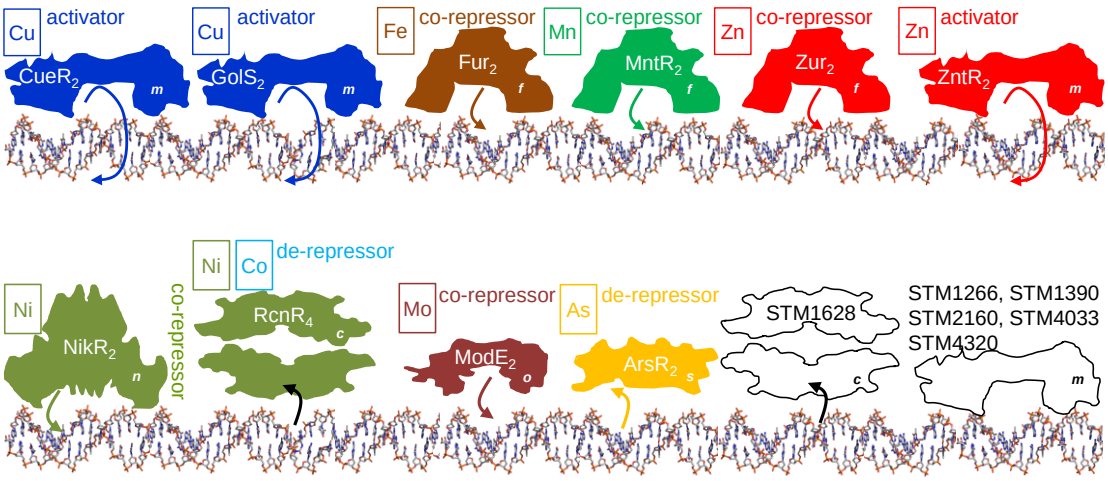


Figure 1. 1 The families are: m = MerR-like activators (CueR, GolS, ZntR, STM1266, STM1390, STM2160, STM4033, STM4320), f= Fur-like co-repressors (Fur, MntR, Zur), n = NikR-like co- repressor (NikR), c = CsoR/RcnR-like de-repressors (RcnR, STM1628), o = ModE-like co- repressor (ModE), s = SmtB/ArsR-like de-repressor (ArsR). At the start of the project, five of the shown regulators have been characterised in *S. typhimurium* (CueR, GolS, Fur, MntR and Zur). For sensors ZntR, NikR, RcnR, ModE, and ArsR were predicted the correct metal operator promoter targets considering the sequence similarity and/or gene context. In addition, five MerR-homologues and one RcnR/CsoR family member have been identified as possible metal-sensing representatives.
This figure was produced by Dr. Jen Cavet for the Industrial Partnership Awards Proposal presented to P&G.

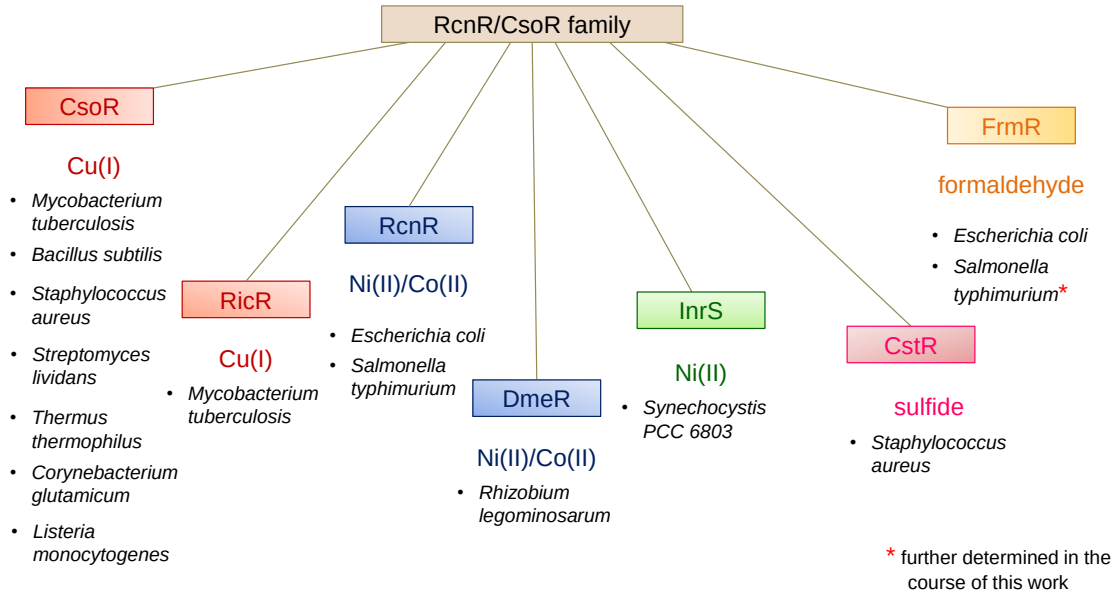


Figure 1. 2 Scheme representing members of the RcnR/CsoR family of DNA-binding transcriptional de-repressors, their cognate metals and the microorganisms in which they have been isolated.

transcriptional regulator was discovered by Liu and collaborators (Liu *et al.* 2007), a novel metalloregulatory family was discovered.

The Cu(I)-sensitive operon repressor CsoR was subsequently characterized in various Gram-positive bacteria, such as *Bacillus subtilis*, *Staphylococcus aureus*, *Thermus thermophilus*, *L. monocytogenes*, and *Streptomyces lividans* (Smaldone & Helmann 2007; Ma *et al.* 2009a; Baker *et al.* 2011; Sakamoto *et al.* 2010; Corbett *et al.* 2011; Dwarakanath *et al.* 2012). CsoR homologues were also identified in cyanobacteria and proteobacteria, however it is not yet clear if they are involved in copper regulation. The gene coding for CsoR, *csoR*, is located upstream of the copper responsive gene *ctpV* which is a putative copper exporter involved in the copper efflux response of *Mycobacterium tuberculosis* to copper stress (Ward *et al.* 2010) The *cso* (copper sensitive operon) operon includes *csoR*, *ctpV* and an uncharacterized gene located between the two, which are co-transcribed (Liu *et al.* 2007).

In addition to RcnR and CsoR, various family members have been characterised, including InrS (internal nickel-responsive repressor), RicR (regulated in copper repressor), DmeR (divalent metal efflux repressor) and the non-metal responding CstR (CsoR-like sulphur transferase repressor) and FrmR (formaldehyde responsive repressor).

1.2.1.1 Ni(II)- and Co(II)-sensing repressor RcnR

In *E. coli*, regulation at the transcriptional level of genes coding for RcnA and RcnB, (responsible for Ni(II) and Co(II) export when the concentration of these ions reaches the toxicity limit) is operated by RcnR. RcnR is a homotetramer of ~ 40 kDa that regulates the transcription of RcnA and RcnB by binding to the promoter in the apo-form. Although RcnR is capable of binding a number of transition metals, only coordination of Ni(II) and Co(II) triggers the conformation change which leads to the disruption of the protein:DNA interactions as the affinity of the holo-form for the promoter region is strongly decreased (Iwig *et al.* 2008, Higgins *et al.* 2012). One Ni(II) or Co(II) ion binds to a RcnR protomer with low-nanomolar affinities (2.5×10^{-8} and 5×10^{-9} M, respectively) which stabilizes the protein to denaturation (Iwig *et al.* 2008). In *E. coli*, when the concentration of free Ni(II) or Co(II) increases above the toxicity level, RcnR binds the metal ions leading to induction of *rcnA* expression Iwig *et al.* 2008) and the correct [Ni(II)] or [Co(II)] will be restored due to the resulting efflux. Both Ni(II)- and Co(II)-RcnR complexes possess a common N-terminal coordination motif (NH_2-Xaa-NH-His) but the remaining coordination environment is different for the two transition metals. As revealed by XAS, XANES and EXAFS analyses, Ni(II) and Co(II) bind in a pseudo-octahedral geometry to a similar six-coordinated metal site (with ligands $Cys(His)_2(N/O)_3$) (Iwig *et al.* 2008). Each metal coordinates a conserved cysteine ligand with distinct M-S distances, 2.54 Å for Ni(II)-RcnR and 2.24 Å for Co(II)-RcnR). The long distance observed by Iwig and collaborators (Iwig *et al.* 2008) can be explained considering that six-coordinate Ni(II)

complexes are high spin (with two unpaired electrons), with a small crystal field splitting energy which renders the complex substitutionally labile, characterised by longer M-L interatomic distances.

Since *E. coli* contains an additional nickel-responsive transcriptional regulator, NikR, Ni(II)-binding sites of the two proteins were extensively compared. When nickel stress is present in the cytoplasm the first to detect Ni(II) is NikR which subsequently binds to its operator region in order to prevent transcription of *nikABCDE* cassette which would otherwise allow production of the NikABCDE uptake permease. Only when full repression of the cassette is achieved, RcnR, which possesses a weaker affinity for Ni(II), starts to detect the metal and does dissociate from *rcnA* operator (Iwig *et al.* 2006).

E. coli NikR is characterized by two Ni(II) binding sites, where the tightest has three His and one Cys involved in Ni(II) coordination (Schreiter *et al.* 2003). This metal site has a high crystal field splitting, resulting in a low-spin complex with a square planar coordination geometry (Carrington *et al.* 2003, Schreiter *et al.* 2003). The lower affinity (comparable to the Ni(II)-site in RcnR) site is high-spin, six-coordinated but unlikely to be metallated due to a greater abundance of RcnR. The intrinsic difference between the high-spin (in RcnR) and the low-spin (in NikR) Ni(II)-binding sites results in the two proteins adopting different ligand exchange mechanisms, in fact RcnR is supposed to undergo a dissociative ligand-substitution (favored in tetrahedral and octahedral complexes) whereas NikR an associative ligand-substitution (favored in square planar complexes) (Iwig *et al*. 2008; Wilkins 1991). The mechanism adopted by NikR means that a conformational change must occur in order to allow metal release as the metal-binding site becomes buried and not easily accessible (Wilkins 1991; Rowe *et al.* 2005). Conversely, RcnR interaction with Ni(II) is readily disrupted as the metal ion level starts to decrease (Iwig *et al.* 2008; Rowe *et al.* 2005). Both the Ni(II)-binding site in RcnR and the high affinity Ni(II)-binding site in NikR contains one cysteine residue coordinates the divalent metal. This amino acid residue is characterized by a thiolate group, which is a soft ligand according to the HSAB theory, and is often used in metal binding (an example are the cysteines involved in Zn(II) coordination in Zinc Finger domains, (Tang *et al.* 2014) or the [NiFe]-hydrogenase which binds Ni(II) through Cys residue). Cysteines, usually deprotonated, bind metals by interacting with a single metal ion or forming a bridge between two, exploiting the two lone electron pairs on the sulfur. Sulfur shows both σ- and π-electron donation as it possesses occupied *p* orbitals, therefore it tends to form a stronger interaction with metals which present unfilled σ-symmetry orbitals. Ni(II)-NikR is a low spin complex (empty σ-orbitals, Ni-Cys 2.13 Å; Carrington *et al.* 2003), whereas RcnR forms a high-spin complex with Ni(II) and the partially occupied σ-orbitals result in a longer M-L bond (Ni-Cys 2.54 Å; Iwig *et al.* 2008). Since Co(II) is a d^7 ion with one vacancy in the π-orbital allowing π-interactions, the M-L interaction will have an intermediate length between the two just described (Co-Cys 2.24 Å; Iwig *et al.* 2008).

Site-directed mutagenesis analysis suggested which residues are involved in Ni(II)- and Co(II)-binding sites, as already proposed by EXAFS data (Iwig *et al.* 2008). Both Ni(II) and Co(II) are coordinated through the primary amine NH_2 of the N-terminal region (Iwig *et al.* 2008). This portion was shown to be fundamental in order for RcnR to distinguish between cognate and non-cognate metals, as only Ni(II) and Co(II) are detected employing this coordination motif (Higgins *et al.* 2012). Although RcnR is able to bind a number of additional metals (Zn(II), Cu(II), Cu(I)), it does so involving three protein residues and one anion from the buffer, not including the N-terminal amine, resulting in these metals inability to elicit an allosteric conformational change (Higgins *et al.* 2012). The side-chain of His 3 appears to be decisive in metal selectivity in RcnR. Both mutation of the His3 into a Leu or insertion of an extra Ala in position 2, disrupt Ni(II)- and Co(II)- binding implying a fundamental role of the N-terminus region in metal responsiveness (Iwig *et al.* 2008; Higgins *et al.* 2012). Since RcnR has a preference for high coordination metal complexes, the introduction of non-coordinating ligands and subsequent formation of a metal binding site with ligands placed with a square-planar geometry, led to no detectable responsiveness of Co(II) and Ni(II) (Iwig *et al.* 2008). Further confirmation of His3 being involved in metal coordination came from mutation of the residue to a potential metal ligand such as Cys or Glu with generation of a residual metal response, although substantially impaired (Higgins *et al.* 2012). Combination of XAS analyses and site-directed mutagenesis suggests that in addition to His3 sidechain also the backbone amide constitutes a Ni(II) and Co(II) ligand (Iwig *et al.* 2008).

In addition to His3 further metal-binding residues were identified in His64 and, as already mentioned, Cys35, which contribute to metal coordination in different ways. Spectroscopic studies conducted by Iwig and collaborators (Iwig *et al.* 2008) with C35ARcnR in the presence of Co(II) or Ni(II) demonstrate that the residue is a key ligand in metal binding. However, substitution of Cys35 to Ala fails to disrupt Ni(II) responsiveness which may be unexpected. An explanation alludes to a more conspicuous role of Cys in metal-selectivity rather than the allosteric conformational change upon metal binding (Iwig *et al.* 2008). In the literature it is possible to find examples of characterized metalloregulator proteins where metal-binding residues were distinguished between those having an allosteric function (i.e. capable to communicate the conformational change upon metal-binding to the DNA-binding regions) and those necessary to maintain the preferred geometry (Eicken *et al.* 2003; Pennella *et al.* 2006). However, a C35ARcnR variant shows a completely abolished Co(II) responsiveness, indicating a structural difference in the Ni(II)/Co(II)-mediated mechanisms regulating DNA release from RcnR (Iwig *et al.* 2008).

His60 appears to be involved in Co(II) coordination but not Ni(II), as shown by spectroscopic and mutagenic studies (Iwig *et al.* 2008), suggesting that RcnR can implement this extra ligand in order to discriminate between the two divalent metals.

Summarizing these findings, Co(II) is coordinated by the NH_2-terminus, His3 backbone amide and His3, Cys35, His60 and His64 sidechains, whereas Ni(II)-coordination differs only for not using His60. The sixth position remains therefore free to be coordinated by a solvent molecule or an additional main chain amide (Iwig *et al.* 2008).

1.2.1.2 Copper(I)-sensing repressor CsoR

CsoR was firstly structurally characterized by Liu and co-workers (Liu *et al.* 2007) showing an α-helical protein constituted by a dimer of dimers where four helices (α1-α2-α1'-α2') are arranged in a four-helix bundle architecture and two additional helices (α3/α3') from the C-terminal region are organized at the tetramer interface. CsoR binds cuprous ions with extremely tight affinity (K_d spans a range of 10^{-18} - 10^{-21} M amongst different organisms; Dwarakanath *et al.* 2012; Ma *et al.* 2009a; Liu *et al.* 2007) at the periphery of the bundle. The protein adopts a trigonal coordination geometry which employs two cysteine residues (in *M. tuberculosis* notation, Cys36 from the α2-helix and Cys65' and His61' from the opposite subunit within the dimer) as shown by X-ray crystallography (Liu *et al.* 2007). XAS (X-ray absorption spectroscopy) and EXAFS (extended X-ray absorption fine structure) spectroscopies also confirmed that *M. tuberculosis* CsoR binds Cu(I) ion via a S_2N ligand set allocated in a trigonal coordination geometry (Liu *et al.* 2007). Although XAS analysis executed on H61A reports a two-coordinate Cu(I), suggesting that His61 is involved in cuprous binding, EXAFS analysis of wild-type CsoR does not exhibit features attributable to imidazole coordination, which can be explained by taking in account a different Cu(I) coordination in solution or a diverse orientation of the imidazole ring with respect to the Cu-N bond (Liu *et al.* 2007). Further investigation conducted by EMSA on C36A and H61A CsoR variants display the loss of copper responsiveness and a significant weakening of Cu(I) affinity (Liu *et al.* 2007).

Analogous structural studies conducted on *B. subtilis* (*Bsb*) and *S. aureus* (*Sa*) CsoR homologues confirmed a 3-coordinate Cu(I) with M-S distance similar to what reported for *M. tuberculosis* (*Mtb*) CsoR (Ma *et al.* 2009a; Grossoehme *et al.* 2011).

Mutation of Glu90 in *Bsb*CsoR (corresponding to Glu81 in *Mtb*CsoR) does not affect Cu(I) binding at least in as much as the protein variant retains an analogous copper affinity (Ma *et al.* 2009a).

Stabilization of copper-CsoR is further accomplished with the aid of a hydrogen-bond network which involves His61', Glu81' (from one subunit) and Tyr35 (from the other subunit) that is important to communicate the Cu(I)-binding to CsoR to the protein regions involved in DNA-interaction (Higgins *et al.* 2007) (see Section 7.3.2 for further discussion). Glu81 has been identified as central in the allosteric response regulated by cuprous coordination (Liu *et al.* 2007). In fact, when Glu81 is mutated to a residue unable to bind metals (E81 → A), the CsoR variant still binds the *cso* operon with an affinity comparable with that of wild-type but Cu(I) coordination does not result in an equivalent weakening of DNA binding affinity (Liu *et al.*

2007). A similar result was found by Ma *et al*. (2009a) in *B. subtilis* CsoR (Glu90 in *Bsu*CsoR notation) corroborating the hypothesis that the hydrogen-bond network is functionally important in driving negative allostery regulation of DNA binding by Cu(I) amongst CsoR homologues .

In a recent study by Chang and collaborators (Chang *et al.* 2014) *Geobacillus Thermodenitrificans* CsoR (*Gt*CsoR) has been used as a model to study CsoR proteins belonging to other mesophilic bacilli such as *B. subtilis*, which differ significantly from the better characterised *M. tuberculosis* CsoR. The interpretation presented in this work is that several Cu(I)-sensing CsoR proteins have evolved within the RcnR/CsoR family, conserving specific residues (other than those belonging to the X-Y-Z motif or the secondary coordination sphere) which allow discrimination amongst these sub-CsoR families (Chang *et al.* 2014). Members of the *Gt*CsoR sub-family employ Arg65 and Lys101 to achieve high affinity DNA-binding; upon Cu(I) binding these residues are either involved in metal-binding or merely reoriented causing the protein to rearrange in a low affinity DNA-binding state (Chang *et al.* 2014)

Use of NMR and SAXS analysis to investigate the allosteric changes induced by Cu(I)-binding to *Gt*CsoR, shows that metal binding elicits a compaction of the tetramer, with the N-terminal tail redirected toward the Cu(I)-binding site and the formation of hydrogen-bonds and electrostatic interactions which stabilise the holo- form of the protein (Chang *et al.* 2014)

Since the N-terminal region is also employed by non-Cu(I) responsive CsoR family members, such as *E coli* RcnR (Iwig *et al.* 2008) and *Synechocystis* InrS (Foster *et al.* 2012; Foster *et al.* 2014), Chang and collaborators propose to use the *Gt*CsoR model to better comprehend how these proteins coordinate the respective cognate metals using the N-terminal region. Since residue Arg18 in *Gt*CsoR is shown to be crucial during the folding process of the N-terminal tail, the aligned residues in *Ec*RcnR (His3) and *Sy*InrS (His21), were proposed to be involved.

1.2.1.3 Regulated in copper repressor RicR

Festa and co-workers (Festa *et al.* 2011) identified a second Cu(I)-inducible pathway involved in copper homeostasis in *Mycobacterium tuberculosis*, in addition to CsoR. The two systems appear to be mutually exclusive but not redundant (Festa *et al.* 2011; Shi *et al.* 2014). RicR, named as "regulated in copper repressor", regulates the expression of five genes, *lpqS, Rv2963, mymT, socAB* and *ricR*, all sharing the palindromic motif 5'-TACCC-N_5-G/AGGTA-3' which is recognized by RicR in the absence of Cu(I) stress. Since the constitutive expression of the five genes was observed upon *ricR* disruption, it was hypothesised that RicR may have a protective role against toxic levels of Cu(I) (Festa *et al.* 2011). *mymT* is the only gene, amongst those regulated by RicR, which has already been characterised (Gold *et al.* 2008) and encodes for a copper metallothionein probably involved in Cu(I) resistance. Rv2963 and LpqS possess histidine-rich domains therefore they could work together in order to export cellular Cu(I).

Moreover, it has been suggested that RicR may regulate the expression of the multi-copper oxidase Rv0846c (divergently expressed from *lpqS*) which oxidizes Cu(I) to the less toxic Cu(II) (like CueO in *E. coli* (Kosman 2010)) and hence could contribute to protect *M. tuberculosis* from Cu(I) noxious effects (Festa *et al.* 2011).

1.2.1.4 Dimer metal efflux repressor DmeR

Cupriavidus metallidurans, dmeF gene encodes for a cation diffusion facilitator (CDF) which is involved in Ni(II)/Co(II) resistance (Munkelt *et al.* 2004). A homologue of this gene (39 % amino acid identity) was identified in *Rhizobium leguminosarum*, a bacterium able to establish a symbiotic relationship with legume plants, fixing atmospheric nitrogen into ammonia (Canfield *et al.* 2010; Rubio-Sanz *et al.* 2013).

A gene sharing high similarity with *E. coli rcnR* was identified upstream *dmeF* and sequence alignment revealed conservation of the W-X-Y-Z motif (H3-C35-H60-H64) (Rubio-Sanz *et al.* 2013) involved in Ni(II)/Co(II) response in RcnR (Higgins *et al.* 2013; Iwig *et al.* 2008). The gene was named *dmeR* (dimer metal efflux repressor). DmeR interacts with a palindromic type-1 region (according to Iwig & Chivers notation (Iwig & Chivers 2009, Sections 7.3.1)) upstream *dmeR*, suggesting a repression mechanism similar to RcnR.

1.2.1.5 Internal nickel-responsive sensor InrS

InrS was first characterised in *Synechocystis* PCC 6803 (hereinafter called *Synechocystis*) by Foster and co-workers in 2012 (Foster *et al.* 2012). This repressor regulates the expression of *nrsD*, which is part of the *Synechocystis* Ni(II)-efflux operon, *nrsBACD* (Garcia-Dominguez *et al.* 2000; Foster *et al.* 2012). NrsS (periplasmic Ni(II) sensor) and NrsR (transcriptional regulator) also participate in *nrs* regulation, allowing Ni(II) efflux in the presence of toxic metal concentration (Garcia-Dominguez *et al.* 2000; Lopez-Maury *et al.* 2002). Since cyanobacteria are ancestors of chloroplasts, their metal homoeostasis is particularly important. Moreover this class of bacteria is characterised by special demands for metals due to the photosynthetic machinery (Keren N. *et al.* 2004; So &, Espie 1998; Tottey *et al.* 2007). Although InrS senses Ni(II) ions, it shares more similarities with CsoR rather than RcnR (Foster *et al.* 2012). Moreover the presence of an N-terminal arm rich in His residues, not identified in RcnR or CsoR, further differentiates this protein from the other family members (Foster *et al.* 2012).

InrS coordinates Ni(II) by adopting a square planar coordination geometry (like Ni(II)-CsoR or Ni(II)-NikR) (Foster *et al.* 2012; Ma *et al.* 2009a; Wang *et al.* 2004). It was shown that InrS Ni(II)-binding affinity is tighter than the other metal sensors present in *Synechocystis* (ZiaR, ZuR, CoaR), explaining why nickel ions do not trigger an allosteric response in these proteins *in vivo* (Foster *et al.* 2012). Site-directed mutagenesis was employed to test the role of specific residues in InrS allosteric regulation. In view of the A-B-W-X-Y-Z motif, InrS possesses a His

in position Y (His78) and a Glu in position B (Glu98) like CsoR, but lacks a tyrosine residue in position A. In addition, a conserved Glu residue (Glu95 in InrS notation) was identified in other InrS homologues and its position was designated as "C" (Foster *et al.* 2012). Glu95 and Glu98 were demonstrated to contribute toward the allosteric mechanism without being absolutely necessary. Moreover, unlike CsoR, the hydrogen bond between His78 (first coordination sphere) and Glu98 (second coordination sphere) is not essential in coupling Ni(II)-binding to the weakened DNA-binding affinity in InrS ((Foster *et al.* 2012).

A recent study reports Cu(I), Cu(II) and Zn(II) ability, in addition to Ni(II), to trigger InrS conformational change *in vitro* (Foster *et al.* 2014a). Moreover, Foster and co-workers proved that if zinc concentration transiently exceeds the set-point for InrS detection (defined by InrS zinc binding constant), the protein is able to sense Zn(II) ions *in vivo* (Foster *et al.* 2014a).

1.2.1.6 CsoR-like-sulfur-transferase repressor CstR

Although CstR is a member of the RcnR/CsoR family of transcriptional regulators, this protein does not sense cellular metal stress but sulfite, S-methylmethanethiosulfonate (MMTS) and persulfide (Grossoehme *et al.* 2011; Luebke *et al.* 2013, Luebke *et al.* 2014). This protein was first identified in *S. aureus,* a Gram positive pathogen (Lowy 1998; Nizet 2007; Conrady et al. 2008) which lacks a biological pathway to assimilate sulfate as source of sulfur to be used in biosynthesis of cysteines and Fe-S clusters (Lithgow *et al.* 2004; Jacob *et al.* 2003). It has been shown that perturbation of cysteine metabolism impacts on the ability of this pathogen to produce the biofilm needed to survive outside and inside the host (Soutourina *et al.* 2009; Soutourina *et al.* 2010). However, *S. aureus* may use inorganic thiosulfate, as sulfur source, which can be produced by the mammalian host (Lithgow *et al.* 2004; Grossoehme *et al.* 2011).

CstR shares the critical features of its homologue CsoR, which is also present in this organism and regulates the proposed Cu(I)-exporter *copA* (Grossoehme *et al.* 2011). As *M. tuberculosis* and *B. subtilis* CsoRs (Ma *et al.* 2009a; Ma *et al.* 2009b), *S. aureus* CstR tightly binds two type 2 operator sites (4 GC base-pairs flanked by AT-rich regions) located within *the cstR-cstA* intergenic region (Grossoehme *et al.* 2011). Although *S. aureus* CsoR binds the same operator sites with an affinity ~ four-fold tighter than CstR K_{DNA}, its inability to adopt a higher order structure as CstR does (four tetramers bound to two operator sites) prevents the regulation of the *cst* genes by CsoR (Grossoehme *et al.* 2011).

Repression of *cst* expression is alleviated in conditions that can alter thiol-disulfide homeostasis (including sulfite-stress), as a pair of intra-protomer cysteine residues (Cys31 and Cys60') interact with the effector (e.g. sulfite) yielding a di- or tri-sulfate bond across the protein which leads to a conformational change and a weakening in the DNA-binding affinity (Grossoehme *et al.* 2011; Luebke *et al.* 2013). Furthermore, CstR was shown to react with chalcogen oxyanions yielding $CstR_2^{(RS\text{-}SR')}$ and $CstR_2^{(RS\text{-}X\text{-}SR')n}$ (n = 1 or 2; X = Se, Te) species triggering *cst* genes

expression (Luebke *et al.* 2013). Fluorescence anisotropy analyses conducted in the presence of SeO_3^{2-} (selenite), TeO_3^{2-} (tellurite) and SeO_3^{2-} (tetrathionate) oxyanions displayed a substantial decrease in the operator-binding affinity, suggesting that CstR does not discriminate amongst the nature of the cross-linking agent (Luebke *et al.* 2013). More studies may be necessary in order to elucidate selenite and tellurite toxicity in *S. aureus*.

1.2.1.7 Formaldehyde repressor FrmR

E. coli FrmR was identified in the course of a study aimed to determine the effects of reading-through amber stop codons. FrmR possesses such a stop and appeared to regulate the expression of *frmRAB* operon (Herring & Blattner 2004). FrmA has formaldehyde dehydrogenase activity and the operon was subsequently shown to respond to exogenous formaldehyde (when cells were treated with formaldehyde it was observed a 215-fold induction increase in *frmR* expression) (Herring & Blattner 2004; Gutheil *et al.* 1992).

Formaldehyde is a highly toxic compound because of its ability to cross-link proteins and nucleic acids. The origin of formaldehyde in biological systems, and its toxicity and detoxification mechanisms will be discussed in details in Section 1.5. Herring and co-workers identified a glutathione-dependent formaldehyde-resistance pathway in *E. coli* that involves a formaldehyde dehydrogenase (FrmA) and an S-formylglutathione hydrolase (FrmB), both regulated by the transcriptional repressor FrmR. First formaldehyde spontaneously reacts with intracellular glutathione yielding an S-hydroxymethylglutathione which will be then oxidised in S-formylglutathione by FrmA. The last step involves conversion of S-formylglutathione to glutathione and formate catalysed by FrmB (Herring & Blattner 2004). Formate is further oxidised to carbon dioxide which can then diffuse to the periplasm.

The *in vitro* biophysical properties of *E. coli* FrmR have been investigated in a PhD dissertation by J. R. Law (Law 2012). The project aimed to further understand how bacteria sense formaldehyde and the connection with formaldehyde detoxification systems. All the attempts to determine the crystal structures of wild-type FrmR failed, probably due to oxidation modifications, whilst C36SFrmR crystals were successfully produced and analysed by X-ray diffraction, although the merohydral twinning of the crystals probably caused the inability to obtain a refined model (Law 2012).

The study revealed that FrmR protomer assembles in helical tetrameric oligomerization state (like RcnR and CsoR), and specifically binds to the promoter of *frmRAB* operon *in vitro* by EMSA analysis. Moreover HCOH was shown to impair DNA-binding, confirming the results by Herring and collaborators (Herring & Blattner 2004). In addition, Cys36 was proved to be essential in the formaldehyde induced derepression mechanism, and the author proposed the formation of a covalent adduct via successive cysteinyl thiol nucleophilic attack to formaldehyde (Law 2012). Cys36 corresponds to the conserved cysteine residue involved in

RcnR and CsoR metal sensing (Iwig *et al.* 2008) (Liu *et al.* 2007). *Mtb* CsoR's Cys36 is located in a loop in the middle of the peptide chain, and Cu(I)-binding is coupled to a change in the loop position initiating the reciprocal movements of the α-helices. As a consequence, the DNA-binding residues change positions resulting in a conformation with weak DNA-binding affinity (Liu *et al.* 2007). Since *E. coli* FrmR's Cys36 is also located in a similar loop, the author suggested an analogous mechanism (Law 2012).

A number of computer programmes were used in order to predict the residues involved in DNA-recognition and three highly conserved candidates, Arg14, Arg46 and Lys91, were identified. Mutation of these residues to an Ala resulted in a diminished or abolished DNA-binding affinity *in vivo*. Arg14 corresponds to Arg15 in *Mtb* CsoR, which has been shown to be essential for CsoR's DNA-binding by creating a positively charged patch that may electrostatically interact with the DNA phosphate backbone (Ma *et al.* 2009a), suggesting a similar recognition path in *E. coli* FrmR. Moreover, a BLAST search performed by Law (2012) in order to identify FrmR orthologues, detected *frmR* in forty-two bacterial species, denoting that, amongst the others, the first 60 residues and the last three C-terminal residues are highly conserved. This finding contributes to the hypothesis that also K91 is involved in DNA-recognition, and further confirmation comes from the inability of the His-tag FrmR version to bind *frmRAB* in EMSA analyses.

A recent work from Denby and collaborators (Denby *et al.* 2015) shows that this operon is induced during anaerobic respiration using trimethylamine-N-oxide (TMAO) as the terminal electron acceptor. TMAO is an osmolyte in marine organisms and can be reduced by the TMAO reductase TorA to yield trimethylamine (TMA) (Mejean *et al.* 1994). When *E. coli* cultures were exposed to [TMAO] ≥ 5 mM, the transient induction of *frmRAB* operon was observed (Denby *et al.* 2015). Moreover, after TMAO addition up to 0.23 mM dimethylamine (DMA) was detected, a sub-millimolar value comparable to the formaldehyde concentration needed to induce the *frmRAB* operon (Denby *et al.* 2015; Herring & Blattner 2004). These findings led the authors to suspect the presence of TMAO demethylase activity which would produce DMA and formaldehyde from TMAO (Denby *et al.* 2015). The endogenous formaldehyde, generated as a by-product of TMAO demethylation, would then trigger *frmRAB* operon expression in response to potentially toxic formaldehyde levels. The authors confirmed this theory by observing the inhibition of growth in an engineered *frmRAB* mutant unable to detoxify formaldehyde after exposure to 40 mM TMAO for 60 min (Denby *et al.* 2015).

In addition to formaldehyde and TMAO, also carbon monoxide and chloride treatments have been shown to trigger expression of the *frm* operon (Wang *et al.* 2009; Nobre *et al.* 2009). Carbon monoxide, CO, is produced from the incomplete oxidation of organic matter and from haem oxygenase activity (Li *et al.* 2007). Its toxicity originates from the preference of CO to bind proteins containing transitional metals, inducing conformational modifications and altering

the biological function (Nobre *et al.* 2009). When the *E. coli* global transcriptome was analysed following treatment with the CO-releasing CORM-2 (tricarbonyldichlororuthenium (II) dimer), the *frmRAB* operon was induced in both aerobic and anaerobic conditions (Nobre *et al.* 2009).

Wang and collaborators tested the effects of NaClO (sodium hypochlorite) and H_2O_2 (hydrogen peroxide) on *E. coli* strains from eight clades to study how the cell copes at the transcriptional level with oxidative stress (Wang *et al*. 2009). They found that the *frmRAB* operon, amongst others, was dramatically induced by NaClO but not by H_2O_2 treatment. They also registered the induction of genes involved in the glutathione-dependent detoxification pathway (Wang *et al*. 2009) which play a vital role during oxidative stress due to the ability of glutathione to scavenge oxidant compounds (Haenen & Aalt Bast 2014).

1.2.2 *Other families of transcriptional regulators*

1.2.2.1 MerR family

The mercury resistance operon repressor MerR regulates the transcription of the gene coding for the mercury reductase MerA (Schelert *et al.* 2004). MerA is involved in the mercury detoxification pathway across a wide range of organisms catalysing the reduction of Hg(II) to the less reactive and therefore less toxic Hg(0), exploiting the consequent oxidation of NADPH. Regulators belonging to this family sense several environmental stimuli such as metal ions, lipophilic drugs and nitric oxide and superoxide (induced by oxidative stress) (Ahmed *et al.* 1994; Pomposiello & Demple 2001; Chen & He 2008; Brown *et al.* 2003). MerR family members possess similar N-terminal helix-turn-helix DNA binding regions and C-terminal effector binding regions that are specific to the inducer recognized.

This family of transcriptional regulators includes CueR (copper export regulator, Stoyanov *et al.* 2001; Espariz *et al.* 2007) and GolS (gold sensor Checa *et al.* 2007 Pontel *et al.* 2007; Perez Audero *et al.* 2010) which coordinate monovalent metals, Cu(I) and Au(I) respectively, employing two conserved Cys residues from the C-terminal region (metal binding loop, MBL) and ZntR (zinc responsive transcriptional regulator, Brocklehurst *et al.* 1999; Outten *et al.* 1999; Binet & Poole 2000) which sense divalent metal ions by using an extra Cys residue from the start of the α5 helix (Khan *et al.* 2002). This position is occupied by a Ser in monovalent metal sensors and it is thought that the steric and hydrophobic restrictions introduced by this residue participate to metal selectivity within the MerR family. The crystal structures of two *E. coli* MerR-like metalloregulators, Zn(II) specific regulator ZntR and Cu(I) sensor CueR, also contributed to this hypothesis (Changela *et al.* 2003). Comparing the different metal coordination sites of the two proteins suggests that the metal specificity of these MerR proteins derives from differences in coordination number and type of metal ligands utilized. For instance, CueR has an unusual metal binding mode where the metal-receptor site is buried and restricts the metal to a linear, two coordinate geometry (Changela *et al.* 2003). However, recent

work conducted by Ibañez and collaborators (Ibañez *et al.* 2015) shows that metal specificity is achieved not only by ligand selection but also through the flexibility of the metal binding cavity. In fact, if the aforementioned Ser is mutated to an Ala the activation pattern triggered by binding of monovalent metal ions is not altered. Moreover, when the extra Cys is introduced in GolS and CueR to replace the Ser residue, the metal binding pocket assumes a "pan-like" shape making both proteins able to sense with high sensitivity monovalent and Hg(II) ions (Ibañez *et al.* 2015).

1.2.2.2 ArsR/SmtB family

Members of this family of transcriptional regulators repress the expression of operons involved in heavy metals detoxification mechanisms. They possess homodimeric helix-turn-helix (HTH) motifs that specifically bind to their operator/promoter DNA binding sites in the metal-free apo state. De- repression of ArsR/SmtB-regulated promoters occurs when the proteins bind the respective metal resulting in a weaker affinity for DNA. The proteins belonging to this family possess one or more structurally distinct metal binding sites commonly derived from α3-helix and α5-helix ligands however many "themes and variations" exist (Busenlehner *et al*. 2003). Examples of the α3N/α3 subgroup are *Staphylococcus aureus* and *Listeria monocytogenes* CadCs homologues and *E. coli* ArsR. These proteins coordinate Cd(II), Pb(II), Bi(III) and As(III) through a set of Cys residues recruited from the α3-helix (in ArsR) and from the N-terminal region (in CadCs). *Synechococcus* SmtB, *Synechocystis* ZiaR, *S. aureus* CzrA and *M. tuberculosis* NmtR belong to the α5C/ α5 subgroup and exploit a mix of histidine residues and carboxylate groups from amino acid residues on the C-terminal α5 helices (Busenlehner *et al*. 2003). The founding members of this family are the Zn(II) repressor *Synechococcus* PCC 7942 SmtB which regulates the expression of the *smtA* gene (SmtA is a class II metallothionein involved in zinc(II) ions sequestration) (Morby *et al*. 1993) and the arsenic/antimony-responsive, *ars* operon repressor, ArsR (characterised in a number of bacteria, including *E. coli*) (Wu & Rosen 1991; Ji & Silver 1992), involved in the regulation of an arsenate reductase (Xu *et al*. 1998).

SmtB binding to its operator/promoter region is negatively affected preferentially by Zn(II) (Cavet *et al.* 2002) although also Co(II) and Cd(II) can be sensed *in vitro* (VanZile *et al.* 2000) and *in vivo* (Huckle *et al.* 1993). Other divalent metal ions can be sensed by members of this family, such as *S. aureus* ZntR (Zn(II) and Co(II)) (Singh *et al.* 1999; Xiong *et al.* 1998), *M. tuberculosis* NmtR (Ni(II) and Co(II)) (Rutherford *et al.* 1999; Cavet *et al.* 2002) and *S. aureus* CadC (Cd(II), Pb(II), Bi(II), Zn(II)) (Yoon *et al.* 1991; Endo & Silver 1995; Busenlehner *et al.* 2002a; Busenlehner *et al.* 2002b).

1.2.2.3 Ferric uptake regulator (Fur) family

In the early anaerobic environment prior the Great Oxygenation Event (GOE) bacteria used Fe(II) ions as co-factors. However, after this event organisms were forced to acquire Fe(III) which is insoluble in aqueous solution by employing siderophores, specialised molecules with high affinity for ferric ions. Moreover both Fe(II) and Fe(III) can generate reactive oxygen species (ROS) by Fenton chemistry (Fenton 1894) that can oxidise macromolecules and lipids (Imlay & Linn 1988; Lloyd *et al*. 1997) leading to cellular death (Von Harsdorf *et al.* 1999) (Hildeman *et al.* 2003). It is therefore evident that the cell must tightly control the level of iron ions.

Members of the Fur family are now known to be involved in sensing a variety of metals and possess various biological functions (Lee & Helmann 2006) including regulation of genes coding for iron uptake proteins (Hantke 1981; Hantke 1984), Fe(II), Zn(II) and Mn(II) transport systems and enzymes able to protect against ROS damage (Anjem & Imlay 2012; Anjem *et al.* 2009). The founding member is Fur (ferric ion uptake regulator) was first described in *E. coli* and controls iron homeostasis in many Gram-negative bacteria most commonly by binding to target regulatory sequences (fur boxes) (Calderwood & Mekalanos 1987; De Lorenzo *et al.* 1988, Calderwood & Mekalanos 1988; DeLorenzo *et al.* 1987; Stojiljkovic *et al*. 1994) in the promoter regions of iron-responsive genes when iron is replete (Ernst *et al.* 1978; Bagg & Neilands 1985; Bagg & Neilands 1987; Neilands 1993; Escolar *et al.* 1997; Escolar *et al.* 1998). When the iron concentration is low the expression of these genes occurs because Fur no longer acts as a repressor.

The first three-dimensional structure of a Fur family member was obtain by X-ray crystallography by Pohl and collaborators in 2003 (*Pseudomonas aeruginosa* Fur complexed with zinc(II) rather than iron(II)) (Pohl *et al.* 2003). Subsequently, many crystallographic structures of Fur-family regulators have been solved such as *Streptococcus pyogeness NS88.2* and *B. subtilis* PerR homologues (Lin *et al.* 2014; Traoré *et al.* 2006; Traoré *et al.* 2009; Jacquamet *et al*. 2009), with the metalloregulatory site either empty or containing Zn(II) and Mn(II) ions, *Streptomyces coelicolor* and *E. coli* Zur (Shin *et al.* 2011; Gilston *et al.* 2014), *Streptomyces coelicolor* Nur (An *et al.* 2009) and many other Fur homologues from *E. coli* (Pecqueur *et al.* 2006), *Helicobacter pylori* (Dian *et al.* 2011), *Vibris cholera* (Sheikh & Taylor 2009), *Campylobacter jejuni subsp. Jejuni* (Butcher *et al.* 2012). The regulators belonging to this family contain a C-terminal ββαβα dimerization domain linked to a typical winged helical domain and two metal sites per subunit.

1.2.2.4 ModE-like regulators

ModE regulates the ModABC molybdnenum transport system in a variety of organisms, although it is not ubiquitous (Wiethaus *et al*. 2006; Studholme & Pau 2003). In addition, *E. coli*

ModE was shown to repress genes coding for molybdoenzymes and molybdopterin synthesis in the absence of Mo (Grunden *et al.* 1996; Anderson *et al.* 1997). Molybdenum is required for the activity of enzymes involved in N, S, C metabolism (Hille 1996; Zhang & Gladyshev 2008), and is often associated to a pterin to yield the Moco co-factor. Pterin is thought to control the redox-chemistry of Mo and to direct the metal toward the active center of molybdenum enzymes (Kisker *et al.* 1997; Mendel & Kruse 2012). *E. coli* ModE coordinates molybdenum through a tandem repeat of the Molybdopterin-binding protein (Mop) domain situated at the C-terminus (Hall *et al.* 1999). The conserved SARNQ sequence was identified in a number of ModE homologues, and suggested to contain the positively charged amino acid residues involved in Mo coordination (Grunden *et al.* 1996).

DNA-binding is achieved by interaction of the N-terminal domain with an 8-bp inverted repeated sequence (TAAC GTTA) (Grunden *et al.* 1996).

1.2.2.5 NikR-like regulators

They are ribbon-helix-helix proteins which mediate the repression of the *nik* operon in Gram-negative bacteria and archaea in response to nickel(II). The first regulator of this family was identified in *E. coli* (De Pina *et al*. 1999; Chivers & Sauer 1999; Schreiter *et al.* 2003; Schreiter *et al*. 2006). Regardless of the presence of nickel(II) the structure is homotetrameric. In this metal-sensing protein the N-terminal domain binds to DNA and the C-terminal domains bind to nickel(II) (Chivers & Sauer 2002; Schreiter *et al.* 2003; Chivers & Sauer 1999). The crystallographic structure of the C-terminal region of *E. coli* NikR visualizes the drastic conformational changes required for operator recognition and shows the formation of a metal-binding site in the presence of DNA (Schreiter *et al.* 2003). In fact tetrameric NikR has two different Ni(II)-binding sites, one of them possessing a high affinity for nickel(II) which activates binding to DNA (Chivers & Sauer 2002; Wang *et al.* 2004). The second Ni(II)-binding site has a lower nickel(II) affinity and enhances the DNA binding affinity of NikR (Bloom & Zamble 2004; Chivers & Sauer 2002). When the protein is not bound to DNA nickel has square-planar coordination geometry whereas, upon binding to the operator DNA, the new metal site undergoes a change in geometry switching to a six-coordinate one. In this new arrangement there are six N/O donors incorporating at least a subset of the original square planar donors. This conformational change alters the orientation of the N-terminal DNA binding domains relative to the core domain and probably leads to the positive regulation of DNA binding by NikR (Schreiter *et al.* 2006; Phillips *et al.* 2008).

1.3 How metal sensors detect the right metal

It has been proposed that metal sensors achieve metal specificity by operating a selection of the metal ion detected on the bases of relative metal-binding affinity, relative allostery and relative

access (Waldron *et al.* 2009; Waldron & Robinson 2009; Foster *et al.* 2014b). In the following sections these thermodynamic and kinetic parameters will be further examined.

1.3.1 *Relative metal affinity*

Foster and collaborators have shown that nickel sensing in a cyanobacterium follows thermodynamics (Foster *et al.* 2012). An example of how relative metal affinity can explain the specific metal detection by a metal sensor is provided by the Ni(II)-sensor InrS, described in Section 1.2.1.5. InrS is characterized by a tighter affinity for Ni(II) when compared to the other metal sensors from *Synechocystis* PCC 6803 (Foster *et al.* 2012). Direct competition experiments between pairs of metal sensors confirmed the partition of Ni(II) ions to InrS (Foster *et al.* 2012).

Since InrS regulates the expression of *nrsD* (part of the Ni(II)-efflux operon *nrsBACD*) (Garcia-Dominguez *et al.* 2000; Foster *et al.* 2012), the level of cellular Ni(II) needed by the other sensors (ZiaR, ZuR and CoaR) to detect the metal will never be reached as the Ni(II) ions in surplus (exceeding the level set by $K_{Ni(II)}$ of InrS) will be expelled by NrsD (Foster *et al*. 2012). Therefore InrS achieves metal specificity by buffering intracellular nickel content such that the other metal sensors do not gain access to nickel (Foster *et al.* 2012).

1.3.2 *Allosteric mechanism triggered by effector binding*

Members of the RcnR/CsoR transcriptional repressors family detect anomalous cellular concentrations of several metals (e.g. Cu(I) or Ni(II)/Co(II) for CsoR and RcnR, respectively) in order to trigger a correct response. Allostery is exploited to link cognate metal-binding to the release of DNA operator-promoter region located upstream of the gene or operon regulated by these proteins.

According to the classical "concerted model", also named after the researchers who first described it (Monod-Wyman-Changeux, MWC model), the regulated protein exists in two different states depending on the presence or absence of the ligand molecule acting as an effector (Monod *et al.* 1965). Ligand-binding to a certain subunit of the protein will trigger the conformational change that, once propagated to the other protein subunits, will result in an alteration of ligand-binding affinities for these subunits (Monod *et al.* 1965).

Coupling allosteric free energy ΔG_c is a parameter which reports the magnitude of the allosteric driving force in a quantitative way (Grossoehme & Giedroc 2009; Reinhart 2004). Depending on the system and protein regulator characteristics, it can be calculated by measuring metal-binding affinity of free and DNA-bound protein or otherwise by measuring DNA-binding affinity of apo- and holo-protein forms (Grossoehme & Giedroc 2009; Reinhart 2004).

Grossoehme and coworkers (Grossoehme & Giedroc 2009) reported the coupled thermodynamic equilibria (assuming a closed system) which describes in a simplicistic way the relationship between protein regulator (P), *n* atoms of metal (*n* M) and a single DNA operator (D) (see Figure 3. 22A). The scheme represents the four end-states that a metalloregulator protein can adopt in the presence of its cognate metals and a limited amount of DNA (Grossoehme & Giedroc 2009). Each of these four equilibria is described by the respective affinity constant. The coupling constant (K_c) (Figure 3. 22B) is a dimensionless quantity which represents the magnitude of allosteric regulation and can be determined either from the ratio K_4/K_3 ($K_D^{Mn\cdot P}/K_D^{P}$) or the ratio K_2/K_1 ($K_M^{D\cdot P}/K_M^{P}$) and used to calculate ΔG_c using the standard thermodynamic equation (Equation 1):

$$\Delta G_c = -RT\ln K_c$$

Equation 1

where R is the ideal gas constant (8.314 $J\,K^{-1}\,mol^{-1}$) and T is 298 K (temperature at which the experiments is performed). In the case of DNA-binding de-repressors such as RcnR and CsoR, where the conformational change induced by cognate metal binding leads to an assembly state with a diminished DNA-binding affinity, the ternary complex (P·M)D is less stable than P·M and free P ($K_3 < K_1$; $K_4 < K_2$) and $K_c > 0$. As a consequence the ligand exchange equilibrium shown in Figure 3. 22B is shifted to the left ($\Delta G_c > 0$) (Grossoehme & Giedroc 2009). This approach suggests that in de-repressors a positive increment of allosteric free energy is necessary in order to link metal-binding to the structural change in protein assembly that will result in the release of the DNA operator-promoter.

An exemplary case is given by ZiaR, the Zn(II)-sensor present in *Synechosystis* PCC6803. Although the organism possesses also the Ni(II)-sensor InrS, which has a $K_{Zn(II)}$ comparable to the tightest sites of ZiaR, the allosteric response elicited by Zn(II) in ZiaR is greater to that observed in InrS. Since both proteins are derepressors this conclusion can be derived from the simple comparison of their $\Delta G_c^{Zn(II)}$. As a result, InrS is less effective than ZiaR in derepressing its target promoter region upon Zn(II)-binding, therefore to achieve a comparable degree of derepression it would need a higher concentration of Zn(II) (Foster *et al.* 2014b).

Another exemplary case occurs in *Mycobacterium tuberculosis*: the DNA-binding transcriptional repressor NmtR, senses surplus cobalt and nickel but has tighter affinity for zinc, while a related protein SmtB, detects zinc but does not detect nickel. Both of them contain a helix-turn-helix motif and the metal-sensing residues are located in the same regions of the respective folds but SmtB has only four binding residues whereas NmtR possesses six metal-binding residues. In this case the specificity for metal sensing is given by the coordination geometry preferred by the metals determining allostery. Nickel and cobalt in NmtR bind to six residues in an octahedral geometry, the geometry necessary to lead to the conformational

change in the sensor that alters its binding to DNA. However, binding of zinc takes place through only four of the six residues in a tetrahedral geometry and although zinc binds to NmtR tightly it does not produce the conformational change required to drive allostery (Cavet *et al.* 2002; Pennella *et al*. 2003) and as a result repression *in vivo* is not reduced. Experiments conducted *in vivo* showed no detection by NmtR of zinc (Cavet *et al.* 2002; Pennella *et al.* 2003).

1.3.3 *Channeling*

Another strategy for controlling metal availability involves protein-protein contact, for example there exist metallochaperones which traffic metals acting as intermediaries providing the right metal ion to the protein. The metal acquired by a protein from the carrier protein depends more on the protein-protein interactions than the affinity of the protein for the metal. For instance in *Synechocystis* sp. PCC 6803 the zinc exporter ZiaA binds copper(I) more tightly than zinc(II) but the copper-chaperone Atx1 does not interact with ZiaA, therefore the specific Atx1 interactions prevent ZiaA from gaining access to copper(I).

Another example of how thermodynamics cannot be applied to explain metal-sensing is provided by *Synechocystis* CoaR. CoaR is a Co(II)-sensing activator belonging to the MerR family of transcriptional regulators (Section 1.2.2.1) and regulates the transcription of CoaT, a P1-type ATPase Co(II) exporter (Rutherford *et al.* 1999). Co(II)-affinity of CoaR tightest site was estimated to be approximately within the range of 8.64 nM < $K_{Co(II)}$ < 1 mM by UV-vis spectroscopy and competition with the ratiometric fluorescent metal chelator Fura-2 (Fura-2; $K_{Co(II)}$ = 8.64 x 10^{-9} M) (Patterson *et al.* 2013). Since CoaR is an activator, and hence binds Co(II) when on its operator promoter, $K_{Co(II)}$ of the DNA-bound form of CoaR was also investigated, however, it did not appear tighter than that of the free-form (Patterson *et al.* 2013).

Since *Synechocystis* PCC 6803 possesses also metal sensors from other families (ZiaR from ArsR/SmtB family, InrS from CsoR/RcnR family, and ZuR from Fur family) which all possess a tighter Co(II)-binding affinity of their tightest sites (Patterson *et al.* 2013), the specificity of Co(II)-binding by CoaR cannot be explained by relative affinity. In the same work, Patterson and colleagues show by fluorescence anisotropy that Co(II) may act as an allosteric effector *in vitro* also for ZiaR, InrS and Zur, in addition to CoaR (Patterson *et al.* 2013).

CoaR possesses a precorrin isomerase-like domain (Rutherford *et al.* 1999), therefore the possibility that Co(II) ions are channeled to the metal binding site of the protein was invoked to explain the specificity of the detection of cellular Co(II) by the sensor. In fact, precorrin isomerases catalyse the methyl isomerization in cobalamin biosynthesis and their substrates are cobalt-binding tetrapyrroles (Moore & Warren 2012). This observation, along with the weak Co(II)-binding affinity, suggests that CoaR may not solely detect free Co(II) ions, but either a tetrapyrrole which may aid Co(II) insertion into the protein by enhancing the $K_{Co(II)}$ of the site,

or a pre-formed Co(II)-tetrapyrrole adduct (Patterson *et al.* 2013). These hypotheses were investigated by Patterson and coworkers by mutating the residues involved in the binding of tetrapyrrole substrates in precorrin isomerases and showing the consequent loss of activation of the Co(II)-dependent expression from the *coaT* promoter (Patterson *et al.* 2013).

1.4 DNA recognition

1.4.1 *Structure of DNA-binding proteins*

The SCOP (Structural Classification of Proteins) database classifies the proteins deposited in the PDB in distinct superfamilies on the bases of their structural properties (Murzin *et al*. 1995). Rohs and collaborators (Rohs *et al.* 2010) grouped the DNA binding proteins available in the SCOP database in 70 superfamilies depending on the secondary structure of the DNA binding domains. This group revised the more representative superfamilies in a review dated 2010 and a brief synopsis will be now reported here.

1.4.1.1 Mainly α motifs

This category includes proteins from 16 SCOP superfamilies (Rohs *et al.* 2010) which use α-helical domains to contact predominantly the DNA major groove, although interactions with the minor groove have also been reported (e.g. Lac repressor (Lewis *et al.* 1996; Schumacher *et al*. 1994). Helix-turn-helix (HTH), winged helix-turn-helix (wHTH), helix-loop-helix (HLH) and leucine-zipper motifs, all belong to this category. They are characterised by a recognition helix, which interacts with DNA by establishing hydrogen bonds and hydrophobic and van der Waals interactions, and a second helix which stabilizes the protein-DNA complex. The presence of one or more "wings" in the wHTH motifs allows extra-stabilization of the interaction usually by contacting the minor groove.

1.4.1.2 Mainly β motifs

β-domain structures can be identified in 7 SCOP families and include the TATA-box binding domain, Immunoglobulin-like β-sandwich, β-trefoil and β-β-β motifs (Rohs *et al.* 2010). TATA-binding proteins (TBP) recognize DNA minor groove using a β-sheet surface (Kim Y. *et al.* 1993; Kim J. L. *et al*. 1993) inducing a large distortion in the double helix which allows the formation of extra contacts between the concave surface of the protein and the edges of the nucleotides in the TATA-box. Runt-domains (Tahirov *et al.* 2001) and p53-like transcriptional factor (Cho *et al.* 1994) show the Immunoglobulin-like (Ig-like) β-sandwich fold. In this fold the domain can be oriented with different angles relative to the double helix axis, and the loops can be involved in the specific interaction with DNA.

1.4.1.3 Mixed α/β motifs

In 48 SCOP superfamilies, α/β domains are employed to recognize target DNA sequences (Rohs *et al.* 2010). An example is given by zinc-finger proteins, which possesses a short recognition helix, used to specifically contact the base pairs on the major groove, and two-stranded anti-parallel β-sheets (Pavletich & Pabo 1991). Classification of zinc-finger proteins is based on the residues coordinating the structural Zn(II) ion (Cys_2His_2, Cys_4, Cys_6, etc.). Further specificity is achieved by contacting the same DNA sequence with multiple zinc-fingers motifs (overlapping pattern).

A Ribbon-helix-helix motif can be identified in Met, Arc and NikR repressors (Somers & Phillips 1992; Raumann *et al*. 1994; Chivers & Sauer 1999). This motif is characterised by a two-stranded antiparallel β-ribbon, which is inserted in the major groove, and two helices which are involved in dimerization allowing interaction with other proteins possessing the same domain.

1.4.1.4 Multi-domain proteins

In order to achieve higher specificity and affinity for DNA, a large number of proteins establish contacts with more than one DNA binding domain. A representative example is MarA, a member of the AraC prokaryotic transcriptional activator family. This protein contains two α-helical sub-domains, each of them containing a HTH motif, and a long C-terminus loop. The two recognition helices (one from each sub-domain) interact with two consecutive major grooves on the same face of DNA inducing a distortion of the double-helix since the distance between the two α-helices is smaller than the pitch of B-DNA (Rhee *et al.* 1998).

1.4.2 *Sequence-dependent variations of DNA-structure*

DNA binding proteins can recognize specific sequences on the double helix exploiting sequence-dependent deviations in the DNA structure from the common B-DNA form (Rohs *et al.* 2009a). These variations of DNA structure can produce different electrostatic potentials that can be distinguished by proteins, and therefore have a fundamental role in DNA recognition. Structural deviations can be divided in global and local shape variations.

1.4.2.1 Global shape variations

The most common conformation of genomic DNA, B-DNA, is right-handed, with the base pairs perpendicular to the double-helix axis. The major groove is wider and shallower compared to the minor groove (Table 1. 1), which presents a more electronegative potential (Rohs *et al.* 2010).The potential difference between AT-rich and GC-rich tracts arises, in both grooves, from the disposition of the polar groups and AT-rich tracts exhibit a more negative potential (Rohs *et al.* 2010; Jayaram, *et al.* 1989; Lavery & Pullman 1981). Under conditions of dehydration it is

possible to observe a reversible conformational change in the B-DNA resulting in a wider and shallower structure named A-DNA (Lu *et al.* 2000). The same conformation can be detected in some protein-DNA complexes and is favored by GC-rich tracts (Shakked *et al.* 1989) A-DNA's major groove structurally resembles B-DNA's minor groove and shares a negative electric potential (Table 1. 1; Shakked & Rabinovich 1986). This occurs because the base-pairs are inclined at around 20º in the direction of the helix axis, narrowing the major groove. Therefore A-DNA causes compression along the helical axis. An A-DNA variant is the TA-DNA (tilted-DNA) which is observed in TATA-boxes where the rotation around the glycosidic bond causes a ~ 50º inclination of the base-pairs relative to the helix (Guzikevich-Guerstein & Shakked 1996).

The left-handed Z-DNA conformation is observed in alternating purine/pyrimidine sequences under high salt conditions (Rohs *et al.* 2010; Wang *et al.* 1979; Arnott, *et al.* 1980). The nucleotides along the sequence alternate in the syn and anti conformations producing a zig-zag backbone (Wang *et al.* 1985). The minor groove is narrow and deep, similar to B-DNA while the major groove is very shallow (Table 1. 1). The presence of salt stabilizes this structure reducing the electrostatic repulsion between the phosphate groups of the opposite filaments.

In addition to these deviations from the classical B-DNA shape, a global curvature of the DNA double-helix is observed in A-tracts (three or more consecutive A·T base pairs without a TpA step) (Rohs *et al.* 2010; Nelson *et al.* 1987; Hizver *et al.* 2001; Haran & Mohanty 2009). The flexibility of such DNA-tract ascribed to variations in the roll angles between the base pairs at the A-tract/ non-A-tract junctions and the rest of the chain (Hizver *et al.* 2001). Moreover Haran and coworkers suggest that the biological role of the A-tract, for example in gene regulation, may result from its distinct structure rather than the global bending induced in these sequences (Haran & Mohanty 2009).

1.4.2.2 Local shape variations

DNA kinks and minor groove narrowing are amongst the local shape variations (Rohs *et al.* 2010). The main difference between DNA kinks and DNA bending is that the first refers to a local deviation from the B-DNA shape, which affects only a single base pair, while the rest of the α-helix usually retains its linear conformation. This local deformation causes the loss of stacking of a base pair step and is likely to involve a pyrimidine-purine step, since these are less stabilized by base stacking. ApT step is the most stabilized by these contacts, whereas the TpA step, also defined as the "hinge" step (Crothers & Shakked 1999; Olson *et al.* 1998), is the most flexible as the small overlap between the base pairs does not allowed stabilization by stacking (Rohs *et al.* 2010).

Local variation of the minor groove width is caused by differences in the interactions made at each dinucleotide step in an attempt to optimize and maximize hydrogen bonds and stacking

	B - DNA	A-DNA	Z-DNA
	Right handed	Right handed	Left handed
bp per turn	10.4	11	12
Vertical rise per bp (Å)	3.4	2.3	3.8
Rotation per bp	+36 º	+33 º	-30 º
Helical diameter (Å)	19	23	18
Major groove	wide, shallow	narrow, deep	flat
Minor groove	narrow, deep	hallow	narrow, deep

Table 1. 1 Structural properties of B-DNA, A-DNA and Z-DNA conformations. Adapted from Freeman 2002.

interactions. (Rohs *et al.* 2009b). It is often observed in A-tracts whilst TpA steps and CG-rich sequences tend to widen the minor groove (Rohs *et al.* 2009b; Haran & Mohanty 2009).

Upon minor groove narrowing, an increase in the negative character of the electrostatic potential can be identified at distinct positions (Joshi *et al.* 2007; Rohs *et al.* 2009a). These evenly spaced positions can be detected by the side chain of basic amino acids such as arginine, contributing to the sequence-specific readout (Rohs *et al.* 2009b).

1.4.3 *Mechanisms of protein-DNA recognition*

Many interactions are involved in the molecular recognition of a target DNA sequence by a protein. Janin and colleagues explored the crystal structures of protein-DNA assemblies deposited in the PDB in order to identify which geometric and chemical interface properties govern the stability and specificity of the interaction (Janin *et al.* 2007). An average of 24 amino-acid residues and 12 nucleotides were found to be employed at the macromolecular interface (Janin *et al.* 2007). The following paragraphs describe the two principal modalities to achieve DNA binding specificity.

1.4.3.1 Base readout

This recognition mechanism consists of a DNA-binding protein that interacts with specific bases in either major or minor groove by reading the physical-chemical properties of the base or base pair (Rohs *et al.* 2010).

In order to create specific connections with the major groove, the protein exploits hydrogen bonds and hydrophobic interactions. In contrast to the minor groove, the four possible base pairs in the major groove possess distinct patterns of hydrogen bond acceptors and donors (Seeman *et al.* 1976; Harrison & Aggarwal 1990) which can be read by proteins using HTH domains, zinc-fingers, leucine-zippers and Ig-like motifs (Garvie & Wolberger 2001; Luscombe *et al.* 2000; Hong *et al.* 2008; Rohs *et al.* 2010). In addition to α-helical domains, also β-sheets can interact with the target DNA sequence by widening the major groove, forming hydrogen bonds with the bases, and exposing a generally more electrostatically positive surface (Tateno *et al.* 1997). Hydrogen bonds in the major groove allow distinction between purines whilst hydrophobic contacts with the protein side chains are adopted to discriminate between pyrimidines (Harrison & Aggarwal 1990; Xiong & Sundaralingam 2001).

Another way for proteins to achieve DNA recognition is through hydrophobic contacts with the minor groove. These often elicit a widening of the DNA conformation and the release of water molecules. The resulting increase in the system entropy is the driving force for DNA binding (Rohs *et al.* 2010; Crane-Robinson *et al.* 2006; Privalov *et al.* 2007; Privalov *et al.* 2009). Displacement of single water molecules can also occur when amino acids intrude in the minor groove (Li *et al.* 1998). Hydrogen bond formation is possible also with bases in the minor

groove although these interactions do not allow the differentiation between AT from TA and GC from CG (Rohs *et al.* 2010; Seeman *et al.* 1976).

1.4.3.2 Shape readout

Since proteins can detect local and global deviations from the B-DNA conformation, shape readout also contributes to DNA binding affinity and specificity (Rohs *et al.* 2009b; Travers 1989). As mentioned in section 1.4.2.2, proteins can contact the minor groove by recognizing the local narrowing, often associated with AT-rich tracts, by employing Arg residues. In addition to arginines, proteins also exploit positively charged side chains of lysine and histidine, although less frequently (Rohs *et al.* 2009b).

DNA kinks (section 1.4.2.2) may be recognized by the protein by insertion of side chains in the local deformation, resulting in a further deviation from the linearity of the double-helix (Rohs *et al.* 2010). Moreover, shape readout mechanism can also involve large sections of DNA chain, where A-DNA, Z-DNA and bent regions can be selectively bound by the target protein. A-DNA (favoured by CG-tracts) are characterised by having the sugar moieties more exposed and therefore able to interact with non polar amino acids (Ala, Leu, Phe, Val) (Kim Y. *et al.* 1993; Tolstorukov *et al.* 2004). But proteins can not only read the A-DNA sequence but also recognize the B → A conformational conversion at the A-DNA/B-DNA junction (Travers 1995). In the A-DNA shape the sugar-phosphates are selectively exposed, therefore their location, in addition to other structural parameters such as *roll*, *twist angles*, etc., contribute to DNA recognition (Lu *et al.* 2000). The distinctive zig-zag disposition of sugar-phosphate backbone in Z-DNA is also the signature code in Z-DNA binding proteins. For example Schwartz and colleagues, by determining the X-ray structure of the editing enzyme dsRNA adenosine deaminase (ADAR1), revealed that this enzyme uses a winged-helix fold (usually adopted to contact the B-DNA) to bind Z-DNA with high affinity (Schwartz *et al.* 2001).

1.4.4 *DNA recognition in RcnR and CsoR homologues*

Proteins sharing DNA binding domains such as helix-turn-helix, winged helix and ribbon-helix-helix typically interact with the major groove because base pair functional groups are here more exposed (Rohs *et al.* 2010; Garvie & Wolberger 2001). Thermodynamically, this type of interaction is usually enthalpically driven (Privalov *et al.* 2007), while when the contact involves the minor groove the driving force is ascribed to the entropy change (ΔS increases) resulting from the release of a significant number of water molecules from the site (Privalov *et al.* 2007). Minor groove interaction typically takes place when proteins lack the DNA binding domains previously described (Section 1.4.1) and involves a pronounced kinking of DNA (Kim J. L. *et al.* 1993). Interestingly, this is the scenario for RcnR/CsoR family members. Transcriptional factors belonging to this family can be divided into three groups according to the operator sequences that they recognize as proposed by Iwig and Chivers (Iwig & Chivers

2009; Tan *et al.* 2013). The first group includes proteins which recognize a single G/C tract (3-8 bp) flanked by inverted repeats rich in A/T (hereafter called type 1 site) (Iwig & Chivers 2009). Proteins belonging to the second group bind to a sequence where two shorter G/C tracts are separated by 2-4 bp and the A/T rich inverted repeats are located outside the G/C tracts (hereafter called type 2 site) (Tan *et al.* 2013). A third group of proteins bind to so defined "tandem" sites, which have two separate sites of type 1, type 2, or both (Iwig & Chivers 2009). *E. coli* RcnR belongs to the first group because two protein tetramers bind a tandem operator site consisting of two type 1 sites (each binding with a 1:1 stoichiometry), where CsoR homologues (from *S. lividans*, *M. tuberculosis* and *B. subtilis)*, bind at the type 2 semi-continuous G/C tract with a 2:1 stoichiometry (Iwig & Chivers 2009; Ma *et al.* 2009; Tan *et al.* 2013). *E. coli* RcnR was shown to recognize both major and minor groove regions by interacting with inverted repeats rich in T/A and with G/C tract respectively, promoting DNA wrapping (Iwig & Chivers 2009). Extensive studies employing isothermal titration calorimetry analysis (ITC) tested the role of the T/A sites in DNA recognition in both *E.coli* RcnR (Iwig & Chivers 2009) and *S. lividans* CsoR (Tan *et al.* 2013). Mutations in these regions had striking effects on the DNA binding affinity on both proteins suggesting an "end-to-end" interaction of both proteins across the site (Iwig & Chivers 2009, Tan *et al.* 2013). Major groove contact is carried out by interactions with the G/C tract in both proteins (Iwig & Chivers 2009; Tan *et al.* 2013). As discussed in Section 1.4.2, poly-(G) sequences were shown to impose A-form characteristics to DNA, resulting in a deepening of the major groove and a widening of the minor groove (Peticolas *et al.* 1988; Basham *et al.* 1995; Werner *et al.* 1995; Jauch *et al.* 2012). These dramatic changes were observed in *E. coli rcnRAB* promoter region by Circular Dichroism analysis (CD), where it was possible to distinguish between A-DNA (in the G/C tract) and B-DNA (in the flanking regions) features (Iwig & Chivers 2009). Combination of the two features results in a significantly different architecture, including the formation of kinks at the juncture between the two different forms of DNA which can aid DNA recognition and wrapping (Iwig & Chivers 2009). An analogous A-DNA feature was observed by CD in the G/C tract from the consensus sequence recognized by *S. lividans* (Tan *et al.* 2013). The A-DNA characteristic of this tract was enhanced upon addition of up to two CsoR tetramers providing a more suitable protein-binding site as tetrameric CsoR width approximately coincides with the length of the consensus sequence (Tan *et al.* 2013).

1.5 Formaldehyde: Toxicity, origins and detoxification mechanism

Formaldehyde (HCOH) is the simplest and the most reactive among the aldehydes. It is a highly toxic and colourless compound.

1.5.1 *Reactivity*

Formaldehyde is characterized by a carbonyl group C=O, which is planar, with σ-bond (where sp^2 orbitals overlap to each other) and a π-bond (overlap of 2*p* orbitals). The carbonyl group is polarized toward the most electronegative atom (oxygen) making the carbon electro-deficient. Formaldehyde, as aldehydes in general, can easily undergo oxidation to yield the corresponding carboxylic acid, formic acid (HCOOH) (McMurry 2011).

The most common reaction involving aldehydes (and ketones) is the nucleophilic addition, where a nucleophile species attacks the electron deficient carbon on the carbonyl group. The hybridization of C changes from sp^2 to sp^3 and a tetrahedral alkoxide ion intermediate is generated (McMurry 2011). After this step the reaction can continue following two main routes yielding an alcohol or a species with a C=Nu (where Nu is a nucleophilic species) double bond (McMurry 2011). Nucleophilic additions are more favourable on aldehyde rather than ketones for steric and electronic reasons. In fact, the carbonyl C is less crowded and more accessible in aldehydes since these organic substances possess only one substituent larger than a proton resulting in a low energy transition state (McMurry 2011). Moreover, secondary carbocations are less reactive than primary ones because they possess two alkyl groups, which are electron donating and therefore can stabilize the positive charge on C more efficiently (Figure 1. 3).

Hence, aldehydes carbonyl group is more polarized than ketones increasing the reactivity of these molecules in nucleophilic additions (McMurry 2011). This applies to a greater extent to formaldehyde which undergoes nucleophilic attack even more readily since it does not have any alkyl group to inductively stabilize the C^+ in the corresponding carbocation.

To further introduce formaldehyde reactivity in this section we will discuss in details the nucleophilic addition reaction between a generic aldehyde and an oxygen (or sulfur) nucleophile (Section 1.5.1.1), with the consequent formation of an acetal (or thioacetal), or a nitrogen nucleophile (Section 1.5.1.2), which yields a Schiff base.

1.5.1.1 Nucleophilic addition of oxygen and sulfur nucleophiles

Figure 1. 4 shows the reaction mechanism of the acetal, $R_2C(OR')_2$, synthesis occurring upon the initial attack of an oxygen (or sulfur) nucleophile to the carbonyl group of the aldehyde. An analogous sequence of steps (not shown) yields thioacetals, $R_2C(SR')_2$ from the initial attack of thiol (RSH) groups.

The reaction can be subdivided in two steps: addition of one molecule of alcohol to produce one molecule of hemiacetal and the subsequent conversion to acetal upon addition of a second molecule of alcohol. Although alcohols are weak nucleophiles, the rate of the reaction can speed up in the presence of basic or acidic conditions. Since alcohols are not strong nucleophiles, the reaction must be catalysed. The synthesis of hemiacetals can occur in both acid and basic

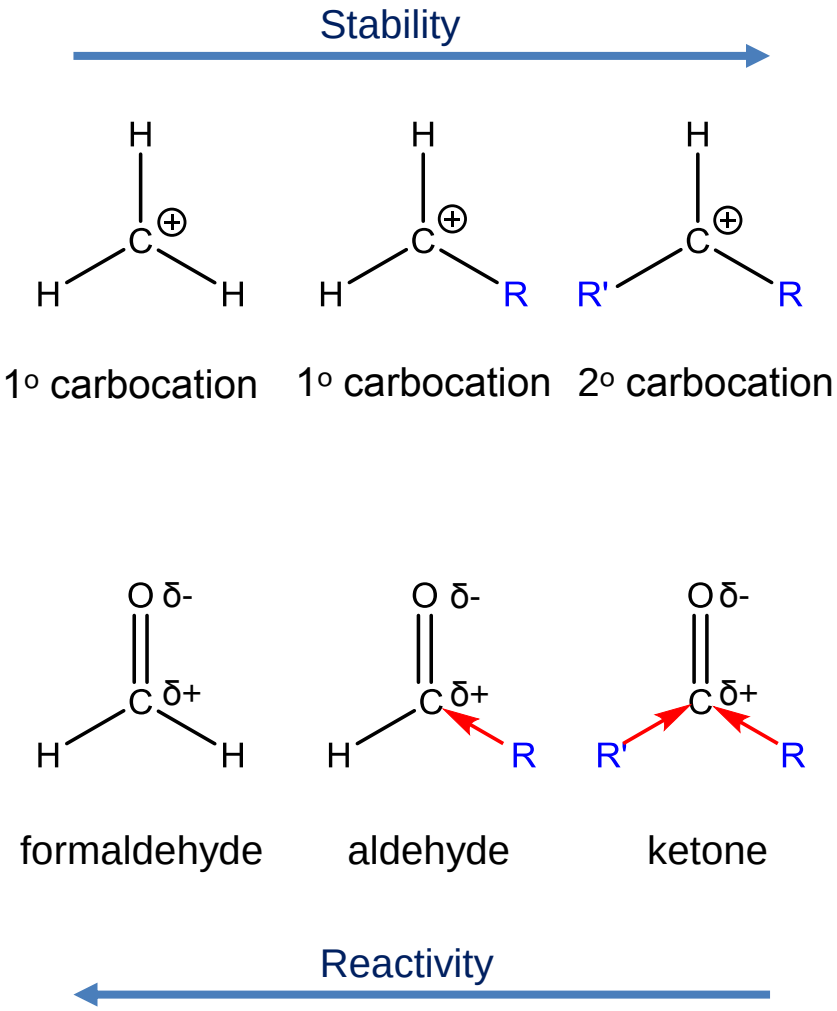


Figure 1. 3 The stability order of carbocations derived from formaldehyde, generic aldehyde and generic ketone compared to the reactivity order of the corresponding carbonyl compounds.

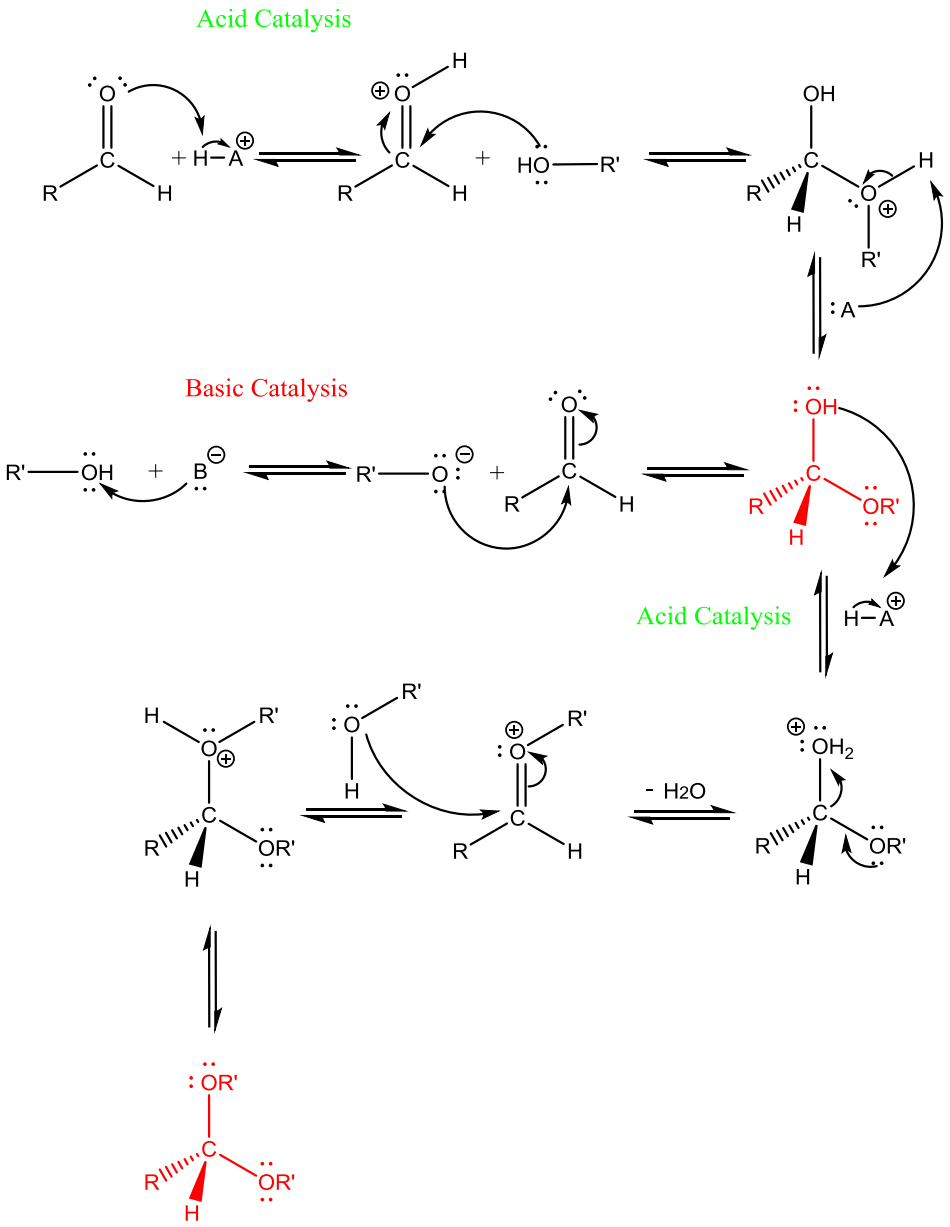


Figure 1. 4 Under acid conditions, the first step involves the protonation of the carbonyl oxygen with the subsequent activation of the carbonyl group for nucleophilic attack of the oxygen from the alcohol species. The formation of the tetrahedral intermediate hemiacetal is achieved by loss of the proton on the alcohol group and restoration of the acid catalyst. Otherwise, in the presence of a base the alcohol is deprotonated and converted to a better nucleophile which attacks the carbonyl group yielding the hemiacetal. The reaction proceeds in the presence of an acid catalyst which protonates the hydroxyl group, converting it in a good leaving group. Loss of a water molecule yields the oxonium ion which undergoes the addition of a second molecule of alcohol. Loss of a proton regenerates the acid catalyst and yields the acetal as final product. Hemiacetal and acetal products are highlighted in red.

conditions (Figure 1. 4), while the subsequent formation of acetal can only take place in acid conditions (McMurry 2011).

1.5.1.2 Nucleophilic addition of nitrogen nucleophiles

Nucleophilic addition of a primary amine to an aldehyde (or a ketone) yields an imine (RC=NR) (Figure 1. 65), whereas secondary amine produces an enamine (R_2N-CR=CR_2) (Figure 1. 6) (McMurry 2011). Both reactions occur in acidic conditions, with the maximum rate observed at pH = 4.5, low enough to protonate the tetrahedral intermediate (carbinolamine) but not too low to also protonate the amine, preventing in such a way the initial nucleophilic attack (McMurry 2011). The two mechanisms are quite similar, and are characterised by the synthesis of the tetrahedral intermediate carbinolamine and the final product with a newly formed C=Nu. However this reaction cannot occur if the aldehyde does not possess a hydrogen in position α (Figure 1. 6), which is the case of formaldehyde. Hence reaction with formaldehyde would stop after the formation of carbinolamine.

1.5.2 *Toxicity*

Formaldehyde is an irritant chemical which can lead to a variety of symptoms, ranging from mild inflammations to severe ulceration. In 2004 it has been classified as a mutagen, carcinogen to humans (Group 1 substances) by the International Agency for Research on Cancer. The World Health Organization (WHO) studied the association between formaldehyde exposure and various types of cancers such as nasopharyngeal, sinonasal and leukaemia. (WHO 2006; Kerns *et al.* 1983). Since formaldehyde carcinogenicity derives from the ability of this aldehyde to cross-link proteins and DNA, the number of these interactions have been exploited as a genotoxicity marker (WHO 2006).

1.5.2.1 Protein-DNA crosslinking with formaldehyde *in vitro*

Formaldehyde toxicity results from its ability to react with amino groups of cytosines, guanines and adenines of single and double stranded DNA and, with higher rates, with imino groups of thymines and guanines of single stranded DNA (Solomon & Varshavsky 1985; Brodolin 2010) (Figure 1. 7). In proteins, formaldehyde usually reacts with lysine, arginine, tryptophan and histidine residues. When long crosslinking times (t > 2 min) are allowed, multiple crosslinks between protein-protein and DNA-protein occur (with both melted and double-stranded regions of DNA). Since the groups involved in formaldehyde crosslinks must be in close contact (the interaction is at the range of Van der Waals radii (Kunkel *et al.* 1981), these crosslinks can be used as a tool to study protein domains in close contact with DNA (Brodolin 2010). One application is the use of relatively mild crosslinking conditions to study specific RNA polymerase-promoter interactions, (Brodolin 2010). The adopted conditions usually involve a concentration of formaldehyde of 2-30 mM and incubation times of 10-30 sec at 37 ºC, while

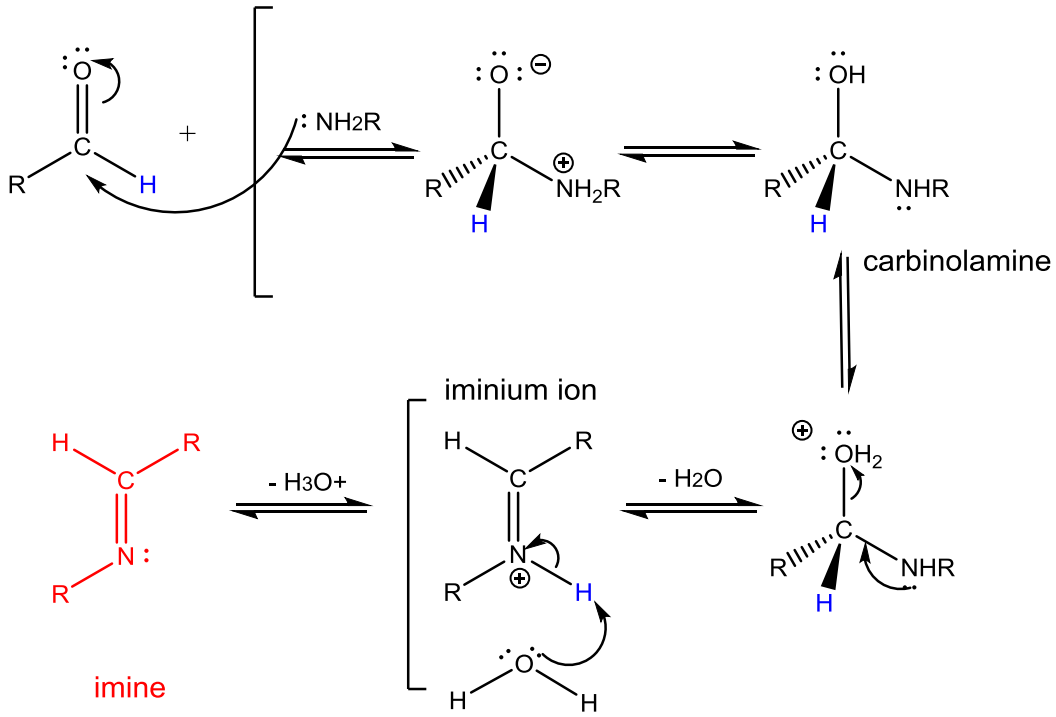


Figure 1. 5 The first step involves the nucleophilic attack of the primary amine to the carbonyl carbon (electrophile) with the production of a dipolar intermediate which is then converted to a neutral carbinolamine by proton transfer. Protonation of the hydroxyl group by the acid catalyst and subsequent loss of water produces the iminium ion. Regeneration of the acid catalyst yields an imine (in red).

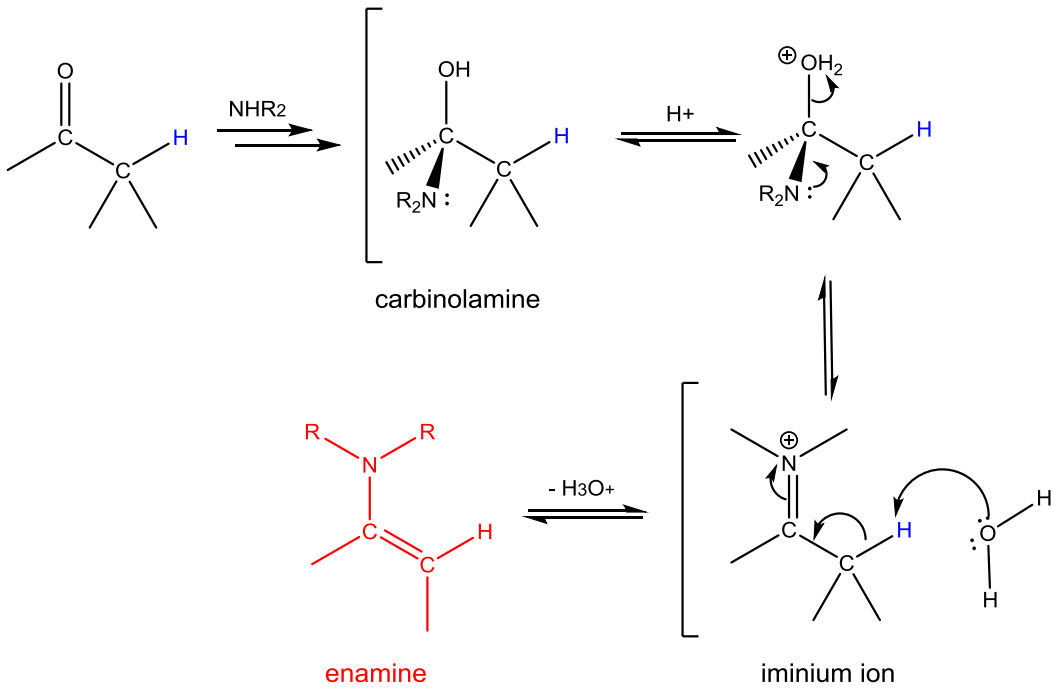


Figure 1. 6 The first part of the mechanism is identical to the imine synthesis. The lack of an additional proton on the secondary amine group now bound to the aldehyde causes a variation in the mechanism. Extraction of a proton bound to the Cα to regenerate the acid catalyst results in the formation of a double bond C=C.

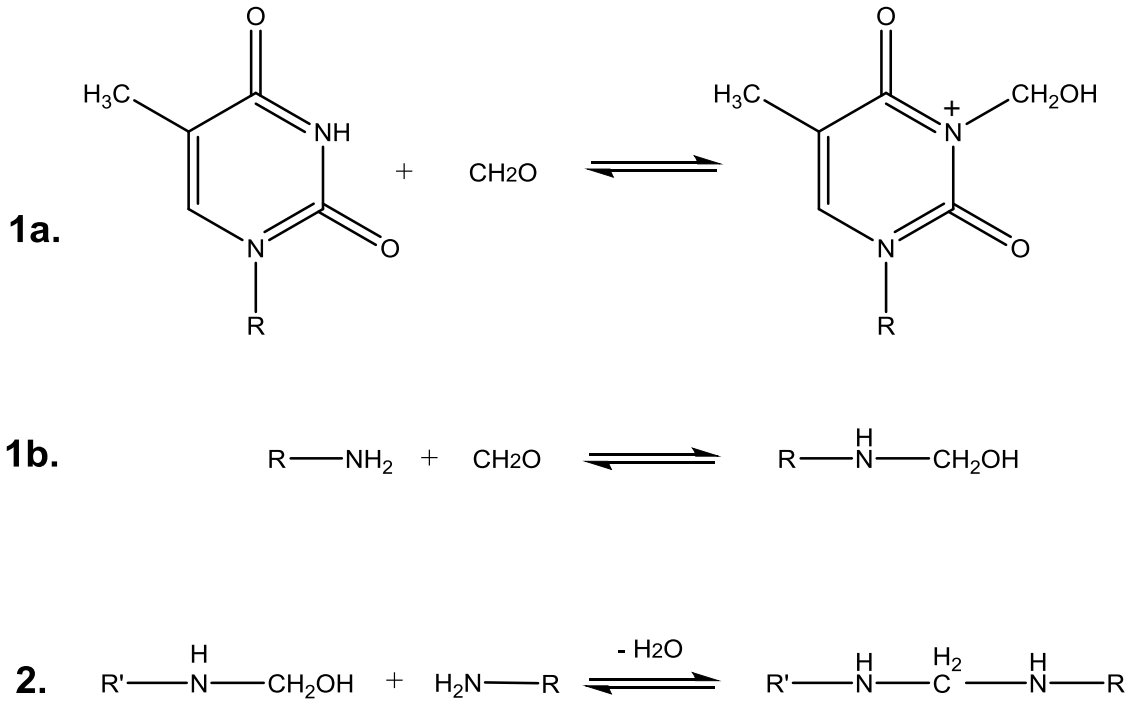


Figure 1. 7 Reactions mediated by formaldehyde. Formation of a methylol derivative on thymine (1a.) or on the amino group of an amino acid side-chain (1b.) is followed by the production of an immine which undergoes the addition of a second amino group yielding a stable condensation product (2.). This scheme has been adapted from "Protein-DNA crosslinking with formaldehyde *in vitro*" by Brodolin K. (Chapter 10 "DNA-Protein interaction" 2010).

higher concentrations could interfere with the formation of specific DNA-protein contacts. The crosslinking reaction is then terminated by addition of SDS which denatures the nucleoprotein complexes and destroys any non –covalent interaction between protein and DNA (Brodolin 2010) and then analysed by SDS-PAGE.

1.5.3 *Origins*

Formaldehyde can be ubiquitously found in all living organisms (Heck *et al.* 1985), in part because it is the by-product of many metabolic pathways. An example is given by the oxidative demethylation of DNA, which is associated with formaldehyde production following the reconversion from methylated to unmodified bases (Trewick *et al*. 2002). Moreover, when liver homogenate oxidizes biological labile methyl groups in sarcosine to yield carbon dioxide, production of formate and formaldehyde can be detected amongst several other intermediates (Handler *et al.* 1941; MacKenzie 1950). Furthermore formaldehyde plays an important role in atmospheric chemistry being a product of photochemical reactions in the troposphere. Aldehydes, and in particular HCOH, act as autocatalytical sources of hydroxyl radicals (·OH) (Monks 2005; Michoud *et al.* 2012). Formaldehyde can be emitted by plants (Lipari *et al*. 1984), from oxidation of Volatile Organic Compounds (VOCs) (Carlier *et al.* 1986) or can be a product of incomplete combustion (Ciccioli *et al.* 1993; Sassine *et al.* 2014).

1.5.4 *Detoxification*

In order to avoid formaldehyde accumulation, living organisms evolved several detoxification pathways. Some organisms have adapted to employ formaldehyde in their energy metabolism, maintaining the formaldehyde level high while tightly modulating it in order to avoid cytotoxicity (Christoserdova *et al.* 2009; Vorholt 2002). Methylotrophic bacteria use C_1 compounds, species without C-C bonds (CH_4, CH_3OH, HCOH etc.), as sources of carbon and energy (Chistoserdova *et al.* 2009). These bacteria may convert formaldehyde (for example to carbon dioxide) adopting several metabolic strategies: the ribulose monophosphate pathway (RuMP) which assimilates formaldehyde at relative low concentrations (Dijkhuizen *et al.* 1992; Ferenci *et al.* 1974; Anthony 1982; Vorholt 2002), metabolic pathways involving the addition of a cofactor to formaldehyde prior to oxidation, and the C_1-metabolism which sees the initial oxidation of C_1 species to yield formaldehyde that is then utilized to produce energy or assimilated into biomass (Vorholt 2002). An example is given by *Methylobacterium sp* MF1, isolated by Mitsui, Omori and colleagues (Mitsui *et al*. 2005), which can use HCOH (or CH_3OH) as the sole source of carbon by exploiting a metabolism that makes use of enzymes similar to those involved in C_1 metabolism. In these microorganisms, disruption of the formaldehyde assimilation pathway may result in the accumulation of formaldehyde up to 100 mM in less than 1 min (Attwood & Quayle 1984; Vorholt *et al.* 2000). The analogous

metabolism has been observed also in plants which, in addition, can assimilate formaldehyde into the Calvin cycle (Song *et al.* 2013; Khadem *et al.* 2012).

Microorganisms have evolved several formaldehyde detoxification pathways, amongst which the glutathione dependent mechanism, present in some prokaryotes and all eukaryotes, is the most commonly found (Harms *et al.* 1996). The second most abundant detoxification pathway is the ribulose monophosphate metabolism (RuMP) which, while adopted by methyltrophic bacteria to fix formaldehyde, is widespread in non-methyltrophic bacteria and archaea to avoid formaldehyde accumulation to toxic levels. In the next paragraphs these pathways will be discussed in details, with a particular focus on the GSH-dependent mechanism.

1.5.4.1 GSH-dependent formaldehyde detoxification pathways

The glutathione-dependent pathway represents the most widespread formaldehyde detoxification strategy adopted by organisms of wide phylogeny (Harms *et al.* 1996) with the exception of archaea (Gonzalez *et al.* 2006). Harms and colleagues investigated this mechanism in *Paracoccus denitrificans* where the oxidation of methanol, methylamine and choline leads to the transient production of formaldehyde (Harms *et al.* 1996). The microorganism faces a regulation problem since it must respond to toxic levels of formaldehyde without slowing its further oxidation.

Methanol and methylamine are converted to formaldehyde by the corresponding periplasmic enzymes, and formaldehyde is subsequently transported to the cytoplasm (Harms & Van Spanning 1991). Oxidation of choline to glycine occurs in several steps and is coupled with the release of formaldehyde in the cytoplasm (Harms *et al.* 1996). Here, formaldehyde spontaneously reacts with reduced glutathione yielding S-hydroxymethylglutathione which is then oxidized to S-formylglutathione by the NAD-dependent formaldehyde dehydrogenase (GD-FALDH) (Van Ophem & Duine 1994). The following step involves the hydrolysis of S-formylglutathione by S-formylglutathione hydrolase (FGH) to glutathione and formate, which can then be further oxidized to carbon dioxide by formate dehydrogenase (Min *et al.* 1988). The gene coding for FGH has been also recognised to be identical to the human esterase D (ESD) in human erythrocytes (Eiberg & Mohr 1986). Expression of *flhA* and *fghA*, which encode for GD-FALDH and FGH respectively, is low but constant and increases upon addition of methanol, methylamine and choline (Harms *et al.* 1996).

1.5.4.2 Other formaldehyde detoxification pathways

The ribulose monophosphate mechanism represents the second most widespread formaldehyde detoxification pathway adopted by living organisms. It has been first observed in methyltrophic bacteria, which utilize this pathway to fix formaldehyde, and in non-methyltrophic bacteria and archaea to avoid formaldehyde accumulation (Yasueda *et al.* 1999; Mitsui *et al.* 2003; Kato *et al.* 2006).

The mechanism involves the aldolic condensation of formaldehyde with ribulose 5-phosphate catalyzed by 3-hexulose-6-phosphate synthase (HPS) to yield D-arabino-3 hexulose-6-phosphate which is then isomerized to fructose 6-phosphate by the 6-phospho-3 hexuloisomerase enzyme (PHI). The hydrolysis reaction product may then be further phosphorylated to form fructose 1,6-biphosphate which can be metabolized to pyruvate by glycolysis (Quayle & Ferenci 1978; Kato *et al.* 2006). Three moles of formaldehyde can therefore yield one mole of pyruvate:

$$3\ HCOH + ADP + NAD^+ \rightarrow pyruvate + ATP + NADPH + H^+$$

It has been shown that expression of HPS and PHI is induced by formaldehyde in bacteria, whereas in archaea these enzymes are constitutively expressed. Archaea also exploit the reverse reaction in the pentose phosphate biosynthesis (Kato *et al.* 2006).

Two other less common detoxification strategies are the glutathione-independent and the tetrahydromethanopterin-dependent (H_4MPT) pathways. The gene coding for the GSH-independent pathway in bacteria share only 24 % identity with those required in the GSH-dependent mechanism. Here, formaldehyde is not condensed with a co-factor prior to be assimilated. The GSH-independent formaldehyde dehydrogenase, *fdhA*, is a member of the zinc-containing medium chain alcohol dehydrogenase and catalyzes the dismutation of two molecules of formaldehyde to formate and methanol (Oppenheimer *et al.* 1997; Ogushi *et al.* 1984):

$$2\ HCOH + H_2O \rightarrow HCOO^- + CH_3OH$$

Moreover, all methylotrophic bacteria possess the H_4MPT-dependent pathway, which has also been identified in methanogenic archaea and in some *Burkholderia* species such as *B. fungorum* LB400 (Marx *et al.* 2003; Vorholt 2002). The metabolism involves a first step where formaldehyde is condensed with the archaeal cofactor H_4MPT by the formaldehyde activating enzyme (FAE) (Vorholt *et al.* 2000), followed by several steps which convert the adduct to formate (Vorholt 2002).

In order to prevent formaldehyde from accumulating to toxic concentrations, organisms often use a combination of the formaldehyde detoxification pathways discussed above, (Marx *et al.* 2004) which in some cases do not appear redundant (Chistoserdova *et al.* 2000*).*

1.5.5 *Regulation of detoxification pathways*

Expression of genes coding for enzymes involved in formaldehyde detoxification pathways must be tightly regulated. In fact, regulators of formaldehyde detoxification pathways must avoid accumulation of this substance above a toxicity threshold but also prevent the wasteful oxidation of formaldehyde if its concentration is too low. These regulatory mechanisms have

been investigated in a few organisms employing the GSH-dependent formaldehyde detoxification pathway.

As already introduced in Section 1.5.4.1, *Paracoccus denitrificans* possesses *flhA* and *fghA* genes which encode GD-FALDH (formaldehyde dehydrogenase) and FGH (S-formylglutathione hydrolase) respectively, involved in the formaldehyde oxidation to formate. *flhRS* genes have been isolated and sequenced by Harms and colleagues (Harms *et al.* 2000) and were found to encode for a two-component regulatory system, FlhRS. This system can both activate the detoxification pathway and control the activity of the quinoprotein methanol dehydrogenase MDH (which produces formaldehyde) relative to the formaldehyde consumers, GD-FALDH and FGH (Harms *et al.* 2000). *Rhodobacter sphaeroides* possesses two-component systems. In addition to the regulatory machinery analogous to *P. denitrificans*'s (AfdS and AfdR proteins share ~ 50 % identity with the corresponding *P. denitrificans*'s FlhR and FlhS) (Barber & Donohue 1998), the RfdSR system was identified by Hickman and collaborators (Hickman *et al.* 2004). However, RfdS and RfdR show a weak homology with AfdS and AfdR and it has been shown that they repress transcription of the GSH-FDH gene independently from formaldehyde levels, suggesting that other effectors may be sensed by this system (Hickman *et al.* 2004). In addition to *Escherichia coli* FrmR, which is involved in the glutathione-dependent formaldehyde detoxification system *frmRAB* (as extensively described in Section 1.2.1.7), a regulatory system has been identified in *Bacillus subtilis* and *Bacillus cereus*. *Bacillus subtilis* possesses both glutathione-dependent and RuMP pathways. The glutathione-dependent system appears to lack the FGH enzyme, suggesting that other metabolic routes must be adopted by the microorganism to catabolize S-formyl-glutathione. The gene coding for GDH-FDH enzyme, *adhA*, is part of the *adhA-yraA* operon (*yraA* encodes a cysteine proteinase) which is regulated by the formaldehyde-sensing transcriptional regulator AdhR. This protein belongs to the MerR family (Section 1.2.2.1) which includes metal-sensing activators but also proteins responding to toxic levels of oxidative stress (e.g. formaldehyde) (Potter *et al.* 2010; Ahmed *et al*. 1994; Pomposiello & Demple 2001; Chen & He 2008; Brown 2003). It has been shown that Cys52 has a key role in formaldehyde sensing by AdhR (Huyen *et al*. 2009). This residue may be modified by formaldehyde, inducing a conformational change in AdhR which can lead to activation of the *adhA-yraA* operon transcription.

B. subtilis contains also *hxlA* and *hxlB* genes coding for the HPS and PHI enzymes required during the metabolic RuMP metabolic pathway. Formaldehyde sensor HxlR is encoded by a gene divergently transcribed relative to *hxlAB* operon (Yasueda *et al*. 1999). HxlR is the founding member of the HxlR-like family of transcriptional regulators and part of the GntR super-family of proteins which bind their promoter regions employing the wHTH motif in the N-terminus region and the effector with the C-terminus region (Hoskisson & Rigali 2009). In the absence of formaldehyde, HxlR binds two promoter regions upstream *hxlAB* operon. In the presence of toxic levels of formaldehyde the expression of the *hxlAB* operon is induced. In order

to explain the activation of the transcription induced by formaldehyde two hypotheses have been proposed by (Yurimoto *et al.* 2005). The first involves a conformational change in the transcriptional regulator, elicited by formaldehyde binding, which stimulates the recruitment of RNA polymerase. The second one considers formaldehyde binding to other factors which can then trigger the activation of HxlR. An analogous effect has been observed also with methylglyoxal raising questions about HxlR selectivity towards formaldehyde-binding. A member of the HxlR-family has also been identified in *Bacillus cereus* where regulates the operon involved in the glutathione-dependent detoxification pathway (Law 2012).

1.6 Aims and objectives

The aims of this project were two-fold:

1. To pursue a deeper understanding of the bacterial systems for sensing and adapting to metal stress in *Salmonella enterica serovar typhimurium*, the formaldehyde-sensing FrmR was examined through a combination of biochemical, biophysical and structural studies. Crucial determinants of metal sensing in cells were tested by creating a FrmR variant (E64HFrmR) capable of sensing cellular Zn(II) and Co(II). The gain-of-function Glu64 → His mutation was originally intended to enhance FrmR Ni(II)- and Co(II)-affinity *in vitro* by creating a RcnR-like metal binding site, although unexpectedly switched the formaldehyde-sensor into a metal-sensor.

Zn(II)-, Co(II)-, Cu(I)-binding affinities, DNA-affinity of apo- and holo-protein, and allosteric coupling free energy values were measured for FrmR and a variety of mutants. These experiments aimed to characterize FrmR and enhance understanding of the microbial metal-regulatory system that somehow discriminates between metals in *Salmonella*.

2. The second part of this work focused on *Salmonella* FrmR cellular role as the formaldehyde-sensing transcriptional regulator of the *frmRA* operon, involved in the cellular formaldehyde-detoxification pathway. By determining the residues involved in formaldehyde detection through site-directed mutagenesis, a potential formaldehyde-sensing mechanism was formulated. In order to investigate the specificity of the formaldehyde-FrmR interaction, DNA-binding in the presence or absence of formaldehyde stress was compared to that of *Salmonella* RcnR. Creation of an *in vitro* formaldehyde-responsive variant of RcnR further corroborated the mechanism here proposed.

Chapter 2

Materials and methods

2.1 Reagents and chemicals

All reagents and chemicals were purchased from standard commercial suppliers (Sigma-Aldrich, AnalaR, Thermo Scientific and Acros Organics) unless otherwise specified. Enzymes were obtained from New England Biolabs and Promega. All solutions were prepared using bacteria- and particle-free double deionised (nanopure) water (nH_2O).

2.2 Molecular biology methods

2.2.1 *Growth conditions*

For all gene cloning steps, the *E. coli* strain DH5α (Genotype: F-, *φ80dlacZΔM15*, Δ(*lacZYA-argF*)*U169*, *deoR*, *recA1*, *endA1*, *hsdR17*(*rk-*, *mk+*), *phoA*, *supE44*, *λ-*, *thi-1*, *gyrA96*, r*elA1*) was used. For overexpression of recombinant proteins, the *E. coli* strain BL21(DE3) (Genotype: F–, *omp*T, *hsd*SB (*rB–*, *mB–*), *dcm*, *gal*, λ(DE3), *pLysS*) was used. All liquid *E. coli* cultures were grown in Luria-Bertani (LB) medium (Sambrook & Russell 2001) at 37 °C with orbital shaking at approximately 180 rpm, unless otherwise stated. Bacterial *E. coli* cultures grown on LB agar plates were incubated overnight at 37 °C to allow colony formation.

Dr. J.S. Cavet (University of Manchester) kindly provided the *S. enterica* sv. *Typhimurium* SL1344 and LB5010a strains that were used as wild-type and as restriction-deficient modification-proficient hosts for DNA manipulations, respectively.

All media for bacterial culturing was sterilised by autoclave prior to use.

2.2.2 *Antibiotics*

Antibiotics (supplied by Sigma and Melford laboratories Ltd.) were used as selectable markers when growing *E. coli* cells transformed to antibiotic resistance with recombinant plasmids. Antibiotics routinely used were kanamycin (50 μg ml^{-1}) for pET29a and carbenicillin (100 μg ml^{-1}) for pGEMT plasmids, unless otherwise stated.

2.2.3 *List of oligonucleotides*

All primers, listed in Table 2. 1, were obtained from Sigma-Aldrich. All subcloned DNA fragments were sequenced (DBS Genomics, Durham University, UK) to verify no mutations/errors in sequences.

Primer name	Sequence (5′-3′)	Description
frmR_F	GAACATATGCCGCATTCACCTGAAGATAAAAAAC	PCR primer (production of pGEM-TfrmR and pETfrmR)
frmR_R	GAAGAATTCTTATTTTAGATAAGCGCGAAGAAGATGGCC	PCR primer (production of pGEM-TfrmR and pETfrmR)
rcnR_F	GAACATATGTCACATACCATCCGGGAC	PCR primer (production of pGEM-TrcnR and pETrcnR)
rcnR_R	GAAGAATTCCTATTTTATATAAGAATCCAGCACCTTTAATATTACG	PCR primer (production of pGEM-TrcnR and pETrcnR)
frmR_E64H_SDM_F	GAAATCCATCTGAAAGATCACCTGGTCAGCGGGGAG	FrmR Glu64→His ‘quickchange’
frmR_E64H_SDM_R	CTCCCCGCTGACCAGGTGATCTTTCAGATGGATTTC	FrmR Glu64→His ‘quickchange’
frmR_C35A_SDM_F	GAGTCTGGCGAACCTGCTCTGGCGATTCTGCAAC	FrmR Cys35→Ala ‘quickchange’
frmR_C35A_SDM_R	GTTGCAGAATCGCCAGAGCAGGTTCGCCAGACTC	FrmR Glu64→His ‘quickchange’
frmR_H60L_SDM_F	GTGAAATGGTTGAAATCCTTCTGAAAGATGAGCTGGTCAG	FrmR His60→Leu ‘quickchange’
frmR_H60L_SDM_R	CTGACCAGCTCATCTTTCAGAAGGATTTCAACCATTTCAC	FrmR His60→Leu ‘quickchange’
frmR_P2S_SDM_F	GGAGATATACATATGTCGCATTCACCTGAAGATAAAAAACGTATCC	FrmR Pro2→Ser ‘quickchange’

frmR_P2S_ SDM_R	GGATACGTTTTTTATCTTCAGGTGA ATGCGACATATGTATATCTCC	FrmR Pro2→Ser 'quickchange'
frmR_E81A_ SDM_F	GGTTCGGATGGCGGCAATCGGCCAT CTTC	FrmR Glu81→Ala 'quickchange'
frmR_E81A_ SDM_R	GAAGATGGCCGATTGCCGCCATCCG AACC	FrmR Glu81→Ala 'quickchange'
rcnR_S2P_ SDM_F	GAAGGAGATATACATATGCCACATA CCATCCGGGACAAAC	RcnR Ser2→Pro 'quickchange'
rcnR_S2P_ SDM_R	GTTTGTCCCGGATGGTATGTGGCAT ATGTATATCTCCTTC	RcnR Ser2→Pro 'quickchange'
pGEMCon3_F	ATCACTAGTGCGGCCGCC	Production of pGEMCon3
pGEMCon3_R	GATAACAATTTCACACAGGAAA CAGC	Production of pGEMCon3
rcnRAProEM_F	GAAGAAGTGTCGAAAATTCACCC	Production of rcnRAProEM
rcnRAProEM_R	GAACCGGATGGTATGTGACATTACA AC	Production of rcnRAProEM
frmRAPro_F	[HEX]TTCTGATAGTATACCCCCCTAT AGTATATGGAG	Production of *frmRA*Pro
frmRAPro_R	CTCCATATACTATAGGGGGGTATAC TATCAGAA	Production of *frmRA*Pro
frmRPro-T/A-mod_F	[HEX]TTCTGGTTCAACACCCCCCTAT AGTATATGGAG	Production of *frmRA*Pro-T/A-mod
frmRPro-T/A-mod_R	CTCCATATACTATAGGGGGGTGTTG AACCAGAA	Production of *frmRA*Pro- T/A-mod
*rcnRA*Pro-site1_F	[HEX]CATTCTACTCCCCCCCAGTAT AGAATAC	Production of *rcnRA*Pro^{1site}

*rcnRA*Pro-site1_R	GTATTCTATACTGGGGGGGAGTAGAATG	Production of *rcnRA*Pro^{1site}
*rcnRA*Pro-sites1,2_F	[HEX]TACTCCCCCCCAGTATAGAATACTACCCCCCAGTA	Production of *rcnRA*Pro^{1,2sites}
*rcnRA*Pro-site1,2_R	TACTGGGGGGTAGTATTCTATACTGGGGGGGAGTA	Production of *rcnRA*Pro^{1,2sites}

Table 2. 1 List of primers used during this work. [HEX] is the fluorescent tag hexachlorofluorescein.

2.3 DNA manipulation

2.3.1 *Amplification of DNA by Polymerase Chain Reaction (PCR)*

Polymerase chain reactions (PCR) were prepared to a final volume of 50 µl containing dNTPs (a mixture of dATP, dCTP, dGTP, dTTP, each at a final concentration of 40 µM), 10 - 100 ng of plasmid or genomic DNA, 0.4 µM of forward and reverse PCR primers (primers frmR_F,R; rcnR_F,R; Table 2. 1), Promega PCR buffer (containing 20 mM $MgSO_4$) and 2-3 units of *Pfu* polymerase enzyme. PCR reactions were performed in a thermocycler using 95 ^oC for the melting step, a variety of temperatures for annealing and 72 ^oC for extension.

2.3.2 *3'-A tailing of PCR products*

The pGEM-T plasmid (Promega) was used for the propagation of cloned DNA fragments. An 'A-tailing' strategy was chosen in order to facilitate ligation into this vector. Immediately following the PCR strategy described in Section 2.3.1, the PCR reaction was heated to 95 ^oC for 20 min in order to degrade *Pfu* polymerase, then 15 µl of 2 mM dATP was added to the reaction mix along with 5 units of *Taq* polymerase. The reaction was then incubated at 70 ^oC for 15 min. In this way a single deoxyadenosine (A) was added to the otherwise blunt 3' ends of the PCR products to complement the single 3'deoxythymidine present in the pGEM-T vector (Section 2.3.4).

2.3.3 *Agarose gel electrophoresis*

DNA fragments were analysed by agarose gel electrophoresis. Agarose concentrations were selected depending on the size of the DNA fragment to be analysed knowing that the greater the percentage of agarose, the smaller the linear DNA that can be resolved. The gels were cast in a Tris/Borate/EDTA (TBE) buffer system which was also used for running the gel (Sambrook & Russell 2001). Ethidium bromide (0.5 µg ml^{-1}) was also added to the gel to allow DNA visualization under UV light. DNA samples containing 6x loading dye (NEB) and a 1 kb DNA marker (NEB) were loaded onto the gel. Agarose gels were typically run at 90 V for approximately 1 h.

DNA was then visualised by exposure to UV light using a UV transilluminator. DNA fragments that were required for further manipulation were excised from the gel and DNA was purified from the gel matrix using a GenElute gel extraction kit (Sigma-Aldrich) or a Qiaquick gel extraction kit (Qiagen) according to manufacturer instructions.

2.3.4 *Ligation into pGEM-T plasmid*

For ligation into pGEM-T, components of the pGEM-T vector system (Promega) were used. A 10 µl ligation reaction was prepared, containing 1 µl pGEM-T plasmid vector stock (50 ng μl^{-1}),

3 µl gel purified PCR product (~ 150 ng μl^{-1}), 5 µl of T4 DNA ligase 2x reaction buffer and 1 µl of T4 DNA ligase enzyme (3 units μl^{-1} stock). Positive and negative controls were also prepared replacing the PCR products with 2 µl of a control DNA insert or with pure water, respectively. Reactions were either incubated at room temperature for 1 h or incubated at 4 ^{o}C overnight. Following incubation, 2 µl from each reaction was used to transform *E. coli* strain DH5α to carbenicillin resistance as described in Section 2.3.10.

2.3.5 *Blue/white identification of transformant cells and colony PCR*

In order to quickly and easily distinguish between recombinant and non-recombinant colonies, the blue/white colony screening was used. The multi-cloning site of pGEM-T is contained within the α-peptide coding region of the enzyme *β-galactosidase*. Successful ligation of PCR inserts interrupts this coding sequence, therefore these clones are unable to convert the colourless substrate X-gal (bromo-chloro-indolyl-galactopyranoside) into the insoluble, blue compound 5-bromo-4-choroindole. A white colony indicates that the α-peptide in the plasmid is disrupted (insert present) whereas a blue colony indicates that the α-peptide is intact (no insert). LB-Agar plates containing carbenicillin were streaked with 40 µl 2 % w/v X-gal and 7 µl 20 % w/v isopropyl β-D-1-thiogalactopyranoside (IPTG) which was allowed 3-4 h to soak into the plate, after which aliquots of DH5α cells transformed with pGEM-T derivatives were then able to be plated.

Plates were incubated overnight at 37 ^{o}C to allow colony growth. White colonies were further analysed by a colony PCR screen to confirm the presence of the desired insert. Colony PCR reaction mixtures were prepared as described in Section 2.3.1 however using 5-10 units of *Taq* polymerase and NEB Thermopol buffer (diluted from a 10x stock containing 20 mM $MgSO_4$). PCR primers were those used to amplify the DNA fragment from genomic DNA (Table 2.1). Colonies approximately 1 mm in diameter were picked with a sterile pipette tip, streaked onto a replica plate and then transferred to the PCR tube as DNA templates. PCR reaction was performed as described in Section 2.3.1 using 1 kb min^{-1} at 68 ^{o}C for extension. PCR-amplified DNA fragments were analysed by agarose electrophoresis (Section 2.3.3) allowing identification of the desired insert. Plasmid DNA was prepared from colonies harbouring the desired PCR product (Section 2.3.7) and sequenced (DBS Genomics, Durham University) to confirm the integrity of the cloned fragment.

2.3.6 *Ligation into pET29a plasmid*

A protocol similar to that employed for ligation into pGEM-T plasmid was used to ligate DNA fragments into the overexpression plasmid pET29a.

The 10 µl ligation reaction contained 1 µl of restriction digested pET29a plasmid (~ 1 µg μl^{-1}), up to 7 µl of the restriction digested DNA insert from pGEM-T, 1 µl of T4 DNA ligase 10x

reaction buffer (Invitrogen) and 1 µl of T4 DNA ligase enzyme (3 units μl^{-1} stock). Reactions were either incubated at room temperature for 1 h or incubated at 4 °C overnight and 100 µl cells of the *E. coli* strain DH5α were transformed using up to 5 µl from each reaction and screened to confirm successful ligation. Cells were transformed (Section 2.3.10) to kanamycin resistance and ~ 100 µl of cells were plated onto LB agar plates containing 50 µg ml^{-1} kanamycin. Following overnight incubation at 37 °C, colonies appeared on the plates. Colonies containing successfully ligated plasmids were identified by diagnostic restriction digestion (Section 2.3.8). Several colonies were picked from the plate with a sterile pipette tip and used to inoculate approximately 5 ml of LB media containing 50 µg ml^{-1} kanamycin. These cultures were then grown overnight at 37 °C with orbital shaking of 150 rpm. Plasmid DNA was isolated from each culture (Section 2.3.7) and each DNA sample was subjected to restriction digestion. To verify the presence of the correct insert in each sample, agarose gel electrophoresis was performed.

2.3.7 *Isolation of plasmid DNA*

For Miniprep DNA isolation, 5 ml *E. coli* cell overnight cultures were centrifuged at 6000 x g for 10 min at 4 °C using a bench-top microcentrifuge (Beckman Coulter Microfuge 18), where for Midiprep DNA isolation a 100 ml culture was grown overnight and pelleted by centrifugation in a Beckman Coulter Allegra X-22R benchtop centrifuge (3082 x g, 10 min, 4 °C). Plasmid DNA was extracted from cells using either a Qiaprep Spin Miniprep Kit (Qiagen) or a GenElute Plasmid Miniprep kit (Sigma-Aldrich) for minipreps and a Qiagen Plasmid Midi kit for midipreps according to manufacturer protocols. Recovery of plasmid DNA was confirmed by agarose gel electrophoresis and using a NanoDrop 1000 spectrophotometer (Thermo Scientific) which also allowed DNA concentration to be determined.

2.3.8 *Restriction digestion*

In order to allow efficient incorporation of fragments into pET29a plasmid, restriction enzymes were used to generate overhanging ends, with protruding terminal nucleotides, and complementary to those generated by analogous digestion of the plasmid. Each digestion reaction contained 5 µl of isolated plasmid, 1 µl of each of the pair of restriction enzymes (NEB), 5 µl of 10 x restriction enzyme buffer and nH_2O up to a final volume of 50 µl. Following 1 h incubation at 37 °C, samples were analysed to verify the presence of insert by agarose gel electrophoresis loading a small aliquot or all of the reaction volume if insert or digested plasmid was to be recovered. In the latter case the desired DNA inserts were isolated from the gel matrix as described in Section 2.3.3.

2.3.9 *Site directed mutagenesis using the 'Quick Change' method*

'Quickchange' site directed mutagenesis protocol, developed by Stratagene, was used to introduce substitutions in the amino acid sequence of proteins. This involves a linear rather than an exponential amplification using a pET29a plasmid as templates for the DNA synthesis step. For this reason the low copy number plasmid was first purified according to the Midiprep protocol (Section 2.3.7) in order to obtain it in higher concentration. Each reaction was prepared in a similar way to that described in Section 2.3.1 except for the use of 5 µl of midiprep DNA instead of genomic DNA. The Stratagene protocol suggests designing a primer with a T_m of at least 78 ^oC, a minimum GC content of 40 % and a length between 25 and 45 bases. The manual also specifies that the primers need to be PAGE or FPLC purified although desalted primers were successfully used in this work. Primers used are listed in Table 2.1. The initial denaturation was achieved incubating the samples at 95 ^oC for 30 sec, and was followed by sixteen cycles at 95^oC for 30 sec (melting), 55 ^oC for 1 min (annealing) and 72 ^oC for 14 sec (extension). A control reaction was set up where *Pfu* polymerase was substituted with nanopure water. The amplified DNA fragments were then incubated with 1 µl *Dpn1* restriction enzyme (20 unit μl^{-1}) at 37 ^oC for 3 h in order to eliminate the template DNA. This restriction enzyme is specific for methylated DNA and the mutated plasmid, which being generated *in vitro* is unmethylated, will be left intact. An aliquot (3µl) of each reaction was used to transform 50 µl of DH5α cells to kanamycin resistance (see Section 2.3.10). Typically no colonies were found on the minus *Pfu* control plate. Around 10 colonies were transferred to a replica plate and used to produce plasmid minipreps, then sequenced in order to confirm correct changes had been introduced.

2.3.10 *Production and transformation of E. coli chemically competent cells*

Competent cells were produced using a variation of the $CaCl_2/MgSO_4$ method. LB medium (500 µl) was added to a previously prepared aliquot of cells kept at -80 ^oC. Following incubation at 37 ^oC with orbital shaking for 1 h, the sample was used to streak a LB plate (no antibiotic added). The following day a colony from this plate was used to prepare a 5 ml culture in LB medium which was grown in the absence of antibiotic at 37 ^oC overnight. An aliquot (2 ml) from this culture was used to inoculate 100 ml of LB medium which was then incubated at 37 ^oC, while shaking at 180 rpm, until an optical density ($OD_{600\ nm}$) of ~ 0.5 (corresponding to mid-log phase) was reached. The cells were spun down for 10 min, at 4 ^oC and 18,595 x g in a Beckman Coulter Avanti J-20 XP centrifuge using a JLA 10.500 rotor. The supernatant was discarded and the cell pellet was gently resuspended in ice-cold 100 mM $MgCl_2$ (to a volume of 25 % the original culture volume) and incubated on ice for 5 min. Cells were then recentrifuged in sterile 50 ml centrifuges tubes in a Beckman Coulter Allegra X-22R benchtop centrifuge (1204 x g, 10 min, 4 ^oC). The cell pellet was resuspended in 5 % the initial culture volume of ice-cold 100 mM $CaCl_2$ sterile solution before addition of a further 45 % and incubation on ice

for 1 h. This cell suspension was recentrifuged in sterile 50 ml falcon tubes in a Beckman Coulter Allegra X-22R benchtop centrifuge (1204 x g, 10 min, 4 °C) and resuspended in 2 % the initial culture volume of ice cold, sterile solution of 85 mM $CaCl_2$, 15 % (w/v) glycerol. The resuspended cells were divided into 50 or 100 µl aliquots (on ice), flash-frozen in liquid nitrogen and stored at -80 °C.

To transform *E. coli* cells to antibiotic resistance, 50 µl was inoculated with 1-5 µl of concentrated, purified DNA construct or ligation product respectively. Samples were incubated on ice for 15 min, heat shocked at 42 °C for 90 sec and incubated again on ice for 4 min. LB medium (1 ml) was added to the mixture that was then incubated with shaking at 37 °C. After at least 1.5 h, cells were pelleted by centrifugation in a Beckman Coulter Microfuge 18 (18,000 x g, 10 min), resuspended in 50 - 70µl LB medium and plated onto a LB agar plate containing the appropriate antibiotic.

2.3.11 *Production and transformation of S. typhimurium electrocompetent cells*

Salmonella typhimurium electrocompetent cells were produced using 500 µl of LB medium added to a previously prepared aliquot of cells kept at -80 °C. Following incubation at 37 °C with orbital shaking for 1 h, the sample was used to streak a LB plate (no antibiotic added). The following day a colony from this plate was used to prepare a 10 ml culture in LB medium which was grown in the absence of antibiotic at 37 °C overnight. An aliquot (500 µl) from this culture were used to inoculate 50 ml of LB medium which was then incubated at 37 °C, while shaking at 200 rpm, until an optical density ($OD_{600\ nm}$) of ~ 0.4 - 0.6 was reached. The cells were transferred into two 50 ml falcon tubes and chilled on ice for 15 min before being spun down for 10 min, at 4 °C and 2,975 x g in a Beckman Coulter Avanti J-20 XP centrifuge using a JLA 10.500 rotor. The supernatant was discarded and the cell pellet was gently resuspended in 10 ml of ice-cold sterile pure water. Cells were then recentrifuged at 2,975 x g for 10 min at 4 °C. The supernatant was poured off, the cell pellet resuspended again in 10 ml sterile water and centrifuged. Supernatant was discarded and cells were resuspended in 10 ml of ice-cold 10 % w/v sterile glycerol. Cells were pelleted at 2,975 x g for 10 min at 4 °C and supernatant discarded. 1.25 ml of 10 % w/v glycerol was added. The resuspended cells were then divided into 50 µl aliquots (on ice), flash-frozen in liquid nitrogen and stored at -80 °C. To transform *S. typhimurium* cells to antibiotic resistance, 50 µl electrocompetent cells were thawed and kept on ice for 10 min. Plasmid DNA was also thawed and kept in an ice box for 10 min. DNA plasmid (4 µl) was aseptically added to competent cells and mixed gently by pipetting. Samples were incubated on ice for 15 min and aseptically transferred to a pre-cooled 2 mm gap width electroporation cuvette and kept on ice. Electroporation was carried out using a BioRad Gene Pulser set to resistance 200 ohms, capacitance 25 µF, voltage 2.5 µV. 500 µl LB medium was then immediately added to the cuvette in an aseptic environment. Samples were incubated for 1

h at 37 °C with orbital shaking prior to plate cells on to selective agar LB-carbenicillin (100 µg/ml) plate.

2.3.12 *Generation of Salmonella Deletion Mutants*

Deletion mutants of strain SL1344 were produced by Dr. Deenah Osman as described previously (Osman *et al.* 2015).

2.3.13 *Generation of Promoter-lacZ Fusion Constructs and β-Galactosidase Assays*

P_{frmRA} and P_{frmRA}-*frmR* were produced by Dr. Deenah Osman who performed *β-galactosidase* assays as described previously (Osman *et al.* 2015).

2.4 Protein production and purification

The composition of the buffers used in the following sections is summarized in the Appendix (Table 8. 1).

2.4.1 *SDS-PAGE analysis*

In order to determine purity and abundance of recombinant proteins, SDS-PAGE analysis (Sambrook & Russell 2001) was carried out during and post purification. For analysis of purified FrmR and RcnR wild type and mutants, gels were made at 18 % w/v acrylamide concentration. Gels were run at 200 V for approximately 1 h. Gels were stained in Instant Blue Coomassie stain (Thermo Scientific) overnight and then destained in pure water (nH_2O).

2.4.2 *Native gel analysis*

Non denaturing conditions were chosen to analyse DNA samples after annealing experiment. A native gel containing 12 % w/v acrylamide was prepared according to Sambrook & Russell, 2001. Gels were run at 90 V for approximately 2 h and then transferred to a tank filled with ~ 50 ml of water and 10 µl of ethidium bromide (stock 10 mg ml^{-1}). Following staining, visualization of the gel was carried out with UV light.

2.4.3 *Overexpression and solubility testing*

During this work pET29aFrmR, pET29aE64HFrmR, pET29aC35AFrmR, pET29aP2SFrmR, pET29aH60LFrmR, pET29aE81AFrmR, pET29aRcnR and pET29aS2PRcnR were produced. Small scale tests of protein expression were performed to find the right conditions to be applied to large-scale overexpression. For each expression test a single *E. coli* transformant colony was added to 5 ml LB medium in presence of 50 µg ml^{-1} kanamycin and incubated overnight at 37 °C with orbital shaking. The following day 250 µl from each culture were used to inoculate 25

ml LB medium and cells were grown until the $OD_{600\ nm}$ reached approximately 0.6. Different IPTG concentrations (200 - 500 µM) were then added to each culture. Aliquots of culture (5 ml) were removed at specific time points (after 1-2-3-4-5-16 h) and centrifuged using a Beckman Coulter Allegra X-22R benchtop centrifuge (1,764 x g, 10 min, 4 °C). The supernatants were discarded and cell pellets frozen at -20 °C. An aliquot of culture was removed also prior to addition of IPTG in order to assess pre-induction protein expression. Cell pellets from each condition were thawed on ice, resuspended in 1 ml of resuspension buffer (10 mM Hepes pH 7.0, 10 mM EDTA, 10 mM DTT, 100 mM NaCl) and lysed by sonication (8 x 20 sec bursts performed on ice). Cell lysates were clarified twice by centrifugation (18,000 x g, 4 °C, 10 min in a microfuge, and 39,191 x g, 4 °C, 20 min using a Beckman JA25.50 rotor). The soluble fraction (supernatant) was decanted while the insoluble fraction (pelleted cell) was resuspended in 1 ml of resuspension buffer. Aliquots were taken from each fraction and loaded onto an 18 % w/v SDS-PAGE run as described in Section 2.4.1. Both FrmR and RcnR proteins were mainly found in the soluble fractions.

2.4.4 *Overexpression conditions for recombinant FrmR and RcnR in E. coli*

1 µl of pET29a plasmid, harbouring the coding sequence for wild-type and mutant versions of FrmR and RcnR, was used to transform 50 µl of BL21(DE3) competent cell to kanamycin resistance (Section 2.3.10). The following day 10 ml of LB medium were inoculated with a single transformant colony and incubated overnight at 37 °C using an orbital shaker. The culture was then diluted 100-fold into 1L pre-warmed fresh LB supplemented with 50 µg ml^{-1} kanamycin and incubated at 37 °C, 180 rpm until OD_{600nm} ~ 0.5 (approximately 4 - 5 h), at which point protein expression was induced by addition of 200 µM IPTG for 3.5 - 4 hours under the same growth conditions. Cells were collected by centrifugation at 3,993 x g, 4 °C, 20 min using a JLA 8.100 rotor. Each pellet was suspended in 20 ml fresh LB, transferred to 50 ml Falcon tube before a second round of centrifugation in a Beckman X-22R benchtop centrifuge at 3,082 x g 4 °C, 15 min and storage at - 20 °C.

2.4.5 *Overexpression conditions for recombinant ZntR and CueR in E. coli*

ZntR and CueR were overexpressed by Dr. Deenah Osman, as described previously (Osman *et al.* 2015).

2.4.6 *Purification of recombinant wild-type and mutated FrmR proteins*

Pellets from 1L culture were thawed at room temperature (approximately 15 min) and resuspended in 50 mM sodium phosphate, pH 7.4, 300 mM NaCl, 5 mM DTT, 1 mM PMSF and 10 mM imidazole (buffer A). Cells were transferred to 1.5 ml microcentrifuge tubes in 750 µl aliquots, lysed by sonification (30 s at 35 % power on ice) and clarified twice by centrifugation (18000 x g, 4 °C, 10 min in a microfuge, and 39,191 x g, 4 °C, 20 min using a

Beckman JA25.50 rotor). The soluble cell lysate was loaded onto a 5 ml HisTrap FF column (GE Healthcare) prepared according to manufacturer's instruction and equilibrated in buffer A at 2.5 ml min^{-1}. Bound protein was washed with 25 ml buffer A and eluted in buffer A with 300 mM imidazole without PMSF (buffer B). The presence of protein in HisTrap eluate was first confirmed by SDS-PAGE analysis (Section 2.4.1) but given the high reproducibility of the protocol this step was then omitted. A quick test using 90 µl of Bradford reagent (Sigma-Aldrich) and 10 µl of sample was enough to confirm the presence of protein routinely in fraction 2. HisTrap eluate (~ 3 ml) was further purified by size exclusion chromatography (HiLoad 26/60 Superdex S75, GE Healthcare) running at 2.5 ml min^{-1} in buffer B300 (with 10 mM Hepes, pH 7.0, 300 mM NaCl, 10 mM DTT and 10 mM EDTA). Fractions containing FrmR (as identified by 18% w/v SDS-PAGE) were pooled, diluted in buffer B100 (NaCl 100 mM, Hepes pH 7.0 10 mM, 10 mM DTT, 10 mM EDTA) and bound to two 1 ml HiTrap Heparin columns connected to each other and pre-equilibrated with buffer B100. Bound protein was washed with 20 ml buffer B100 and eluted in a single step with buffer B500 containing 10 mM Hepes, pH 7.0, 500 mM NaCl, 10 mM DTT and 10 mM EDTA.

2.4.7 *Purification of recombinant wild-type and mutated RcnR proteins*

Pellets from 1L culture were thawed at room temperature (approximately 15 min) and resuspended in buffer B300 with inclusion of 1 mM PMSF. Cells were transferred to 1.5 ml eppendorfs in 750 µL aliquots, lysed by sonication (30 s at 35 % power on ice) and clarified twice by centrifugation (18000 x g, 4 ^{o}C, 10 min in a microfuge, and 39,191 x g, 4 ^{o}C, 20 min using a Beckman JA25.50 rotor). The soluble cell lysate was loaded onto an equilibrated 5 ml HiTrap Heparin column at ~ 2.5 ml min^{-1}. Bound protein was washed with 25 ml buffer B300 and eluted in buffer B800 (800 mM NaCl, 10 mM EDTA, 10 mM DTT, 10 mM Hepes pH 7.0) collecting four fractions (4.5 ml, 5.5 ml, 5 ml, 5 ml). Fractions containing RcnR (exclusively Fraction 2) were further purified by size-exclusion chromatography using a HiLoad 26/60 Superdex 75 running at 2.5 ml min^{-1} in buffer B300. Fractions enriched for RcnR (usually F28-F33, but fractions with significant contaminating proteins as observed in SDS PAGE were avoided) were pooled, diluted to buffer B100 and loaded onto an equilibrated 5 ml HiTrap SP column at ~2.5 ml min^{-1}. Bound protein was washed with 25 ml buffer B200 (200 mM NaCl, 10 mM EDTA, 10 mM DTT, 10 mM Hepes pH 7.0) and eluted in buffer B300 collecting six fractions (4.5 ml and 5 x 5 ml). Fractions 2 and 3 routinely contained RcnR, as confirmed by 18 w/v % SDS-PAGE.

2.4.8 *Purification of recombinant wild-type ZntR and CueR*

ZntR and CueR were purified by Dr. Deenah Osman, as described previously (Osman *et al.* 2015).

2.4.9 *Estimation of protein concentration*

For FrmR, RcnR, ZntR and CueR, concentrations were estimated by measurement of absorbance at 280 nm and by use of the experimental extinction coefficients determined by quantitative Amino Acid Analysis (AAA) (AltaBioscience Ltd). The analysis consists of an initial hydrolysis (24 hours at 110 °C) of the protein sample, followed by ion exchange separation of the amino acids and by post column photometric detection. The determination of the number of moles per ml of each amino acid allows the calculation of the experimental concentration of the given protein sample. Lambert-Beer equation (Equation 2, Section 3.3) was then applied in order to determine the experimental extinction coefficient. These were 1951 $M^{-1}cm^{-1}$ for FrmR proteins, 2422 $M^{-1}\,cm^{-1}$ for RcnR proteins, 11505 $M^{-1}\,cm^{-1}$ for ZntR and 5136 $M^{-1}\,cm^{-1}$ for CueR (Osman *et al.* 2015).

2.4.10 *Detection of Met1 cleavage in FrmR and E64HFrmR by liquid chromatography (LC)*

The following analyses have been performed by Dr. Huggins T. G. and Dr. Chen J. at the Procter and Gamble Mason Business Centre, Cincinnati (Ohio).

An aliquot (5 µg) of FrmR or E64HFrmR was digested in 200 µl 50 mM NH_4HCO_3 buffer by adding 1 µl trypsin (0.5 mg ml^{-1}). Digestion mixture was mixed with an orbital shaker incubator (1000 rpm, 37 °C, 16 h,) and stopped with 5 µl 15 v/v % formic acid. The digested sample was separated by gradient elution at 0.3 ml min^{-1} using a Zorbax Eclipse Plus C18 column (2.1 x 150 mm, 3.5 µm particles; Agilent Technologies) at RT. Mobile phase A and B consisted of 0.1% formic acid in water and 0.1% formic acid in acetonitrile, respectively. A volume of 10 µl was loaded onto a 6500 triple quadrupole mass spectrometer (AB Sciex) operating in positive ionization mode. Acquisition methods used the following parameters: 5500 V ion spray voltage, 25 psi curtain gas, 60 psi source gas, 550 °C interface heating temperature, 40 V declustering potential, 26 V collision energy, and 27 V collision cell exit potential. Scheduled multiple reaction monitoring (MRM) was carried out with a 90 s MRM detection window and 1.00 s target scan time.

2.5 Anaerobic manipulation of protein

2.5.1 *Production of Chelex-treated, anaerobic buffers*

In order to avoid any metal contamination of protein samples prepared in the absence of EDTA, all buffers were ion exchanged with Chelex-100 matrix (Sigma-Aldrich) prior use. The Chelex-100 resin is a styrene-divinylbenzene co-polymer containing iminodiacetic groups, able to sequester transition metal ions. The column method involves pouring a column with the Chelex resin and passing the sample through to achieve the separation. The following protocol was used

in order to prepare the matrix according to manufacturer instructions: 2 bed volumes in 1 M HCl, 5 bed volumes water rinse, 1 M NaOH, 5 bed volumes water rinse. The sample was slowly added to the column, taking care to not disturb the resin bed. The first 50-100 ml (void volume) passed though the column was discarded prior to collecting the metal-free effluent. Buffers to be used under anaerobic conditions were purged with N_2 for 3 - 4 h before transferring into an anaerobic chamber (Belle Technology).

2.5.2 *Production of anaerobic, metal-free protein samples*

All *in vitro* experiments, unless otherwise stated, were carried out under anaerobic conditions using Chelex-treated and N_2-purged buffers as described previously (Section 2.5.1).

In order to study metal- and formaldehyde-binding properties of proteins in the absence of added reducing agents, all the analysis were carried out in an anaerobic chamber. FrmR was buffer-exchanged on 1 ml Heparin column (GE Healthcare) equilibrated in buffer B100 outside and washed with at least 10 ml 20 mM NaCl, 80 mM KCl, 10 mM Hepes pH 7.0 (buffer C100) inside the anaerobic chamber. Proteins was then eluted with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 (buffer C500) collecting three fractions (~ 900 µl and 2 x 1 ml). RcnR was transferred inside the anaerobic chamber in a similar way except for using an overall salt concentration (1 NaCl : 4 KCl) of 300 mM (buffer C300) for the washing step and 1 M (buffer C1000) for the elution. Proteins were stored at 4 ^{o}C for up to 1 month.

2.5.3 *Determination of reduced sulphur content of proteins*

5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) is a symmetrical aryl disulphide which readily undergoes the thiol-disulphide interchange reaction in the presence of free thiols. Treatment of the protein with DTNB can yield the number of oxidized cysteines Cys-S-S-Cys. A solution of protein (1 ml) and DTNB (72 µg ml^{-1}) was prepared anaerobically in a gas tight 1 ml quartz cuvette (Hellma). First the DTNB blank A_{412} was determined in the absence of protein using a Perkin Elmer λ35 UV-visible spectrophotometer. The protein was then added to the sample in order to obtain a cysteine concentration of 30 µM. Absorbance at 412 nm was read after 30 min. The number of thiols modified was calculated using the maximum absorbance at 412 nm measured, Lambert-Beer's law and DTNB extinction coefficient (ε_{412nm} = 14,150 M^{-1} cm^{-1}, Riddles *et al.* 1979). Only protein preps with reduced cysteine content above 85 % were used.

2.5.4 *Measurement of metal content of purified proteins*

Inductively coupled plasma mass spectrometry (ICP-MS) was used to determine the residual metal content in purified protein stock after removal of chelator agent (Section 2.5.2). Protein samples were prepared anaerobically and diluted to 20 - 30 µM in the elution buffer (C500 for FrmR and C1000 for RcnR). An aliquot of this sample was then diluted to 10 % into 2 % w/v

nitric acid and analysed for metal content. Only protein preps with metal content below 5 % were used.

2.6 Experimental procedures

2.6.1 *Preparation of metal stocks*

All metal stocks (except for CuCl) were prepared by dissolving the metal salt into nanopure water. Serially diluted stocks of metal solutions were analysed by ICP-MS in order to assess the experimental metal concentration. Solid CuCl was weighed aerobically, transferred inside the anaerobic chamber and then dissolved in oxygen-free 100 mM HCl and 1 M NaCl. The concentration of copper was determined by ICP-MS and the concentration of Cu(I) by titration against an excess of bathocuproinedisulfonic acid (BCS, SigmA-Aldrich). Before being used, every day CuCl stock was confirmed to be > 95 % Cu(I) by titration against BCS (Barry *et al.* 2011; Dainty *et al.* 2010).

2.6.2 *Fractionation of protein-bound metal complexes by size-exclusion chromatography*

For experiments to resolve bound and free protein-metal complexes, a Sephadex G-25 matrix column (PD10 column, GE Healthcare) was used. The PD10 column was equilibrated according to manufacturer's instructions by washing with 2 column volumes of pure water, 500 µl of 0.5 M EDTA, followed by 2 column volumes of pure water. The PD10 column was then transferred inside the anaerobic chamber and washed with 2 column volumes of buffer C500. Where stated the final buffer was supplemented with metal. Proteins were incubated (t > 60 min) with a three-fold molar excess of $ZnCl_2$, $CoCl_2$, $NiCl_2$ or CuCl (verified to be > 95 % Cu(I)) in buffer C500 and an aliquot (500 µl) was applied to the PD10 column. Alternatively, apo-protein (no metal added) was resolved as described above except with addition of 20 µM $NiCl_2$ or $CoCl_2$ to the chromatography buffer. Fractions (500 µl) were analysed for metal by ICP-MS and protein by Bradford assay using known concentrations of the analysed protein as standards.

2.6.3 *UV-visible absorption spectroscopy*

Experiments were carried out anaerobically in 1 ml gas tight cuvettes (Hellma). Buffer C500 was used for FrmR, E64HFrmR, ZntR, and CueR, with addition of 5 % v/v glycerol for RcnR. Protein was titrated with $CoCl_2$, $NiCl_2$ or CuCl. In the case of Zn(II), protein was first pre-equilibrated with $CoCl_2$ and then titrated with $ZnCl_2$. Each sample was left to reach equilibrium prior to use of a Perkin Elmer λ35 UV-visible spectrophotometer to record absorbance spectra.

2.6.4 *Monitoring quenching of protein intrinsic fluorescence upon metal-binding*

Experiments were carried out anaerobically in 1 ml gas tight cuvettes (Hellma). Both FrmR and RcnR possess 1 residue, a tyrosine, of the three aromatic amino acids which may contribute to protein intrinsic fluorescence. Therefore it was possible to detect a weak fluorescence emission using a Cary Eclipse Fluorescence Spectrophotometer. The instrument excitation and emission wavelengths were set to λ_{ex} 280 nm and λ_{em} 300 - 400 nm. Spectra were recorded anaerobically diluting the protein in buffer C500. Changes in fluorescence were recorded at 20 ^{o}C at 5 min intervals to equilibrium.

2.6.5 *Measurement of the metal-binding affinities of recombinant FrmR proteins and RcnR*

Experiments were carried out anaerobically in 1 ml gas tight cuvettes (Hellma). The buffer system was 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 (C500). A Cary Eclipse Fluorescence Spectrophotometer was used for fluorescence spectroscopy and a Perkin Elmer λ35 for UV-vis spectroscopy. For competition with fura-2 the fluorescence spectrophotometer excitation and emission wavelengths were set to λ_{ex} 360 nm and λ_{em} 510 nm as described previously (Patterson *et al.* 2013). Fura-2 and quin-2 were quantified using extinction coefficients obtained from the supplier (ε_{363nm} 28,000 $M^{-1}cm^{-1}$ for fura-2 and ε_{261nm} 37,000 $M^{-1}cm^{-1}$ for quin-2).

Mag-fura-2 was quantified via its extinction coefficient ε_{369nm} 22,000 $M^{-1}cm^{-1}$ (Golynskiy *et al.* 2006).

2.6.5.1 Measurement of the Zn(II) binding affinities of FrmR and E64HFrmR by competition with quin-2 and mag-fura-2

The metallochromic indicator quin-2 forms a Zn(II)-quin-2 complex with an association constant of 2.70 x 10^{11} M^{-1} (Jefferson *et al.* 1990) while the affinity of mag-fura-2 for Zn(II) is 2.2 x 10^{-8} M (Reyes-Caballero *et al.* 2010). A mixed solution of chelator and protein was titrated with $ZnCl_2$ and absorbance recorded to equilibrium (after each metal addition reactions were left to equilibrate until no further change was observed in absorbance values). A control experiment was always included where the metallochromic indicator alone was titrated with metal salt aliquots. An increase in absorbance at 325 nm was monitored with mag-fura-2 (Reyes-Caballero *et al.* 2010), while a decrease in absorbance at 261 nm was detected when quin-2 was used (Jefferson *et al.* 1990). Data were fit to the models described in Chapters 3-4 and in the Appendix, using Dynafit, to determine Zn(II) binding constant.

2.6.5.2 Estimation of the Cu(I) affinity of FrmR and E64HFrmR by competition with BCS

CuCl was produced anaerobically as described in Section 2.6.1. Competition experiments were carried out in a single step addition and/or titration fashion. In the first case BCS (10 µM) was incubated with 0.5 molar equivalent of Cu(I). The absorbance at 483 nm was monitored to equilibrium, determining the [Cu(I)-BCS]. Addition of protein (10 µM) caused the bleaching of the spectral feature at 483 nm, allowing determination of the amount of Cu(I) migrated to the protein. For competition titration assays, Cu(I) was titrated into BCS (25 µM) in the absence or presence of FrmR (20 µM) and the absorbance at 483 nm was monitored.

2.6.5.3 Estimation of the Cu(I) affinity of FrmR and E64HFrmR by competition with BCA

CuCl (produced anaerobically as described in Section 2.6.1) was titrated into a solution of BCA (40 µM) and protein (10 µM) in buffer C500 (100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0) and the absorbance at 562 nm was recorded. Data were fit using Dynafit to a model describing a total of eight Cu(I) binding sites per FrmR or E64HFrmR tetramer (see Appendix). $\beta_{2Cu(I)}$ $1.58 \times 10^{17.2}$ M^{-2} at pH 7.0 was used for BCA using the Schwarzenbach's α-coefficient method (Xiao & Wedd 2010).

2.6.5.4 Estimation of the Co(II) affinity of FrmR and E64HFrmR by competition with BisTris and Fura-2

Competition between FrmR proteins and Fura-2 was performed by Dr. Deenah Osman. A mixed solution of protein and Bis-Tris or protein and Fura-2 was titrated with $CoCl_2$ in buffer C500. Competition with fura-2 was monitored at 20 ºC using a fluorescence spectrophotometer with excitation and emission wavelengths set to λ_{ex} 360 nm and λ_{em} 510 nm as described previously (Patterson *et al.* 2013). Competition with Bis-Tris was monitored by recording $Abs_{336\ nm}$ at equilibrium using a UV-visible spectrophotometer. Data were fit to the model described in the Appendix (four sites with equal affinity per tetramer) using Dynafit to determine $K_{Co(II)}^{protein}$. Fura-2 and Bis-Tris cobalt binding constants ($K_{Co(II)}^{Fura\text{-}2} = 8.64 \times 10^{-9}$ M at pH 7.0, Kwan & Putney 1990, and $K_{Co(II)}^{Bis\text{-}Tris} = 2.26 \times 10^{-2}$ M at pH 7.0, calculated using the Schwarzenbach's α-coefficient method, Xiao & Wedd 2010) were used.

2.6.5.5 Estimation of the Ni(II) affinity of FrmR and E64HFrmR by competition with mag-fura-2

A solution of the chromophore 10 µM mag-fura-2 ($K_{Ni(II)} = 2 \times 10^7$ M^{-1} at pH = 7.0, Reyes-Caballero *et al.* 2010) was titrated with Ni(II) in buffer C500 (100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0) and the absorbance at 366 nm was recorded using a UV-vis spectrometer, observing the formation of a 1:1 complex. For competition assay, Ni(II) was titrated into a mixed solution of mag-fura-2 (10 µM) and protein (10 µM) and Abs_{366} was monitored.

2.6.6 *Interprotein metal-exchange*

2.6.6.1 Metal-exchange between *Salmonella* ZntR, CueR, RcnR and FrmR

In order to assess relative Cu(I)- and Zn(II)-affinities, competition experiments with *Salmonella* cognate metal sensors were carried out by Dr. Deenah Osman, as described previously (Osman *et al.* 2015).

2.6.6.2 Metal-exchange between *Salmonella* ZntR and E64HFrmR

In order to assess relative Zn(II)-affinity, 56 μM of E64HFrmR was anaerobically pre-equilibrated with 14 μM $ZnCl_2$ in C500 buffer prior to addition of 28 μM of ZntR in a final volume of 1 ml. Reaction mixtures were then diluted to C100 buffer (salt content 20 mM NaCl, 80 mM KCl) and applied to a 1 ml heparin affinity chromatography column. E64HFrmR was eluted with C500 buffer, whereas ZntR was recovered in the loading buffer as it does not bind the column. Fractions (1 ml) were analysed for protein by SDS-PAGE and metal by ICP-MS.

2.6.7 *Analysis of protein-DNA interaction by fluorescence anisotropy (FA)*

2.6.7.1 Production of fluorescently labelled annealed, double stranded DNA oligonucleotide

Complementary single stranded oligonucleotides (frmRAPro_F,R; frmRAPro-T/A-mod_F,R; rcnRAPro-site1_F,R; rcnRAPro-sites1,2_F,R; Table 2. 1) containing the identified FrmR or RcnR binding site and flanking nucleotides were obtained from Sigma-Aldrich. One of the strands was purchased with the 5' end labelled with the fluorescent tag hexachlorofluorescein (HEX). DNA primers were annealed by heating 10 μM of each strand in 10 mM Hepes pH 7.0, 150 mM NaCl to 95 °C before being cooled to room temperature overnight. The successful annealing was confirmed by Native PAGE analysis using a 12 % w/v gel with a TBE buffer system (Sambrook & Russell 2001).

2.6.7.2 Analysis of protein-DNA binding by fluorescence anisotropy

Experiments were carried out anaerobically in 1 ml gas tight cuvettes (Hellma). The fluorescently labelled, annealed probe was diluted to 10 nM in C300 with inclusion of 5 mM EDTA or 5 μM $ZnCl_2$ as required. EDTA (5 mM) was also included when experiments were conducted in presence of different concentrations (10 μM – 20 μM – 50 μM – 100 μM – 0.36 M) of formaldehyde, in order to prevent metal contamination. The 0.36 M value was chosen as correspond to 1 % v/v formaldehyde (d = 1.09 g/ml), a concentration used for random protein – DNA and protein – protein crosslinking (Brodolin 2000). All experiments were performed under anaerobic conditions. Protein was prepared in C500 and 5 mM EDTA. In fluorescence anisotropy (FA) experiments where metal effects on DNA-binding were to be analysed, the protein was incubated with 1.2 molar equivalents of $ZnCl_2$ or CuCl (verified to be > 95% Cu(I)) and EDTA was omitted. For dissociation experiments EDTA was also omitted from the buffer.

A modified Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies) fitted with polarising filters (λ_{ex} 530 nm, λ_{em} 570 nm, averaging time 20 s, replicates 5, 25°C) was used to measure change in anisotropy (Foster *et al.* 2014a). Upon each addition data was recorded after 5 min to allow equilibration. Dynafit was used to fit data to models described throughout the work and in the Appendix. When experiments were performed in the presence of formaldehyde or metals and DNA-binding did not saturate, the response (Δr_{obs}) was obtained from maximum response value determined from apo experiments. The standard thermodynamic equation, $\Delta G_C = -RT\ln K_C$, was used to calculate the coupling free energy (ΔG_C) as will be further described in Chapter 3 (Foster *et al.* 2014a; Guerra & Giedroc 2012). Each experiment was performed in triplicate allowing calculation of mean K_1 ($K_{DNA}^{apo\text{-}protein}$) or K_3 ($K_{DNA}^{effector\text{-}protein}$) and standard deviations. Mean ΔG_C values were calculated by permuting pair-wise these association constant values.

2.6.7.3 Assessment of FrmR-, E64HFrmR- and RcnR- DNA stoichiometry by fluorescence anisotropy

Production of fluorescently labelled annealed, double stranded DNA oligonucleotides was carried out as already described (Section 2.6.7), however, the annealing reaction was conducted using 200 µM of each strand in order to obtain a highly concentrated DNA stock. Protein aliquots (1 molar equivalent steps) were added to 2.5 µM DNA diluted in buffer C300 and 5 mM EDTA in 1 ml gas tight cuvettes (Hellma). The samples were thoroughly mixed by pipetting and allowed to equilibrate. Again, a modified fluorescence spectrophotometer was used to record change in anisotropy, as described previously (Section 2.6.7.2).

2.6.8 *Analysis of DNA binding by Electrophoretic mobility shift assay (EMSA)*

2.6.8.1 Production of probe DNA fragments by PCR

Protocols described in Sections 2.3.1-2.3.8 were used to produce the probe DNA fragments. The entire *rcnA-rcnR* intergenic region was amplified from genomic DNA using appropriate primers (*rcnRA*ProEM_F,R; Table 2. 1), then ligated to pGEMT to create pGEM*rcnRA*Pro. The specific *rcnRA*ProEM DNA fragment was sequenced and amplified from pGEM*rcnRA*Pro using the same primers and *Pfu* polymerase. The non-specific DNA fragment, consisting of sequence from the pGEM-T multi-cloning site and flanking regions, was amplified from recircularised empty pGEM-T using appropriate primers (pGEMCon3_F,R; Table 2. 1) and designated pGEMCon3 (141 bp).

A single gel extraction column was used in order to collect the product of multiple PCR reactions and to obtain a highly concentrated stock. DNA stock concentration was measured using a NanoDrop 1000 spectrophotometer (Thermo Scientific).

2.6.8.2 Binding of RcnR to DNA monitored by EMSA

A purified stock of RcnR was buffer exchanged by loading onto a 1 ml Heparin column pre-equilibrated with buffer B300. The column was then washed with 10 column volumes of buffer B300 before elution in 1.2 M NaCl, 40 mM Hepes pH 7.8, 8 mM EDTA, 8 mM DTT (buffer E). Protein concentration was determined by measuring UV absorbance at λ 280 nm.

RcnR was incubated for 30 min at room temperature with *rcnRA*ProEM and nsDNA (100 nM of each) with varying [RcnR] in 10 mM Hepes, pH 7.8, 300 mM NaCl, 2 mM DTT, 2 mM EDTA, 3 % v/v glycerol, and 0.05 mM spermidine, then resolved by native PAGE using a TBE buffer system (Sambrook & Russell 2001) and bands visualised with ethidium bromide.

2.7 Crystallography

2.7.1 *Crystal trials*

Protein was purified as described in Section 2.4.6 but instead of being moved inside the anaerobic chamber was stored aerobically in a buffer containing a reducing agent (500 mM NaCl, 10 mM Hepes pH 7.0, 1 mM DTT, 1 mM EDTA). Protein was quantified by measuring the absorbance at 280 nm and subsequently concentrated to ~ 10 mg/ml (~ 1000 µM) using a Vivaspin2 (3kDa MWCO) and diluted back to 400 mM NaCl, 10 mM Hepes pH 7.0, 1 mM DTT, 1 mM EDTA (chelex-treated buffer). When crystal trials were designed to include metal ions, protein was eluted from the last purification step and subsequently diluted in buffer not containing any chelator. Desired final concentration was ~ 500 µM (equivalent to ~ 5mg/ml since MW_{FrmR} 10185.8 Da). Samples were stored at 4 ^{o}C for up to two weeks and protein purity and folded state was checked by 18% w/v SDS-PAGE daily prior to use. When crystals of protein-metal complex were desired, protein was first incubated with 1.2 molar equivalent of $ZnCl_2$ prior to crystallization assays.

Pure samples of FrmR and E64HFrmR were screened against various commercial screens obtained by Molecular Dimensions (Hampton Research) (Table 2. 2) using the sitting-drop vapor diffusion method. Crystal trials were set up in a 96-well crystallization plate format. An aliquot of screening solution (100 µL) was deposited into the reservoir wells of the plate and 200 nL of protein was dispensed into each microbridge of the tray with a Screen Maker robot (Innovadyne technologies). The robot then dispensed 200 nL of the reservoir in the corresponding microbridge well already containing the protein sample. Trays were sealed and kept at 20 ^{o}C.

Hanging-drop technique was subsequently used to screen the most promising crystallization conditions (Table 2. 3). Sample solution was spun down (18,000 x g) for 5 min in order to remove any particulate such as dust or precipitated protein. Grease was applied with a 20 ml

SITTING DROP KITS
JCSG-*plus*™
Clear Strategy™ Screen I
Clear Strategy™ Screen II
PACT *premier*™
Structure Screen 1

Table 2. 2 Hampton Research crystallization kits used for a preliminary screening of crystallization conditions of FrmR and E64HFrmR. These kits allow the screening of several conditions, such as various salt concentrations, different organic and PEG precipitants, effect of ΔpH (4.0-10.5) or addition of a number of cryoprotectants.

	HANGING DROP CONDITIONS
1.	PEG4000 15% to 25% w/v (2 % step), BisTris pH 6.0 – 6.5
2.	PEG3350 15% to 25% w/v (2 % step), BisTris pH 4.9 – 5.0 – 5.1 – 5.2
3.	PEG1150 15% to 25% w/v (2 % step), BisTris pH 4.9 – 5.0 – 5.1 – 5.2
4.	PEG2000-MME 15% to 25% w/v (2 % step), BisTris pH 4.9 – 5.0 – 5.1 – 5.2
5.	PEG200 40% to 50% w/v (2 % step), BisTris pH 6.0 – 6.5
6.	$(NH_4)_2SO_4$ 1.5 to 2 M (2 % step), $NaCH_3COO$ pH 4.0 – 4.5 – 5.0
7.	$(NH_4)_2SO_4$ 1.5 to 2 M (2 % step), BisTris pH 5.5 – 6.0
8.	PEG1500 15% to 25% w/v (2 % step), $NaCH_3COO$ pH 4.0 – 4.5 – 5.0

Table 2. 3 Crystallization cocktails adopted for the optimization screening of crystallization conditions of FrmR and E64HFrmR.

syringe around each well of a 24-well hanging-drop tray. An aliquot of reservoir solution (500 µL) was then manually pipetted inside each well and increasing concentrations of NaCl (5 % v/v to 20 % v/v) were added to each row. Micro-drops containing different protein:ratios (1:1, 2:1) were pipetted on a 0.5 mm glass coverslip which was then placed upside down on the greased well. Trays were stored at 20 °C. Crystals were checked with an illuminator microscope (Leica KL1500 LCD) every day for 7 days and every couple of days afterwards.

2.7.2 *Data collection*

Crystals successfully grown as described in Section 2.7.1 were treated with a cryo-protectant solution (glycerol, sorbitol) and then fished out of the cryo-solution with the appropriate loop. At this step we encountered some difficulties as crystals appeared to disintegrate immediately upon treatment with glycerol and, to a lesser extent, sorbitol. Crystals were therefore rapidly fished out, resulting in an inefficient protection against freezing. Samples mounted on the loops were flash-frozen in liquid nitrogen and sent to the Diamond Light Source (UK's national synchrotron science facility) for X-ray diffraction analyses. Beamline (ID: IO3) wavelength was 0.97625 Å. Images were recorded by a PILATUS detector (Broennimann *et al.* 2006) through 180° with an individual oscillation angle of 0.5°.

At the time of writing, data processing is still ongoing and is being carried out by Dr. Ehmke Pohl (Durham University) therefore it will not be described in this work.

2.8 Bioinformatics

To identify sequences displaying homology with the input sequence, BLAST searches were performed using the NCBI Protein BLAST tool (Altschul *et al.* 1990). EMBOSS Needle (Rice *et al.* 2000) was used to compare amino acid sequences and determine the optimal global sequence alignment while ClustalW2 (Thompson *et al.* 1994) was chosen when three or more sequences were to be aligned. ProtParam tool (Gasteiger *et al.* 2005), an algorithm available at the Expasy Proteomics Server (Gasteiger *et al.* 2003), was used to obtain various physical and chemical parameters (such as molecular weight, theoretical pI and amino acid composition) of the input protein.

Software used by Dr. Ehmke Pohl (Durham University) in order to process X-ray diffraction data collected at Diamond Light Source are not listed or described in this work.

Chapter 3

Characterisation of *Salmonella* FrmR

3.1 Aims and objectives

This chapter presents the identification of the gene encoding *Salmonella* FrmR, member of the RcnR/CsoR-family of transcriptional repressors, followed by protein characterization, analyses of metal binding properties and Zn(II)-, Cu(I)-, Co(II)- and Ni(II)-binding affinities. Moreover, the region to which FrmR binds has been identified upstream of the *frmRA* operon, and the protein:DNA interaction has been investigated by fluorescence anisotropy. Site-directed mutagenesis was employed to identify the amino acid residues involved in DNA-binding and to further understand metal-binding.

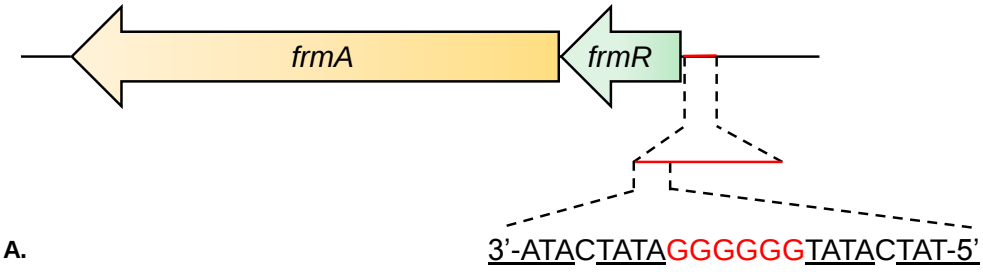
3.2 Identification of recombinant FrmR

The *frmRA* operon is displayed in Figure 3. 1A, which also shows the putative promoter region to which FrmR, transcriptional regulator of this operon, is predicted to bind. This region is located eleven bp upstream of *frmR* and contains a palindromic motif, a likely candidate transcription factor binding site. This sequence has analogy to CsoR/RcnR binding sites (Figure 3. 1B), which are centered with a GC rich region (Iwig & Chivers 2009; Grossoheme *et al.* 2011). FrmR in *Salmonella* is predicted to act as a DNA-binding transcriptional repressor, interacting with the palindromic sequence and preventing the binding of the RNA polymerase, thus impeding the expression of *frmA.*

The primary sequence of FrmR was aligned with *Mycobacterium tuberculosis* CsoR (CsoR_Mtb), *Synechocystis* PCC 6803 InrS (InrS_Syn), *E. coli* RcnR (RcnR_Ec) and *S. typhimurium* RcnR (RcnR_Sty) (all belonging to CsoR/RcnR family), using ClustalW2 (Figure 3. 2), whereas their similarity and amino acid identity were compared with FrmR using EMBOSS Needle (Table 3.1).

3.3 Production and purification of recombinant FrmR

Salmonella FrmR was overexpressed in *E. coli* (~ 5 mg/l culture) as a recombinant, tag-free protein and a purification procedure was developed based on the inherent biochemical properties of FrmR (Figure 3. 3). Since FrmR is likely to bind metals because it retains at least two potential metal-ligands, (residues X and Y in the WXYZ motif), the first step in the purification protocol involved a nickel affinity column, which was loaded with crude cell lysate and eluted in a single step with an elution buffer containing 0.3 M imidazole. A highly concentrated sample was routinely recovered in fraction 2 (although protein was also found in fractions 1 and 3) (Figure 3. 3A) and then loaded onto a size exclusion chromatography column (HiLoad 26/60 Superdex S75 GE Healthcare) which was used to separate FrmR from other proteins and molecules in the cell lysate depending on molecular weights (MW). The size exclusion chromatography column was previously equilibrated with a buffer containing 0.3 M NaCl. The same buffer was used to elute FrmR which was recovered



A.

Mtb_CsoR	3'-GTTCGACTCCTTGGGTAGCCCACCCCCAGTGGGGTGGGATACCATGAACGGGTG-5'	-19bp
Bs_CsoR	3'-AATCATAAAGCGTTTTTTATTGTAATACCCTACGGGGGTATGGTAGGATGAAAA-5'	-28bp
Ec_RcnR	3'-ATTAATCTACTGGGGGGTAGTATCAGGTACTGGGGGGGAGTAGAATCAGATTGC-5'	-26bp
Sy_InrS	3'-CTCATCAATATCCCCCCTGGGGGCATAGAATAGAGATCAATTTTCTACCCCAAA-5'	-8bp
Ec_FrmR	3'-TATATACTATAGGGGGGTATGCAN8TAGAATACCCCCCTATAGTATATTGCATG-5'	-10bp
Sty_FrmR	3'-ATATACTATAGGGGGGTATACTATCAGAATTTTATGTTTATGTAATTGTAAAAT-5'	-9bp

B.

Figure 3. 1 A Genomic region of *frmR* and *frmA* (to scale) and deduced transcriptional regulator binding site (palindromic region rich in CG) upstream of *frmR*. The G tract is highlighted in red and the T/A-rich tracts are underlined. **B** Comparison of DNA binding sites of RcnR/CsoR proteins. G/C tracts are highlighted in red and T/A-rich tracts are underlined. The value reported next to each sequence refers to the number of bp until the predicted translation start.

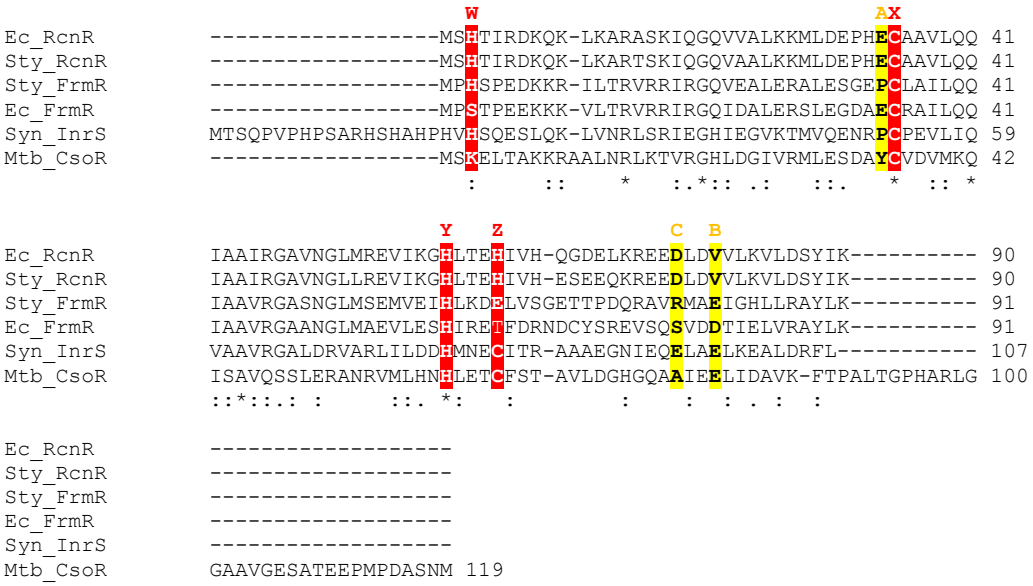


Figure 3. 2 Alignments of *S. Typhimurium* FrmR (Sty_FrmR) with *Mycobacterium tuberculosis* CsoR (Mtb_CsoR), *Synechocystis* PCC 6803 InrS (Syn_InrS), *E. coli* FrmR (Ec_FrmR), *E. coli* RcnR (Ec_RcnR) and *S.* Typhimurium RcnR (Sty_RcnR) by Clustalw2. The residues involved in the primary [W, X, Y, Z] and secondary [A, B, C] co-ordination sphere are highlighted in red and yellow, respectively.

S. typhimurium FrmR	**Similarity (%)**	**Identity (%)**
E. coli FrmR	74.7	52.7
E. coli RcnR	64.1	39.1
S. typhimurium RcnR	64.8	39.6
M. tuberculosis CsoR	35.2	21.9
Synechocystis InrS	27.8	13.6

Table 3. 1 Similarity and identity of amino acid sequences of *Mycobacterium tuberculosis* CsoR (Mtb_CsoR), *Synechocystis* PCC 6803 InrS (Syn_InrS) , *E. coli* RcnR (Ec_RcnR) and *S. Typhimurium* RcnR (Sty_RcnR) were compared to *S. Typhimurium* FrmR (Sty_FrmR) using EMBOSS NEEDLE.

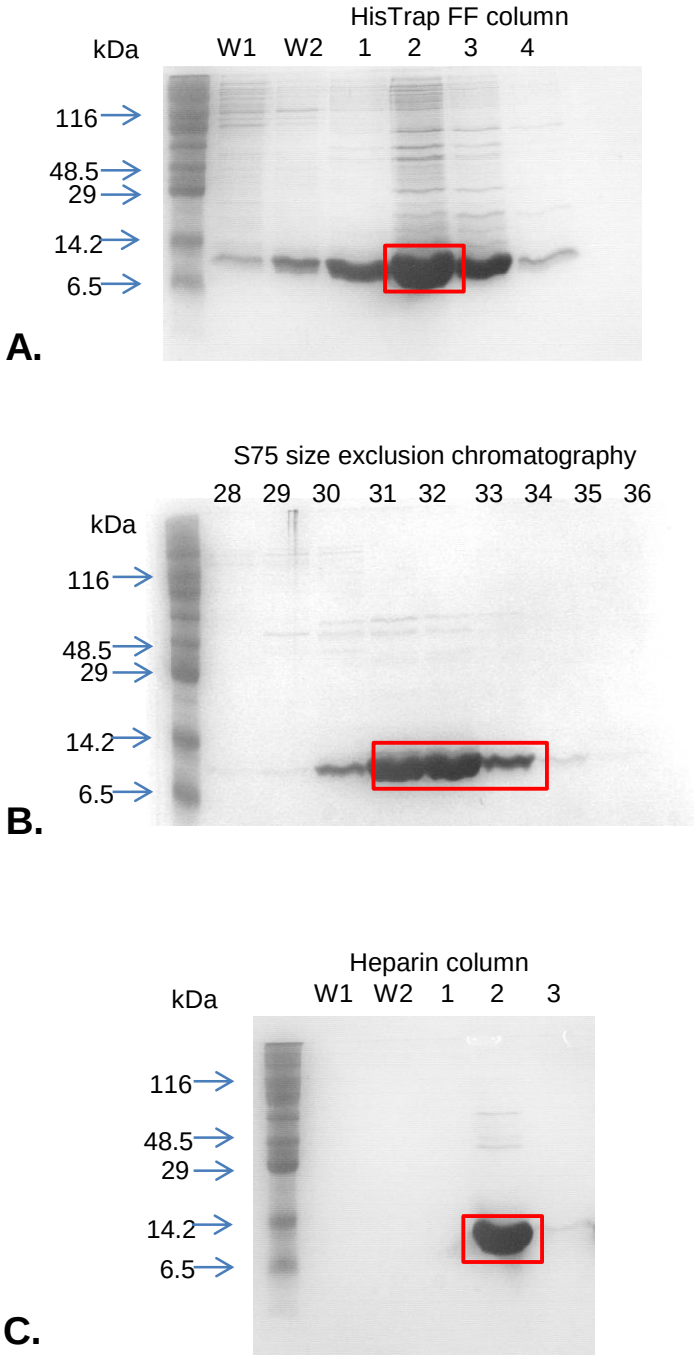


Figure 3. 3 Purification of FrmR by nickel affinity chromatography, size exclusion and heparin affinity chromatography. **A** SDS-PAGE analysis (18 % w/v acryl-bis) of fractions eluted from a 5ml Ni(II)-affinity column with buffer A containing 0.3 M imidazole. Fraction 2 was routinely found to contain FrmR in the highest concentration. **B** SDS-PAGE analysis (18 % w/v acryl-bis) of fractions 28-36 eluted from a Superdex S75 column loaded with fraction 2 (3 ml) from the previous purification step (A.) **C** SDS-PAGE analysis (18 % w/v acryl bis) of fractions eluted at 0.5 M NaCl from two 1 ml Heparin columns linked together loaded with size exclusion fractions enriched for FrmR.

in fractions 30-34 (Figure 3. 3B) although only the 2-3 most concentrated samples were pooled and further purified by heparin-Sepharose affinity chromatography (Figure 3. 3C). In these columns agarose beads are coupled with heparin and, due to its polyanionic nature, this chromatography methodology is used to bind cationic macromolecules (Farooqui 1980), such as DNA-binding proteins. This last step was also added in order to concentrate the protein sample. FrmR purity was assessed after each purification step by 18 % w/v SDS PAGE confirming ≥ 95 % purity for the final sample. For use in the anaerobic chamber, protein was loaded onto a 1 ml heparin affinity column to be moved inside the anaerobic glove box. Since we intended to investigate FrmR metal binding properties, it was necessary to remove EDTA and DTT present in the buffers up to this step by washing the column with Chelex-treated, oxygen-free buffers inside the chamber. 5-5'-dithiobis-(2nitrobenzoic acid) (DTNB or Ellman's Reagent) was used to determine FrmR reactive thiol content. FrmR possesses one cysteine, Cys35, (Figure 3. 2) therefore only one sulfhydryl group is able to react with DTNB. Sulfhydryl group concentration was assayed using the extinction coefficient of TNB (14,150M^{-1} cm^{-1} at 412 nm) confirming the production of almost completely reduced protein samples (only samples with a reduced thiol content ≥ 85 % were accepted, although routinely the value was in the 95-100 % range). Prior to be utilized for *in vitro* characterisation, protein samples were analysed for metal content by ICP-MS verifying a contamination < 5 %.

This preliminary characterization of apoFrmR showed a single species roughly corresponding to a homotetramer as deduced by the elution volume of FrmR (peak fractions volume = 155-160 ml) from the Superdex S75 column. The size exclusion column used for FrmR (HiLoad 26/60 Superdex S75, GE Healthcare) was calibrated with standards of known molecular weight by Dr. A. W. Foster obtaining a linear relationship between *k* (gel phase distribution coefficient) and $\log_{10}$MW (k = -0.52 $\log_{10}$MW + 2.6499; Foster 2012). The *k* coefficient was then calculated for FrmR using the formula $k = (V_e - V_0)/(V_t - V_0)$, where V_e is the elution volume of FrmR (157.5 ml), V_0 is the void volume of the column (100 ml) and V_t is the total volume of the column (318 ml) (Foster 2012). k^{FrmR} (0.26) can be converted to $\log_{10}$MW which corresponds to MW^{FrmR} = 39,436 Da. This value is consistent with FrmR adopting a tetrameric state under micromolar concentrations (monomer MW = 10185.8 Da) and this has subsequently been confirmed by X-ray crystallography (Chapter 6). Moreover it is consistent with what was observed for *E. coli* FrmR by Law using multi angle light scattering (MALS) and size exclusion chromatography (SEC) techniques (Law 2012). Other RcnR-CsoR family members are tetramers at micromolar concentrations (*E. coli* RcnR, Iwig *et al.* 2008 and CsoR from different organisms, Ma *et al.* 2009a; Sakamoto *et al.* 2010; Dwarakanath *et al.* 2012).

Amino-acid sequencing (or Amino Acid Analysis, AAA) was carried out on the protein to calculate experimental extinction coefficient ε, by determining the experimental concentration of a purified FrmR sample previously analysed by UV-vis spectroscopy and by applying the Lambert-Beer equation (Equation 2):

$$A = \varepsilon bc$$

Equation 2

where A is the absorbance value at 280 nm, ε (M^{-1} cm^{-1}) is the extinction coefficient (to be determined), b is the path length (here 1 cm) and c (M) is the protein concentration (here determined by AAA) (Lambert 1760; Beer 1852). The UV-vis absorbance spectrum and the calculations showing how the extinction coefficient was obtained will be presented in the Appendix (Figure 8. 1). Furthermore, this analysis suggested the cleavage of the N-terminal methionine, resulting in a proline being the first amino acid residue, as will be further discussed in Section 5.3.1. The numbering used in this work will still refer to the full *frmR* gene sequence, therefore the above mentioned proline will be named Pro2.

3.4 Analysis of FrmR metal binding properties

3.4.1 *Co(II)- and Zn(II)-binding properties*

Since *Salmonella* FrmR shares similarity with other metal sensors belonging to RcnR/CsoR family (see Table 3.1), particular interest is raised in regard of FrmR metal binding properties, especially since it contains at least two potential metal-ligands of the WXYZ fingerprint (Figure 3. 2).

Titrations of protein with Co(II) (Figure 3. 4) leads to absorption bands present in the near-UV region at 294 and 336 nm, indicative of a S-Co(II) Ligand-Metal-Charge-Transfer (LMCT) absorption. These high energy features are a direct probe of ligand-metal bonding and indicate that the protein ligand valence orbitals (electron donor) are close in energy to metal *d* orbitals (electron acceptor) (Holm *et al*. 1996). The intensity of the LMCT features allows us to quantify the extent of the ligand: metal interactions. (Holm *et al*. 1996). An intense low-energy CT transition such as those observed for Co(II)-FrmR are consistent with a metal site containing a thiolate or a phenolate. In addition to the LMCT features Co(II)-FrmR also shows a broad *d-d* transition envelope centred at ~ 600 nm (peak maximum at 614 nm). *d-d* electronic transitions are very valuable because they are diagnostic of the Co(II) coordination sphere in the metal binding site. FrmR cobalt-dependent *d-d* transition features have an $\varepsilon \sim 0.3 \times 10^3$ M^{-1} cm^{-1} which is consistent with one Co(II) ion bound in a tetrahedral coordination environment (VanZile *et al.* 2002b). This is distinguishable from cobalt binding with symmetrical coordination geometry, such as octahedral, where the *d-d* transitions would have much lower intensity ($\varepsilon < 0.05 \times 10^3$ M^{-1} cm^{-1}) because theoretically forbidden by the Ligand field theory selection rule (Roe *et al.* 2007). According to this theory, before a ligand approaches a metal ion, all *d* orbitals are equivalent in energy (Roe *et al.* 2007).

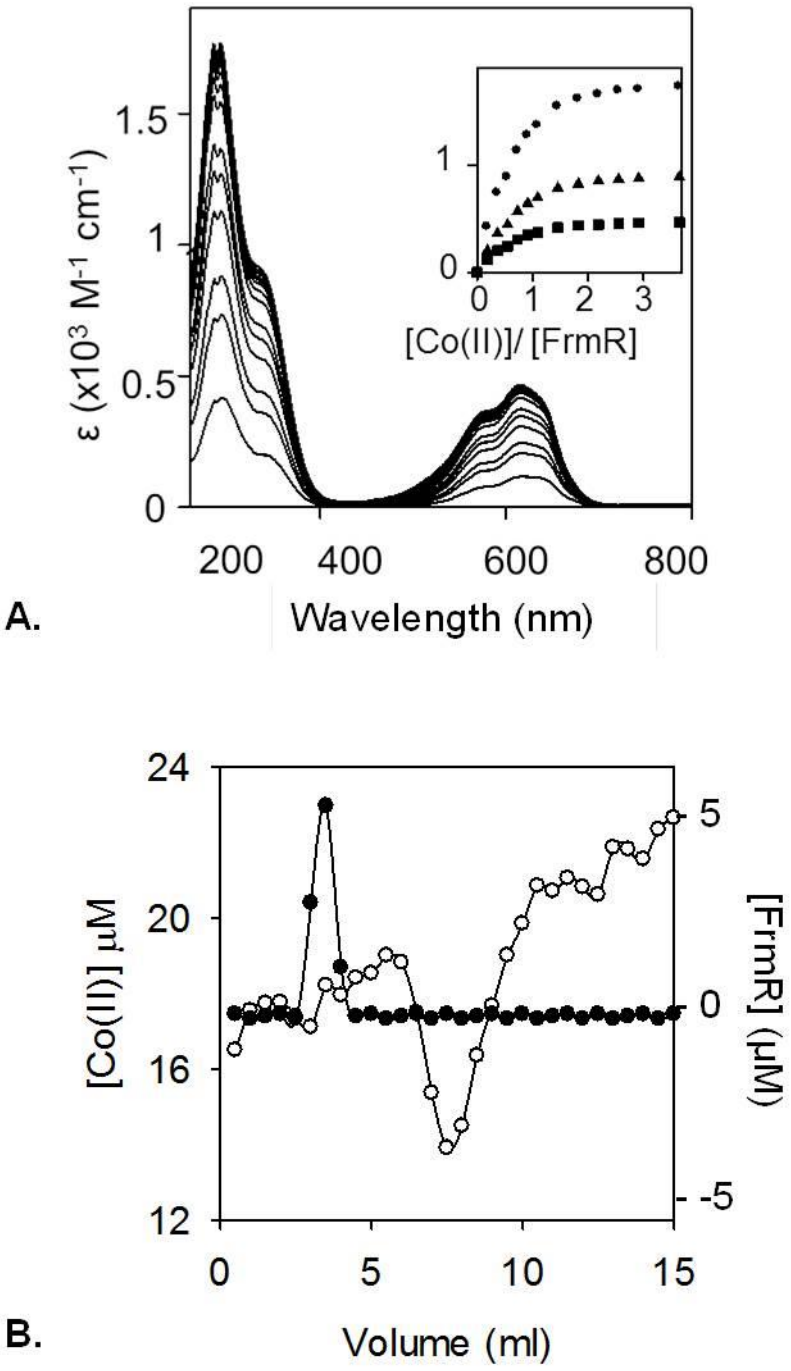


Figure 3. 4 A Representative apo-subtracted UV-visible difference spectra of FrmR (83.9 µM protomer) upon titration with $CoCl_2$ and binding isotherms (inset) at absorbance maxima 294 nm (circles), 336 nm (triangles), and 614 nm (squares). **B** Size-exclusion fractionation of protein-bound and free Co(II). FrmR (50 µM) was applied and resolved anaerobically on a Sephadex G-25 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 with the addition of 20 µM $CoCl_2$. Bound and free metal were eluted in the same buffer. Fractions (500 µl) were analysed for protein by Bradford assay (closed circles) and cobalt by ICP-MS (open circles).

Upon ligand approach, the energy of the metal d orbitals increases in an uneven fashion resulting in a split in two levels (t_{2g} and e_g) separated by the ligand-field splitting parameter Δ_o (where "o" refers to the octahedral field) (Roe *et al.* 2007). In a strictly octahedral complex, the ligand arrangement has inversion of symmetry therefore the *d* orbitals retain their centres of symmetry. Therefore all the *d* orbitals retain the same symmetry and, according to Laporte selection rule, electronic transitions between levels of the same symmetry are forbidden (i.e. *gerade* → *gerade*, *ungerade* → *ungerade*) (Roe *et al.* 2007). We would actually still detect low intensity *d-d* transitions, because the approaching ligand is vibrating asymmetrically causing a distortion in the symmetry of *d* orbitals.

However, in tetrahedral complexes the coordination environment is non-symmetric therefore transitions between split-orbitals are not Laporte-forbidden and so electronic transitions can take place. For tetrahedral complexes we observe high intensity *d-d* transitions but at a lower energy because the ligand-field splitting parameter Δ_t (where "t" refers to the tetrahedral field) is lower than Δ_o (Roe *et al.* 2007). The intensity of the LMCT feature in the region of 330 nm of Co(II)-FrmR, $\varepsilon \sim 0.9 \times 10^3\ M^{-1}\ cm^{-1}$, is consistent with a single Cys thiolate ligand (ε between 0.8 and $1.2 \times 10^3\ M^{-1}\ cm^{-1}$ for each Co(II)-thiolate bond (VanZile *et al*. 2002b). FrmR is expected to be a tetramer under these conditions, therefore there are four metal sites per tetramer. Both LMCT and *d-d* transition spectral features saturate after 1 molar equivalent of cobalt suggesting a 1:1 Co(II)-FrmR monomer stoichiometry although binding curves are linear implying $K_{Co(II)}$ too tight to estimate by this method (insets Figure 3. 4A).

To further elucidate stoichiometry, apoFrmR was resolved anaerobically on a Sephadex G25 matrix previously equilibrated with buffer containing 20 μM $CoCl_2$. Samples were analysed for metal by inductively coupled plasma mass spectrometry (ICP-MS), and for protein by Bradford assay. FrmR does not migrate with cobalt implying a weak cobalt affinity of FrmR (further investigated in Section 3.5.2) (Figure 3. 4B).

In order to study Zn(II) binding to FrmR, a method involving a competition [Zn(II) vs Co(II)] experiment was employed (Figure 3. 6). Many examples can be found in the literature showing that Zn(II) addition quenches Co(II)-dependent spectral features (VanZile *et al.* 2000; Ma *et al.* 2011a; Foster *et al*. 2012). In fact Zn(II) has a $3d^{10}$ configuration so there is not a free t_{2g} orbital to be filled with an electron excited from an e_g orbital. ApoFrmR was titrated as described with Co(II) to a stoichiometry excess (2.2 molar equivalent) and optical spectra were then measured upon addition of $ZnCl_2$. Quenching of spectral features in the near-UV (Figure 3. 5) occurred after ~ 1 molar equivalents of zinc. This outcome suggests that Zn(II) and Co(II) occupy the same binding site on FrmR and adopt the same coordination geometry (tetrahedral with one thiol in the inner coordination sphere). Since 1 molar equivalent of zinc is able to displace 2.2 molar equivalents of cobalt it can be hypothesized a greater affinity for zinc than for cobalt although it was not possible to determine a K_{Zn} from these data because the excess of cobalt

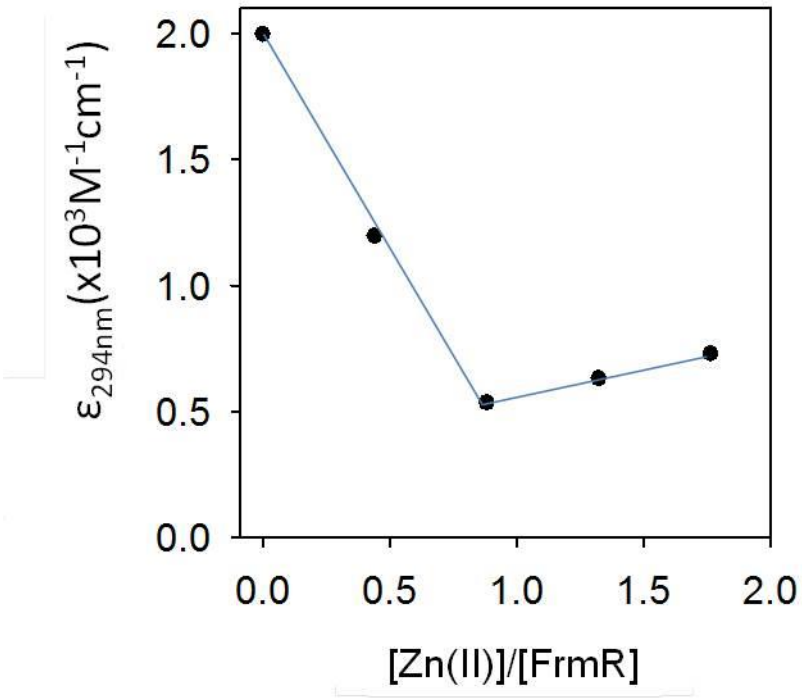


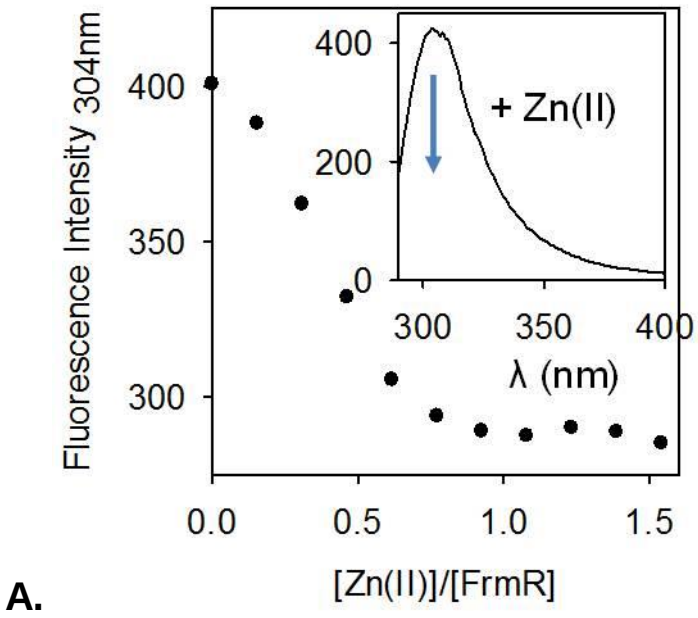
Figure 3. 5 Apo-subtracted UV-visible difference spectra of Co(II)-FrmR (58.97 µM, monomer; equilibrated with excess of $CoCl_2$) and decrease upon addition of 0.5 and 1 molar equivalent of $ZnCl_2$ per FrmR monomer. Quenching of the Co(II)-dependant spectral features indicates displacement of Co(II) from FrmR by Zn(II). The absorbance of Co(II)-dependant LMCT is completely quenched by 1 molar equivalent of Zn(II) per FrmR monomer. The final values do not return to 0 because of the upshift in the baseline after Co(II) addition before zinc quenching.

required in this analysis led to protein precipitation. The slightly upshifted baseline results in zinc not being able to completely quench the cobalt dependent feature. However, if subtraction of the difference in absorbance at 294 nm resulting from protein precipitation is performed, the aforementioned spectral feature is completely quenched.

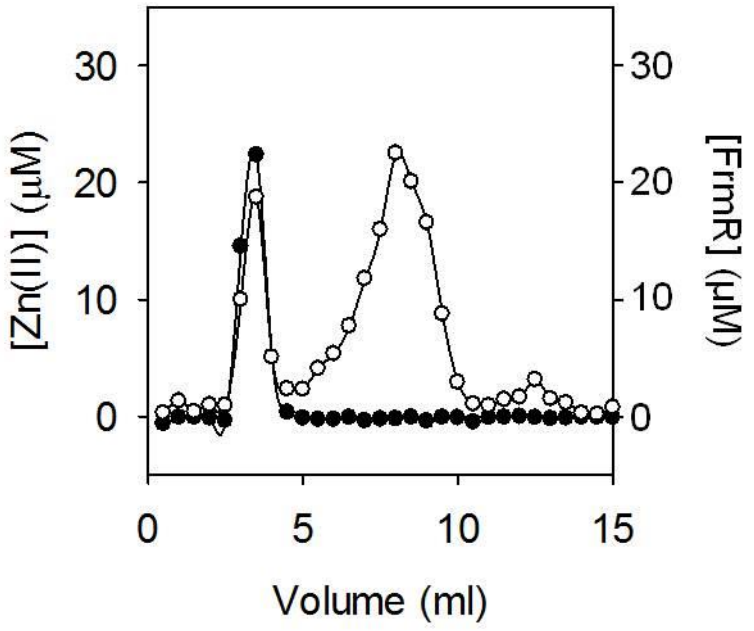
UV-vis absorbance spectroscopy cannot be used directly to assess Zn(II)-binding by FrmR and so an alternative strategy was used to determine Zn(II):FrmR stoichiometry. Figure 3. 6A exhibits FrmR tyrosine natural fluorescence quenched upon zinc titration. Fluorescence spectroscopy is a very powerful method to study protein assembly and interactions because it is broadly applicable since almost all proteins possesses intrinsic fluorophores (e.g. tyrosine and tryptophan) (VanZile *et al.* 2002b; Munishkina & Fink 2007; Waldron *et al.* 2009). Tryptophans are commonly used to probe protein folding and dynamics since they have a higher quantum yield than tyrosines and energy transfer to Trp residues usually quenches the Tyr fluorescence (Munishkina & Fink 2007). FrmR is a Trp-lacking protein, possessing only a single tyrosine residue (Y89). The Tyr emission is not as sensitive to the polarity of the environment as the Trp emission but can still be a valuable tool for studying the complex mechanisms of metal protein binding (Munishkina & Fink 2007). Another great advantage of using fluorescence emission techniques is the high signal to noise ratio which allows the use of small protein concentrations. Consistent with this, a titration of Zn(II) into a solution containing FrmR (13.1 μM, pH 7.0) was carried out, monitoring the protein fluorescence emission spectrum following excitation at 280 nm.

As shown in Figure 3. 6A inset, apo-FrmR Tyr emission has a maximum at ~ 305 nm, which was then quenched after addition of 1 molar equivalent of zinc (Figure 3. 6A). This outcome suggests the formation of a 1:1 zinc:protein complex and it is in accord with what was observed for Co(II)-quenching by UV-visible spectroscopy.

In support of a 1:1 Zn(II)-binding stoichiometry, FrmR migrates with approximately one molar equivalent of zinc when a sample of FrmR, anaerobically incubated with excess metal for 2 h, was resolved on Sephadex G25 media (Figure 3. 6B). Since Zn(II) displaced Co(II) in the metal binding site, as shown by UV-vis spectroscopy (Figure 3. 6A), FrmR presumably has a higher affinity for zinc than for cobalt therefore the matrix and/or buffer competition effect here should be negligible. This stoichiometric binding allows a minimum estimate of Zn(II)-binding affinity of FrmR of $< 10^{-7}$ M.



A.



B.

Figure 3. 6 A Fluorescence emission at 304 nm (λ_{ex} = 280 nm) of FrmR (13.1 μM, monomer) and following titration with $ZnCl_2$. Fluorescence is quenched after addition of 1 molar equivalent of Zn(II) per FrmR monomer, consistent with a Zn(II)-dependant conformational change of FrmR which alters the environment of the tyrosine residue. **B** Size-exclusion fractionation of protein-bound and free Zn(II). FrmR (50 μM) was incubated anaerobically with 150 μM Zn(II) for 2 h. Bound and free metal were resolved on Sephadex G-25 equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 and eluted in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and zinc by ICP-MS (open circles).

3.4.2 *Cu(I)- and Ni(II)-binding properties*

Reduced Cu(I) stocks were prepared in a strictly anaerobic environment by dissolving CuCl salt in N_2-purged buffers inside an anaerobic glove-box. The solution was then validated to be ~100 % Cu(I) by calibrated BCS assay (Figure 3. 7). BCS (bathocuproine disulfonate) is a Cu(I) binding metallochromic indicator which forms a 2:1 complex (Section 2.6.1).

To determine if FrmR could bind Cu(I) *in vitro,* Cu(I) titration into FrmR was monitored by UV-visible spectroscopy (Figure 3. 8A). A single spectral feature with a maximum at 240 nm was observed. An inflection after 1 molar equivalent of Cu(I) is detectable although the feature continues to rise after this point (Figure 3. 8A inset). Since the spectral feature appears as a straight line up to one equivalent, tight binding of at least one Cu(I) can be inferred. No other spectral features were detected at lower energies, confirming the reduced status of the copper used in this analysis. This LMCT spectral feature possesses a molar absorptivity at 240 nm ~ 8 x 10^{-3} M^{-1} cm^{-1} and is diagnostic of S-Cu(I) bond (Angeletti *et al.* 2005). A similar feature at 240 nm was previously assigned to distorted tetrahedral coordination geometry in Hg-proteins (Tamilarasan & McMillin 1986; Choudhury S. 1983). The same band was then attributed to a trigonal planar complex with strong support for a three-coordinate environment coming from EXAFS (extended X-ray absorption fine structure) data for Hg-MerR (Watton *et al.* 1990). Attributing Cu(I)-FrmR spectral feature at 240 nm to Cys thiolate-Cu(I) trigonal coordination is in line with what is reported in the literature for other Cu(I) binding proteins: *B. subtilis* CsoR (Ma *et al.* 2009a), *S. lividans* CsoR (Dwarakanath *et al.* 2012), *M. tuberculosis* CsoR (Liu *et al.* 2007), *S. aureus* CsoR (Grossoehme *et al.* 2011), *E. coli* RcnR (Higgins *et al.* 2012b) and *Synechocystis* InrS (Foster *et al*. 2012).

CsoR homologues from *B. subtilis* (Ma *et al.* 2009a) and *M. tuberculosis* (Liu *et al.* 2007) exhibit a maximum in the absorbance spectrum at 240 nm indicative of the formation of the thiolate-copper coordination bond. EXAFS were used to determine that *B. subtilis* CsoR binds Cu(I) forming a trigonal S_2N coordination site (Ma *et al.* 2009a). In addition *O. brevis* BxmR is capable of forming low coordination number (n =2 or 3) metal complexes with Ag(I) and Cu(I) ions (both d^{10}). Theoretically, these metal ions have closed-shell d^{10} centers which cannot be filled by electrons from the lower energy level. However an interaction, attributed to the relativistic effect, has been observed between d^{10} metals (Yam 2001). We can speculate that the LMCT observed for Cu(I) complexes could be explained in a similar way considering a configuration mixing of the filled *d* orbitals with the empty ones at higher energies.

Furthermore one equivalent of Cu(I) binds sufficiently tightly to co-migrate with FrmR in size exclusion chromatography (Figure 3. 8B). Reduced FrmR (50 µM) was incubated with three-fold

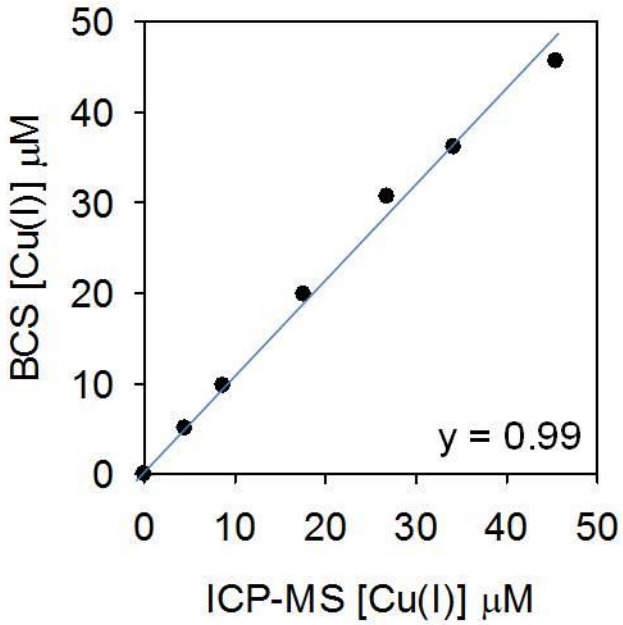
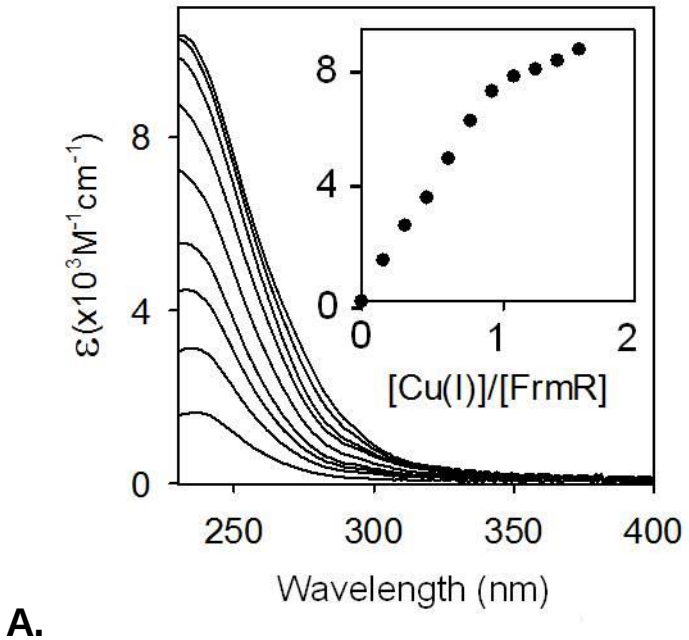
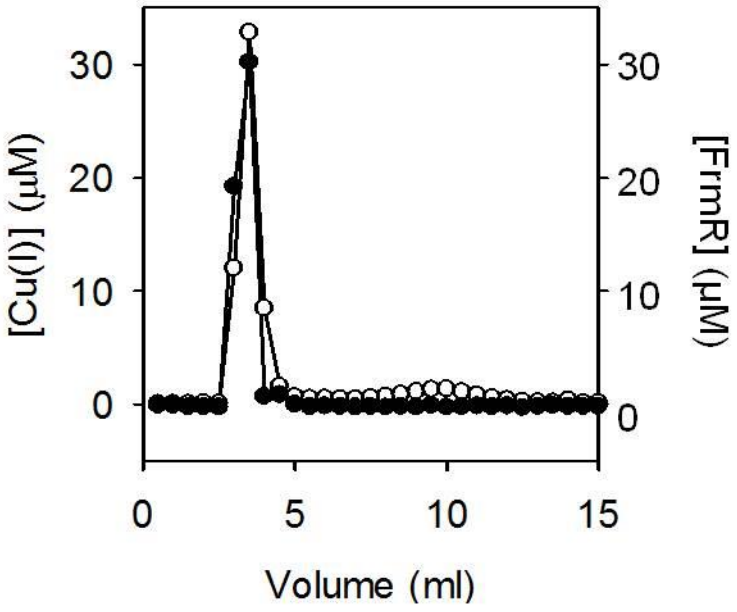


Figure 3. 7 Stocks of Cu(I) were prepared anaerobically and analysed for total dissolved Cu(I) by ICP-MS. The proportion of Cu(I) in solution was verified as approximately 100% by titration against the Cu(I)- specific metallochromic indicator BCS.



A.



B.

Figure 3. 8A FrmR (21.3 μM) upon titration with CuCl (oxidation state verified as Cu(I)) and binding isotherm (inset) at 240 nm. **B** Size-exclusion fractionation of protein-bound and free Cu(I). FrmR (50 μM) was incubated anaerobically with 150 μM Cu(I) for 1 h. Bound and free metal were resolved on Sephadex G-25 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 and eluted in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and copper by ICP-MS (open circles).

excess of CuCl and applied to a column. After elution only [Cu] ~ 50 µM (migrated with the protein) was recovered suggesting a possible competition from matrix.

Ni(II)-binding properties of FrmR were assessed by UV -visible spectroscopy and size exclusion chromatography. Ni(II)-dependent spectral features were observed in the sub 300 nm region (Figure 3. 9A). In particular an LMCT feature, which rises as a straight line and saturates at ~ 0.8 equivalent (ε ~ 1500 M^{-1} cm^{-1}) (Figure 3. 9A inset), can be detected at 280 nm suggesting a minimum estimate of Ni(II) binding affinity of FrmR of < 10^{-7} M. The presence of non-specific light scattering on Ni(II) additions make the spectra interpretation more difficult, however a weak band at ~ 500 nm can be identified. This weak absorbance feature has a molar absorptivity of ~ 200 M^{-1} cm^{-1} and it is indicative of FrmR binding Ni(II) with a square planar geometry (Chen *et al.* 2000). The presence of non-specific light scattering caused by protein precipitation could be indicative of a not entirely formed Ni(II) binding site where adventitious ligands may be recruited to coordinate the metal or of additional weak sites.

A square planar configuration presents the four ligands equidistant from the central metal ion with the ligands arranged at the corners of a square in a plane that include the central ion (Roe *et al.* 2007). When Ni(II) is bound in a square planar geometry we observe a lack of any electronic transition at lower energies ($\lambda \geq 700$ nm) indicating a large crystal-field splitting and a single absorption feature at ~ 500 nm ($50 < \varepsilon < 500$ M^{-1} cm^{-1}) (Lever 1984; Chen *et al.* 2000; Ünver & Hayvali 2010). Ni(II)-binding proteins such as *Synechocystis* InrS (Foster *et al.* 2012), *E. coli* NikR (Wang *et al.* 2004, Chivers & Sauer 2000) and *H. pylori* NikR (Abraham *et al* 2006) show a similar spectra featuring the peak at ~ 300 nm and the weak band at ~ 475 nm although it is difficult to find a correlation between the number of thiolates involved in the bond and the intensity of the LMCT peak which appears significantly weaker in intensity for FrmR.

To further establish the stoichiometry of the Ni(II):FrmR complex, the protein was applied and eluted from Sephadex G25 media which was equilibrated with buffer containing 20 µM $NiCl_2$. However, Ni(II) did not co-migrate with the protein (Figure 3. 9B). The column matrix or buffer may compete for Ni(II). The tyrosine fluorescence emission spectrum shows no quenching upon nickel titration (Figure 3. 10). One explanation may be that tyrosine is not perturbed by the ligands rearranging to coordinate Ni(II).

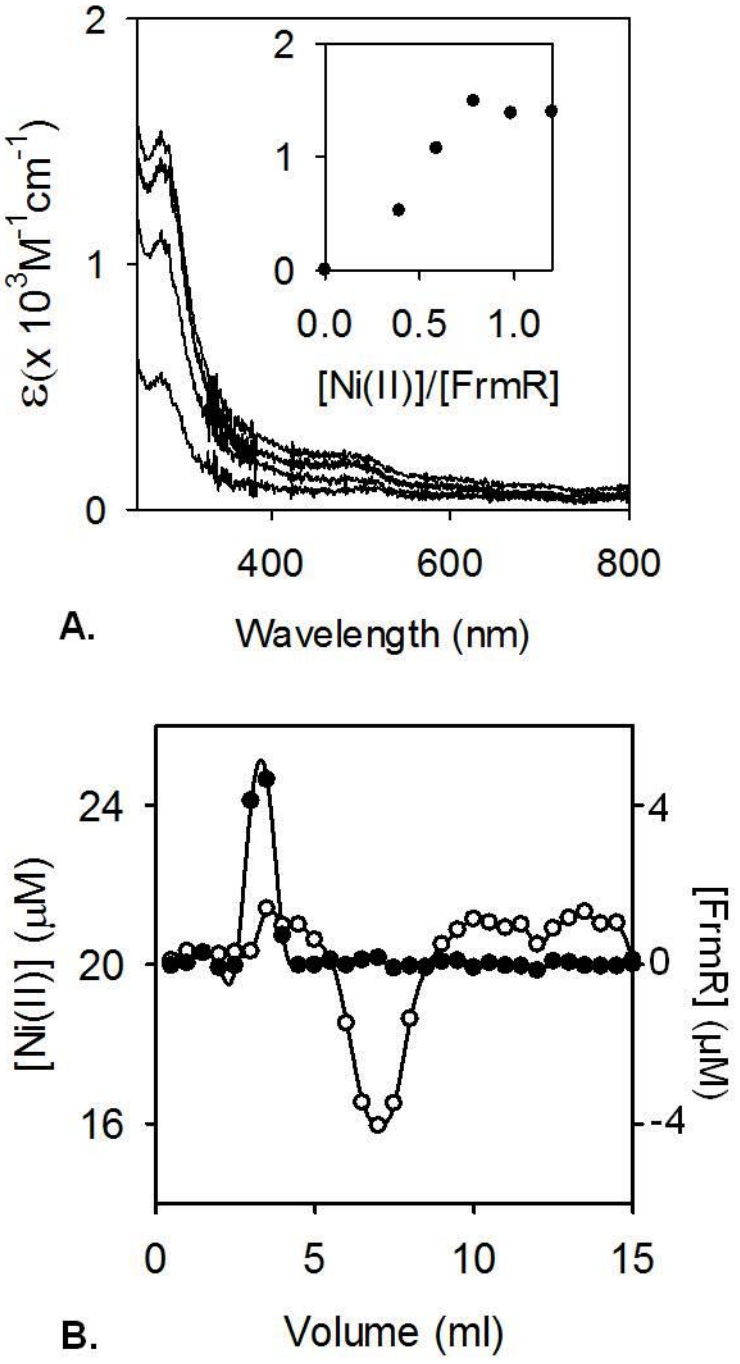


Figure 3. 9 Ni(II)-binding properties of FrmR. **A** Apo-subtracted UV-visible difference spectra of FrmR upon titration with $NiCl_2$ (15.2 μM monomer). Inset: Binding isotherm of the spectral feature at 280 nm. **B** Size-exclusion fractionation of protein-bound and free Ni(II). FrmR (50 μM) was applied and resolved anaerobically on a Sephadex G-25 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 with the addition of 20 μM $NiCl_2$.Bound and free metal were eluted in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and copper by ICP-MS (open circles).

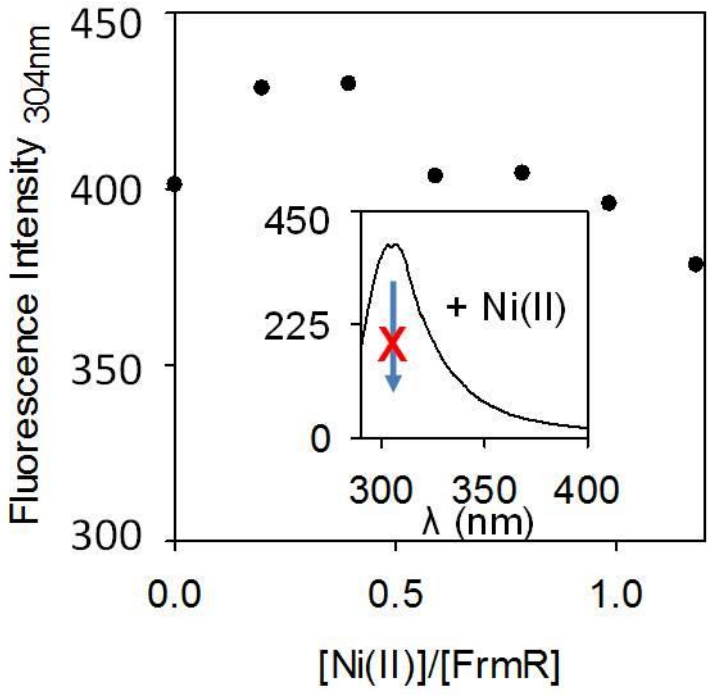


Figure 3. 10 Fluorescence emission at 304 nm (λ_{ex} = 280 nm) of FrmR (13.1 µM) and following titration with $NiCl_2$. Fluorescence is not quenched by nickel, suggesting that FrmR binds the metal too weakly to be detected by this assay or Ni(II)-binding does not elicit a ligand rearrangement which involves Tyr.

3.5 Metal binding affinities of FrmR

We revealed that FrmR binds Co(II), Zn(II), Cu(I) and Ni(II) (Figures 3. 4-3. 10). Moreover, FrmR is shown to be able to retain Zn(II) and Cu(I) upon resolution by size exclusion chromatography. Exploring if this protein would be capable to bind those metals with affinities close to those of *Salmonella* metal sensors could provide valuable insights in order to understand metal selectivity and homeostasis.

3.5.1 *Measurement of the Zn(II)-binding affinity of FrmR by competition with mag fura-2 and quin-2*

Saturation of Co(II)-FrmR absorbance peaks and co-migration of FrmR with ~ 1 molar equivalent of zinc during size exclusion chromatography, indicate a stoichiometric binding of Zn(II) (Figure 3. 6A-B). In order to measure Zn(II)-binding affinity, FrmR was competed against mag fura-2 and quin-2, chromophoric chelators which form 1:1 complexes with Zn(II) with known affinities ($K_{Zn(II)}^{\text{mag fura-2}} = 2.0 \times 10^{-8}$ M and $K_{Zn(II)}^{\text{quin-2}} = 3.7 \times 10^{-12}$ M).

Mag-fura-2 and quin-2 are valuable tools for measuring divalent metal affinities since the absorbance diagnostic of apo-mag-fura-2 undergoes a blue shift (from 366 to 325 nm) upon metal binding whereas metal binding to apo-quin-2 causes quenching of the spectral feature at 265 nm. Metal-binding to the protein can therefore be monitored by disappearance of the apo mag-fura-2 absorbance at 366 nm or the apo quin-2 absorbance at 265 nm, allowing estimation of protein $K_{Zn(II)}$ (Foster *et al.* 2014; Lisher *et al.* 2013; Golynskiy *et al.* 2006; Jefferson *et al.* 1990, Simons 1993; Kwan & Putney 1990; Xiao & Wedd 2010; VanZile *et al.* 2002b).

Mag fura-2 (12.2 μM) was titrated with Zn(II) in the presence of FrmR (20.4 μM, monomer). Absorbance at 366 nm did not change upon addition of up to 0.5-0.75 equivalents of Zn(II) per FrmR monomer, suggesting competition between protein and chromophore for Zn(II) (Figure 3. 11A). As mentioned earlier, at micromolar concentrations FrmR exists as a tetramer with one metal-binding site per monomer, comparable to CsoR/RcnR family members which also show negative cooperativity between sites (Ma *et al.* 2009a; Iwig *et al.* 2008; Foster *et al.* 2012; Chang *et al.* 2014)). As shown in Figure 3. 11A, competition for Zn(II)-binding is complete after ~ 27.5 μM Zn(II), suggesting that only three sites in FrmR can compete with mag-fura-2. DynaFit (Kuzmic 1996) was exploited to fit the data using a model describing tight binding of three molar equivalents of Zn(II) per protein tetramer, with dashed lines representing simulated curves describing $K_{Zn1\text{-}3}$ ten-fold tighter or ten-fold weaker than FrmR (Figure 3. 11A). Fitted data points approach the $K_{Zn1\text{-}3}$ ten-fold tighter simulated curve, suggesting that FrmR $K_{Zn1\text{-}3}$ may be outside the range of this assay.

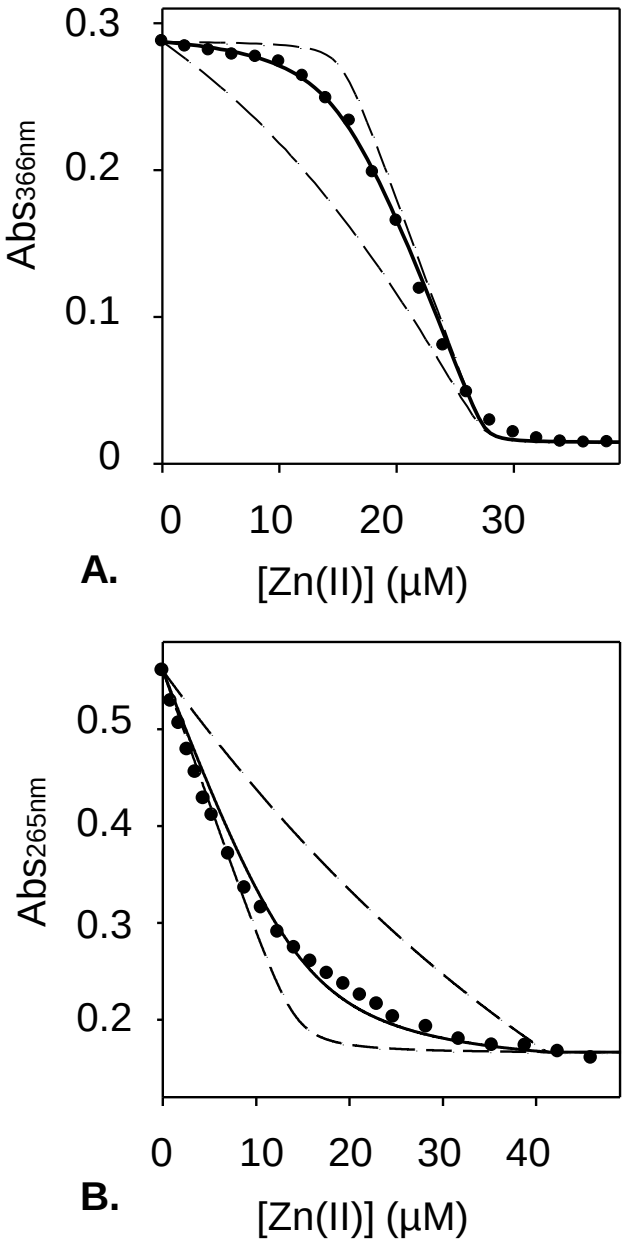


Figure 3. 11 A Representative (n =3) of mag-fura-2 (12.2 μM) titrated with $ZnCl_2$ in the presence of FrmR (20.04 μM). The absorbance at 366 nm was measured by UV-vis spectrophotometer. **B** Representative (n = 3) of quin-2 (14.1 μM) with $ZnCl_2$ in the presence of FrmR (39.9 μM). The absorbance at 265 nm was measured by UV-vis spectrophotometer. In both cases, the model describes competition of FrmR monomer with the chelator for 0.75 molar equivalents of Zn(II) (3 sites per tetramer, $K_{Zn1\text{-}3}$). Dashed lines represent simulated curves describing $K_{Zn1\text{-}3}$ 10-fold tighter and 10-fold weaker.

A similar approach was adopting using quin-2, which possesses a tighter affinity for Zn(II). Quin-2 (14.1 μM) was titrated with Zn(II) in the presence of 39.9 μM FrmR monomer (Figure 3. 11B). FrmR' binding of three molar equivalents of Zn(II) was fitted by DynaFit. Dashed lines in Figure 3. 11B describe simulated curves for K_{Zn1-3} ten-fold tighter or ten-fold weaker than FrmR. K_{Zn1-3} = 1.7 ± 0.7 x 10^{-10} M is within the range of this competition assay (Figure 3. 11B, Table 8.3 Appendix).

3.5.2 *Measurement of Co(II)-binding affinity of FrmR by competition with Fura-2 and BisTris*

As previously discussed, FrmR does not migrate with cobalt upon resolution by size exclusion chromatography (Figure 3. 4B), although it does bind the metal as monitored using UV-vis spectroscopy (Figure 3. 4A), suggesting a possible competition from the matrix or the buffer and thereby a weak cobalt affinity of FrmR. Hypothesizing an affinity in the nM range, the ratiometric fluorescent metal chelator Fura-2 (Fura-2; $K_{Co(II)}$ = 8.64 x 10^{-9} M) was used as previously reported (Patterson *et al.* 2013; Iwig *et al.* 2008). Co(II)-affinity of FrmR (41.9 μM) was analysed by competition with 9.8 μM Fura-2 (Figure 3. 12A) following quenching of Fura-2 fluorescence emission in the absence (open circles) and presence (filled circles) of protein. As the two sets of data points overlap, both showing an inflection after ~ 10 μM of Co(II), we can conclude that FrmR does not compete with Fura-2.

In contrast, FrmR (83.9 μM) showed competition with a large excess (50 mM) of Bis Tris (Figure 3. 12B). BisTris ($K_{Co(II)}$ = 2.26 x 10^{-2} M) was chosen because it weakly coordinates Co(II) therefore it is suitable for competition experiments with proteins which form weak complexes with this metal (Scheller *et al.* 1980, Xiao & Wedd 2010). Co(II) titration into FrmR was carried out as previously discussed (Figure 3. 4A) except for the presence of BisTris. Increase in absorbance of the deduced LMCT feature at 336 nm was monitored by UV-visible spectroscopy (Figure 3. 12B). The model used to fit the data by DynaFit describes binding of four Co(II) ions with equal affinity per tetramer. Fitted curve significantly diverges from simulated upper and lower limits curves (upper dashed line represents a ten-fold weaker K_{Co1-4} whereas lower dashed line represents a ten-fold tighter K_{Co1-4}) allowing determination of K_{Co1-4}^{FrmR} = 7.59 (± 0.4) x 10^{-6} M from triplicate assays (Table 8.3, Appendix).

3.5.3 *Measurement of the Cu(I)-binding affinity of FrmR by competition with BCS and BCA*

Cuprous binding affinity was examined by competition assays using bathocuproine disulfonate (BCS, β_2 = 6.01 × $10^{19.8}$ M^{-2}, Xiao & Wedd 2010) and bicinchoninic acid (BCA, β_2 = 1.58 ×

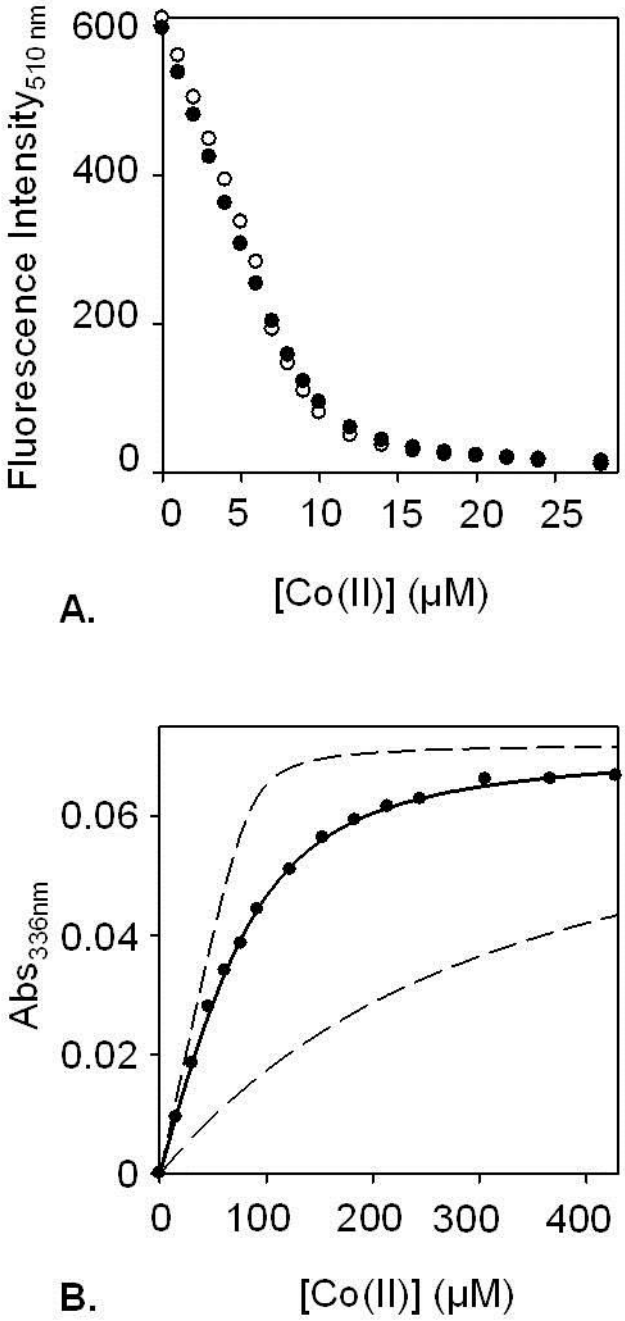


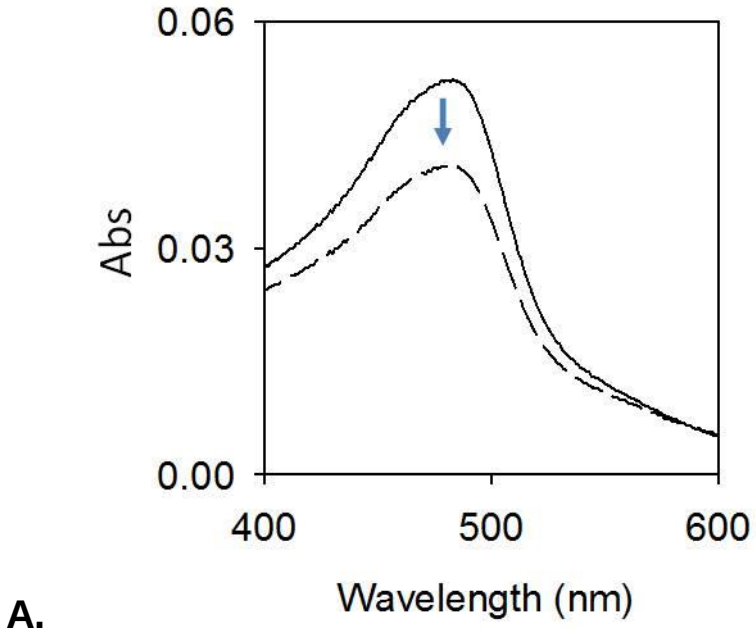
Figure 3. 12 A Representative (n = 3) of fura-2 (10.3 μM) titration with $CoCl_2$ in the presence of FrmR (41.9 μM). Fluorescence emission (λ_{ex} = 360 nm; λ_{em} = 510 nm) was measured by a fluorescence spectrophotometer. This experiment was performed by Dr. Deenah Osman. **B** Representative (n = 3) of FrmR (83.9 μM) titration with $CoCl_2$ in the presence of 50 mM BisTris. Variation of the absorbance at 336 nm (diagnostic of the LMCT feature) was monitored by a UV-vis spectrophotometer. The model (solid line) describes a FrmR monomer competing for 1 molar equivalents of Co(II) (4 sites per tetramer, $K_{Co1\text{-}4}$). Dashed lines represent simulated curves describing $K_{Co1\text{-}4}$ 10-fold tighter and 10-fold weaker.

$10^{17.2}$ M^{-2}, (Xiao & Wedd 2010) as competitor ligands (Figure 3. 13A-B and Figure 3. 14A-C). Figure 3. 13A represents BCS (7.8 µM) absorbance spectrum upon incubation with ~ 0.5 molar equivalent of Cu(I). Formation of a spectral feature at 483 nm is diagnostic of the formation of $[Cu(I)BCS]_2$ complex (solid line) and BCS can be quantified using the extinction coefficient (ε_{483} = 13500 M^{-1} cm^{-1}) for $[CuBCS_2]$ = 3.88 µM. FrmR (10.3 µM) was then added to the preformed complex (incubation time = 2 h), causing partial bleaching of the spectral feature (Figure 3. 13A). Calculating again $[CuBCS_2]$ and subtracting this value, 3.08 µM, from the one obtained in absence of FrmR, it is possible to determine the amount of cuprous ions associated with the protein (c = 0.8 µM). FrmR was capable of quenching ~ 21 % of the $CuBCS_2$ spectral feature. To further establish if BCS can be used as a probe to measure FrmR Cu(I) binding affinity, 25 µM BCS was titrated with Cu(I) in the absence (empty circles) and in the presence of FrmR (20.6 µM) (filled circles) as shown in Figure 3. 13B. An inflection at ~ 12.5 µM and ~ 14 µM was detected in the control and the competition experiments respectively (less sharp in latter case) revealing weak or absent competition by FrmR.

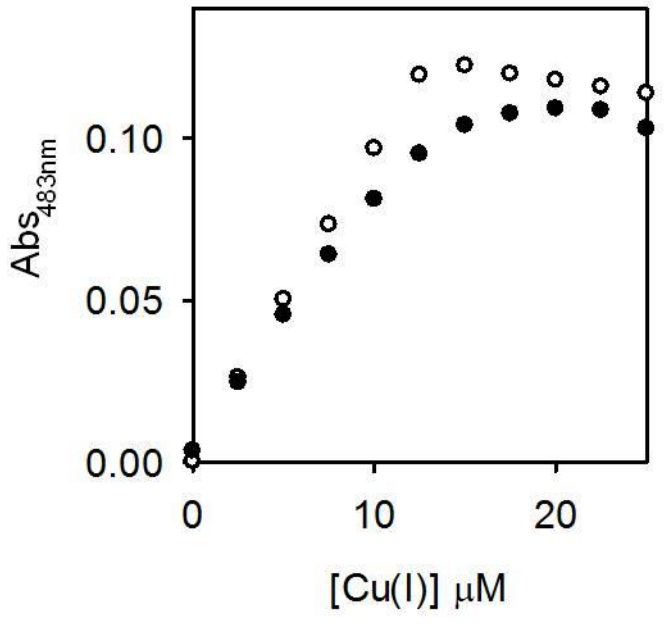
For this reason the cuprous affinity of FrmR was examined using BCA, which also forms a 2:1 complex with Cu(I) but possesses a weaker affinity for the metal. Increase of the absorbance at λ = 562 nm was monitored during titration of BCA (40 µM) with Cu(I) in presence of FrmR (10 µM). An inflection was observed at ~ 40 µM which reveals competition for 2 molar equivalents of Cu(I) per monomer ([BCA]/2 = 20 µM, [FrmR protomer] = 10 µM; 20 µM + 10 µM + 10 µM = 40 µM) (Figure 3. 14A-C). Data were fit, using Dynafit, to a model describing competition from FrmR for 2 molar equivalents of Cu(I) (8 sites per tetramer, with $K_{Cu1-2} < K_{Cu3-4} < K_{Cu5-8}$). Dashed lines represent simulated curves describing K_{Cu1-2} 10-fold tighter and 10-fold weaker than the optimised value with K_{Cu3-4} and K_{Cu5-8} fixed to their optimised values (Figure 3. 14A), K_{Cu3-4} 10-fold tighter and 10-fold weaker than the optimised value with K_{Cu1-2} and K_{Cu5-8} fixed to their optimised values (Figure 3. 14B) and K_{Cu5-8} 10-fold tighter and 10-fold weaker than the optimised value with K_{Cu1-2} and K_{Cu3-4} fixed to their optimised values (Figure 3. 14C). Fitted data depart from simulated curves describing K_{Cu1-2} or K_{Cu3-4} 10-fold tighter and 10-fold weaker, whereas K_{Cu5-8} is too weak to be determined by this assay, and only a lower limit can be inferred. Data from triplicate assays: K_{Cu1-4} = 4.90 ± 1.6 x 10^{-15} M, K_{Cu3-4} = 1.72 ± 0.7 x 10^{-12} M, $K_{Cu5-8} \geq$ 8.36 ± 1.9 x 10^{-11} M (Table 8.3, Appendix).

3.5.4 *Probing Ni(II)-binding affinity of FrmR by competition with mag-fura-2*

Ni(II)-binding affinity was investigated by competition with the chromophore mag-fura-2 which has $K_{Ni(II)}$ = 2 x 10^7 M^{-1} at pH = 7.0 (Reyes-Caballero *et al.* 2010). Titration of FrmR (10.04 µM) and mag-fura-2 (8.1 µM) with Ni(II) (filled circles, Figure 3. 15) gave a negligible difference in absorbance at 366 nm compared to mag-fura-2 alone titrated with Ni(II) (empty circles,



A.



B.

Figure 3. 13 A Representative (n = 3) titration of BCS (10μM) with Cu(I) (10μM) in the absence of FrmR showed a spectral feature at 483 nm, diagnostic of the formation of $CuBCS_2$ complex (solid line). Addition of FrmR (10μM) showed a decrease in absorption at 483 nm (dashed line) after an incubation of 2 h. **B** Titration of 25 μM BCS with Cu(I) in the absence of FrmR showed an inflection at around 12.5 μM consistent with the formation of $CuBCS_2$ complex (open circles). The same experiment carried out in presence of 20.6 μM FrmR showed a shift of 2-3 μM in the inflection suggesting a weak competition from the protein (filled circles).

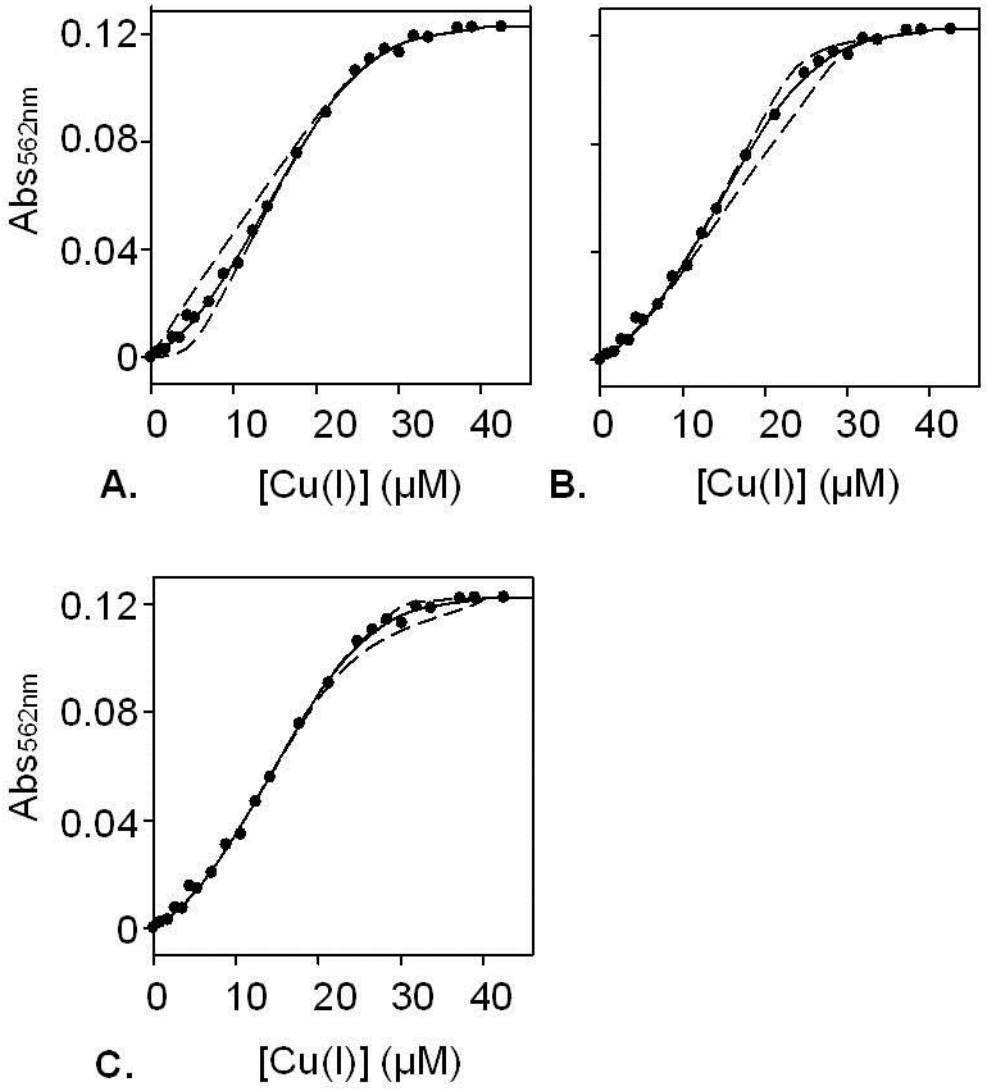


Figure 3. 14 A-C Representative (n = 3) BCA absorbance upon titration of BCA (40 μM) with Cu(I) in the presence of FrmR (10 μM). Addition of Cu(I) gave an increase in absorbance at 562 nm (diagnostic of the $BCA_2Cu(I)$ complex) with an inflection at ~ 40 μM total Cu(I). Solid line represents a fit to a model describing competition from FrmR for 2 molar equivalents of Cu(I) (see Appendix for details). Combination of $K_{Cu5\text{-}6}$ and $K_{Cu7\text{-}8}$ resulted in an improved fit. Each panel shows the fit to a model where each K_{Cu} is fixed to its optimised values except for one, for which simulated curves representing K_{Cu} 10-fold tighter and weaker than optimised value are shown. **A** $K_{Cu1\text{-}2}$ (dashed line_10-fold tighter and 10-fold weaker; **B** $K_{Cu3\text{-}4}$ (dashed lines_10-fold tighter and 10-fold weaker); **C** $K_{Cu5\text{-}8}$ (dashed lines_10-fold tighter and 10-fold weaker).

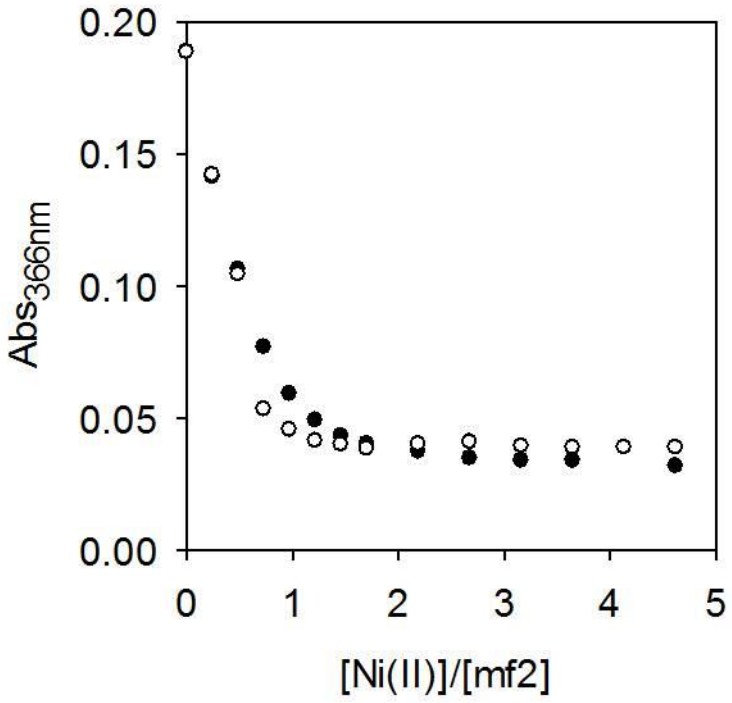


Figure 3. 15 Titration of 8.1 µM mag-fura-2 with Ni(II) in the absence of FrmR showed an inflection after the addition of 1 molar equivalent of nickel consistent with the formation of 1:1 complex (empty circles). The same experiment carried out in the presence of 10.04 µM FrmR showed a shift of ~2 µM in the inflection suggesting a weak competition from the protein.

Figure 3. 15), suggesting none or very little competition from FrmR. Thus it was not possible to determine FrmR' nickel-binding affinity with this assay and no further analyses were carried out to resolve $K_{Ni(II)}$ due to *in vivo* results discussed in a later section (Section 4.2.1).

3.6 Identification of FrmR's DNA binding site and characterisation of the DNA-protein interaction

3.6.1 *FrmR:DNA interactions*

The *frmR* gene is convergently co-transcribed with the *frmA* gene (Figure 3. 1A), which encodes the deduced class III alcohol dehydrogenase FrmA. As discussed in Chapter 1, *E. coli* FrmR binds the *frmRAB* operator region obstructing the site for RNA polymerase binding which cannot start transcription of the downstream genes.

However, in the presence of formaldehyde, FrmR releases the DNA allowing transcription of *frmR, frmA* and *frmB* (the latter has not been identified in *Salmonella*), implying a derepression role for FrmR upon formaldehyde-binding (Herring *et al.* 2004; Gonzalez *et al*. 2006).

E. coli FrmA is a formaldehyde dehydrogenase (class III alcohol dehydrogenase) which oxidizes the glutathione-formaldehyde adduct (S-hydroxymethylgluthatione), following the spontaneous reaction between formaldehyde and glutathione. The resulting product is S-formylglutathione which is then hydrolysed to formate and glutathione by FrmB (S-formylglutathione hydrolase) (Chapter 1) (Gonzalez *et al*. 2006; Stover *et al*. 2005; Moulis *et al.* 1991; Sanghani *et al.* 2002) *Salmonella* and *E. coli* FrmR homologues possess amino acid identity = 52.7 % and similarity = 74.7 % (determined by use of EMBOSS NEEDLE, see Table 3.1). *Salmonella* possesses *frmR* and *frmA* genes, suggesting a similar detoxification pathway to that adopted by *E. coli*. In order to investigate the FrmR interaction with the *frmRA* operator promoter region, a DNA fragment incorporating the region upstream of *frmRA* was produced. Within this fragment is an ATAGTATAC$_6$TATAGTAT sequence with a high degree of similarity to characterized promoter elements recognised by other RcnR/CsoR family members (i.e. *E. coli* RcnR, CsoR homologues, *Synechocystis* InrS) (Figure 3. 1B). The region upstream of *Salmonella frmR* includes a conserved palindromic sequence (ATA-X-TATA-C$_6$-TATA-X-TAT) (see Figure 3. 1B) which corresponds to a type 1 site according to Iwig & Chivers (2009) classification. These sites consist of a single G/C tract (3-8 bp) flanked by inverted repeats rich in A/T. DNA binding by FrmR is explored in the next paragraph by fluorescence anisotropy (FA).

3.6.2 *Fluorescence Anisotropy (FA) principles*

Fluorescence anisotropy has been used to study protein:DNA interactions, and in particular those involving metalloregulators, in a more quantitative way than EMSA (VanZile *et al.*

2002a; Harvie *et al.* 2006; Andoy *et al.* 2009; Humbert *et al.* 2013; Chang *et al.* 2014). A modified Cary Eclipse fluorescence spectrophotometer (Agilent Technologies) fitted with polarising filters (λ_{ex} = 530 nm, λ_{em} = 570 nm, averaging time = 20 s, replicates = 5, T = 25 °C) was used to measure the difference in the tumbling rates during the protein-DNA interaction. Since the molecular size of the complex is sufficiently different from free DNA a DNA binding curve can be obtained. In this assay we exploited the extrinsic fluorophore Hexachlorofluorescein (HEX) (ε_{535} = 96,000 $M^{-1}cm^{-1}$, as stated by the supplier) bound to the 33 bp DNA fragment in order to monitor the signal between the free and bound DNA states. The use of an extrinsic fluorophore with a high absorptivity allows measurements in the nanomolar concentration range. When the fluorophore attached on the 5' end of DNA is excited with plane polarised light, it will emit a light that is also polarised. Depolarisation of the emitted light will occur through rotational diffusion of the fluorophore and if the rate of tumbling is slower than the rate of emission, the emitted light will stay polarised causing intrinsic anisotropy r_o. Upon protein binding to DNA the tumbling will be much slower because of the bigger size of the complex and the anisotropy will increase ($\Delta r_{obs} = (r_{obs} - r_o) > 0$).

The anisotropy value is calculated via Equation 3:

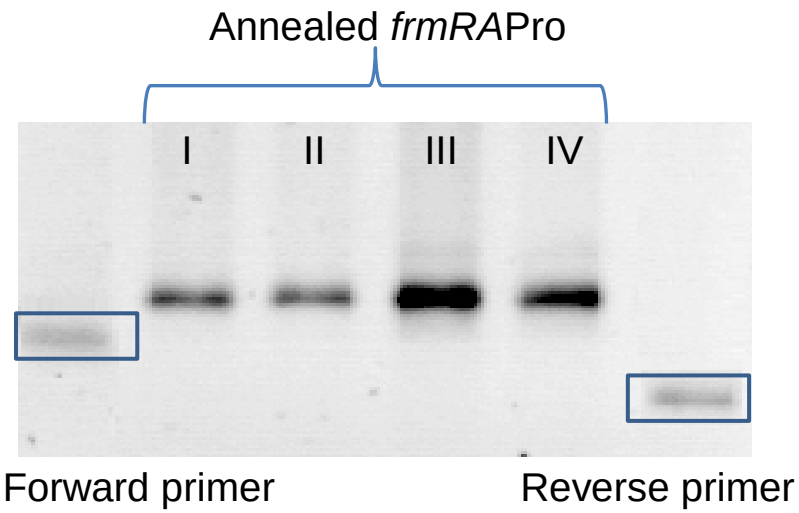
$$r_{obs} = \frac{(I_{VV} - I_{VH})}{(I_{VV} + 2I_{VH})}$$

Equation 3

where I_{VV} and I_{VH} are the fluorescence intensity parallel (measured in the vertical plane) and perpendicular (measured in the horizontal plane) to the excitation polarization (which is made with vertically polarised light) (Andoy *et al*. 2009). Fitting the anisotropy values using a model describing the stoichiometry of the event will yield the association constant of the interaction (Grossoehme *et al.* 2012).

3.6.3 *FA oligonucleotides production*

Complementary oligonucleotides (33 bp) incorporating the predicted FrmR binding site (Figure 3. 16) and 5/6 flanking nucleotides either side of the motif were obtained commercially. One of the pair of primers was labelled at its 5' end with the fluorophore Hexachlorofluorescein (HEX) (Figure 3. 16). The native PAGE shown in Figure 3. 16 confirms successful annealing of the two oligonucleotides (described in Chapter 2) to produce the double stranded sample hereafter called *frmRA*Pro (Chapter 2, Table 2. 1). As expected the unlabelled oligonucleotide (referred in the figure as "reverse primer") migrated through the gel matrix slightly faster than the HEX-labelled ("forward primer") because of the bigger size of the latter. Likewise the double-stranded *frmRA*Pro samples were observed to migrate slower than single stranded primers.



5'-[HEX]TTCTG**ATAGTATACCCCCCTATAGTAT**ATGGAG

Figure 3. 16 Production of a double stranded, 5'-[HEX]-labelled, oligonucleotide containing the FrmR recognition site for use in fluorescence anisotropy assays. DNA sequence (33 nucleotides) composed of the identified FrmR recognition site and flanking nucleotides. The HEX-labelled primer was annealed with its unlabelled reverse complement. The native PAGE shown here confirms the successful annealing of the two oligonucleotides to produce the double stranded *frmRA*Pro. Sample concentrations: 10 µM (lanes 1-2), 200 µM (lanes 3-4).

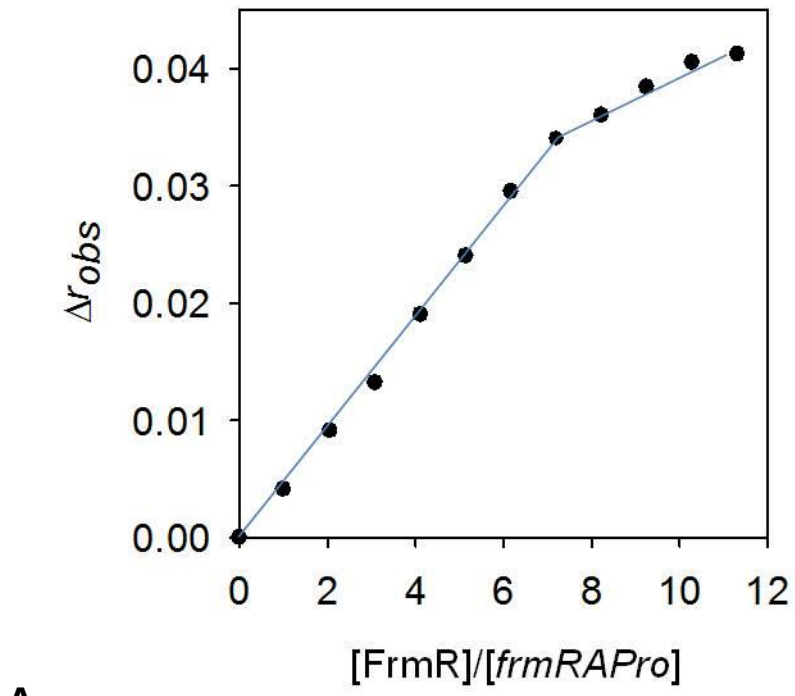
3.6.4 *FrmR:DNA stoichiometry studied by FA*

Stoichiometry of FrmR binding to *frmRA*Pro was carried out anaerobically in buffer containing 60 mM NaCl, 240 mM KCl, 10 mM Hepes pH 7.0, 1 mM DTT and 5 mM EDTA to prevent metal contamination. The 1:4 ratio between NaCl and KCl was intended to emulate Na^+/K^+ inside the cell. Fluorescence anisotropy was used to monitor binding of FrmR to the 33 bp DNA (2.5 µM) under these conditions (Figure 3. 17A). The change in anisotropy (r_{obs}) after each addition of FrmR (one molar equivalent steps) is linear and saturates at about 8 molar equivalents of FrmR (20 µM), Figure 3. 17B shows a hypothetical model describing one FrmR tetramer binding to each side of the DNA recognition site, as will be further discussed in Section 7.3.1, suggesting a stoichiometry of two FrmR tetramers binding to one DNA molecule, since free FrmR is a stable tetramer in solution (Section 3.3). To further study the role of T/A sites in DNA recognition, a new fluorescently labelled double stranded oligonucleotide (named *frmRA*Pro-T/A-mod) was produced in a similar way to *frmRA*Pro. However in *frmRA*Pro-T/A-mod, the first T/A site was randomly modified (ATAGTATA → GTTCAACA) (Figure 3. 18B). Stoichiometry of FrmR binding to *frmRA*Pro-T/A-mod was performed under identical conditions to those just described for *frmRA*Pro, but the point of inflection was observed after addition of 4 equivalents of protein monomer suggesting one tetramer binding to *frmRA*Pro-T/A-mod. This finding may indicate that binding of one of the two protein tetramers was abolished. However, the inflection after four molar equivalents of protein is not fully conclusive and the overall response (Δr_{obs} at the inflection point) is very low compared to what was obtained for the unmodified DNA recognition site, therefore it remains unclear if each FrmR tetramer needs to interact simultaneously with both T/A sites.

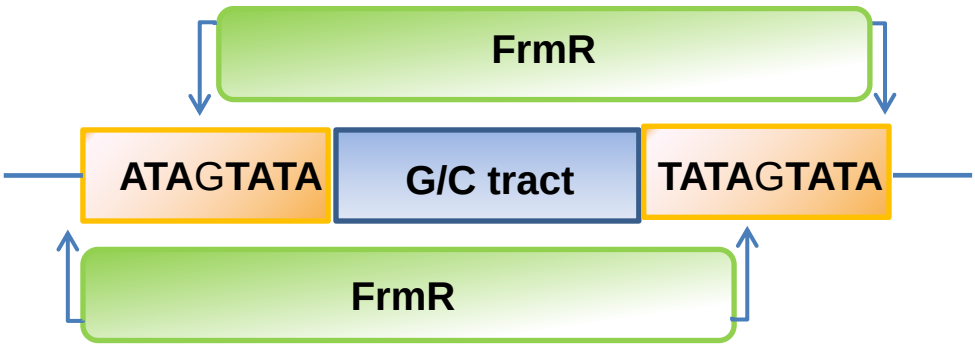
3.6.5 *FrmR:DNA interaction studied by FA*

In order to measure the affinity of FrmR for its promoter operator region, the protein was directly titrated into a solution of a limiting concentration of *frmRA*Pro probe (10 nM) under anaerobic conditions. Buffer conditions were the same as those used in Section 3.6.4. Changes in anisotropy (Δr_{obs}) were measured as described previously (Figure 3. 19A). Upon each addition, the cuvette was allowed to equilibrate at 25 °C for 5 min before recording data. Data were fit to the model describing two non-dissociable tetramers binding to *frmRA*Pro with equal affinity (n = 3) using Dynafit.

The protein:DNA 2:1 stoichiometry has been determined in Section 3.6.4. As commonly found in the literature the DNA-binding constant value refers to models using monomer protein concentrations. K_{DNA} was found to be 99.4 ± 30 nM, which is within the range of DNA affinities available in literature for metalloregulators (~30 - 200 nM) (Ma *et al.* 2009a; Iwig & Chivers 2009; Foster *et al.* 2012) and confirms the binding of FrmR on its own promoter operator region. The same experiment was performed using *frmRA*Pro-T/A-mod (Figure 3. 19B)

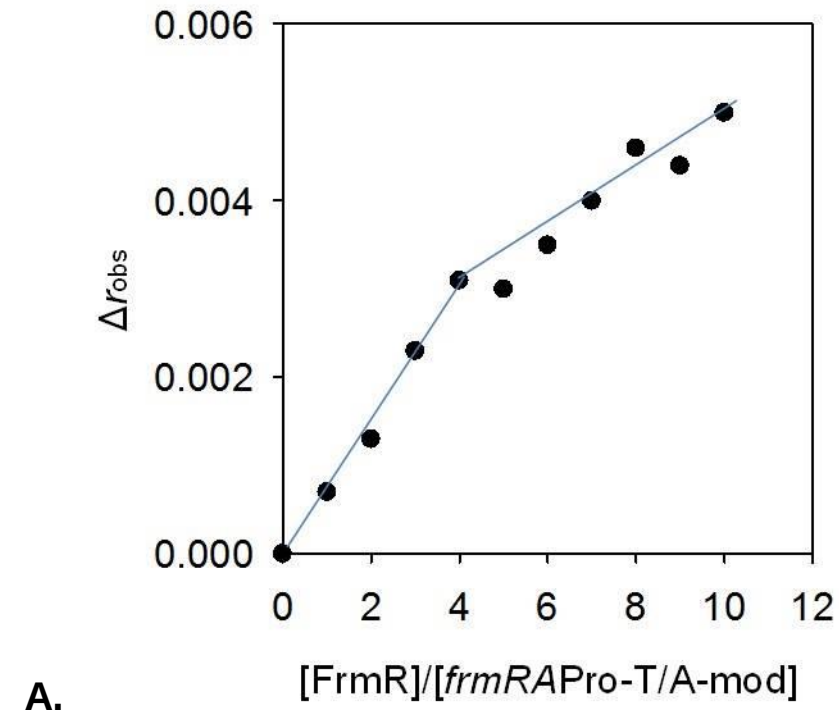


A.



B.

Figure 3. 17 DNA binding stoichiometry (n=1) of FrmR. **A** Anisotropy change upon titration of *frmRA*Pro (2.5µM) with FrmR (2.5 µM monomer with each addition). A point of inflection is observed after addition of ~ 8 equivalents of protein monomer indicating a protein:DNA stoichiometry of 8:1 (two tetramers per *frmRA*Pro). The further increase in anisotropy following this inflection may represent further binding of FrmR to *frmRA*Pro with substantially weaker affinity. Another hypothesis may be that the effect of tumbling changes above the 8:1 stoichiometry. **B** Proposed schematic of FrmR interacting with its operator with the region of the G/C tract in blue and the two T/A sites in yellow. Results shown in panel A suggest a model where two FrmR tetramers bind to this sequence.



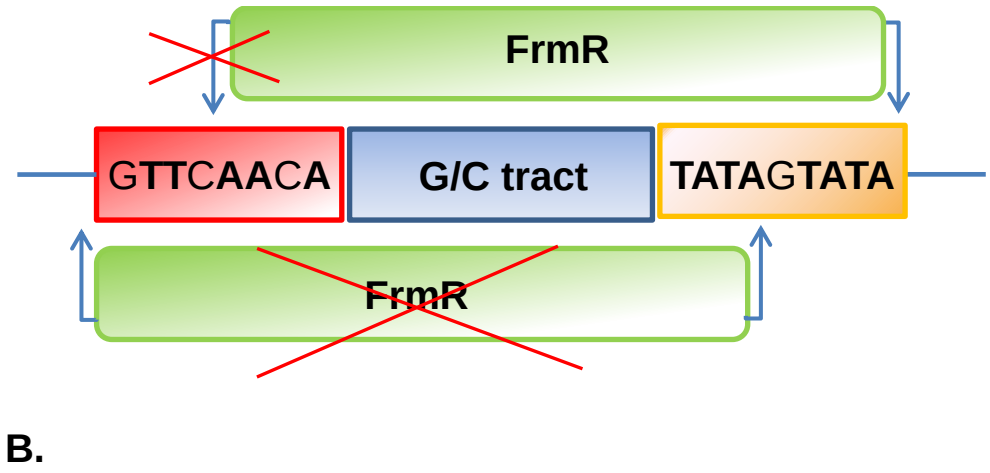
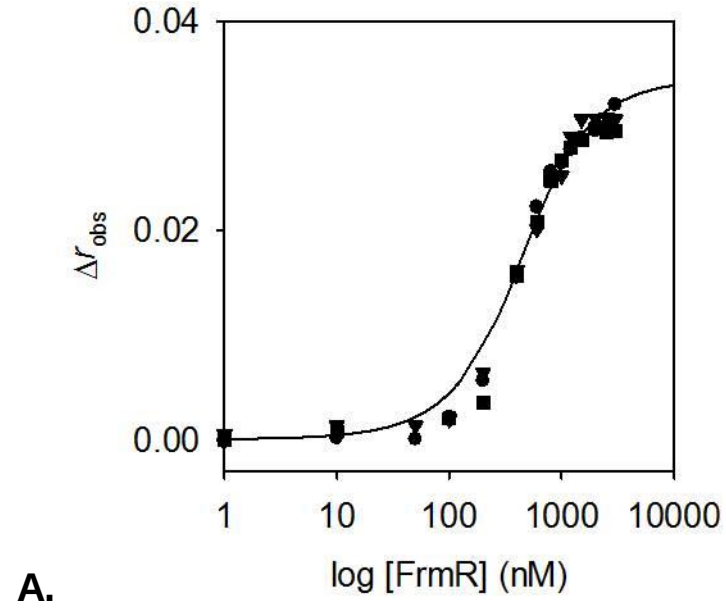


Figure 3. 18 DNA binding stoichiometry (n=1) of FrmR using operator sequence with first T/A site modified (herein called *frmRA*Pro-T/A-mod). **A** Anisotropy change upon titration of *frmRA*Pro-T/A-mod (2.5µM) with FrmR (2.5 µM monomer with each addition). *frmRA*Pro-T/A-mod contains the same promoter region as *frmRA*Pro except for the introduction of a random nucleotide sequence in place of the first T/A site). A point of inflection is observed after addition of ~ 4 equivalents of protein monomer indicating a protein:DNA stoichiometry of 4:1 (one tetramer per *frmRA*Pro-T/A-mod). The overall response is very low so it remains unclear if FrmR interacts simultaneously with both halves of the inverted repeat of the recognition sequence. **B** Proposed schematic of FrmR interacting with its operator with the region of the G/C tract in blue, the first T/A site (modified) in red and the second T/A site (unmodified) in yellow. Only one FrmR tetramer would bind to this sequence with an extremely weakened K_{DNA}.



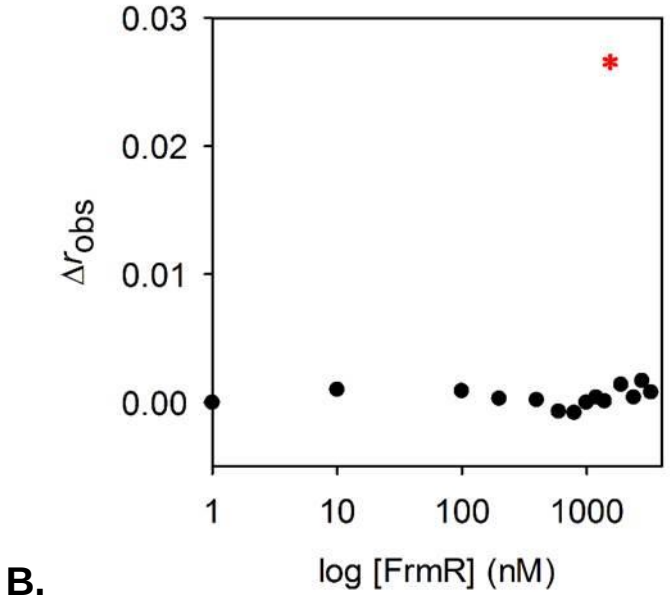


Figure 3. 19 Titration of *frmRA*Pro with apoFrmR. **A** *frmRA*Pro (10 nM) was anaerobically titrated with FrmR in the presence of 5 mM EDTA. DNA binding was monitored by fluorescence anisotropy. Solid line represents simulated curves produced from the average K_{DNA} determined across the experimental replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 FrmR tetramer (non-dissociable):DNA stoichiometry. **B** As A except for the use of *frmRA*Pro-T/A-mod.

showing no binding (the titration was carried out up to [FrmR] = 4000 nM), reinforcing the hypothesis suggested in Section 3.6.4, according to which FrmR tetramers need to interact simultaneously with both T/A sites.

3.6.6 *FrmR:DNA interaction in the presence of Zn(II) and Cu(I)*

In vitro experiments presented in Section 3.4 showed that FrmR is able to bind Zn(II), Cu(I), Co(II) and Ni(II) (the first two with high affinity and the last two with a significant weaker affinity, albeit the binding of Ni(II) was not further investigated). In order to test if these metals would be able to trigger the allosteric mechanism leading to a conformational change in FrmR and the subsequent disruption of protein:DNA complex, FA association analyses have been employed in the presence of metals. Figure 3. 20A shows titration of *frmRA*Pro (10 nM) with a sample of FrmR incubated with a zinc salt in 20 % excess (1.2 molar equivalents $ZnCl_2$). Given FrmR Zn(II) affinity for the tightest sites is $K_{Zn(II)1-3} = 1.7 \times 10^{-10}$ M, 5 µM $ZnCl_2$ was also added to the reaction buffer which should be well in excess of the concentration required in order to maintain a metalled form of the protein throughout the experiment. A notable shift in the FrmR association curve reveals a significantly weakened binding to DNA, which implies that Zn(II) acts as an effector *in vitro* causing conformational change in FrmR which, in the metallated form, releases its operator region. A similar result, although less prominent, was obtained with CuCl (verified to be ≥ 95 % Cu(I)) (Figure 3. 20B). In this experiment no additional cuprous ions were added to the buffer since FrmR possesses an even tighter affinity for the metal ($K_{Cu(I)1-2} = 4.90 \pm 1.6 \times 10^{-15}$ M $K_{Cu(I)3-4} = 1.72 \pm 0.7 \times 10^{-12}$ M; $K_{Cu(I)5-8} \geq 8 \times 10^{-11}$ M). Data from both experiments were fitted to the same model used for apo-association (Section 3.6.2) and $K_{DNA}^{Zn(II)FrmR}$ and $K_{DNA}^{Cu(I)FrmR}$ were calculated to be 3110 ± 400 nM (Figure 3. 20A) and 654 ± 130 nM (Figure 3. 20B) respectively.

For these experiments where DNA binding did not saturate, the average fitted Δr_{obs} maximum value from apo-protein experiments was used in the script as found in literature (Reyes-Caballero *et al*. 2011; Foster *et al.* 2014). Figure 3. 20C displays a representative data set from experiments discussed in this section in comparison with those for the apo-FrmR:DNA association (Section 3.6.5, Figure 3. 19A) in order to better appreciate the effect these metals have on DNA binding. The scheme in Figure 3. 21A was presented by Grossoehme and Giedroc in 2009 and represents the thermodynamic cycle describing the allostery response that a metalloregulatory protein for its DNA operator region will encounter upon metal binding (see Chapter 1). The protein, here FrmR, is in its tetrameric state (P_4) and the assumption that monomeric FrmR does not have affinity for DNA is made (Giedroc & Arunkumar 2007; Grossoehme & Giedroc 2009, 2012).

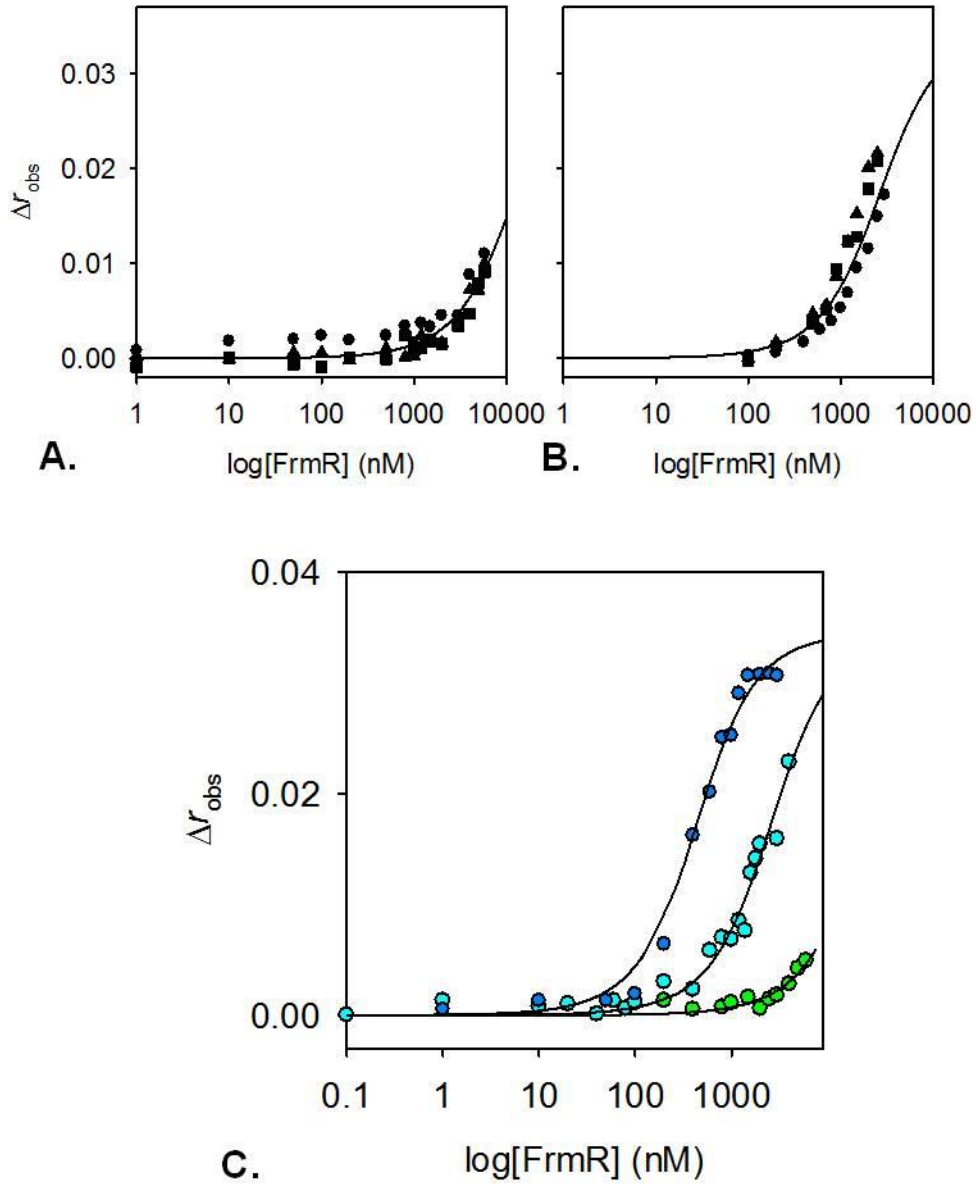


Figure 3. 20 A *frmRA*Pro (10 nM) was anaerobically titrated with FrmR in the presence of 5 µM $ZnCl_2$. The protein was incubated with 1.2 molar equivalents of $ZnCl_2$ and EDTA was omitted. **B** *frmRA*Pro (10 nM) was anaerobically titrated with FrmR incubated with 1.2 molar equivalents of CuCl (verified to be > 95% Cu(I)) and EDTA was omitted. **C** Comparison of the anisotropy change upon titration of *frmRA*Pro with apo-FrmR (blue circles), Cu(I)-FrmR (pale blue circles) and Zn(II)-FrmR (green circles). DNA binding was monitored by fluorescence anisotropy. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 FrmR tetramer (non-dissociable):DNA stoichiometry. Solid line represents simulated curves produced from the average K_{DNA} determined across the experiment replicates shown.

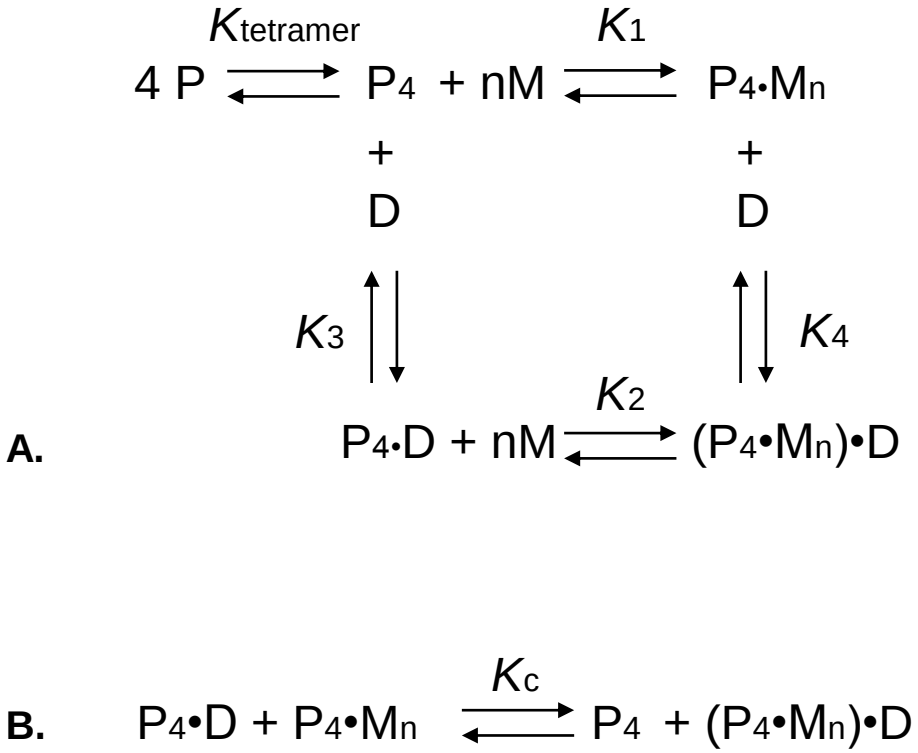


Figure 3. 21 Allosteric coupling scheme as described by Grossoehme and Giedroc (2012) applied to homotetrameric FrmR A Generalized thermodynamic cycle accounting for the four allosteric "end" states of a homotetrameric metalloregulatory protein (P_4) can hypothetically adopt: apo (P_4), metal-bound (P_4·Mn), DNA-bound apoprotein (P_4·D), and a "ternary" protein–metal–DNA complex ((P_4·Mn)·D). Each equilibrium (K_1, K_2, K_3, K_4) describes a direct transition from one configurational state to another as shown. Note also that P_4 and thus the entire scheme is in equilibrium with free P monomer, defined by $K_{tetramer}$, which has no affinity for the DNA. B Ligand exchange equilibrium, defined by the adimensional parameter *K*c, dictates the degree of allostery between the metal binding and DNA binding sites.

The values obtained from the experiments described in the current and in the previous section coincide with the constants in the vertical equilibria in Figure 3. 21A (K_{DNA}^{FrmR} is K_3 and $K_{DNA}^{metal\text{-}FrmR}$ is K_4). The ligand exchange equilibrium constant, K_c, represents the magnitude of allosteric regulation and is expressed by the ratios K_4/K_3 and K_2/K_1 (Grossoehme & Giedroc 2009, 2012). The first ratio corresponds to the difference in metal affinity between the protein bound to DNA and in its free form, while the second can be thought of as the difference in DNA binding affinity between the metallated and the apo-protein (Grossoehme & Giedroc 2012). Having measured K_3 and K_4 by fluorescence anisotropy, we can now determine the coupling equilibrium constant K_c and therefore obtain the allosteric coupling free energy ΔG_c via the standard thermodynamic function described in Section 1.3.2 (Equation 1). Mean ΔG_c values (and standard deviations) were calculated from the full set of (equally weighted) possible pair-wise permutations of K_c. This yields $\Delta G_c^{Zn(II)FrmR}$ = + 2.03 (± 0.08) kcal mol^{-1} and $\Delta G_c^{Cu(I)\text{-}FrmR\cdot DNA}$ = + 1.10 (± 0.10) kcal mol^{-1} (Table 8. 2, Appendix).

This approach revealed that Zn(II) is more allosterically effective when binding to FrmR, because $\Delta G_c^{Cu(I)\text{-}FrmR\cdot DNA}$ is smaller. In fact for repressors in which metal binding induces dissociation of the repressor of the DNA operator, the tertiary state (P_4·Mn)·D is destabilized compared to D and P_4·Mn therefore the difference in free energy associated with the ligand exchange reaction (Figure 3. 21B) will be positive and the reaction will not proceed spontaneously (Grossoehme & Giedroc 2012). Since in the presence of metal the promoter will not be occluded anymore, access for RNA polymerase will be possible.

The ability of Zn(II) and Cu(I) to disrupt the interaction between FrmR and its operator was also tested by titrating the preformed protein:DNA complex with metal aliquots. FrmR (2.5 µM, chosen because FA apo-FrmR curves showed complete associations at this [protein]) was incubated with 10 nM *frmRA*Pro for 30 min before titration with $ZnCl_2$ or CuCl. Figure 3. 22A shows that dissociation of the protein:DNA complex occurs upon addition of 0.25 - 0.5 molar equivalents of Zn(II) suggesting that the allosteric mechanism allowing FrmR to release the DNA is mainly trigged during the filling of the second site of the tetrameric protein. Titration of *frmRA*Pro with Cu(I) results in dissociation after 2 molar equivalents of metal (Figure 3. 22B), consistent with what is described in Section 3.5.1 where competition of eight atoms of Cu(I) per tetramer (2:1 stoichiometry) was observed. Thus FrmR binds specifically to *frmRA* operator-promoter as an inverse function of [Zn(II)] or [Cu(I)]. The weak affinities of FrmR for Co(II) ($K_{Co(II)1\text{-}4}$ = 7.59 ± 0.4 x 10^{-6} M) and Ni(II) (as suggested by competition experiment with mag-fura-2, Figure 3. 15) prevented us to carry out dissociation experiments with these metals.

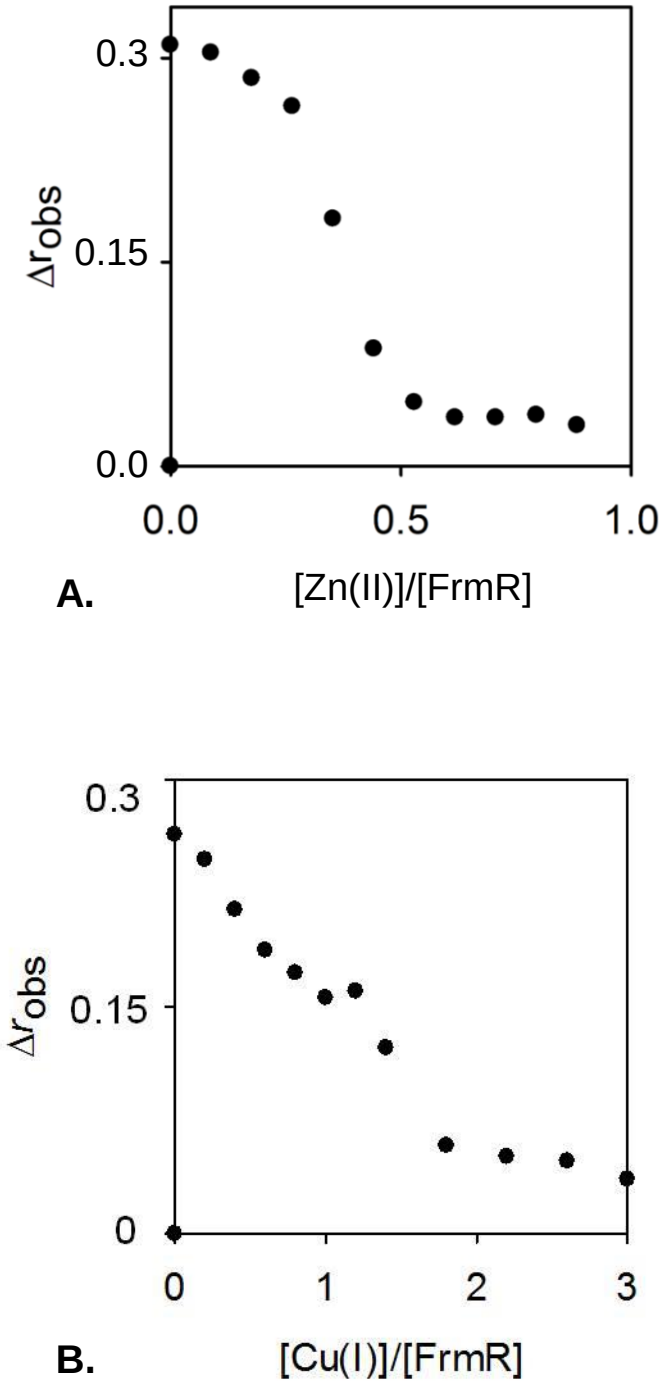


Figure 3. 22 Titration of pre-formed FrmR:DNA complexes with $ZnCl_2$ and CuCl. *frmRA*Pro (10 nM) was pre-incubated with FrmR (2.5 µM) before titration with **A** $ZnCl_2$ and **B** CuCl. Dissociation of protein:DNA complexes was monitored by fluorescence anisotropy. Experiments were performed anaerobically at pH 7.0.

3.7 Identification of residues required for DNA-binding

In addition to the W-X-Y-Z residues, already identified as involved in the first coordination sphere (see Section 1.2.1), two additional residues have been identified as contributing toward allostery. *Synechocystis* InrS possesses two glutamate residues in position C and B (Glu95 and Glu98, respectively) (Figure 3. 23B). The presence of a glutamate in position C is a common feature in multiple family members that lack the secondary shell hydrogen bond network found in CsoR homologues (Foster *et al*. 2014). *M. tuberculosis* and *B. subtilis* CsoRs also have a Glu in position B which is important for the regulation of DNA binding (Ma *et al*. 2009; Liu *et al*. 2007). In common with CsoR *Synechocystis*, InrS conserves H78 (position Y) and E98 (position B) but lacks the tyrosine residue in position A which is substituted with a Pro (Foster *et al*. 2014). E98 has been shown to propagate the allosteric response upon metal binding, while the abovementioned Tyr (position A) seems to modulate the magnitude of the response (Ma *et al*. 2009; Corbett *et al*. 2011). In InrS, in the absence of nickel, α3 and α2 helices are kept parallel allowing the protein backbone to interact with DNA (Foster *et al*. 2014). α3 helix is in fact rich of residues negatively charged and a hypothesis suggests that upon metal binding, the helix moves toward the metal binding site, approaching the also negatively charged phosphate groups on DNA resulting in electrostatic repulsion. This conformational change would cause the protein to release DNA. A hydrogen bond between the two helices involves E98 and H78 (position B and Y respectively) and was tested in order to comprehend its contribution to allostery (Foster *et al*. 2014). Fluorescence anisotropy analyses carried out with E98AInrS (mutant lacking the strong electronegative oxygen in position B), displays how the hydrogen bond between H78 and E98 is not absolutely required for allostery although it does have a small effect (Foster *et al*. 2014). This may be due to the presence of an extra connection between the two helices, which is the salt bridge between R85 and E95 (Foster *et al*. 2014). Since mutation of E95→A showed a smaller allostery coupling free energy compared to wild-type, the residue in position C seems to also contribute towards coupling metal binding and DNA binding in InrS (Foster *et al*. 2014).

Salmonella FrmR possesses a similar set of residues in these positions (Figure 3. 23A) although the Glu-Arg salt bridge that further stabilizes the two helices in InrS (E95-R85) is not present in FrmR.

An analogous pair of Glu-Arg can be identified by examining FrmR sequence (E64 on helix α2 and R78 on helix α3) but the two residues are presumably located too far apart to allow such a weakly interacting salt bridge to form. This assumption needs to be validated by investigation of the X-ray structure of apo-E64HFrmR produced during this work (See Chapter 6). However, at the time of writing, the processing of helix α3 electron density is still ongoing. The only apparent connection between the two helices is therefore the hydrogen bond between H60 and E81 (FrmR notation, position Y and B respectively) which is equivalent to that between H78

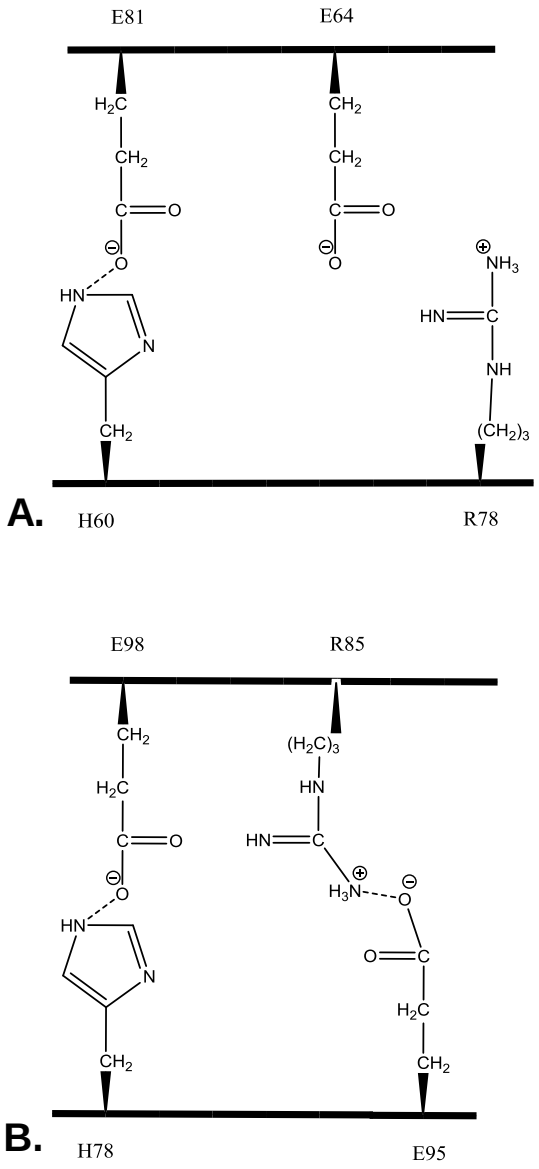


Figure 3. 23 Schematic representation of the interactions between α2 and α3 helices which are theorized to keep the two helices parallel in **A** FrmR and **B** InrS.

and E98 (InrS notation, position Y and B respectively) found in InrS (Figure 3. 23B). If this is true, mutation of one of these residues should result in weakened DNA binding affinity. To test this hypothesis H60LFrmR and E81AFrmR mutants were produced and DNA binding affinities measured by fluorescence anisotropy.

3.7.1 *Production and purification of recombinant H60LFrmR and E81AFrmR*

H60LFrmR and E81AFrmR were produced by site-directed mutagenesis and purified as previously described for wild-type FrmR (Section 2.4.6, Figure 3. 24A-B), and protein samples were transferred inside the anaerobic glove box and subsequently reduced status of cysteine residue and metal contamination assays were performed.

3.7.2 *H60LFrmR and E81AFrmR DNA binding properties studied by FA*

Fluorescence anisotropy analyses were performed on apoH60LFrmR and apoE81AFrmR mutants in an analogous way to that described for wild-type FrmR in Section 3.6.5. As we conjectured, the substitution of residues supposedly vital for DNA binding with amino acid residues unable to form a hydrogen bond, led to a weakened interaction with the operator (Figure 3. 25A-B).

This may be explained as the elimination of the only identified connection between α2 and α3 resulting in the two helices being no longer parallel leading to a conformation that is not optimized for high affinity DNA binding. Figure 3. 25C shows the H60LFrmR and E81AFrmR association curves in comparison with wild-type FrmR to emphasize how shifted the mutants curves appear. Data were fit to the same model used for FrmR, describing two non-dissociable protein tetramers associating with *frmRA*Pro, and K_{DNA} were calculated.

H60LFrmR has DNA affinity weakened over two-fold compared to E81AFrmR (K_{DNA} = 1918 nM vs 870.5 nM) and this could be explained hypothesizing that other residues on α3 helix can interact with His60 (α2) in E81AFrmR, partially stabilizing the DNA binding protein conformation. Experiments must be replicated in order to assign a standard deviation. Further experiments must be conducted in order to truly understand the mechanism connecting metal binding to impaired DNA-binding.

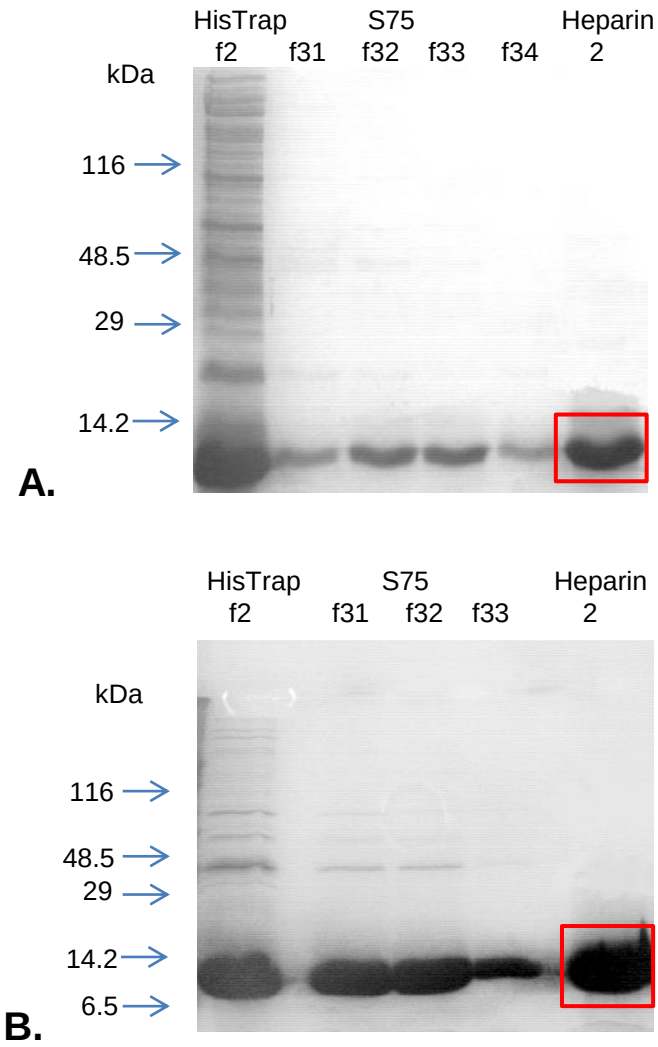


Figure 3. 24 Purification of H60LFrmR and E81AFrmR by nickel affinity chromatography, size exclusion and heparin affinity chromatography. **A** SDS-PAGE analysis (18 % w/v acryl-bis) showing fractions containing H60LFrmR from each purification step. "HisTrap f2" refers to the second fraction eluted from a 5 ml Ni(II)-affinity column with buffer A containing 0.3 M imidazole. Fractions "S75 f31-f34" were eluted from a Superdex S75 column loaded with fraction "HisTrap f2". Fraction "Heparin 2" was eluted at 0.5 M NaCl from two 1 ml Heparin columns linked together loaded with size exclusion fractions f32-f33 combined. Presence of only one band at around 10 kDa confirms the purity of the protein. **B** SDS-PAGE analysis (18 % w/v acryl-bis) showing fractions containing E81AFrmR from each purification step. "HisTrap f2" refers to the second fraction eluted from a 5 ml Ni(II)-affinity column with buffer A containing 0.3 M imidazole. Fractions "S75 f31-f33" were eluted from a Superdex S75 column loaded with fraction "HisTrap f2". Fraction "Heparin 2" was eluted at 0.5 M NaCl from two 1 ml Heparin columns linked together loaded with size exclusion fractions f31-f32 combined. The presence of only one band at around 10 kDa confirms the purity of the protein.

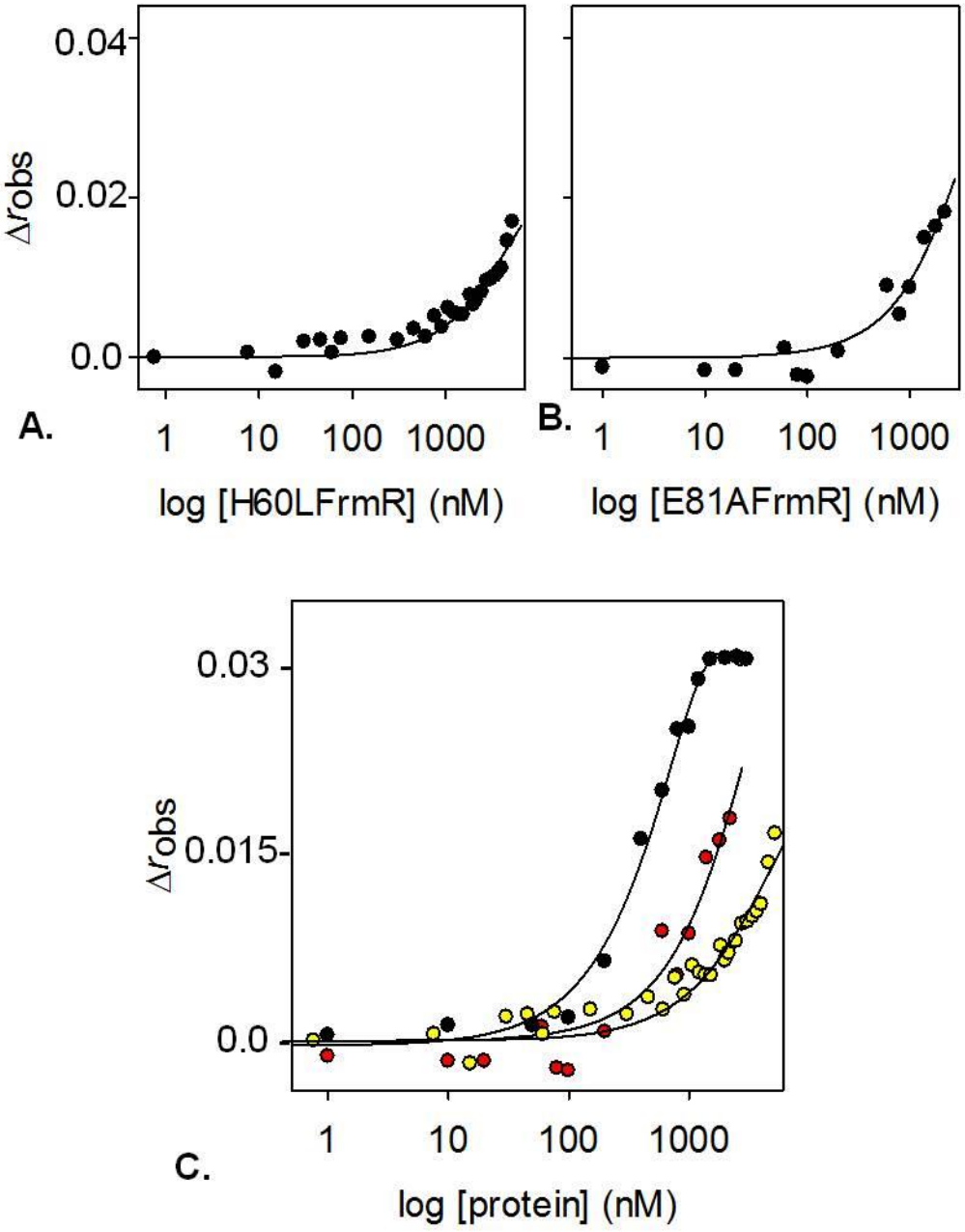


Figure 3. 25 Titration of *frmRA*Pro with H60LFrmR and E81AFrmR. **A** *frmRA*Pro (10 nM) was anaerobically titrated with H60LFrmR in the presence of 5 mM EDTA. **B** Same as A but using E81AFrmR to titrate *frmRA*Pro. **C** Comparison of the anisotropy change upon titration of *frmRA*Pro with wtFrmR (black circles), E81AFrmR (red circles) and H60LFrmR (yellow circles). DNA binding was monitored by fluorescence anisotropy. Data were fit to a model describing a 2:1 protein tetramer (non-dissociable):DNA stoichiometry. Solid line represents simulated curves produced from the K_{DNA} determined for the experiments shown.

3.8 Identification of residues required for metal-binding

As previously discussed in Section 1.2.1, all RcnR/CsoR family members possess a signature fingerprint, W-X-Y-Z, which contains four ligands consistent with metal-sensing (Iwig *et al.* 2008; Ma *et al.* 2009b). As presented in the sequence alignment in Figure 3. 2 the fingerprint in CsoR is x-C-H-C (x is any amino acid) and in RcnR is H-C-H-H. *Salmonella* FrmR contains three of four ligands consistent with binding of Ni(II)/Co(II) in RcnR, in positions W-X-Y. In addition, CsoR conserves second coordination sphere residues, A and B, which propagate the allosteric response occurring upon cognate metal binding (Ma *et al.* 2009b; Liu *et al.* 2007). RcnR does not possess the A-B motif suggesting a different allosteric mechanism relative to CsoR (Iwig *et al.* 2008). *Salmonella* FrmR conserves the position B in the extended fingerprint motif but, as InrS, has a Pro instead of Tyr in position A. During the course of this work, metal binding properties of a number of FrmR single point mutants (C35→A, H60→L, E64→H), each lacking one of the residue conserved in the RcnR/CsoR X-Y-Z fingerprint, have been tested in order to identify the residues contributing to metal binding. H60LFrmR production has already been described in section 3.8 and E64HFrmR production and metal binding properties will be extensively explored in Chapter 4 therefore the next paragraph will describe only production of C35AFrmR variant and metal binding properties of C35AFrmR and H60LFrmR.

3.8.1 *Production and purification of recombinant C35AFrmR*

C35AFrmR was produced by site-directed mutagenesis and purified as previously described for wild-type FrmR (Section 3) (Figure 3. 26). Since C35 is the only cysteine residue, work with this mutant was carried out under aerobic conditions. Metal contamination assay was performed as routine.

3.8.2 *Zn(II) binding properties of C35AFrmR investigated by size exclusion chromatography*

By using a combination of XAS, site-directed mutagenesis and EXAFS analyses it has been shown that Cys35 (*Salmonella* FrmR nomenclature, position X in the fingerprint motif) is a metal ligand in RcnR/CsoR family members (Iwig *et al.* 2008; Higgins *et al.* 2013; Chaplin *et al.* 2015). This residue is highly conserved as reported by Iwig and colleagues (2008) where alignment of 48 predicted Ni/Co-dependent regulators (retrieved using a BLAST search with *E.coli* RcnR) resulted in Cys35 having an identity ≥ 75 % (Iwig *et al.* 2008). Here we aim to study if this residue plays a vital role in metal (zinc) binding as much as found in other family members in cognate metal binding. As shown in Figure 3. 27A, Zn(II) does not migrate with C35AFrmR when the protein was incubated with an excess of zinc, applied and resolved on a Sephadex G25 matrix.

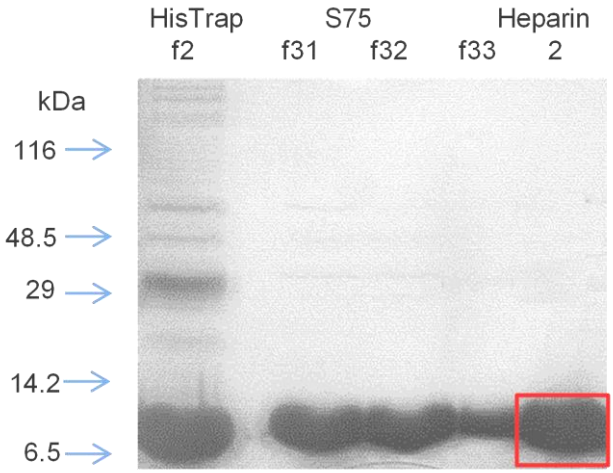


Figure 3. 26 Purification of C35AFrmR by nickel affinity chromatography, size exclusion and heparin affinity chromatography. SDS-PAGE analysis (18 % w/v acryl-bis) showing fractions containing C35AFrmR from each purification step. "HisTrap f2" refers to the second fraction eluted from a 5 ml Ni(II)-affinity column with buffer A containing 0.3 M imidazole. Fractions "S75 f31-f33" were eluted from a Superdex S75 column loaded with fraction "HisTrap f2". Fraction "Heparin 2" was eluted at 0.5 M NaCl from two 1 ml Heparin columns linked together loaded with size exclusion fractions f31-f32 combined. Presence of only one band at around 10 kDa confirms the purity of the protein.

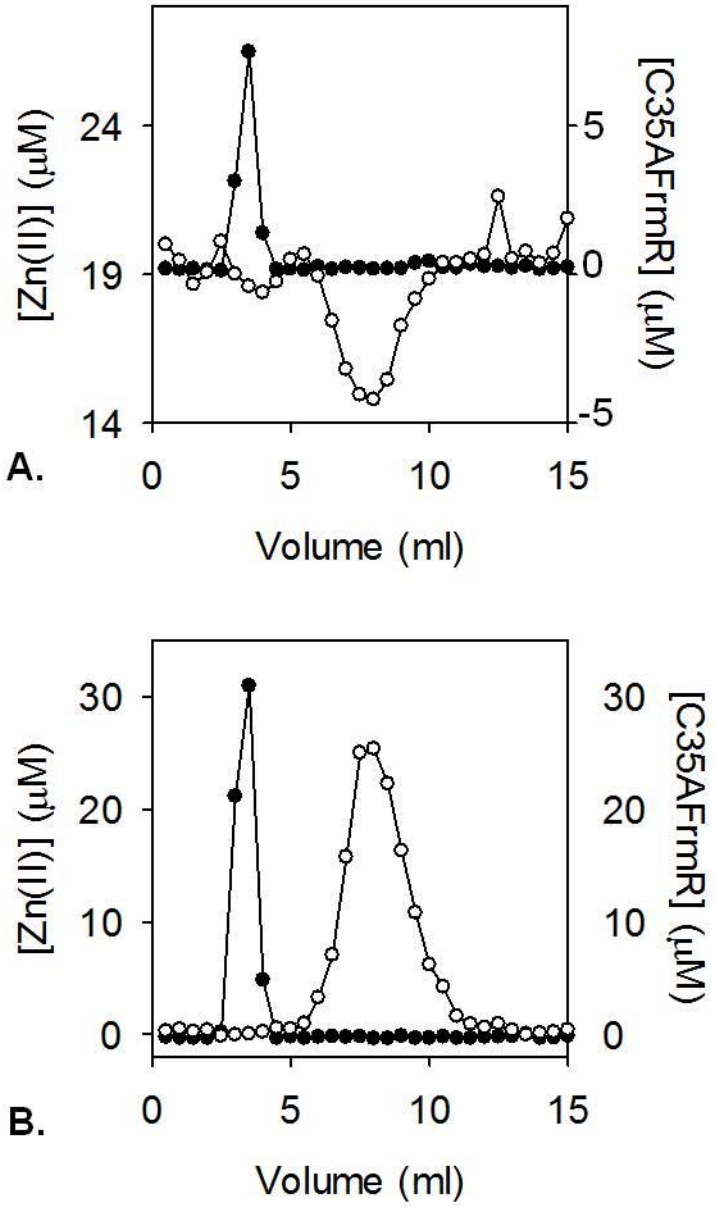


Figure 3. 27 Size-exclusion fractionation of protein-bound and free Zn(II). A C35AFrmR (50 μM) was incubated anaerobically with 150 μM Zn(II) for 2 h. Bound and free metal were resolved on Sephadex G-75 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 and eluted in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and zinc by ICP-MS (open circles). B C35AFrmR (10 μM) was anaerobically loaded onto a Sephadex G-75 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 and 20 μM Zn(II). Elution was carried out in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and zinc by ICP-MS (open circles).

To determine if this was due to weakened zinc binding leading to competition from buffer and/or matrix, the experiment was repeated without incubating FrmR with metal but equilibrating the Sephadex G25 matrix with buffer containing 20 µM $ZnCl_2$ (Figure 3. 27B). Again, no Zn(II) was found in fractions corresponding to FrmR. These data strongly suggest that Cys35 is essential for Zn(II)-binding.

3.8.3 *Apo- and Zn(II)-C35AFrmR:DNA interactions studied by fluorescence anisotropy*

In order to assess if Cys35 is also required for Zn(II)-mediated allostery, fluorescence anisotropy analyses was used. Titration of *frmRA*Pro with C35AFrmR shows that this mutant interacts with the FrmR DNA recognition site with a slightly weaker DNA affinity when compared to wild-type (K_{DNA} = 158.2 ± 17.6 nM) (Figure 3. 28A). The same titration was repeated in the presence of 5 µM $ZnCl_2$ and with C35AFrmR samples incubated with 1.2 molar equivalent of zinc displaying a DNA affinity indistinguishable from the apo-form (K_{DNA} = 147.4 ± 53.8 nM, $\Delta G_c^{Zn(II)\text{-}C35AFrmR}$ -0.07 ± 0.21 kcal mol^{-1}) (Figure 3. 28B). This outcome was further confirmed by assessing the ability of Zn(II) to dissociate the pre-formed C35AFrmR:DNA complex. Zn(II) cannot titrate C35AFrmR off DNA even at a 5-fold molar excess of zinc over protein (Figure 3. 28C). These results suggest that mutation of the residue in position X in the WXYZ motif negatively affects FrmR zinc-binding and responsiveness *in vitro*. It is therefore possible to conclude that Cys35 plays a prominent role in zinc sensing.

3.8.4 *Zn(II) and Cu(I) binding properties of H60LFrmR investigated by size exclusion chromatography and fluorescence emission*

H60LFrmR mutant was prepared as described in Section 3.8. The ability of H60LFrmR to bind Cu(I) and Zn(II) was investigated by resolving the protein on a Sephadex G25 matrix after incubation with a large excess of these metal salts (Figure 3. 30 and Figure 3. 29A). Binding of one molar equivalent of Cu(I) is detected with this assay (Figure 3. 29) however, in contrast with what was observed with wild-type FrmR (Figure 3. 6B), H60L variant is not able to retain binding of Zn(II) under these conditions and the totality of the metal elutes in its unbound form (Figure 3. 30A). In order to minimize metal-binding competition from the column matrix the experiment was carried out without incubating the protein with metal excess but simply by applying H60LFrmR to the column previously equilibrated with a buffer containing 20 µM of $ZnCl_2$ (which was also used to elute the protein). As shown in Figure 3. 30B, zinc co-elutes with the protein showing a 1:1 stoichiometry consistent with what was observed for FrmR.

These preliminary experiments suggest a diminished zinc(II)-binding affinity by H60LFrmR, which could be indicative of His60 being involved in the Zn(II) coordination sphere.

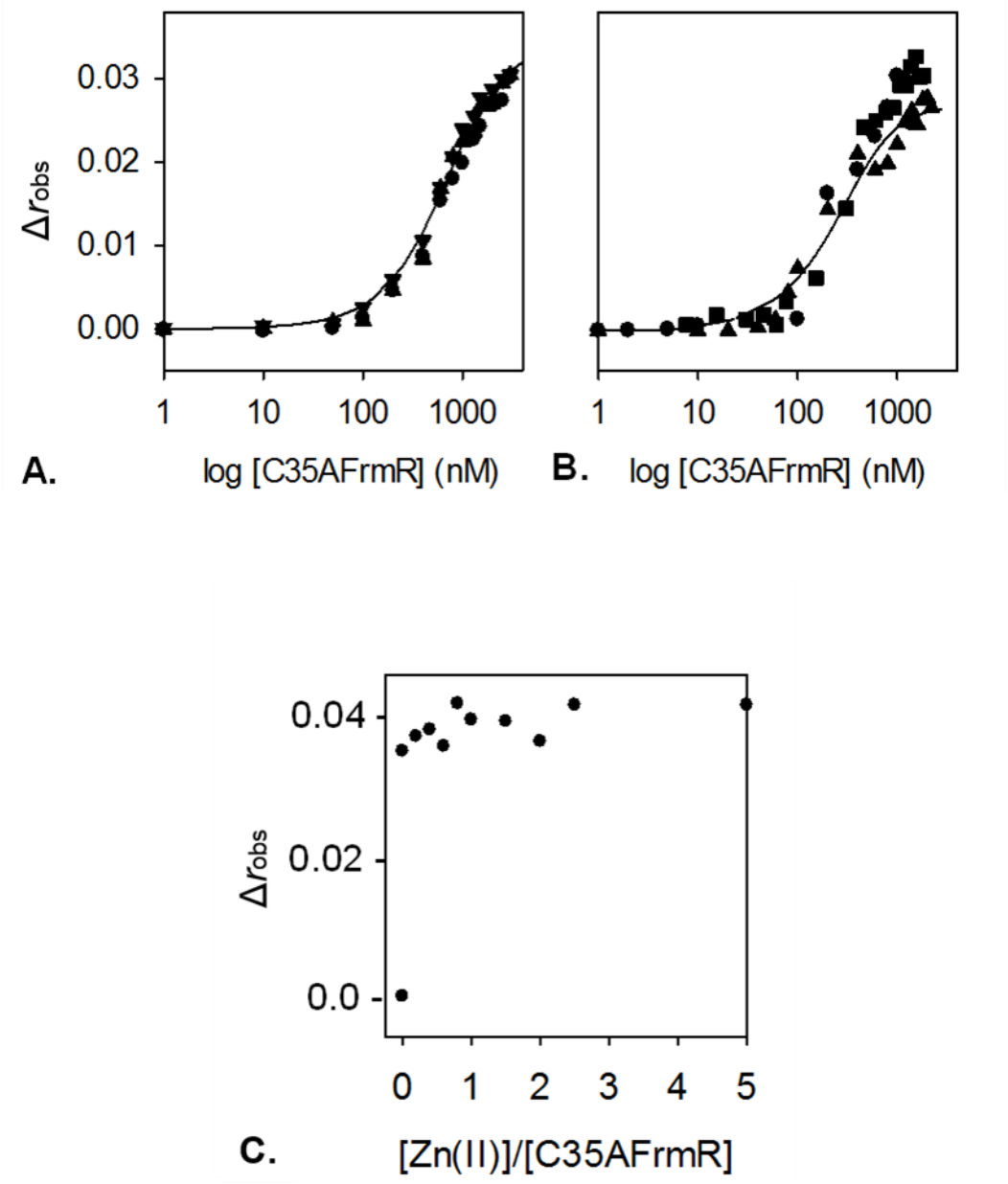


Figure 3. 28 Titration of *frmRA*Pro with apo-C35AFrmR and Zn(II)-C35AFrmR. **A** *frmRA*Pro (10 nM) was anaerobically titrated with C35AFrmR in the presence of 5 mM EDTA. DNA binding was monitored by fluorescence anisotropy. **B** As A. except for the use of 5 µM $ZnCl_2$ in the reaction buffer and the omission of EDTA. C35AFrmR samples were also incubated with 1.2 molar equivalent of $ZnCl_2$. Solid lines represent simulated curves produced from the average K_{DNA} determined across the experiment replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 C35AFrmR tetramer (non-dissociable):DNA stoichiometry. **C** Titration of pre-formed C35AFrmR:DNA complex with $ZnCl_2$ monitored by FA. *frmRA*Pro (10 nm) was pre-incubated with C35AFrmR (2.5 µM) before titration with $ZnCl_2$. Fluorescence anisotropy of protein:DNA complex was monitored during the titration. All experiments were performed anaerobically at pH 7.0.

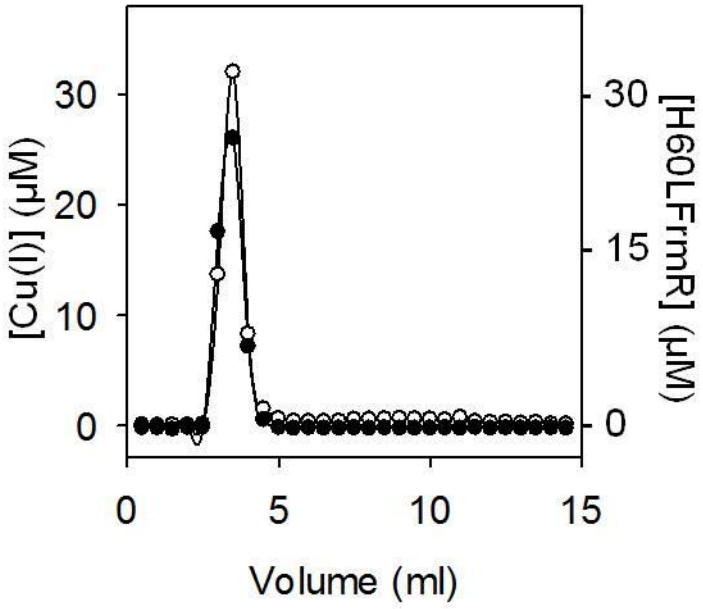


Figure 3. 29 Size-exclusion fractionation of protein-bound and free Cu(I). H60LFrmR (50 μM) was incubated anaerobically with 150 μM Cu(I) for 2 h. Bound and free metal were resolved on Sephadex G-75 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 and eluted in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and copper by ICP-MS (open circles).

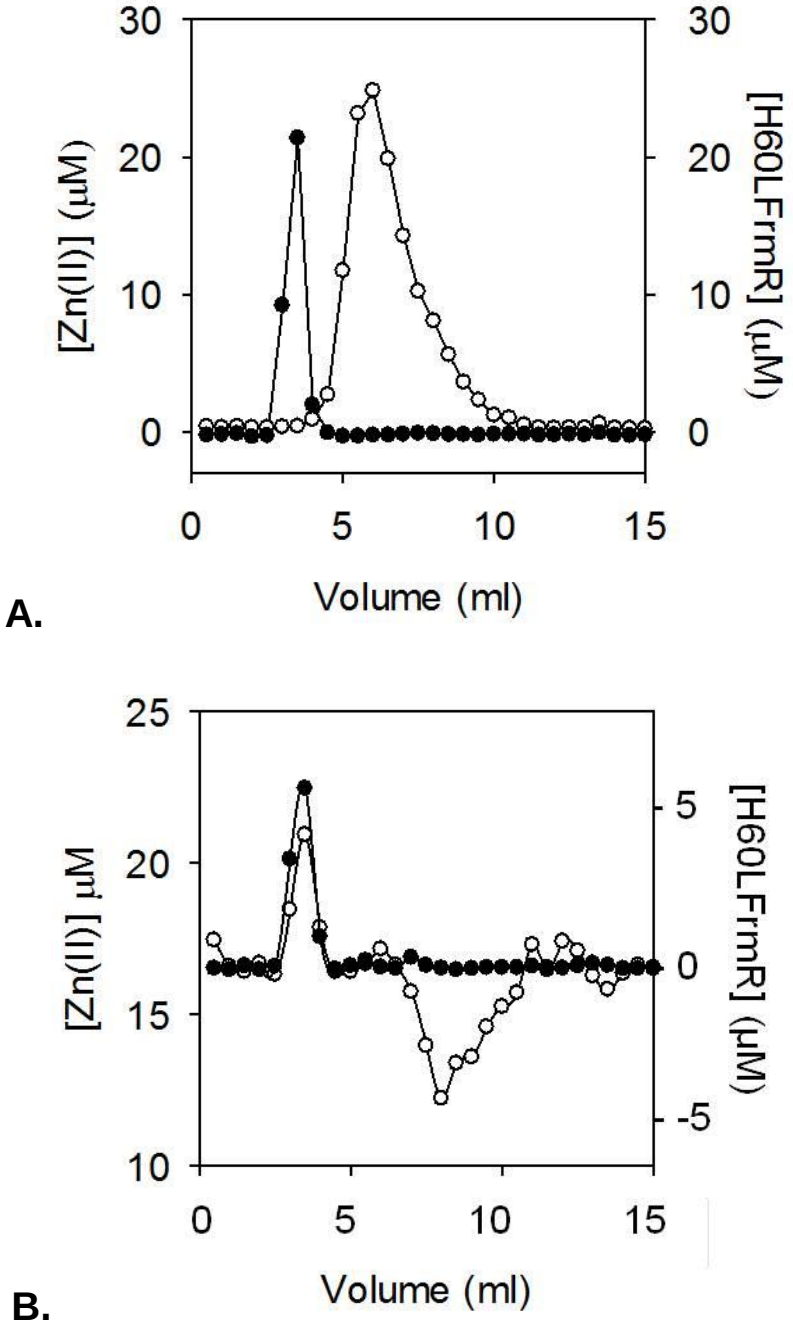


Figure 3. 30 Size-exclusion fractionation of protein-bound and free Zn(II). **A** H60LFrmR (50 μM) was incubated anaerobically with 150 μM Zn(II) for 2 h. Bound and free metals were resolved on Sephadex G-75 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 and eluted in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and zinc by ICP-MS (open circles). **B** H60LFrmR (10 μM) was anaerobically loaded onto a Sephadex G-75 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 and 20 μM Zn(II). Elution was carried out in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and zinc by ICP-MS (open circles).

In the absence of this residue FrmR could recruit other amino acids in close proximity to the metal-binding site in order to replace histidine in metal coordination and therefore retaining an ability to bind Zn(II) although with weaker $K_{Zn(II)}$. Tyrosine fluorescence emission spectrum of H60LFrmR appears very similar to wild-type FrmR's, with an inflection in the signal observed upon addition of ~ 1 molar equivalent of zinc (Figure 3. 31).

3.9 Concluding remarks

Genomic location of the gene coding for *Salmonella* FrmR was identified upstream *frmA*, which encodes the formaldehyde dehydrogenase FrmA. FrmR was shown to bind *in vitro* Zn(II) and Cu(I) and, more weakly, Co(II) and Ni(II), adopting a variety of coordination geometries and displaying different metal-binding affinities. Zn(II) and Co(II) bind the same pocket created by C35, H60 and two additional ligands, and adopting a tetrahedral geometry (Section 7.2). However, Zn(II) is detected by FrmR with a tighter affinity than Co(II), in accord with predictions made by the Irving-Williams series (Section 7.2). The role of two of the WXYZ residues involved in metal-binding in RcnR/CsoR family members was investigated. Cys35 (residue X) appears to be essential for Zn(II)-binding as shown by size-exclusion chromatography and fluorescence anisotropy, whereas His60 (residue Y) retains the ability to bind Cu(I) and Zn(II), although some evidence suggests that the latter is coordinated with a weaker affinity. These findings will be further discussed in Section 7.2.

Interaction of FrmR with the type-1 palindromic region ATA-X-TATA-C_6-TATA-X-TAT was analysed by fluorescence anisotropy and K_{DNA}^{FrmR} was calculated to be 99.4 ± 3.0 nM, which is within the range of DNA-binding affinities documented for RcnR/CsoR family members. This contact can be negatively affected by Zn(II) and Cu(I) ($K_{DNA}^{Zn(II)\text{-}FrmR}$ = 3110.0 ± 40.1 nM, $K_{DNA}^{Cu(I)\text{-}FrmR}$ = 654.3 ± 130.2 nM) (Section 7.3.2).

The role of Glu81 (α3-helix) and His60 (α2-helix) in the propagation of the conformational change upon effector-binding was studied by producing single point mutants each lacking one of these residues (E81AFrmR and H60LFrmR). Both mutants showed an impaired DNA-binding affinity. The most striking effect was observed in H60LFrmR ($K_{DNA}^{H60LFrmR}$ = 1918 nM) suggesting that His60 can still form a hydrogen bond with the electronegative atom of another residue on the α2-helix or, alternatively, other residues can be recruited to form weak contacts between the two helices. See Section 7.3.2 for further discussion.

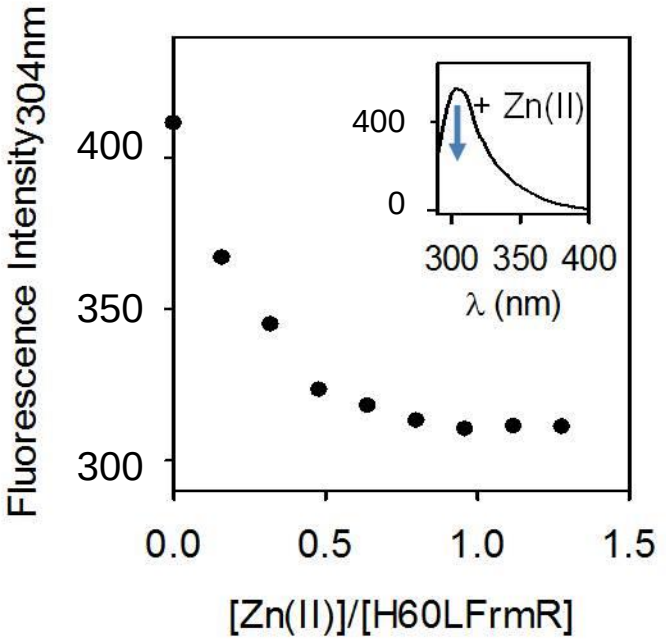


Figure 3. 31 Fluorescence emission at 304 nm (λ_{ex} = 280 nm) of H60LFrmR (13.1 µM, monomer) and following titration with $ZnCl_2$. Fluorescence is quenched after addition of approximately 1 molar equivalent of Zn(II) per FrmR monomer, consistent with a Zn(II)-dependant conformational change of FrmR which alters the environment of the tyrosine residue probably changing access to water.

Chapter 4

Generation of a metal-sensing transcriptional regulator by gain of function point mutation

The contents of parts of this chapter has been published in, and hence adapted from, Deenah Osman[1], Cecilia Piergentili[1], Junjun Chen, Buddhapriya Chakrabarti, Andrew W. Foster, Elena Lurie-Luke, Thomas G. Huggins, and Nigel J. Robinson (2015). *In vivo* assays and *in vitro* competition experiments performed by Dr. Deenah Osman have been noted in figure legends and text. Interpretation is my own.

4.1 Aims and objectives

The concentration of cellular metals is believed to be controlled by the actions of DNA-binding, metal-sensing transcriptional regulators which in turn aids correct protein-metallation (Foster *et al.* 2014; Reyes-Caballero *et al.* 2011): tight affinities of these proteins for a given set of metals results in low [buffered metal] (Foster *et al.* 2014).

Sequence similarity between *Salmonella* FrmR and other CsoR/RcnR family members, suggested that FrmR may coordinate transition metals. This hypothesis was confirmed in Chapter 3. FrmR is able to bind Zn^{2+}, Cu^{+}, Co^{2+} and Ni^{2+}, with $K_{Zn(II)1\text{-}3}{}^{FrmR}$ 1.7 ± 0.7 x 10^{-10} M and $K_{Cu(I)1\text{-}2}{}^{FrmR}$ 4.9 ± 1.6 x 10^{-15} M for the tightest sites.

Recent studies have inferred by correlation that a combination of relative affinity, relative-allostery and relative-access determine the ability of metal-sensors to respond selectively *in vivo* (Foster *et al.* 2014). In this Chapter we test these theories via a mutation of non-metal sensing FrmR. By creating an RcnR-like helix α2′ HxxxH motif (E64HFrmR), FrmR gains responsiveness to cobalt and Zn(II) *in vivo*. Relative-properties which, in combination, enable metal-sensing are identified by comparing the biochemical properties of *Salmonella* FrmR with E64HFrmR, and then relating these parameters to endogenous sensors for cobalt, Zn(II) and Cu(I).

4.2 Wild-type FrmR responds only to formaldehyde *in vivo*

4.2.1 *In vivo expression from pRSfrmR following exposure to metals and formaldehyde*

Despite similarity between FrmR and metal sensing RcnR and CsoR, exposing *Salmonella* cultures to maximum non-inhibitory concentrations (MNIC) of $MnCl_2$, $C_6H_5FeO_7$, $CoCl_2$, $NiSO_4$, $CuSO_4$ or $ZnSO_4$ (Figure 4. 1A) and Na_2SeO_3 or Na_2TeO_3 (Figure 4. 1B) does not induce expression from P_{frmRA}-*frmR* fused to *lacZ* in Δ*frmR* cells, as shown by *β-galactosidase* assay. However, exposure to MNIC formaldehyde does alleviate repression by *Salmonella* FrmR (Figure 4. 1A-B), confirming FrmR as a formaldehyde-sensing transcriptional regulator, as it will be further explored in Chapter 5. These experiments have been performed by Dr. Deenah Osman.

4.2.2 *In vivo expression from pRSfrmRE64H following exposure to metals and formaldehyde*

A single point mutation of FrmR (His64→Glu) generates an RcnR-like WXYZ motif. The ability of E64HFrmR to respond to metals *in vivo* was tested by a *β-galactosidase* assay in Δ*frmR* cells containing *pRSfrmRE64H* fused to *lacZ* in the absence and presence of MNIC of

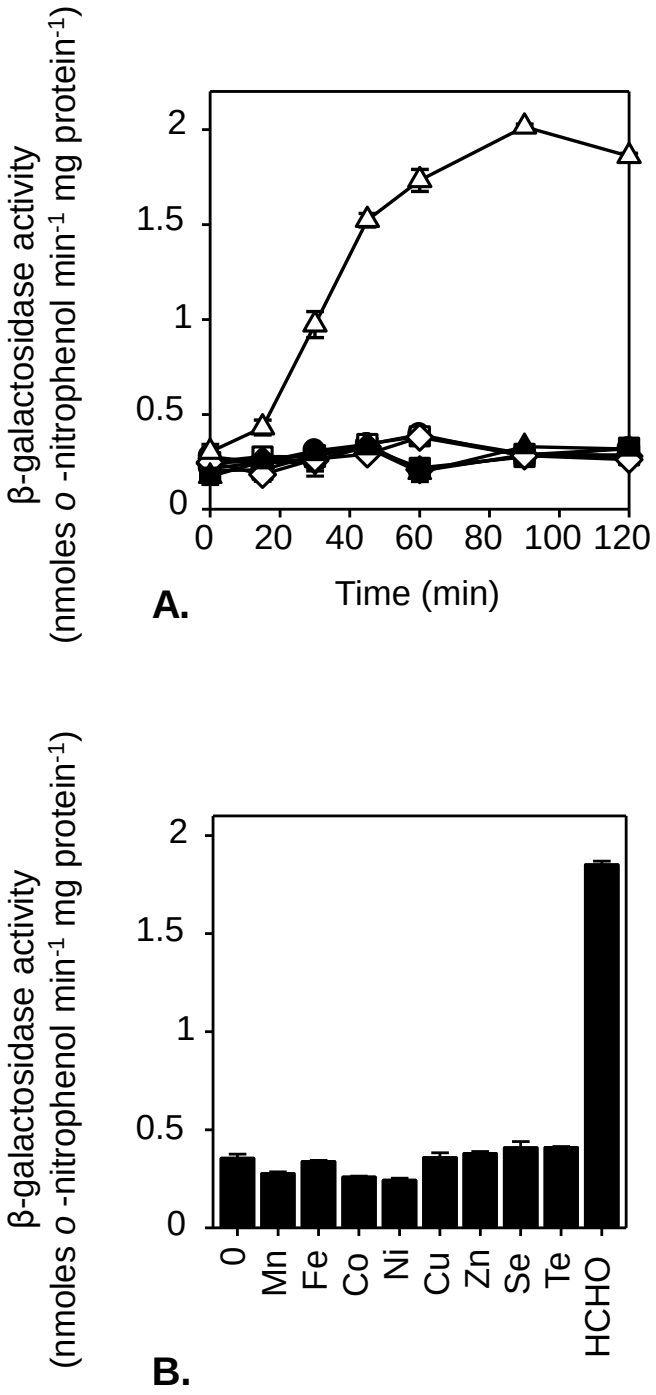


Figure 4. 1 A *β*-galactosidase activity measured in *ΔfrmR* containing P_{frmRA}-*frmR* fused to *lacZ* following growth to mid-exponential phase in M9 minimal medium and continued incubation in M9 with no metal supplement (filled circles) or maximum permissive concentrations of Mn(II) (open circles), Fe(III) (filled squares), Co(II) (open squares), Ni(II) (filled diamonds), Cu(II) (open diamonds), Zn(II) (filled triangles) or formaldehyde (open triangles). **B** *β*-galactosidase activity measured in Δ*frmR* containing P_{frmRA}-*frmR* fused to *lacZ* following growth to mid-exponential phase in M9 minimal medium in A, the absence or presence of MNIC of $MnCl_2$, Fe(III)citrate, $CoCl_2$, $NiSO_4$, $CuSO_4$, $ZnSO_4$, Na_2SeO_3, Na_2TeO_3 or formaldehyde. These experiments have been performed by Dr. Deenah Osman.

$MnCl_2$, $C_6H_5FeO_7$, $CoCl_2$, $NiSO_4$, $CuSO_4$, $ZnSO_4$ or formaldehyde (Figure 4. 2B). Surprisingly, expression was induced by $CoCl_2$ and $ZnSO_4$, therefore, replacement of residue 64 (glutamate) with histidine, generated a metal-sensing variant of FrmR. Other metals tested in this assay, Fe(II), Ni(II) and Cu(II), did not affect expression from P_{frmRA}-*frmRE64H*, while formaldehyde-responsiveness was retained. This experiment has been performed by Dr. Deenah Osman.

A single residue mutant of RcnR gains the ability to detect cellular Zn(II) (Higgins *et al.* 2012) but this work describes the first successful attempt to turn a non-metal sensor into a metal sensor. This creates an opportunity to directly test the inferred determinants of metal-selectivity via gain-of-function rather than solely by correlation. The basal expression from the *frmRA* promoter in cells containing E64HFrmR is greater than that of wild-type FrmR (Figure 4. 1 and Figure 4. 2). This is further discussed in Chapter 7.

4.3 Production and purification of recombinant E64HFrmR

To identify the properties which enable metal-sensing by E64HFrmR, recombinant E64HFrmR was generated and purified and its biochemical properties recorded. These parameters were then compared with those of the endogenous sensors for Co(II), Zn(II) and Cu(I). Glu64 → His substitution was performed by site-directed mutagenesis of pETFrmR, and E64HFrmR was overexpressed in *E. coli* as a recombinant, tag-free protein and purified using its inherent biochemical properties. Crude lysate was applied to a nickel affinity column and eluted in a single step with 300 mM imidazole (Figure 4. 3A).

E64HFrmR was further purified by size-exclusion chromatography. The protein was routinely found in fractions 30-33 (Figure 4. 3B). This elution volume (157.5 ml) corresponds to E64HFrmR adopting a tetrameric state as previously discussed for FrmR (Section 3.3). Following size exclusion chromatography the pooled protein was diluted and applied to a heparin affinity column and eluted with a step-wise NaCl gradient (Figure 4. 3C). SDS-PAGE analysis confirmed the protein to be > 95 % pure (Figure 4. 3C). E64HFrmR was moved into an anaerobic glovebox by application to a Heparin affinity column followed by elution in the glove box using N_2-purged, Chelex-treated, EDTA- and DTT-free buffer. Quantification of E64HFrmR was performed by measuring A_{280} and using the experimentally determined extinction coefficient for FrmR (ε_{280} = 1951 M^{-1} cm^{-1}). The absorbance spectrum of apo-E64HFrmR (Figure 4. 4A) routinely differed from that of FrmR (Figure 8. 4A, Appendix), exhibiting a shoulder at ~ 300 nm (although two early preparations of E64HFrmR did resemble the apo-FrmR spectrum, Figure 4. 4B). Reduced thiol and metal content assays revealed the protein was fully reduced and metal-free ruling out a possible link between the appearance of this feature and oxidation or metal contamination. Furthermore, the hypothesis that this spectral feature may be informative of the formation of an adduct with organic molecules available in the cellular environment was explored. Formaldehyde and glutathione were considered as

Sty_E64HFrmR

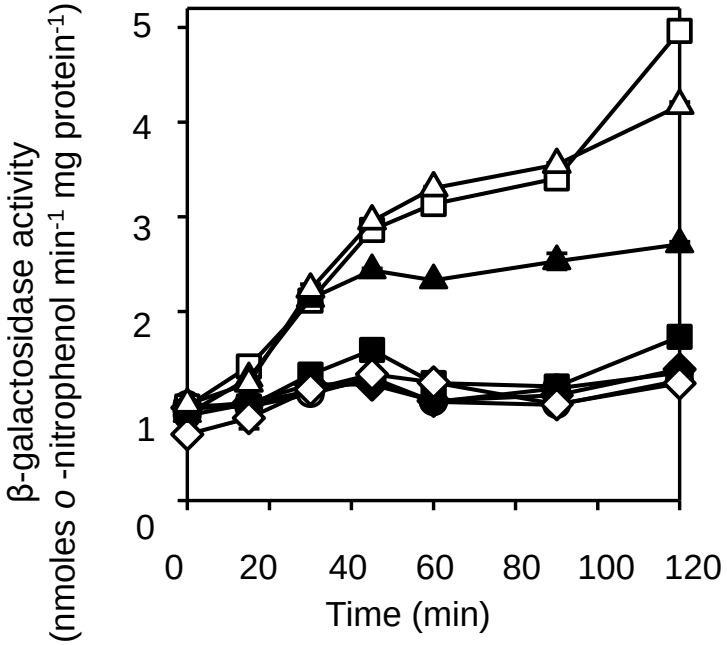
W X

MPHSPEDKKRILTRVRRIRGQVEALERALESGEPCLAILQQIAAVR

Y Z

GASNGLMSEMVEIHLKDHLVSGETTPDQRAVRMAEIGHLLRAYLK

A.



B.

Figure 4. 2 A *S. typhimurium* E64HFrmR (Sty_E64HFrmR) sequence. The residues at positions of the W-X-Y fingerprint are high-lighted in red. Residue Z (now mutated Glu→His) is highlighted in green. **B** *β*-galactosidase activity measured in Δ*frmR* containing P_{frmRA}-*frmRE64H* fused to *lacZ* following growth to mid-exponential phase in M9 minimal medium, and continued incubation in M9 with no metal supplement (filled circles) or maximum permissive concentrations of Mn(II) (open circles), Fe(III) (filled squares), Co(II) (open squares), Ni(II) (filled diamonds), Cu(II) (open diamonds), Zn(II) (filled triangles) or formaldehyde (open triangles). A single residue change renders FrmR responsive to cobalt and zinc, in addition to formaldehyde. Experiment conducted by Dr. Deenah Osman.

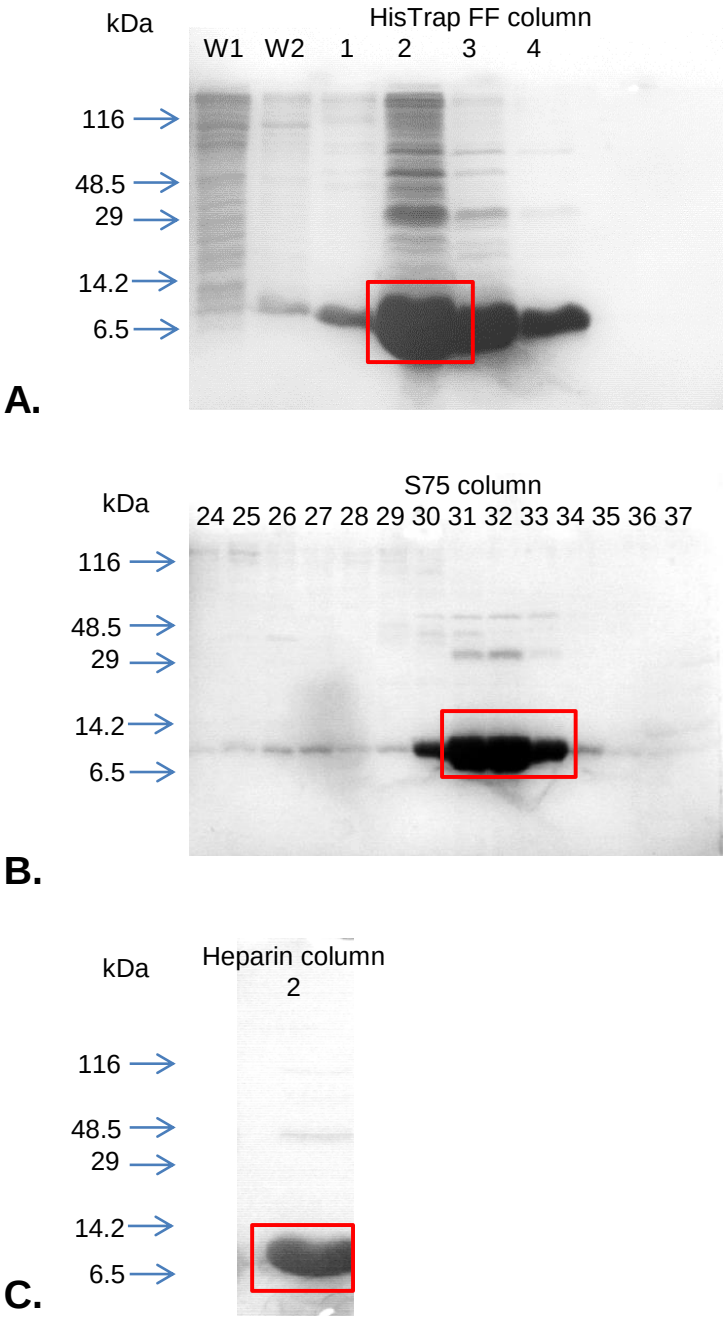


Figure 4. 3 Purification of E64HFrmR by nickel affinity chromatography, size exclusion and heparin affinity chromatography. **A** SDS-PAGE analysis (18 % w/v acryl-bis) of fractions eluted from a 5ml Ni(II)-affinity column with buffer A containing 0.3 M imidazole. Fractions W1 and W2 contain material not bound to the column and flow through during wash step with 25 ml buffer A respectively. Fraction 2 was routinely found to contain E64HFrmR in the highest concentration. **B** SDS-PAGE analysis (18 % w/v acryl-bis) of fractions 24-37 eluted from a Superdex S75 column loaded with fraction 2 (3 ml) from the previous purification step (A). **C** SDS-PAGE analysis (18 % w/v acryl bis) of fractions eluted at 0.5 M NaCl from two 1 ml Heparin columns linked together loaded with size exclusion fractions enriched for E64HFrmR.

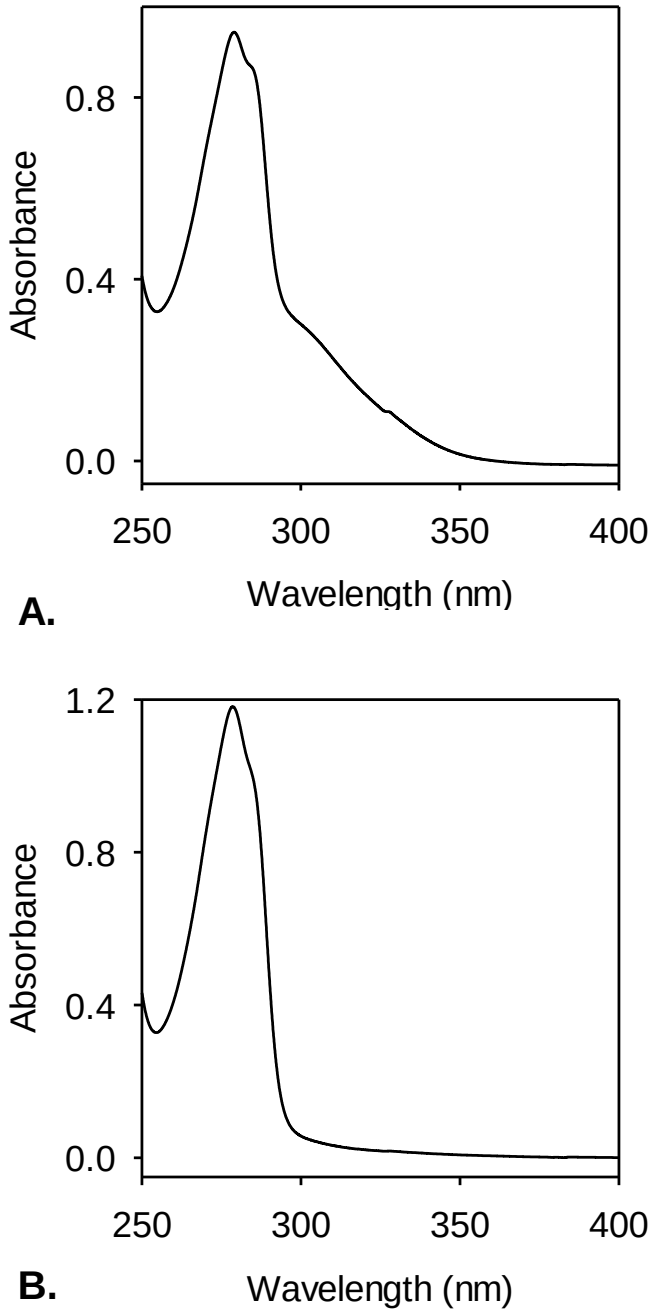


Figure 4. 4 A large majority (all except two) of the apo-E64HFrmR proteins produced during this work showed a second peak at around 300 nm in the UV-visible absorbance spectrum (**A**). Figure **B** shows apo-E64HFrmR UV-visible spectrum obtained in two extracts, lacking the second shoulder and showing no difference compared to apo-wtFrmR UV-visible spectrum (Figure 8. 2A, Appendix).

plausible candidates since the first is detected by FrmR and the second spontaneously reacts with formaldehyde to yield S-hydroxymethylglutathione, FrmA substrate. However, incubation (16 h) with formaldehyde or glutathione in anaerobic conditions or incubation (16 h) of the protein in aerobic conditions did not lead to an increase in the intensity of the feature in E64HFrmR preparations where the shoulder was already present, nor to development of the aforementioned feature in early preparations of E64HFrmR in which the feature was absent (Figure 8. 4B-D). Since the majority of purified E64HFrmR stocks exhibited this spectral feature (although its nature remains unclear) the following data, unless otherwise stated, have been obtained with this protein form.

4.4 *In vitro* analysis of E64HFrmR metal binding properties

As described in Section 4.3, the introduction of a His residue in position Z of the WXYZ motif of FrmR generates a metal sensing mutant, likely through the creation of a metal-binding site.

E64HFrmR metal-binding properties will be explored in this section in order to ascertain if a metal binding site has been generated and to detect any dissimilarity with wild-type FrmR (Section 3.4).

4.4.1 *Co(II)-binding properties*

Titration of E64HFrmR (87.0 µM) with Co(II) results in the appearance of the spectral features observed for FrmR (Figure 4. 5), consistent with Co(II) binding to both proteins with the same coordination geometry. The two spectral features in the near-UV region are indicative of S^- →Co(II) ligand-to-metal charge transfer (LMCT) and the intensity at saturation of the band at 336 nm (~ 0.9 x 10^3 M^{-1} cm^{-1}) is consistent with a single thiolate ligand (Figure 4. 5) (VanZile *et al.* 2000). The intensities of a second set of Co(II)-dependent features in the region of 600 nm, indicative of *d-d* transitions (VanZile *et al.* 2000), suggest tetrahedral coordination geometry. The binding curves (294, 336 and 614 nm) are linear up to one equivalent of Co(II) implying $K_{Co(II)}$ is too tight to be estimated by direct titration (inset Figure 4. 5).

4.4.2 *Zn(II)-binding properties*

Zn(II) is spectrally silent due to its filled $3d^{10}$ orbitals however Zn(II)-binding can lead to a change in intrinsic protein fluorescence (VanZile *et al.* 2002b; Larissa *et al.* 2007; Waldron *et al.* 2010). Titration of Zn(II) into 10.65 µM E64HFrmR was performed and tyrosine intrinsic emission, following excitation at 280 nm, was monitored (Figure 4. 6A). As shown in Figure 4. 6A inset, apo- E64HFrmR Tyr emission has a maximum at ~ 305 nm, as does wild-type FrmR, however $ZnCl_2$ additions resulted in an unexpected increase, rather than quenching observed in FrmR (Figure 3. 7A), of fluorescence intensity. The result suggests that Tyr89 residue is located in a distinct environment compared to FrmR.

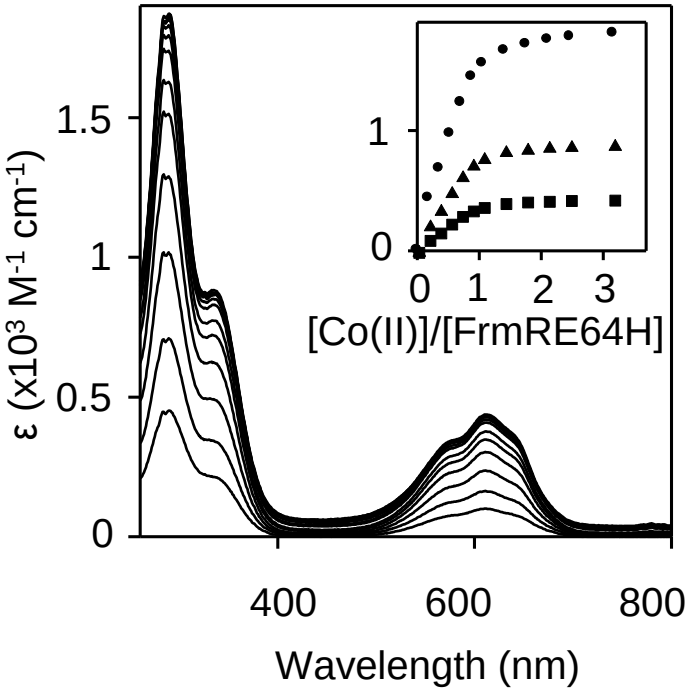


Figure 4. 5 Representative apo-subtracted UV-visible difference spectra of FrmR (87.0 µM protomer) upon titration with $CoCl_2$ and binding isotherms (inset) at absorbance maxima 294 nm (circles), 336 nm (triangles), and 614 nm (squares).

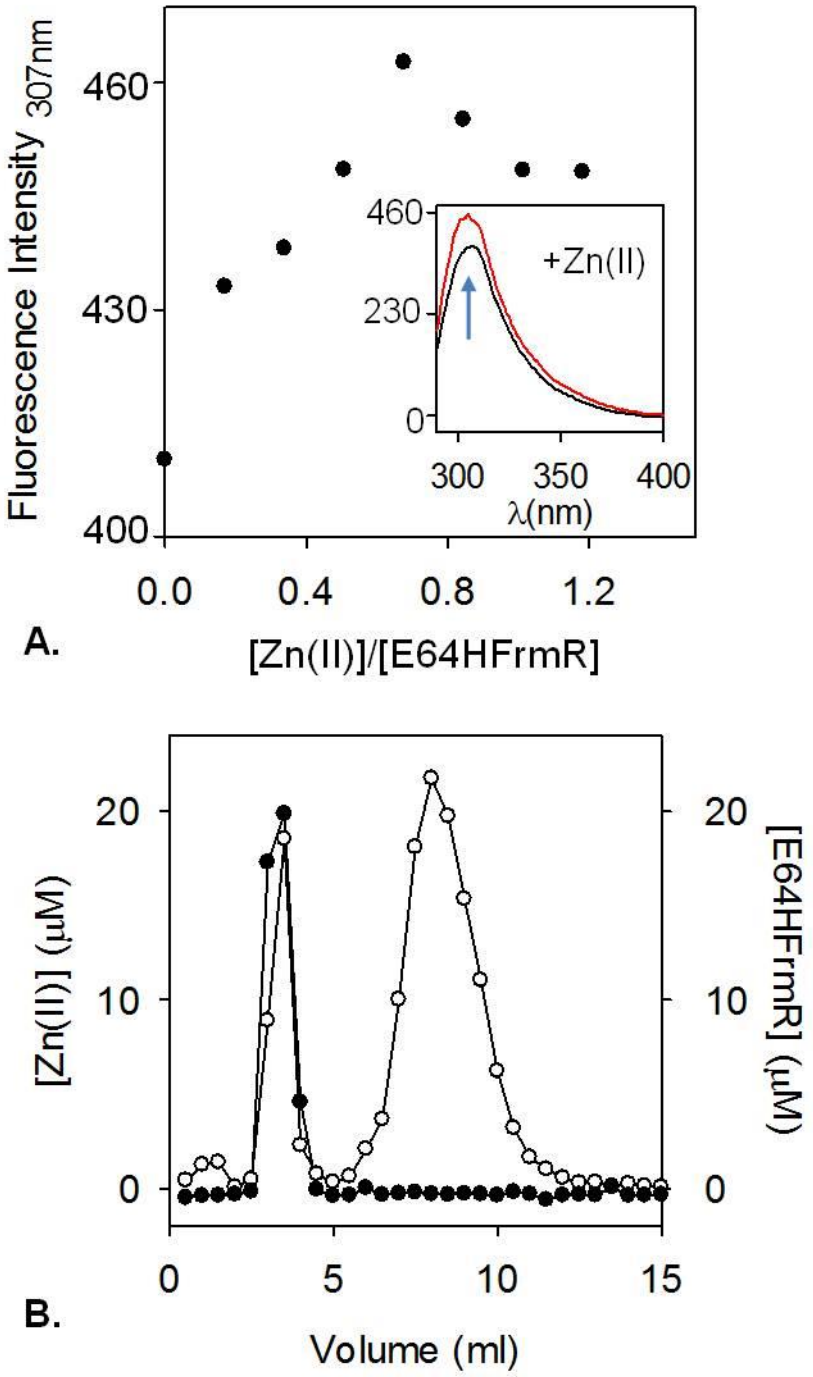


Figure 4. 6 A Fluorescence emission at 307 nm (λ_{ex} = 280 nm) of E64HFrmR (10.65 μM, monomer) and following titration with $ZnCl_2$. Fluorescence is not quenched upon zinc addition, suggesting a different environment of the tyrosine residue in the E64HFrmR structure compared to the wild-type. **B** Size-exclusion fractionation of protein-bound and free Zn(II). E64HFrmR (50 μM) was incubated anaerobically with 150 μM Zn(II) for 2 h. Bound and free metal were resolved on Sephadex G-25 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 and eluted in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and zinc by ICP-MS (open circles).

One equivalent of Zn(II) binds sufficiently tightly to co-migrate with E64HFrmR during size-exclusion chromatography (performed with Sephadex G25 medium) as shown in Figure 4. 6B. This supports a 1:1 Zn(II) binding stoichiometry as already confirmed to be the case for FrmR (Figure 3. 7B).

4.4.3 *Cu(I)- and Ni(II)-binding properties*

In addition to Zn(II) and Co(II), which are competent to alleviate *frmRA* expression by E64HFrmR, Cu(I)- and Ni(II)-binding were also tested as these are the metals sensed by CsoR and RcnR, respectively.

Titration of E64HFrmR (21.5 μM) with a reduced Cu(I) stock showed a Cu(I)-dependent absorption band in the near UV range (240 nm) indicative of the formation of the thiolate-copper coordination bond (Angeletti *et al.* 2005). The inflection in intensity of this feature indicates tight binding of at least one equivalent of Cu(I) (Figure 4. 7A). Again, this finding is in line with what was presented and discussed for FrmR (Figure 3. 8A).

One equivalent of copper also co-migrates with E64HFrmR (50 μM) during size-exclusion chromatography (Figure 4. 7B), analogous to what was shown for FrmR. Although the protein was incubated with a large excess (three molar equivalents) of Cu(I), not all the copper was recovered at the end of the experiment most likely due to the inherant affinity of the size-exclusion matrix for Cu(I). Cu(I)-binding stoichiometry, as inferred from spectroscopy and size exclusion chromatography, and coordination geometry (suggested by UV-vis spectroscopy) do not appear different from that of FrmR, suggesting an unaltered Cu(I) binding site in E64HFrmR.

When E64HFrmR (15.3 μM) is titrated with Ni(II) and monitored by UV-visible absorbance spectroscopy, a spectral feature at 280 nm can be detected (Figure 4. 8A). The band shows an inflection at ~ 1 molar equivalent indicating a 1:1 Ni(II)-binding stoichiometry as previously observed for FrmR (Figure 3. 9A). Although the presence of non-specific light scatter throughout the experiment makes difficult the determination of an extinction coefficient for the LMCT feature, a (probably overestimated) value of $\varepsilon_{280} = 0.8 \times 10^3$ can be made. The retention of the A_{280} nm feature, although of reduced intensity, suggests that this mutant can still bind Ni(II). Furthermore, the spectra appear different from those observed with FrmR because of the absence of the wide band at ~ 500 nm, although the presence of non-specific light scattering may have concealed this feature (Figure 4. 8A). An analogous spectrum was observed with *E. coli* Ni(II)-RcnR (Iwig *et al.* 2008), where only peaks at 235 nm (present also in FrmR and E64HFrmR but not shown in the figures) and 280 nm were detected. The RcnR spectrum is indicative of a six-coordinate Ni(II)-coordination site involving a thiolate, a result confirmed by X-ray absorption and electron paramagnetic resonance spectroscopies (Iwig *et al.* 2008). Since E64HFrmR was originally produced to re-create a RcnR-like Ni(II) binding site, these

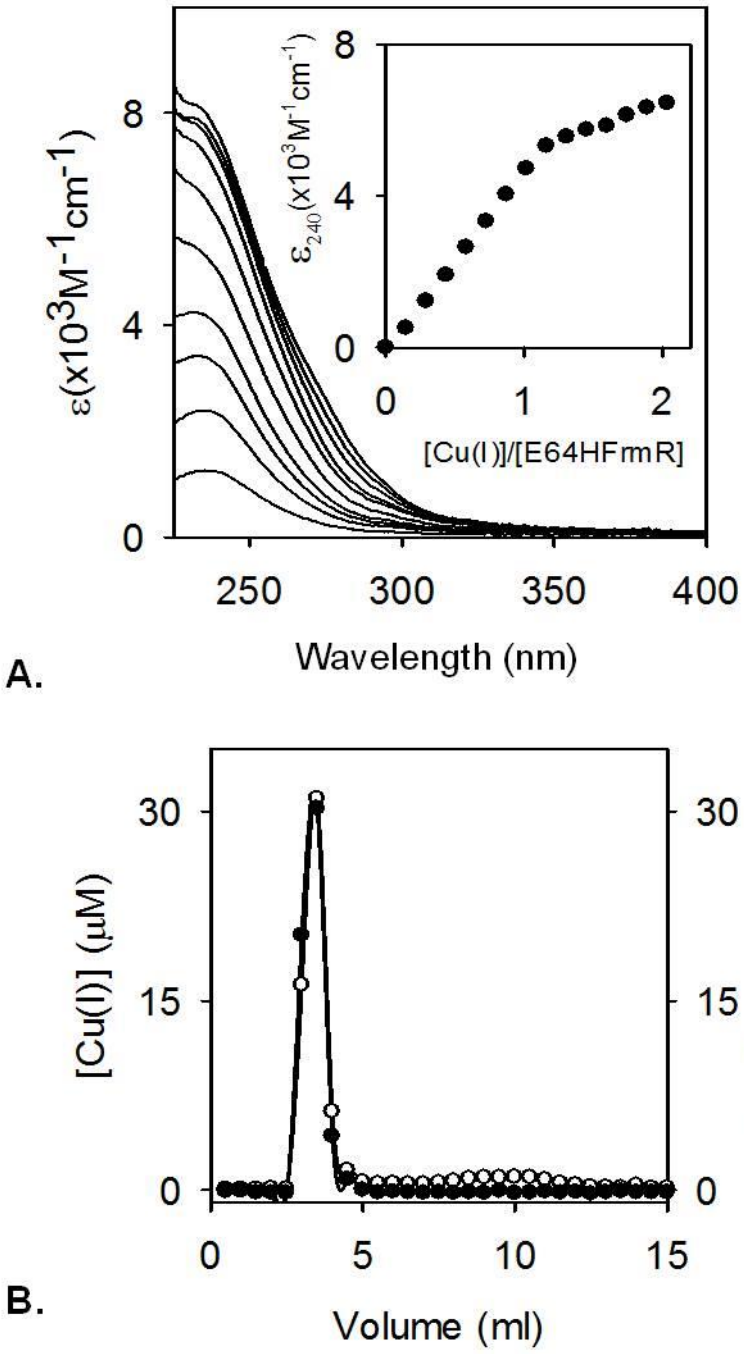


Figure 4. 7 A E64HFrmR (21.5 μM) upon titration with CuCl (oxidation state verified as Cu(I)) and binding isotherm (inset) at 240 nm. **B** Analysis of fractions (0.5 ml) for protein by Bradford assay (filled circles) and copper by ICP-MS (open circles) following size exclusion chromatography of E64HFrmR (0.5 ml at 50 μM) incubated with 150 μM CuCl and eluted with chelex-treated 10 mM HEPES pH 7.0, 100 mM NaCl, 400 mM KCl.

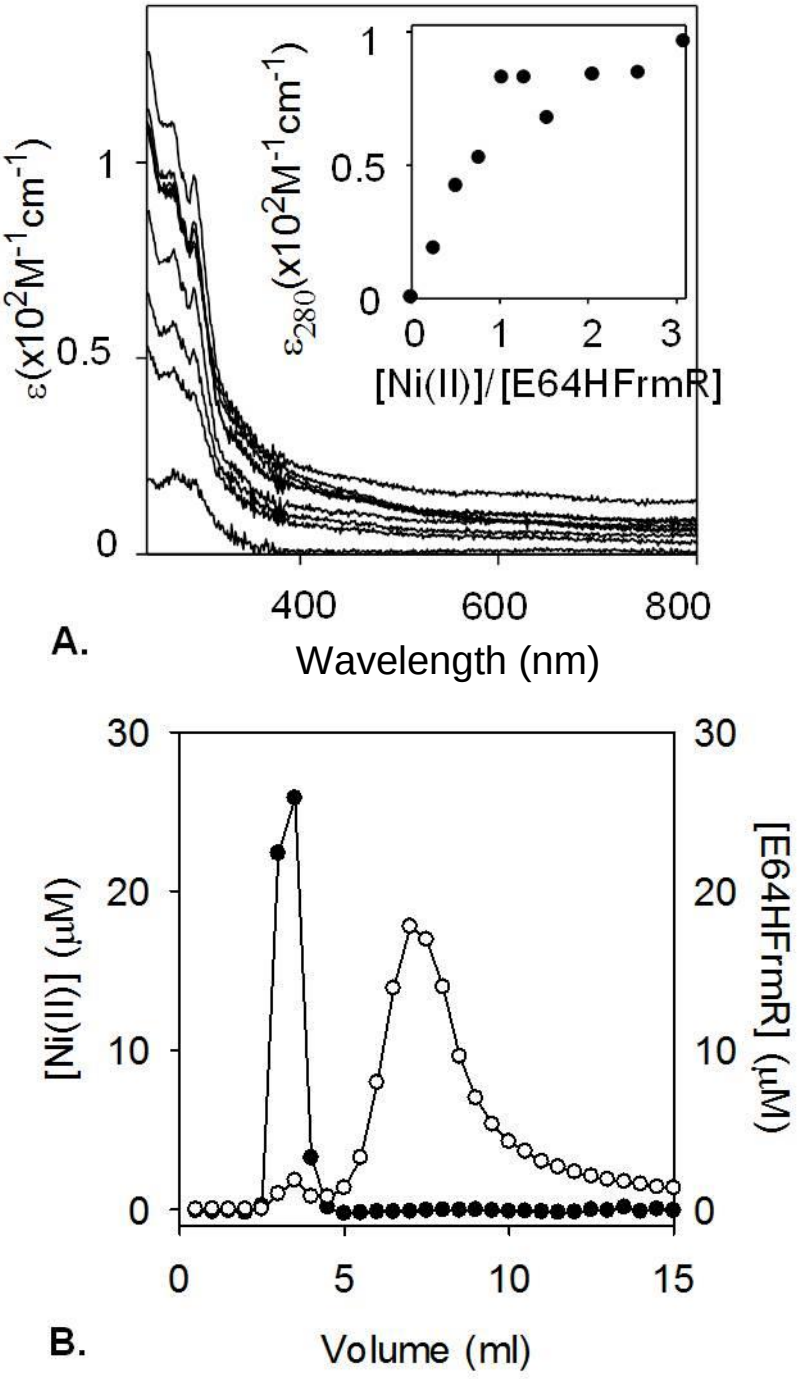


Figure 4. 8 A Apo-subtracted UV-visible difference spectra of E64HFrmR upon titration with $NiCl_2$ (15.3 μM monomer). Inset: Binding isotherm of the spectral feature at 280 nm. **B** Size-exclusion fractionation of protein-bound and free Ni(II). E64HFrmR (50 μM) was incubated anaerobically with 150 μM Ni(II) for 2 h. Bound and free metal were resolved on Sephadex G-25 matrix equilibrated with 100 mM NaCl, 400 mM KCl, 10 mM Hepes pH 7.0 and eluted in the same buffer. Fractions (500 μl) were analysed for protein by Bradford assay (filled circles) and nickel by ICP-MS (open circles).

similarities (albeit the lower intensity of the LMCT feature) are not surprising.

Although the metal binding site may be now more suitable to bind nickel, E64HFrmR still displays a weak Ni(II) affinity as shown in Figure 4. 8B, which shows migration of only ~ 0.1 equivalents of nickel with the protein when resolved on a size-exclusion chromatography column. Again it is possible that the chromatography matrix complexes with Ni(II).

4.5 Metal binding affinities of E64HFrmR

As the E64HFrmR variant, but not FrmR, responds to Zn(II) and Co(II) in the cell, a preliminary hypothesis was that this substitution had succeeded in increasing the affinity for these metals. Zn(II), Co(II) and Cu(I) binding affinities have therefore been determined by use of mag-fura-2, quin-2, fura-2, BCA and BCS chromophores and values have been compared to those measured for FrmR (Section 3.4).

4.5.1 *Measurement of the Zn(II) binding affinity of E64HFrmR by competition with mag-fura-2 and quin-2*

The low-affinity chromophore mag-fura-2 forms a 1:1 complex with Zn(II) and undergoes a concomitant change in absorbance upon metal-binding which can be used to monitor competition with proteins and hence to estimate protein $K_{Zn(II)}$ (Foster *et al*. 2014a; Lisher *et al*. 2013; Simons *et al.* 1993). Titration of 10.1 µM mag-fura-2 ($K_{Zn(II)}^{mag\text{-}fura\text{-}2}$ = 2.0 x 10^{-8} M) with Zn(II) in the presence of 18.8 µM E64HFrmR was monitored by UV-visible spectroscopy and resulted in the increase in intensity of a spectral feature at 325 nm (not reported here) and a decrease in the intensity of a feature at 366 nm (Figure 4. 9A). The two curves intersect at the isosbestic points, 342 nm and 276 nm, indicating a clean 1:1 reaction (Xiao & Wedd 2010). No Zn(II)-dependent change in the spectral features of mag-fura-2 was observed up to ~ 0.75 molar equivalent of zinc per protein monomer, analogous to that previously observed with FrmR. Since change in absorbance is complete after addition of ~ 24.2 µM, three of the four metal binding sites per tetramer are capable of competition with mag-fura-2, whereas the fourth site is too weak. Data were fit using Dynafit (Kuzmic 1996) to a model describing tight binding of three molar equivalent of Zn(II) per tetramer, with dashed lines representing simulated curves describing $K_{Zn1\text{-}3}$ ten-fold tighter and 10-fold weaker than the calculated affinity (Figure 4. 9A). Further inspection of the 10-fold tighter simulated curve, which is close to the calculated one, suggests that $K_{Zn1\text{-}3}$ approaches the upper limit of the assay.

Competition was then conducted with quin-2 ($K_{Zn(II)}^{quin\text{-}2}$ = 3.7 x 10^{-12} M) which has been reported to produce UV difference spectra with isosbestic wavelengths at 342, 317 and 252 nm (Jefferson *et al*. 1990) Formation of 1:1 Zn(II):quin-2 complex can be monitored at 265 nm. Titration of 13.4 µM quin-2 with $ZnCl_2$ in the presence of 42.7 µM E64HFrmR (concentration of protein monomer) showed saturation of the signal upon addition of ~ 45 µM Zn(II), which

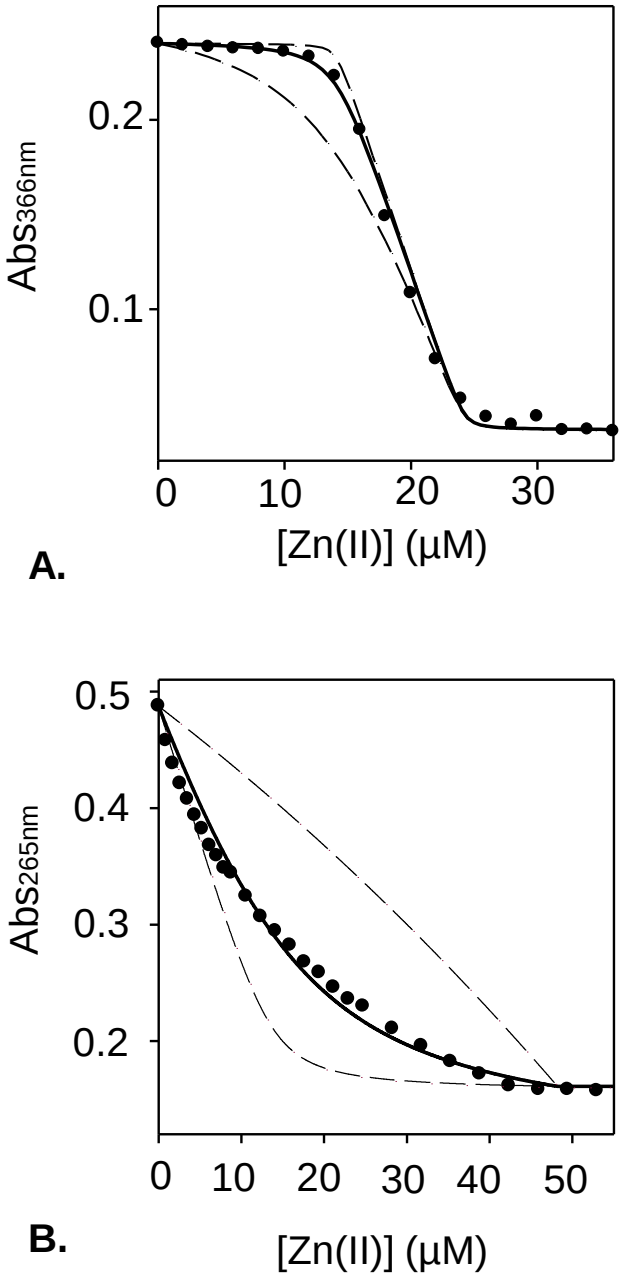


Figure 4. 9 A Representative (n = 3) of mag-fura-2 (10.1 μM) titrated with $ZnCl_2$ in the presence of E64HFrmR (18.8 μM). The absorbance at 366 nm was measured by UV-vis spectrophotometer. **B** Representative (n = 3) of quin-2 (13.4 μM) with $ZnCl_2$ in the presence of E64HFrmR (42.7 μM). The absorbance at 265 nm was measured by UV-vis spectrophotometer. In both cases, the model describes competition of FrmR monomer with the chelator for 0.75 molar equivalents of Zn(II) (3 sites per tetramer, $K_{Zn1\text{-}3}$). Dashed lines represent simulated curves describing $K_{Zn1\text{-}3}$ 10-fold tighter and 10-fold weaker.

indicates competition of three sites per protein tetramer (~ 32 µM + 13.4 µM = 45.4 µM) (Figure 4. 9B). Again, the fourth site does not compete with quin-2. Data were fit to a model describing competition of three molar equivalents of Zn(II) per tetramer and simulated curves describing a ten-fold tighter and ten-fold weaker K_{Zn1-3} (dashed lines) were produced using DynaFit. Since both simulated curves now greatly depart from the calculated curve, the resulting K_{Zn1-3} = 2.33 ± 0.3 x 10^{-11} M it appears to be within the limits of the assay. Therefore E64HFrmR Zn(II) affinity appears to be ten-fold tighter than wild-type FrmR (K_{Zn1-3} = 1.7 ± 0.7 x 10^{-10} M).

The difference in Zn(II) binding affinity between mutant and native protein can be further visualised in Figure 4. 10A -B where comparisons between mag-fura-2 (Figure 4. 10A) and quin-2 (Figure 4. 10B) zinc(II) competition in presence of FrmR (yellow symbols) or E64HFrmR (blue symbols) is shown. Control experiments were performed in the absence of protein and change in both mag-fura-2 or quin-2 spectral features saturated upon addition of 1 molar equivalent of Zn(II) confirming the formation of a 1:1 chelator:Zn(II) complex (empty symbols in Figure 4. 10A and B, respectively).

4.5.2 *Measurement of the Co(II)-binding affinity of E64HFrmR by competition with fura-2*

To measure E64HFrmR Co(II) affinity, competition was initially carried out with the fluorescent metal chelator fura-2 which has a K_{Co}^{fura-2} = 8.64 x 10^{-9} M and therefore has been previously used to determine protein Co(II) affinities in this range (Iwig *et al.* 2008; Patterson *et al.* 2013). E64HFrmR (49.3 µM) was competed with fura-2 (9.8 µM) for Co(II)-binding and quenching of fura-2 fluorescence emission at 510 nm was monitored as shown in Figure 4. 11A.

Since quenching occurred upon addition of ~ 20-22 µM Co(II), consistent with only one site per protein tetramer being filled, the data were fit to a model describing competition from a single site per tetramer (see Appendix, for details). The fitted curve significantly departs from simulated curves describing K_{Co1} ten-fold tighter and ten-fold weaker (Figure 4. 11A). K_{Co1} was calculated to be 2.56 ± 0.4 x 10^{-7} M (from triplicate assays).

In Chapter 3 it was discussed that FrmR is not capable of competing with fura-2 for Co(II) binding, in fact the data points from the competition assay completely overlay those from Co(II) titration of fura-2 alone (Figure 3. 12A). To test the overall order of magnitude of Co(II) affinity for the four binding sites in order to compare it with that calculated for FrmR, competition with a large excess of BisTris (50 mM, $K_{Co(II)}^{BisTris}$ = 2.26 x 10^{-2}M) was performed (Figure 4. 11B). The fitted curve describing binding of four Co(II) ions with equal affinity per tetramer (see Appendix, for details) significantly departs from simulated curves describing K_{Co1-4} 10-fold weaker, but not from the curve describing the upper limit (K_{Co1-4} tighter). For this reason only a weaker limit of $K_{Co1-4} < 10^{-6}$ M can be determined. This outcome, in combination with the

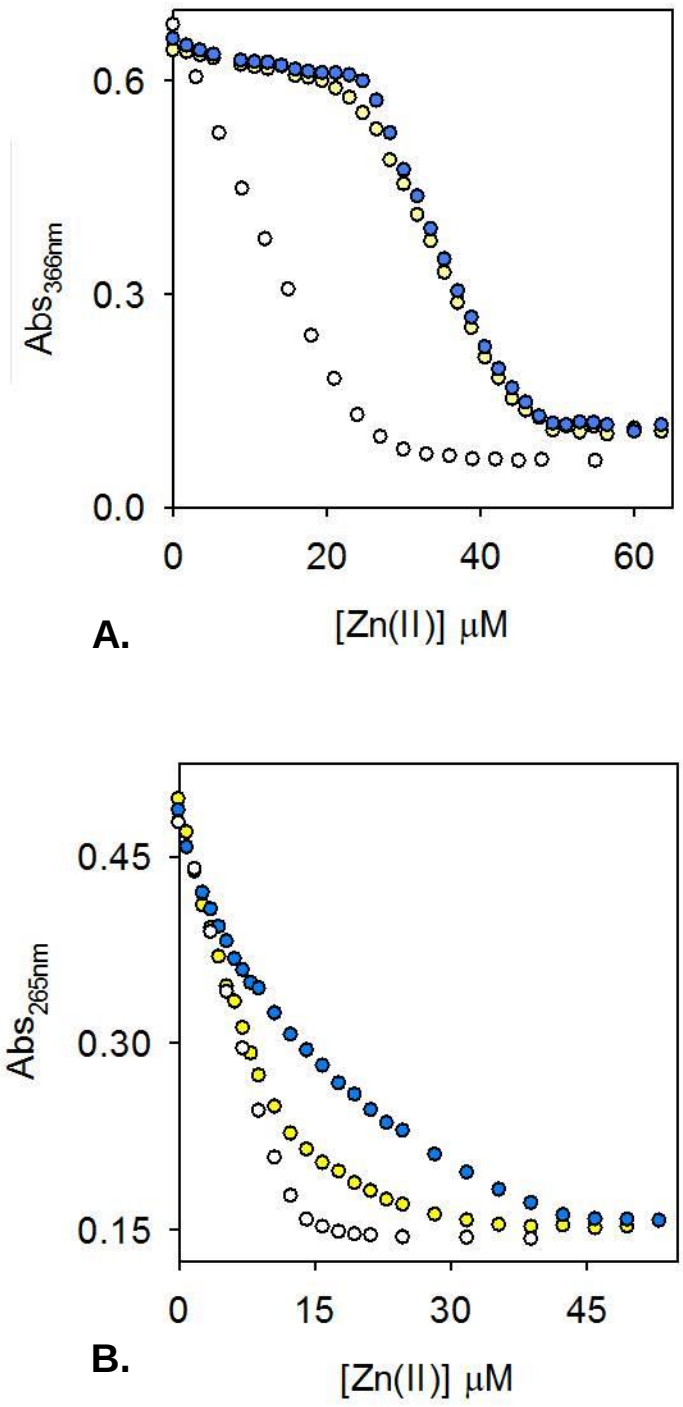


Figure 4. 10 A Comparison between mag-fura-2 absorbance upon titration of mag-fura-2 (30.64 µM) with $ZnCl_2$ in the presence of E64HFrmR (blue circles, c = 32.56 µM) and FrmR (yellow circles, c = 33.03 µM). Control: white circles represent mag-fura-2 titrated with Zn(II) in the absence of protein. **B** Comparison between quin-2 absorbance upon titration of quin-2 (13.4 µM) with $ZnCl_2$ in the presence of E64HFrmR (blue circles, c = 47.32 µM) and FrmR (yellow circles, c = 39.9 µM) as previously shown in Figure 4. 9B and Figure 3. 12B. Control: white circles represent quin-2 titrated with Zn(II) in the absence of protein.

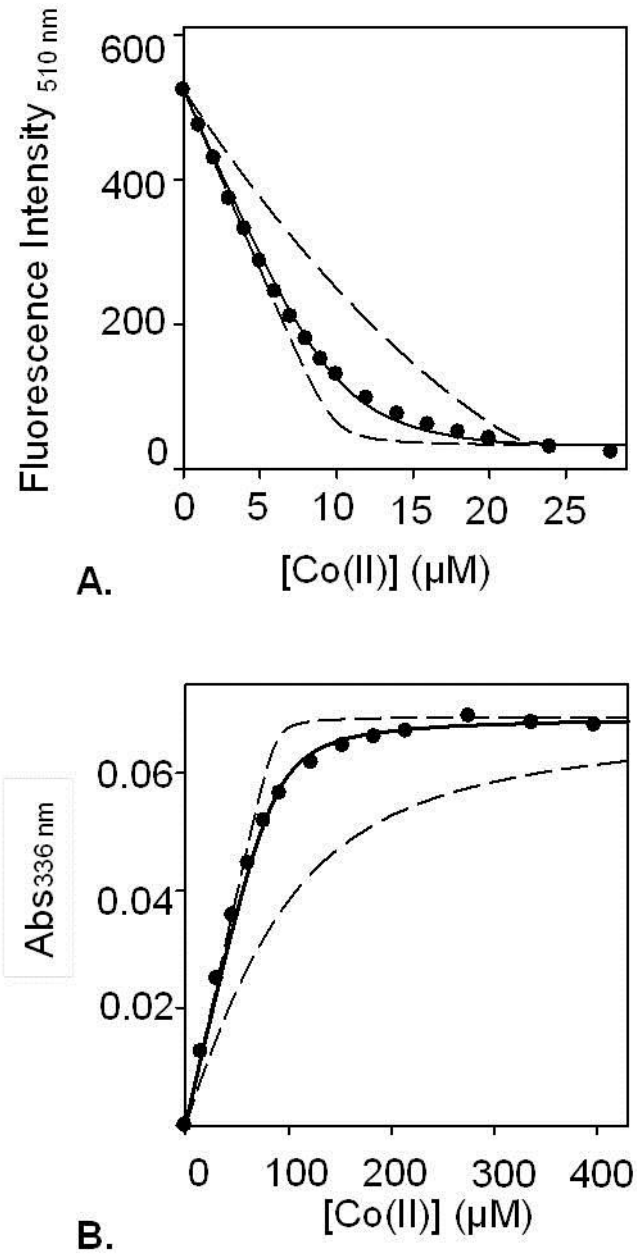


Figure 4. 11 A Representative (n = 3) of fura-2 (9.8 μM) titration with $CoCl_2$ in the presence of E64HFrmR (49.3 μM). Fluorescence emission (λ_{ex} = 360 nm; λ_{em} = 510 nm) was measured by a fluorescence spectrophotometer. The model (solid line) describes a E64HFrmR monomer competing for 0.25 molar equivalents of Co(II) (1 site per tetramer, K_{Co1}). Simulation of K_{Co1} 10-fold tighter and weaker is represented by dashed lines. **B** Representative (n = 3) of E64HFrmR (87.0 μM) titrated with $CoCl_2$ in the presence of 50 mM BisTris. Variation of the absorbance at 336 nm (diagnostic of the LMCT feature) was monitored by a UV-vis spectrophotometer. The model (solid line) describes a E64HFrmR monomer competing for 1 molar equivalents of Co(II) (4 sites per tetramer, K_{Co1-4}). Simulation of K_{Co1} 10-fold tighter and weaker is represented by dashed lines. These data have obtained by Dr. Deenah Osman.

~10^{-7} M value calculated for the tightest site, suggests an increase of ~ one order of magnitude of Co(II) binding affinity in E64H mutant.

4.5.3 *Measurement of the Cu(I) binding affinity of E64HFrmR by competition with bicinchoninic acid (BCA) and bathocuproine disulfonate (BCS)*

Cuprous copper affinity was examined using BCA (bicinchoninic acid $\beta_2 = 10^{17.2}$ M^{-2}) and BCS (bathocuproine disulfonate, $\beta_2 = 10^{19.8}$ M^{-2}) as done for wild-type FrmR (Figures 3. 13-3. 14). BCA is a derivative of 2-2'-bipyridyl and titration with Cu(I) induces an intense absorbance at 562 nm upon formation of $[Cu^{I}BCA_2]^{3-}$ complex (Xiao & Wedd 2010).

Cuprous copper (> 95 % Cu(I)) titration of BCA (40 µM) in the presence of E64HFrmR (11 µM, protomer) revealed negligible increase of absorbance up to ~ 6 µM Cu(I) (Figure 4. 12A-D). After this point, subsequent addition of Cu(I) gave an increase of absorbance at 562 nm with saturation reached at ~ 40 µM consistent with competition from E64HFrmR for 2 molar equivalents of Cu(I). Consequently, data were fit to a model describing binding of eight Cu(I) ions per tetramer, with best fit obtained by describing $K_{Cu1\text{-}2} < K_{Cu3\text{-}4} < K_{Cu5\text{-}6} < K_{Cu7\text{-}8}$. $K_{Cu1\text{-}2}$ appears too tight to be measured by this assay as implied by the close proximity of fit and simulated 10-fold tighter and weaker curves (black line and red solid lines, respectively (Figure 4. 12A). The dashed red line in Figure 4. 12A represents a simulated curve describing $K_{Cu1\text{-}2}$ 100-fold weaker than the calculated value when the remaining K_{Cu} are fixed to their optimised values. Figure 4. 12B-C-D show solid and dashed lines, the first representing the optimised value of $K_{Cu3\text{-}4}$, $K_{Cu5\text{-}6}$ and $K_{Cu7\text{-}8}$ respectively, whereas the second describes affinities 10-fold tighter and 10-fold weaker than the calculated value where the remaining K_{Cu} stay fixed to their optimised value (see Figure 4. 12 footnotes for details). The red dashed lines in Figure 4. 12D also describe $K_{Cu7\text{-}8}$ 100-fold tighter than the calculated value. These data demonstrate that $K_{Cu3\text{-}4}$ and $K_{Cu5\text{-}6}$ depart from simulated upper and lower limits whereas $K_{Cu1\text{-}2}$ and $K_{Cu7\text{-}8}$ are too tight and too weak, respectively, to be determined by this assay, and only limits can be defined.

Competition was then carried out between E64HFrmR (29.7 µM, protomer) and BCS (10µM), which binds Cu(I) more tightly ($\beta_2 = 10^{19.8}$ M^{-2}), and monitored by change in absorbance at 483 nm. However, inflection of the signal was observed at $[Cu^{I}]:[L] \approx 0.5$, indicating formation of the $[Cu^{I}BCS_2]^{3-}$ complex and suggesting little or no competition from E64HFrmR (Figure 4. 13). This finding implies that the E64H mutant can only marginally depart from the value estimated using BCA ($K_{Cu1\text{-}2}$ ~ 5 x 10^{-16} M).

Control experiments (copper titrated in chelator sample in the absence of protein) were routinely conducted on the day of the experiment, with the same concentration of both BCA and BCS used in the competition assay, always confirming formation of the $[Cu^{I}L_2]^{3-}$ complex

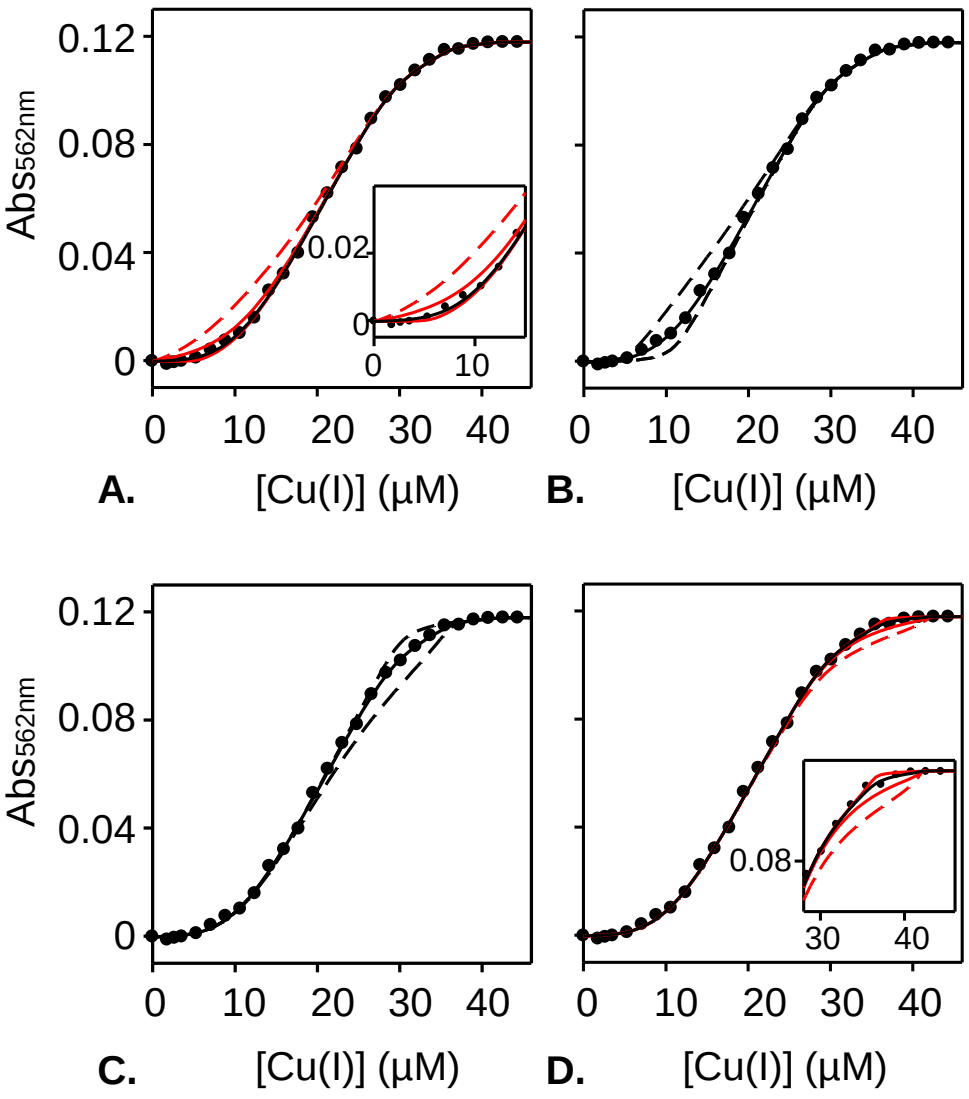


Figure 4. 12 A-D Representative (n = 3) BCA absorbance upon titration of BCA (40 μM) with Cu(I) in the presence of E64HFrmR (11 μM). Increase in absorbance at 562 nm (diagnostic of the $BCA_2Cu(I)$ complex) was observed only after addition of ~ 6 μM Cu(I), with an inflection at ~ 42 μM total Cu(I). Solid black line represents a fit to a model describing competition from E64HFrmR for 2 molar equivalents of Cu(I) (see Appendix for details). Best fit were obtained by pairing the eight sites in the following way: $K_{Cu1\text{-}2} < K_{Cu3\text{-}4} < K_{Cu5\text{-}6} < K_{Cu7\text{-}8}$. Each panel shows the fits to a model where each K_{Cu} is fixed to its optimised values except for one, for which simulated curves representing K_{Cu} 10- or 100- fold tighter and/or weaker than optimised value are shown. **A** $K_{Cu1\text{-}2}$ (solid red line_10-fold tighter and 10-fold weaker; dashed red line_ 100-fold weaker); **B** $K_{Cu3\text{-}4}$ (dashed black lines_10-fold tighter and 10-fold weaker); **C** $K_{Cu5\text{-}6}$ (dashed black lines_10-fold tighter and 10-fold weaker); **D** $K_{Cu7\text{-}8}$ (solid red lines_10-fold tighter and 10-fold weaker; dashed red line_100-fold tighter). Insets: enlargement of specific titration area. These data have been collected by me and analysed in conjunction with Dr Deenah Osman.

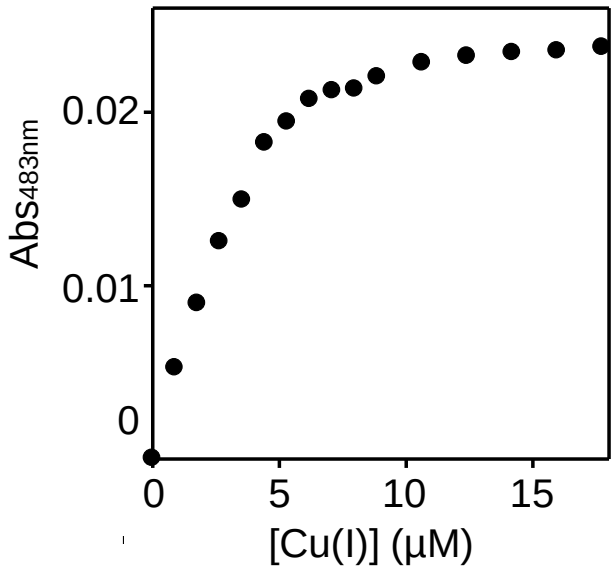


Figure 4. 13 Representative (n=4) BCS absorbance upon titration of BCS (10 µM) with CuCl [> 95% Cu(I)] in the presence of E64HFrmR (29.7 µM, protomer). Addition of Cu(I) gave an increase in absorbance at 483 nm (diagnostic of the $BCS_2Cu(I)$ complex) with an inflection at ~ 5 µM total Cu(I) indicating little, or no competition from E64HFrmR for Cu(I).

(Figure 8. 4, Appendix).

4.6 Relative metal binding affinities of wild-type FrmR and E64HFrmR

Wild-type FrmR and E64HFrmR relative metal affinities have been tested across a range of *Salmonella* metal sensors with the aim of elucidating why FrmR does not sense cellular metals and if E64HFrmR can compete with the Zn(II)-sensor ZntR for Zn(II), in order to explain the observed de-repression of *β-galactosidase* expression previously discussed in Section 4. 2 (Figure 4. 2B).

4.6.1 *Metal-binding competition between FrmR and Salmonella Zn(II)-, Co(II)- and Cu(I)-sensors ZntR, RcnR and CueR*

The experiments presented in this section were performed by Dr Deenah Osman. If metal-sensing is dictated by relative affinity within the set of *Salmonella* metal-sensors, the affinity of E64HFrmR for Zn(II) and Co(II) would need to become comparable to cellular sensors for these metals. Conversely, Cu(I) affinity would need to remain weaker than Cu(I)-sensing CueR making Cu(I) still undetectable (Osman *et al.* 2010, 2013; Espariz *et al*. 2007). The Zn(II) and Co(II) sensors in *Salmonella* are ZntR and ZuR (Zn(II)), and RcnR (Co(II)) (Ammendola *et al.* 2014; Petrarca *et al.* 2010).

To confirm, or otherwise, that FrmR K_{metal} is weaker than CueR $K_{Cu(I)}$, ZntR $K_{Zn(II)}$ and RcnR $K_{Co(II)}$, pair-wise competitions were conducted for the tightest metal-binding site. FrmR (40 µM, protomer) was incubated with CuCl (10 µM, in order to fill only the tightest site) and resolved by heparin affinity chromatography. Cuprous-FrmR solution was applied to a 1 ml Heparin column equilibrated with buffer C100 (Table 8. 1, Appendix) and eluted in buffer C500 (Table 8. 1, Appendix) in fractions F7-F9. Protein and metal elution profiles confirm co-migration with copper (Figure 4. 14C). An analogous result was obtained when Cu(I)-CueR ([CueR] = 20 µM [Cu(I)] = 10 µM) was applied to the column and eluted in C300 (Table 8. 1, Appendix) although CueR was found in a broader range of fractions, ~F8-F15 (Figure 4. 14B). When Cu(I)-FrmR was then mixed with apo-CueR, all the copper migrated with CueR demonstrating that, under these conditions, FrmR is not able to compete (Figure 4. 14A). Likewise, after mixing Zn(II)-FrmR ([Zn(II)] = 10 µM, [FrmR] = 40 µM) with apo-ZntR (20 µM), the elution profile was consistent with zinc predominantly migrating with ZntR (Figure 4. 15A-B-C).

Since RcnR and FrmR share the same elution profile when resolved by heparin affinity chromatography, but bind Co(II) assuming different coordination geometries, appearance or otherwise of *d-d* transitions upon Co(II) addition was monitored by UV-visible spectroscopy in order to discriminate between metal-binding to the two proteins (Figure 4. 16). FrmR (56 µM)

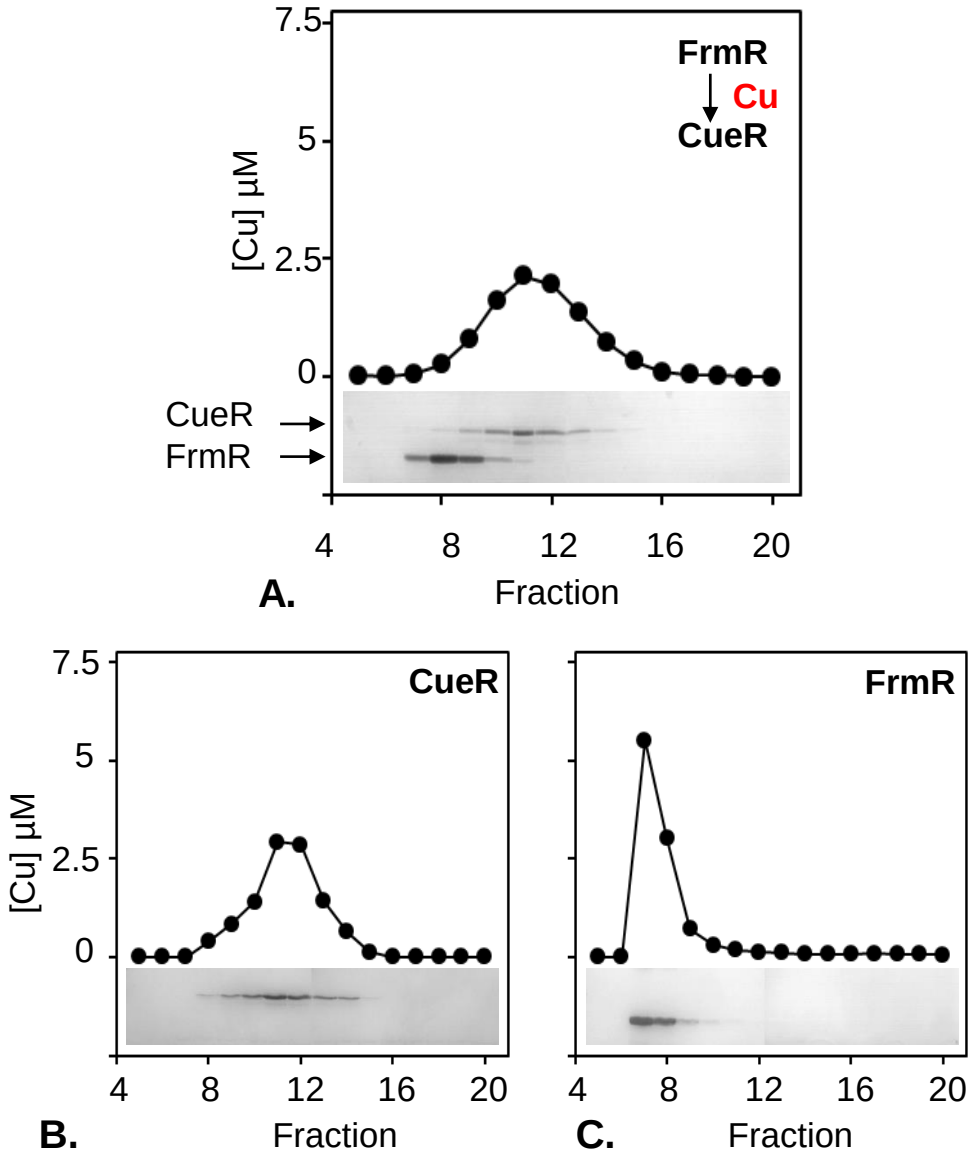


Figure 4. 14 A Representative (n =3) of Cu(I)-binding competition between FrmR and CueR. FrmR (40 µM) equilibrated with Cu(I) (10 µM) and after subsequent addition of apo-CueR (20 µM). Proteins were separated by heparin affinity chromatography and fractions (1 ml) assayed for copper by ICP-MS and protein by SDS-PAGE. FrmR was present predominantly in F7-F9 whereas CueR eluted over a broader range of fractions. **B** Cu(I) loading and chromatography of CueR as described in A except without addition of apo-FrmR. **C** Cu(I) loading and chromatography of FrmR. These data have been obtained by Dr. Deenah Osman.

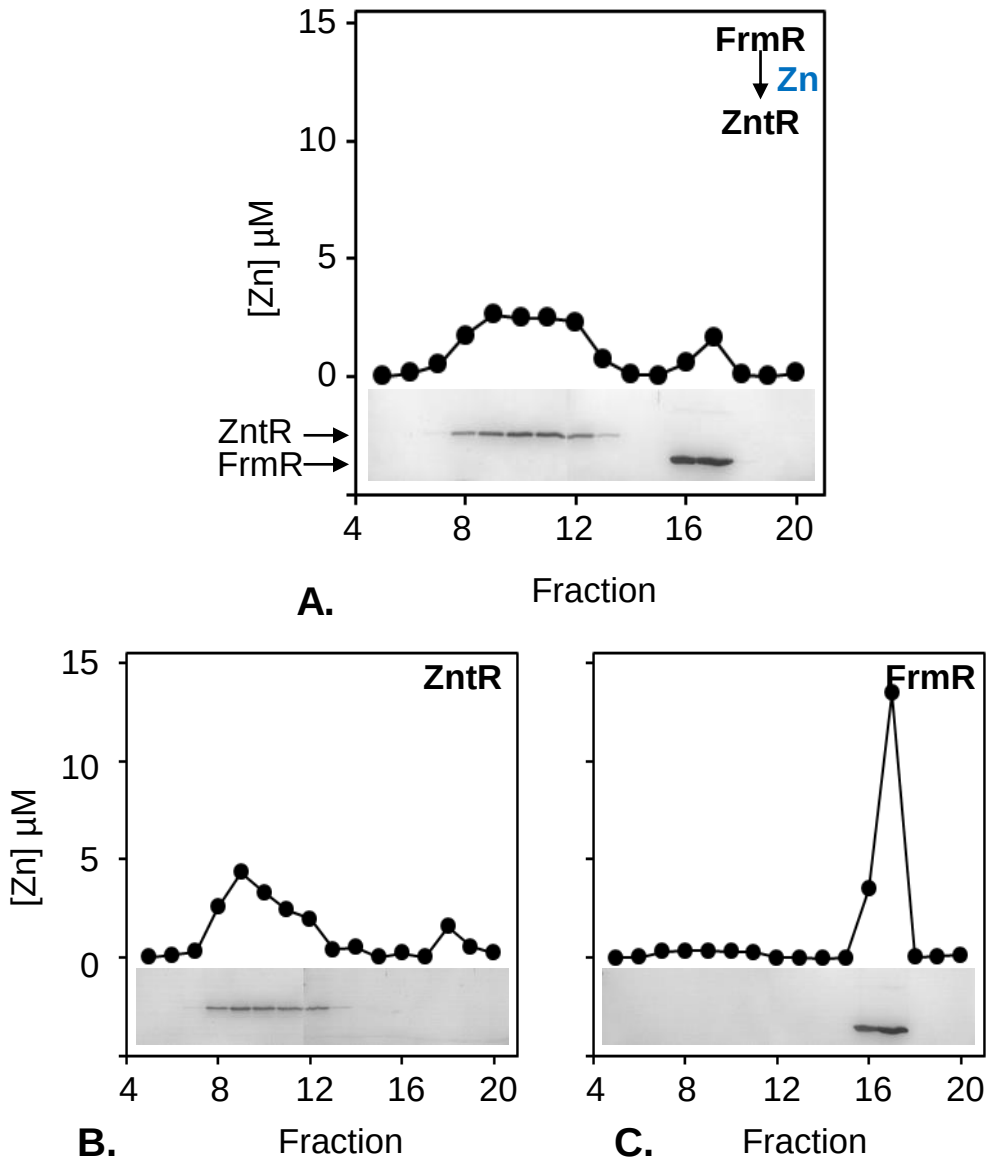


Figure 4. 15 A Representative (n =3) of Zn(II)-binding competition between FrmR and ZntR. FrmR (40 µM) equilibrated with Zn(II) (10 µM) and after subsequent addition of apo-ZntR (20 µM). Proteins were separated by heparin affinity chromatography and fractions (1 ml) assayed for zinc by ICP-MS and protein by SDS-PAGE. ZntR does not bind the column and is present in the flow-through (F8-F12) and FrmR elutes predominantly in F16 and F17. **B** Zn(II) loading and chromatography of ZntR as described in A except without addition of apo-FrmR. **C** Zn(II) loading and chromatography of FrmR. These data have been obtained by Dr. Deenah Osman.

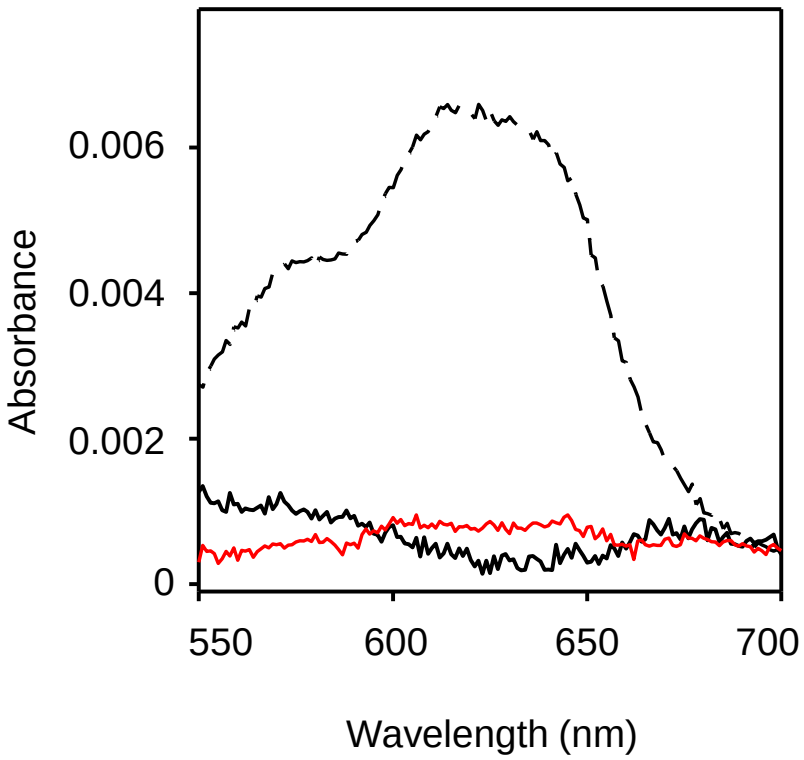


Figure 4. 16 Representative (n =3) of Co(II) (II)-binding competition between FrmR and RcnR. The apo-subtracted difference spectra of RcnR (62 μM; solid black line), FrmR (56 μM; dashed black line), or a mixture (red line) of RcnR (62 μM) and FrmR (56 μM) upon addition of 14 μM Co(II). These data have been obtained by Dr. Deenah Osman.

was incubated with enough Co(II) to fill the first site (14 µM) and formation of a tetrahedral complex was confirmed by the presence of spectral features in the *d-d* transitions region (apo-subtracted difference spectrum of the region is represented as a black dashed line in Figure 4. 16). When RcnR (62 µM) was added to the reaction mixture, the aforementioned spectral feature was lost, indicating formation of Co(II)-RcnR octahedral complex (red line in Figure 4. 16) and therefore establishing that RcnR outcompeted FrmR.

4.6.2 *Salmonella Zn(II)-sensors ZntR competition with E64HFrmR for zinc binding*

To test if the ~ 10-fold increase in Zn(II) affinity would enable E64HFrmR to compete with the *Salmonella* Zn(II) sensor ZntR, the same approach adopted in Section 4.6.1 with FrmR was adopted here. The tightest site of E64HFrmR was anaerobically metallated with Zn(II) ([E64HFrmR] = 56 µM, [$ZnCl_2$] = 14 µM) in buffer C500 (Table 8. 1, Appendix). Zinc-loaded protein was then diluted in buffer C100 (Table 8. 1, Appendix) and applied to a 1 ml Heparin column pre-equilibrated with the same buffer under anaerobic conditions. Elution was carried out with buffer C500 and zinc(II)-E64HFrmR eluted predominantly in fraction F17 (as FrmR) (Figure 4. 17C). ZntR does not bind to Heparin and exhibits a broad elution profile centered at around F11 (Figure 4. 17B). When apo-ZntR (28 µM) is mixed with cuprous-E64HFrmR (56 µM of E64HFrmR equilibrated with 14 µM Zn(II)), the resulting elution profile indicates migration of Zn(II) from E64HFrmR to ZntR (Figure 4. 17A). Although the $K_{Zn(II)}$ of E64HFrmR approaches that of cognate sensors (ZntR and ZuR), ZntR still outcompetes E64HFrmR for zinc binding *in vitro*, suggesting that relative affinity alone might not be sufficient to explain the gain-of-function *in vivo*.

4.7 Contribution of glutathione to metal-sensing by FrmR and E64HFrmR

In addition to responding to cellular zinc, E64HFrmR responds to Co(II), although its $K_{Co(II)}$ is ~ 500-fold weaker than the endogenous cobalt sensor RcnR, which outcompetes the FrmR variant as shown in Figure 4. 16. An ~ 10-fold increase in $K_{Co(II)}$ alone cannot explain why this mutant has become responsive to cobalt. Recent studies conducted on metal sensors from a cyanobacterium showed that Ni(II) partitions to the sensor of tightest affinity whereas Zn(II) sensing is based on relative allostery, therefore detection of these metals follows thermodynamic predictions (Foster *et al*. 2012, 2014). Cobalt detection by CoaR, however, does not follow relative affinity or relative allostery (Patterson *et al.* 2013). To explain why this metal is detected by the weakest Co(II)-binding sensor in the set a substantial kinetic component was invoked. In fact the study suggested that Co(II) ions, free and/or inserted in tetrapyrroles, may be channeled to the protein (Patterson *et al.* 2013).

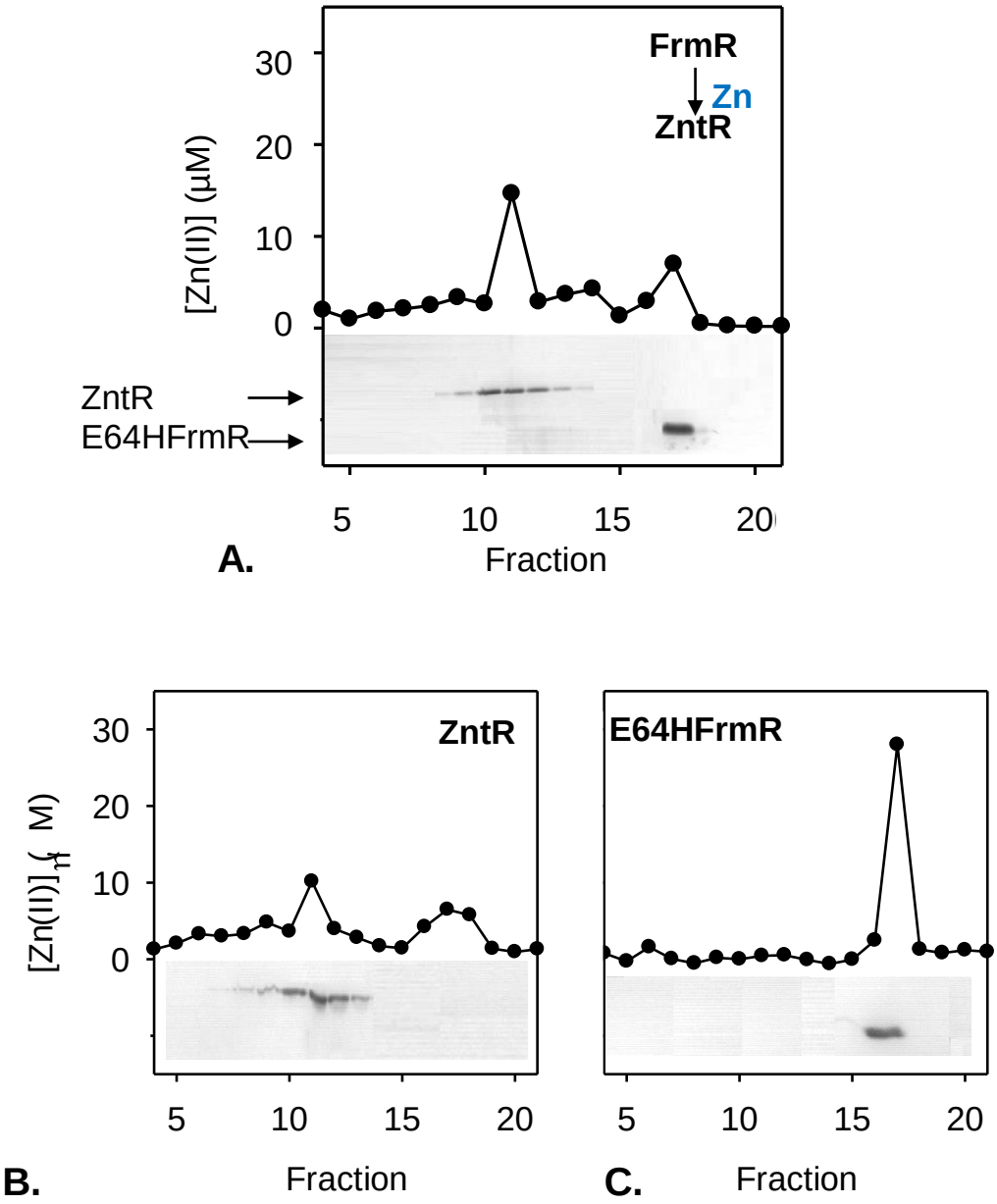


Figure 4. 17 A E64HFrmR (56 μM) equilibrated with Zn^{2+} (14 μM) and after subsequent addition of apo-ZntR (28 μM). Proteins were separated by heparin affinity chromatography and fractions (1 ml) assayed for zinc by ICP-MS and protein by SDS-PAGE. ZntR does not bind the column and is present in the flow-through (F8-F13) and E64HFrmR elutes predominantly in F17. **B** Zn^{2+} loading and chromatography of ZntR as described in A except without addition of apo-FrmR. **C** Zn^{2+} loading and chromatography of E64HFrmR.

Since the tripeptide glutathione (GSH) is well known for its role in metal homeostasis (see Section 7.4.3) (Wang & Ballatori 1998), the following sections will explore its contribution to metal detection by E64HFrmR *in vitro*.

4.7.1 *Assessing GSH contribution to zinc sensing by FrmR and E64HFrmR*

Since $K_{Co(II)}$ is weak even for E64HFrmR (2.56 ± 0.4 x 10^{-7} M for the tightest site), the contribution of glutathione to metal binding by FrmRE64H was analysed only for Zn(II) by competition with mag-fura-2. Migration of zinc(II) from mag-fura-2 (100 µM), previously incubated with a 1:1 molar ratio of Zn(II), to FrmR or E64HFrmR (Figure 4. 19A-B, respectively) was carried out monitoring increase in the spectral feature at 366 nm by UV-visible spectroscopy. The experiment was conducted in the presence (red line) or absence (black line) of 20 µM glutathione (Figure 4. 19A-B). "TimeDrive" software was employed in order to detect any difference in Zn(II) acquisition rates between FrmR and E64HFrmR. Since protein was added to the Zn(II)-mag-fura-2 sample inside the anaerobic glove-box and followed by scanning with an UV-visible spectrophotometer outside, it was possible to monitor change in absorbance only after ~ 30 sec from protein addition (Figure 4. 19A-B).

Mag-fura-2 titration with Zn(II) in the absence of protein produced a sharp inflection upon addition of one molar equivalent of metal, confirming formation of a 1:1 [M:L] complex, both in the presence or absence of glutathione (Figure 4.18A). If protein is added to mag-fura-2 fully loaded with Zn(II) (which does not possess any spectral feature at 366 nm), it is possible to monitor the formation of Zn(II):protein complex by observing rising of the peak at 366 nm. Since upon addition of 5 mM EDTA at the end of the time-course experiment the absorbance of the spectral feature still increases in both wild-type and mutant FrmR, not all the Zn(II) seems to migrate from the chelator to the protein (Figure 4. 18B-C).

When the experiment was carried out in the presence of glutathione, a negligible difference, if any, was noticed in the curve slope if compared to that obtained in the absence of glutathione, suggesting that glutathione does not aid transfer of Zn(II) from mag-fura-2 inferring that proteins receive Zn(II) directly from the chelator (Figure 4. 19A-B).

However, if comparison is made between experiments with wild-type FrmR and E64H variant, in the presence or absence of GSH, in both cases Zn(II) ions migrate faster to E64HFrmR than to wild-type (Figure 4. 20A-B). An explanation could be that E64HFrmR conformation in the apo-form is more open, with ligands already placed in a better position to receive the metal or, possibly, mag-fura-2 under these conditions acts as glutathione in the cell, handing Zn(II) to the E64H variant more promptly than to wild-type for reasons not yet clear. Both speculations could be valid and do not exclude each other.

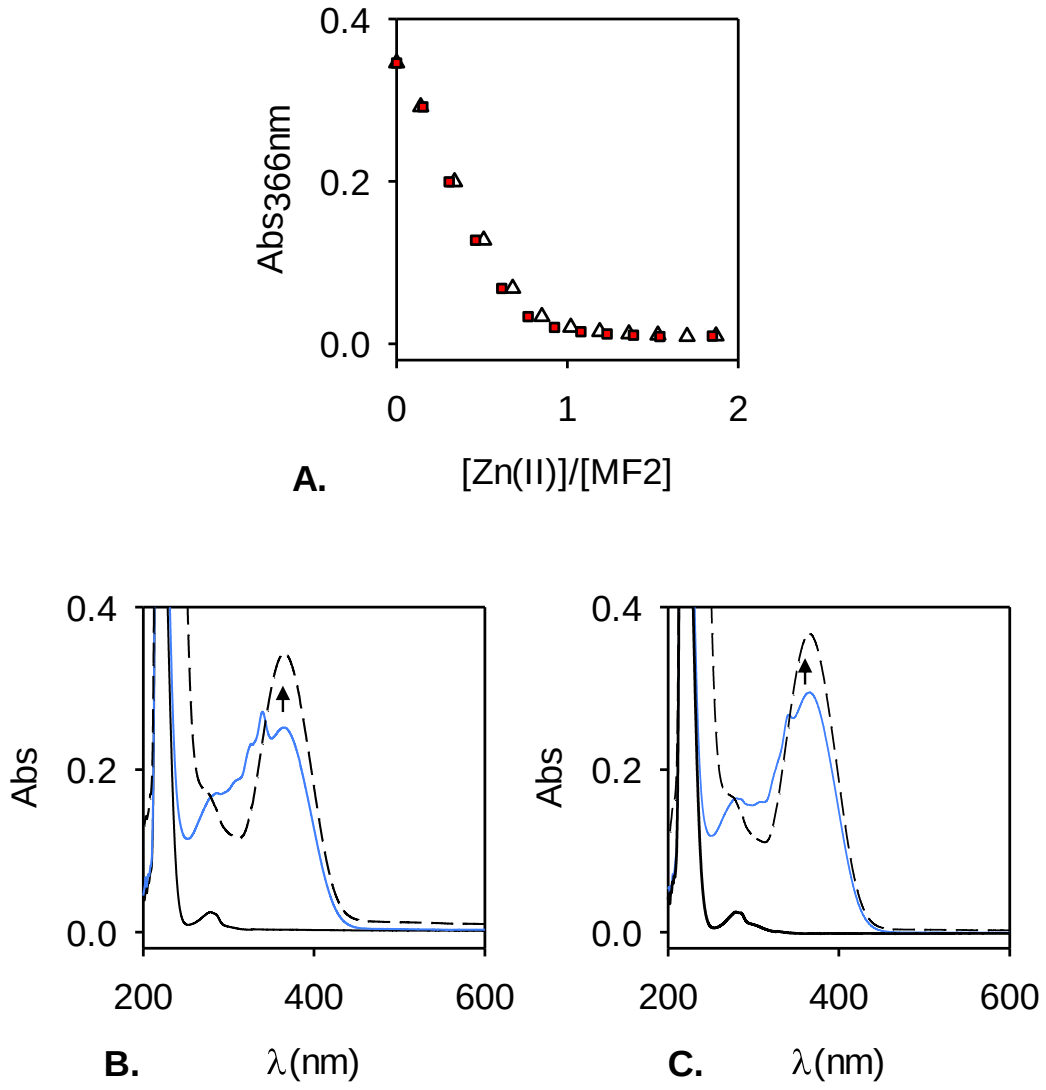


Figure 4. 18 A Mag-fura-2 titrated with $ZnCl_2$ in the presence (red squares) or absence (open triangles) of 20 µM glutathione. **B-C** UV-vis spectra of FrmR (B) or E64HFrmR (C) (solid black line), end of competition between mf2 and protein for Zn(II)-binding (blue solid line) and upon addition of 5 mM EDTA (dashed black line).

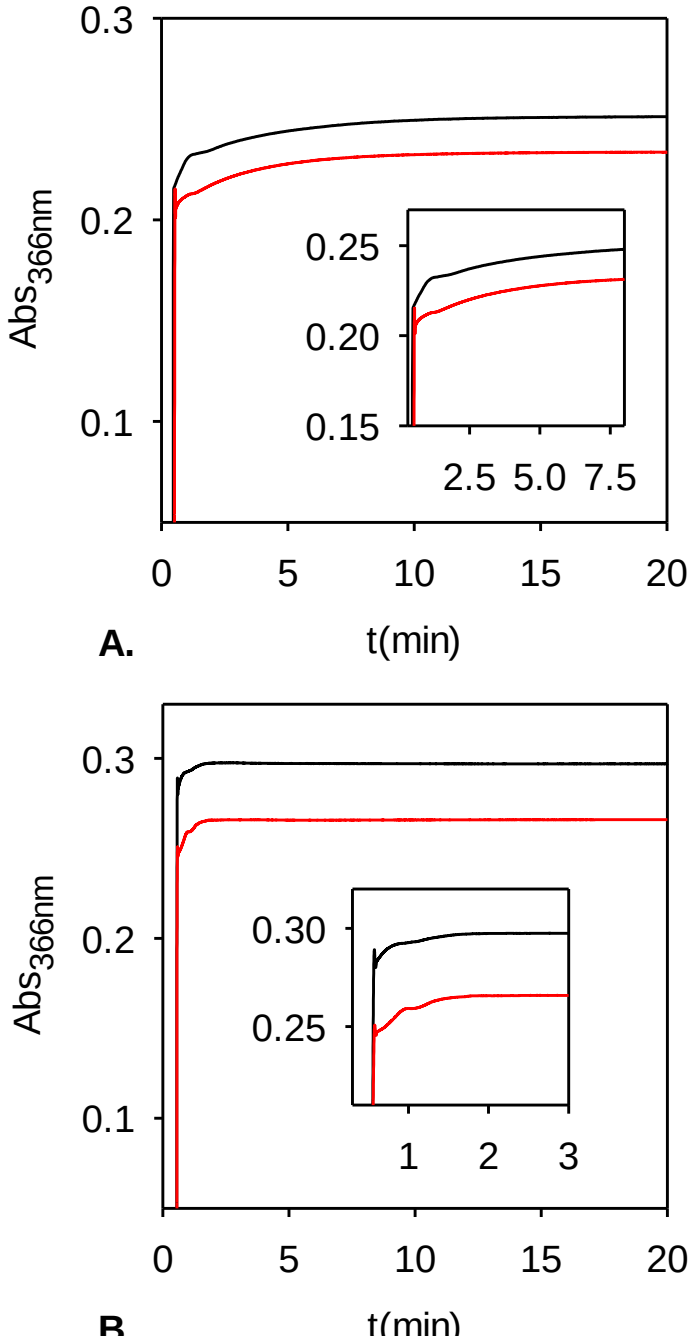


Figure 4. 19 A-B Time course experiments monitoring migration of zinc from mag-fura-2 to FrmR (A) or E64HFrmR (B) in the presence (red lines) or in absence (black lines) of glutathione (20 µM). UV-visible absorbance was recorded at 366 nm using "TimeDrive" software available on the instrument. Mag-fura-2 (100 µM) previously incubated with $ZnCl_2$ (100 µM) was added to FrmR (13.67 µM, monomer) (A) or E64HFrmR (13.73µM) (B). In each case sample preparation and addition were carried out in the anaerobic chamber therefore the change in absorbance, indicative of metal sequestration by the protein, was monitored after 28 sec in A and 33 sec in B.

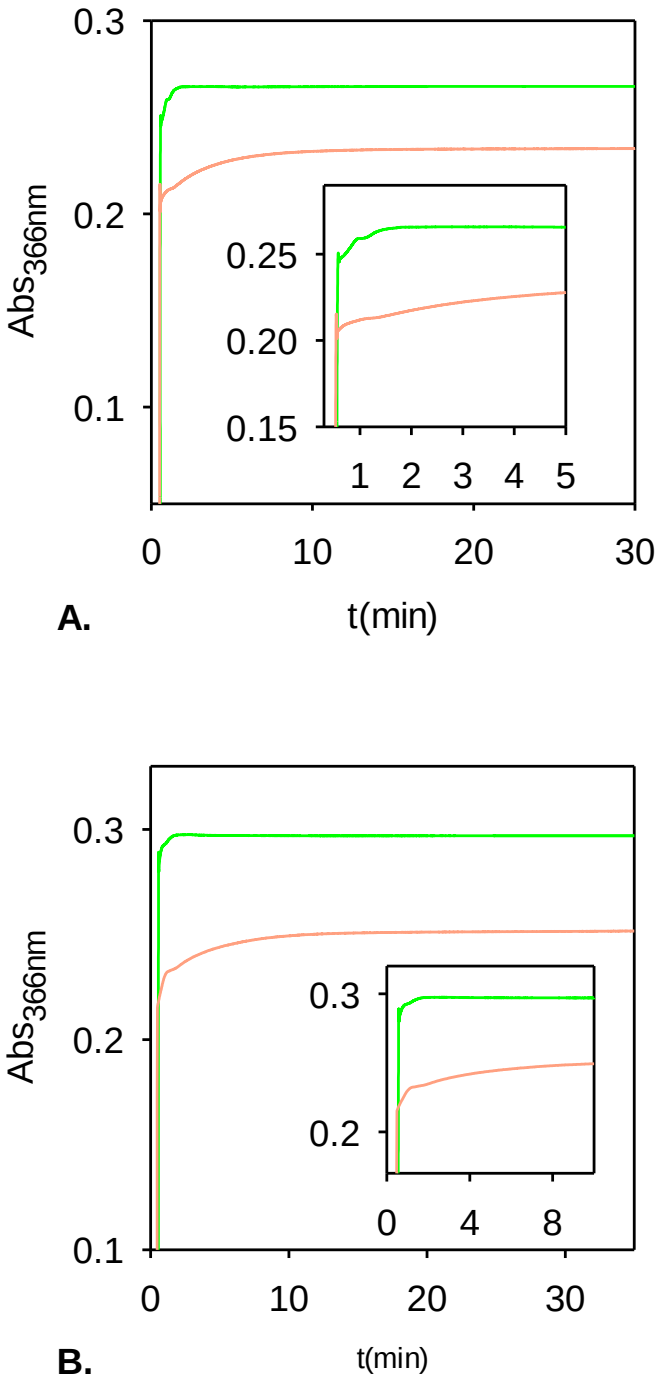


Figure 4. 20 A-B Time course experiments monitoring zinc migration from mag-fura-2 to FrmR (orange line) or E64HFrmR (green line) in the presence (A) or absence (B) of glutathione (20 µM).

4.8 *In vitro* interaction between E64HFrmR and *frmRA* promoter

In order to pursue the identification of the parameter (or set of parameters) regulating E64HFrmR gain-of-function, interaction with the target *frmRA* operator-promoter region was examined by fluorescence anisotropy. As briefly mentioned in Section 4.3, two distinct forms of purified E64HFrmR were discovered, the first characterised by an absorbance spectrum undistinguishable from wild-type (Figure 4. 4A) and the second, more frequently obtained, showing a spectral feature at ~ 300 nm which could not be related to oxidation nor to formaldehyde or glutathione modification (Figure 4. 4B). Preliminary studies conducted on the first form yielded metal affinities virtually identical to wild-type (data not shown). This singular and sporadic E64H form (~ 20 % of purified protein samples), named here as E64HFrmR*, was used to perform experiments to assess DNA binding affinity.

Figures 8. 6A-B and 8. 7A-B (Appendix) report findings obtained with this preparation, which exhibit analogous K_{DNA} to wild-type (106.6 ± 6.8 nM) and which is allosterically inhibited by zinc-binding by the same degree ($\Delta Gc^{Zn(II)E64HFrmR*}$ = 1.92 ± 0.09 kcalmol^{-1}). As it became clear that a more prevalent protein form was routinely purified, E64HFrmR* was set aside. Consequently, the following sections will discuss data obtained with this second and most frequently purified form of E64HFrmR.

4.8.1 *E64HFrmR:DNA stoichiometry determined by fluorescence anisotropy*

The fluorescently labelled, double-stranded DNA fragment used for wild-type FrmR, *frmRAPro*, was employed to determine E64HFrmR:DNA stoichiometry.

Titration of E64HFrmR into 2.5 µM *frmRAPro* was carried out anaerobically and monitored by fluorescence anisotropy. Inflection was visible upon addition of ~ 8 molar equivalents of protein confirming the model where two E64HFrmR tetramers bind to the target DNA region as observed with wild-type FrmR (Figure 4. 21A). The experiment was also conducted using *frmRA*Pro-T/A-mod, which differs from the original target sequence by having the first T/A site randomly altered (Figure 3 .18B). As with wild type FrmR, only one mutant tetramer binds this DNA, as indicated by the sharp inflection at 4 molar equivalents (Figure 4. 21B). However, it must be noted that the response in the latter case is significantly (almost 10-fold) smaller than that observed for *frmRAPro* which complicates interpretation of data.

4.8.2 *E64HFrmR:DNA interaction monitored by fluorescence anisotropy*

Titration of 10 nM *frmRAPro* with E64HFrmR was monitored by fluorescence anisotropy, reporting a weakened affinity for the target DNA sequence compared to FrmR (Figure 4. 22A). The DynaFit script, describing the binding of two non-dissociable protein-tetramers per DNA

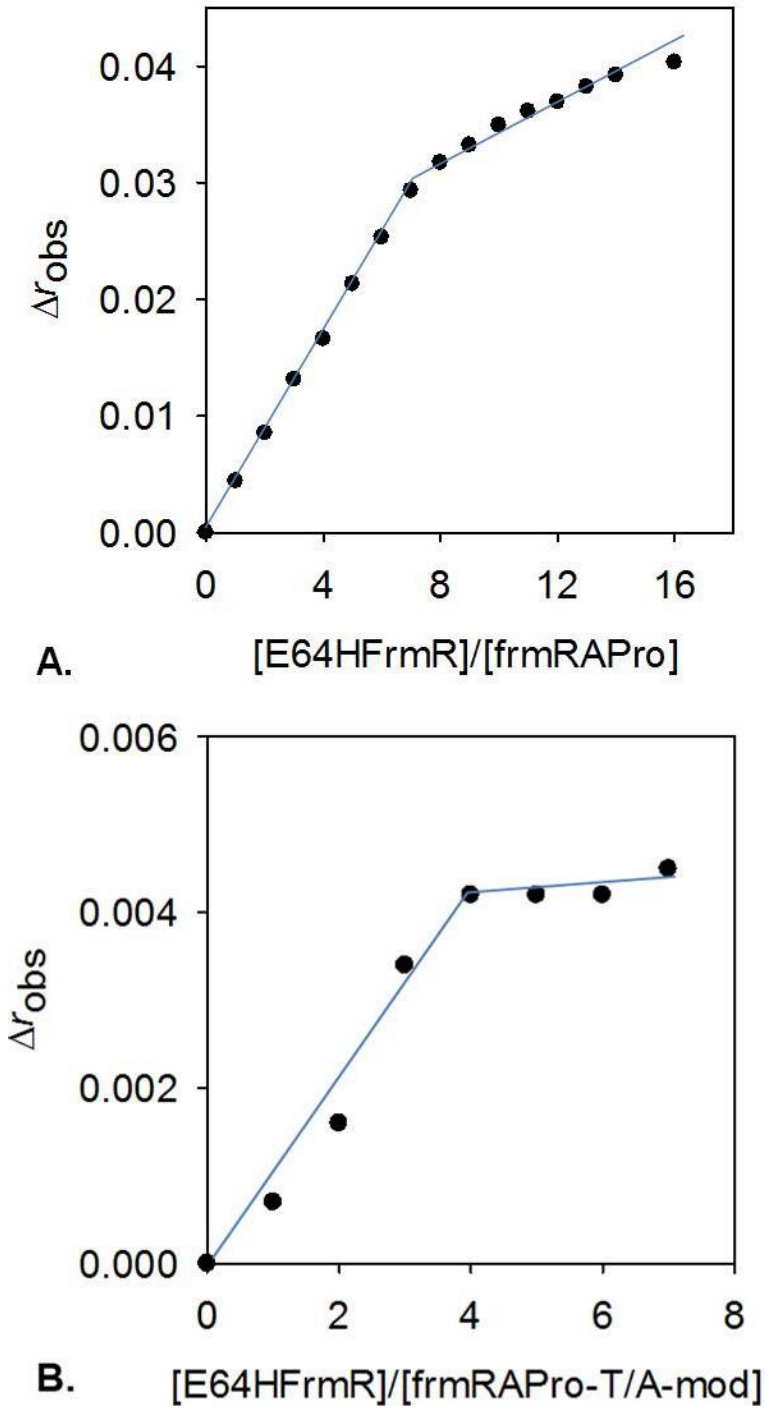


Figure 4. 21 A Anisotropy change upon titration of *frmRA*Pro (2.5µM) with E64HFrmR (2.5 µM monomer with each addition). A point of inflection is observed after addition of ~ 8 equivalents of protein monomer indicating a protein:DNA stoichiometry of 8:1 (two tetramers per *frmRA*Pro). The further increase in anisotropy following this inflection may represent further binding of FrmR to *frmRA*Pro with substantially weaker affinity. **B** As A except for the use of *frmRA*Pro-T/A-mod (DNA sequence containing the same promoter region as *frmR*Pro except for the introduction of a random nucleotide sequence in place of the first T/A tract). Inflection was observed after ~ 4 equivalents of protein monomer indicating a protein:DNA stoichiometry of 4:1 (one tetramer per *frmRA*Pro-T/A-mod).

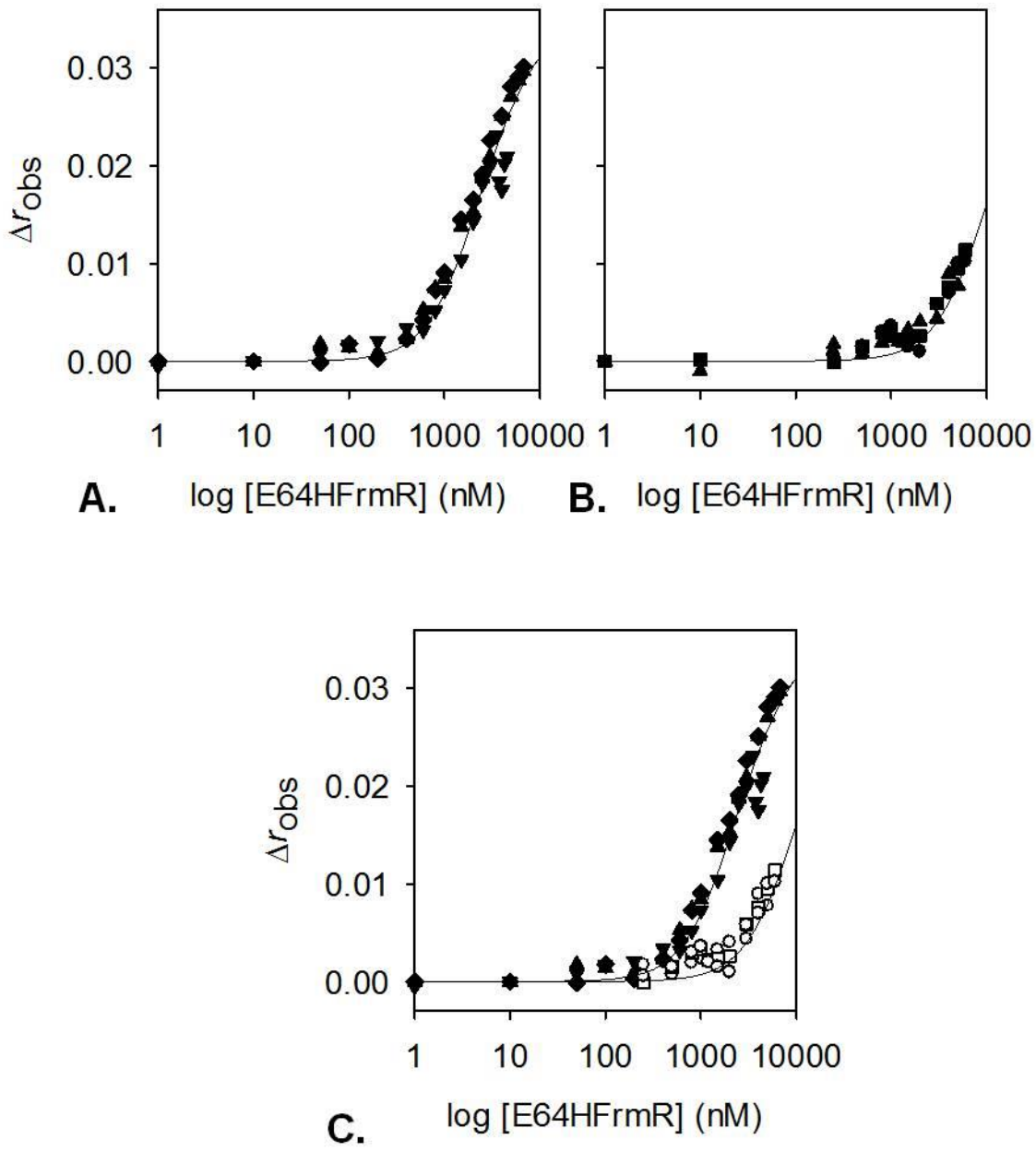


Figure 4. 22 A *frmRA*Pro (10 nM) was anaerobically titrated with E64HFrmR in the presence of 5 mM EDTA. **B** *frmRA*Pro (10 nM) was anaerobically titrated with E64HFrmR in the presence of 5 µM ZnCl. The protein was incubated with 1.2 molar equivalents of $ZnCl_2$ and EDTA was omitted. In A-B DNA binding was monitored by fluorescence anisotropy. Solid line represents simulated curves produced from the average K_{DNA} determined across the experiment replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 E64HFrmR tetramer (non-dissociable):DNA stoichiometry. **C** Data in A (filled symbols) and B (open symbols).

molecule, gave $K_{DNA}^{E64HFrmR}$ 426 ± 40 nM. This result might explain (at least partially) the enigmatic increase observed in the *frmRA* basal expression when regulated by FrmRE64H (Figure 4. 2B). If the mutant possesses a weaker affinity for its operator region than wild-type, this could result in the loss of repression observed in *β-galactosidase* assay.

4.8.3 *E64HFrmR:DNA interaction in the presence of Zn(II)*

The ability of Zn(II) to trigger the allosteric mechanism which alters DNA-binding was tested *in vitro* by fluorescence anisotropy. The experiment was carried out as discussed in the previous section except for the absence of EDTA and the addition of 5 µM $ZnCl_2$ to the reaction buffer and 1.2 molar equivalent of Zn(II) to the protein sample (to ensure fully metallated state throughout the experiment) (Figure 4. 22B). The association of E64HFrmR with *frmRA*Pro was greatly inhibited, consistent with previous observations with FrmR, ($K_{DNA}^{Zn(II)E64HFrmR}$ = 3510 ± 700 nM, $\Delta Gc^{Zn(II)E64HFrmR}$ = 1.24 ± 0.16 kcal/mol). The magnitude of allosteric regulation of Zn(II)-E64HFrmR is smaller than that observed for wild type FrmR ($\Delta Gc^{Zn(II)FrmR}$ = 2.03 ± 0.08 kcal/mol), which is a counterintuitive result considering that *in vivo* Zn(II)- and Co(II)-binding to E64HFrmR inhibits binding to DNA more effectively than for wild type FrmR (Figures 4. 1-4. 2).

However, it must be noted that, despite a smaller ΔG_c, $K_{DNA}^{E64HFrmR}$ (DNA-binding affinity of apoE64HFrmR) is more than four-times weaker than K_{DNA}^{FrmR} whereas Zn(II)-E64HFrmR has a DNA affinity comparable to Zn(II)-FrmR (Table 8.2, Appendix). Hence at equivalent Zn(II) saturation, DNA occupancy by wild type FrmR will be higher than E64HFrmR, with the result that the variant will be more sensitive to de-repression (Osman *et al.* 2015). This outcome is valid only for the FrmR variant which displays a different absorption spectrum to wild-type, whereas is not correct for the E64HFrmR* form (Table 8.2, Appendix). Nevertheless, if our experience in E64HFrmR production reflects the real proportion between the two variant forms in the cell, then most of the mutant has a weaker DNA-binding affinity when metallated with Zn(II) and will be more allosterically effective. Disruption of the pre-formed E64HFrmR:DNA complex upon Zn(II) titration was monitored by fluorescence anisotropy as previously described for FrmR (Section 3.6.6). Figure 4. 23 shows one experimental replicate where full dissociation occurs upon addition of 0.5 - 1 molar equivalents of Zn(II), however great variability was observed amongst multiple replicates precluding the possibility of determining how many metallated sites are needed to elicit DNA release.

4.9 Concluding remarks

This chapter has presented the first successful attempt to turn a non-metal sensor into a metal-sensor by a gain-of-function mutation. The WXYZ motif of FrmR was modified in order to reproduce an RcnR-like consensus sequence (HCHH). *B*-galactosidase assay shows that

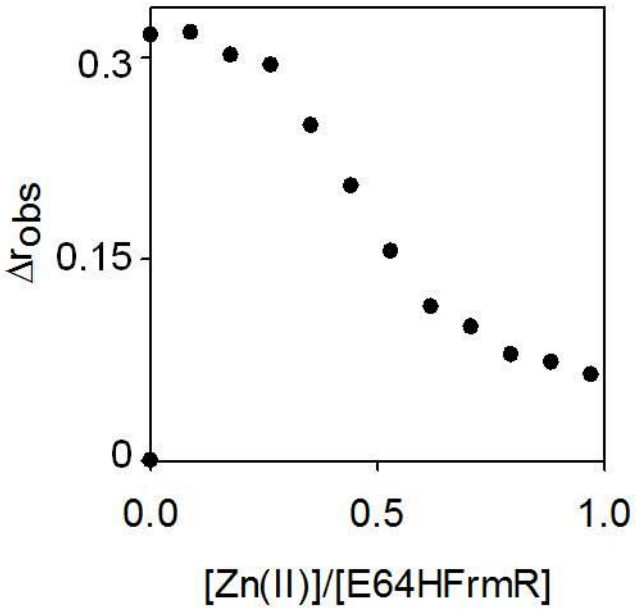


Figure 4. 23 *frmRA*Pro (10 nm) was pre-incubated with FrmR (2.5 μM) before titration with $ZnCl_2$. Dissociation of protein:DNA complexes was monitored by fluorescence anisotropy. Experiments were performed anaerobically at pH 7.0.

E64HFrmR is able to detect cellular Zn(II) and Co(II), in addition to formaldehyde. In order to explore the determinants underlying metal-specificity in FrmR proteins, the biochemical properties of E64HFrmR were analysed and compared to those of FrmR in Chapter 3 (see Section 7.4 for further discussion).

In addition to Zn(II)- and Co(II)-binding properties also Ni(II)- and Cu(I)-coordination was explored since RcnR and CsoR sense these metal ions. E64HFrmR binds Zn(II), Co(II) and Cu(I) with ~ one order of magnitude tighter affinity than wild-type and the inability of Zn(II) to quench Tyr89 intrinsic fluorescence suggests a difference in the spatial disposition of this residue and, maybe, of the ligands involved in Zn(II)-coordination. However, a slightly tighter Zn(II)-binding affinity alone may not be sufficient to allow E64HFrmR to adequately compete for Zn(II)-binding with *Salmonella* Zn(II)-sensor ZntR *in vivo*. Likewise, E64HFrmR' detection of cellular Co(II) cannot be explained by a tighter Co(II)-binding affinity since its $K_{Co(II)}$ is still ~ 500-fold weaker than the endogenous cobalt sensor RcnR (Section 7.4.1). The hypothesis that a kinetic component may be involved in metal-sensing was tested by exploring glutathione (one of the major components of the cellular buffer) role in Zn(II)-detection *in vitro* (Section 7.4.3).

Unexpectedly, E64HFrmR has a ten-fold weaker DNA-binding affinity (compared to FrmR), resulting in a smaller $\Delta Gc^{Zn(II)}$ and yet, is more allosterically effective. This result seemingly contradicts previous knowledge of metal-sensor transcriptional de-repressors (Grossoheme & Giedroc 2012) and will be further discussed in Section 7.4.2. Hence, gain of Zn(II)-binding function appears to be the result of two distinct contributions, a ten-fold tighter metal-binding affinity and ten-fold weaker K_{DNA}. As a consequence DNA occupancy of E64HFrmR is a hundred-fold smaller than wild-type (Section 7.4.2).

Chapter 5

Formaldehyde sensing by *Salmonella* FrmR and RcnR

β-galactosidase assays were planned, performed and interpreted by Dr. Deenah Osman. These contributions are noted in figure legends and/or in text.

5.1 Aims and objectives

FrmR has been shown to respond to formaldehyde *in vivo* (Chapter 4), consistent with *E. coli* FrmR (Law 2012; Herring & Blattner 2004). However, the mechanism of sensing of any FrmR remains unexplored. This chapter will explore further the FrmR formaldehyde responsiveness, *in vitro* and *in vivo*, by *β-galactosidase* assay and fluorescence anisotropy. Site-directed mutants in proposed formaldehyde binding residues will be generated to test a hypothetical model of formaldehyde sensing. Furthermore, a paralogous sensor RcnR, which is homologous to *E. coli* Ni(II)/Co(II)-sensing RcnR, has been employed to test the formaldehyde specificity of FrmR *in vitro*.

5.2 Identification of formaldehyde-sensing characteristics

Based on homology to *E. coli* FrmR, it was anticipated that *Salmonella* FrmR would also responds to formaldehyde and this has already been confirmed *in vivo*. The effect of formaldehyde on DNA-binding by FrmR was tested by fluorescence anisotropy (FA) and *β-galactosidase* assay.

5.2.1 *Formaldehyde weakens FrmR K_{DNA}*

To investigate FrmR ability to sense formaldehyde, we explored the interaction between FrmR and its promoter operator region in the presence of various formaldehyde concentrations. As described in Chapter 3, fluorescence anisotropy (FA) can be used to assess DNA-protein interactions. Figure 5. 1A-B shows the anaerobic titration of *frmRA*Pro with FrmR in the presence of EDTA and formaldehyde. Formaldehyde is an extremely reactive molecule used as a cross-linker (non-specific) for protein and DNA conjugation at concentrations ranging between 2 mM and 360 mM (1 % v/v, d = 1.09 g/ml) (Section 1.5.2.1) (Brodolin 2000). Hence, in order to prevent any unspecific crosslinks within protein residues and/or DNA bases, the concentrations of formaldehyde used in the *in vitro* experiments were much lower (ranging from 10 µM to 100 µM).

The proposed binding mechanism indicates a stoichiometric interaction between formaldehyde and FrmR (Section 5.3.1), suggesting that a specific response may be observed at much lower formaldehyde concentrations. For this reason experiments were carried out with 10 and 20 µM formaldehyde.

Data were fit with Dynafit as previously described for apo-FrmR generating a $K_{DNA}^{HCOH\text{-}FrmR}$ = 637 ± 16 nM (if [HCOH] = 10 µM) and 7089 ± 593 nM (if [HCOH] = 20 µM). Figure 5. 1C shows the comparison between these data and the apo-FrmR association curve shown in Chapter 3, highlighting weakening of FrmR K_{DNA} in the presence of formaldehyde. ΔGc values were calculated, as described previously (Table 8.2, Appendix), to be 1.10 ± 0.02 kcal mol^{-1}

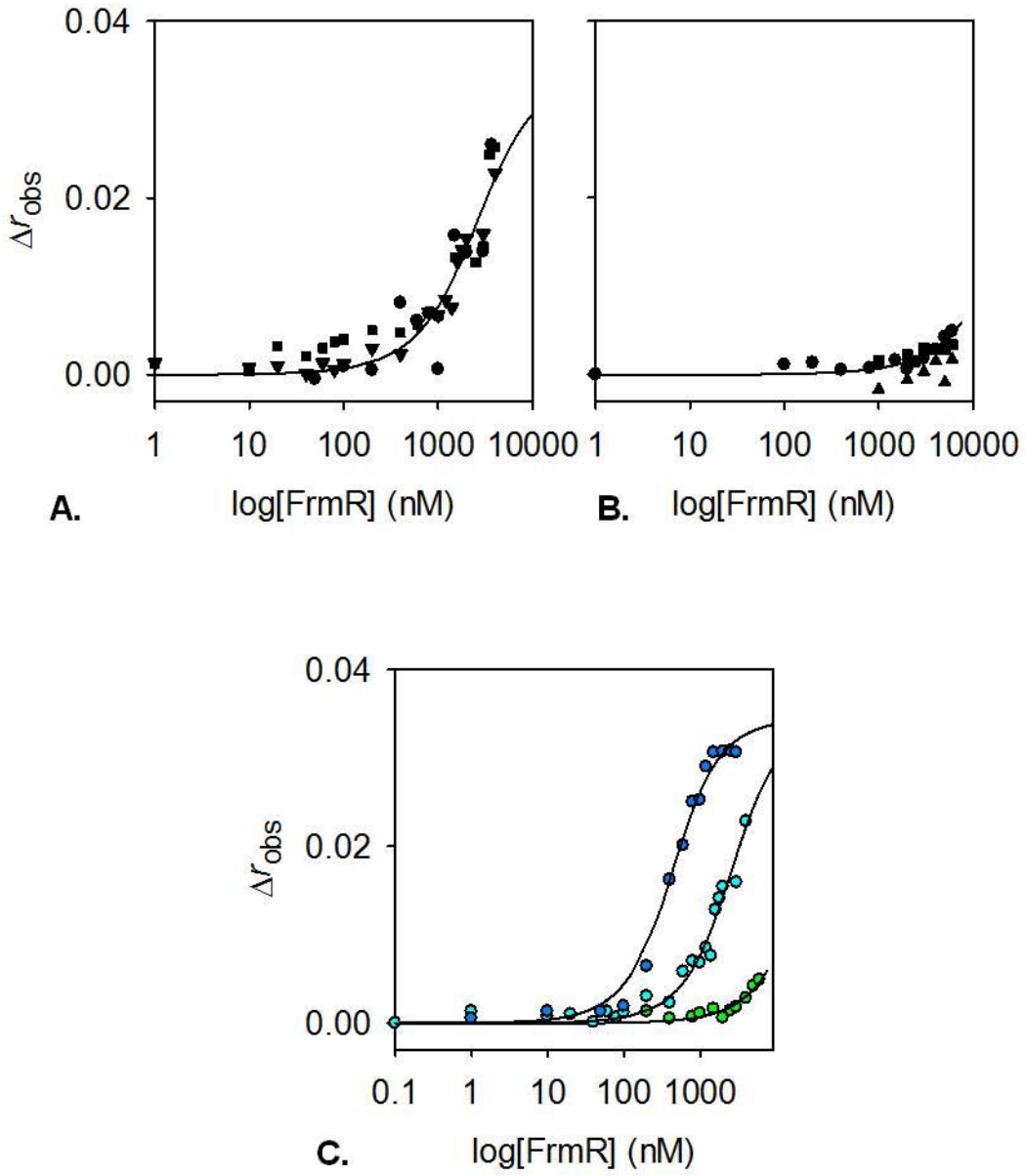


Figure 5. 1 *frmRA*Pro (10 nM) was anaerobically titrated with FrmR in the presence of 5 mM EDTA in presence of 10 µM (**A**) and 20 µM (**B**) of formaldehyde. Figure **C** shows the comparison of data from the same FrmR:DNA association in the absence (blue circles), or in the presence of 10 µM (pale blue) and 20 µM (green circles) of formaldehyde. DNA-binding was monitored by fluorescence anisotropy. Solid line represents simulated curves produced from the average K_{DNA} determined across the experiment replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 FrmR tetramer (non-dissociable):DNA stoichiometry.

(with 10 µM HCOH), and 2.52 ± 0.05 kcal mol^{-1} (with 20 µM HCOH).

Furthermore, in order to explore if FrmR responds specifically upon formaldehyde stress, the same experiment has been performed in the presence of 20 µM of ethanol (C_2H_5) or acetaldehyde (C_2H_4O) (Figure 5. 2). FrmR affinity for *frmRA*Pro was calculated by DynaFit using the analogous script previously adopted for apo-FrmR analyses (Appendix) which gave a $K_{DNA}^{EtOH\text{-}FrmR}$ = 92.06 nM and $K_{DNA}^{C_2H_4O\text{-}FrmR}$ = 85.55 nM (more replicates of these experiments will be needed to calculate standard deviations). These findings suggest that FrmR is either not able to interact with any of the alcohols and aldehydes tested or these interactions are not coupled with an allosterically active conformational change. These results do not contradict what was found with *E. coli* FrmR where small but significant activity *in vivo* was observed for small aldehydes but not for bulkier such as furaldehyde and tribromoacetaldehyde (Law 2012). This is informative of the specificity of formaldehyde detoxification since it suggests that the size of the binding pocket within the protein can accommodate only a small molecule such formaldehyde.

5.2.2 *In vivo study of FrmR function*

In Chapter 4 it was shown by *β-galactosidase* assay that *Salmonella* FrmR responds to MNIC of formaldehyde. Moreover, *β-galactosidase* assays on *SL1344*Δ*frmR* harbouring either P_{frmRA}-*::lacZ* or P_{frmRA}-*frmR::lacZ* construct, show that FrmR is necessary for repression of P_{frmRA} in *Salmonella* Figure 5. 3A. In fact, the formaldehyde response is lost and basal expression elevated, in cells harbouring the construct devoid of *frmR* (P_{frmRA}-*::lacZ*) (Figure 5. 3A). Thus, in common with *E. coli* FrmR (Herring & Blattner 2004; Law 2012), the *Salmonella* homologue represses expression from the *frmRA* operator promoter with repression alleviated by formaldehyde.

To further analyse the specificity of P_{frmRA} operator induction upon formaldehyde stress, *β-galactosidase* assay was employed to assess *frmR* activity in the absence or presence of minimum non-inhibitory concentrations (inhibited growth ≈ 10 %) of a range of alcohols and acetaldehyde in addition to formaldehyde, used as a control (Figure 5. 3B). Activity was induced only by formaldehyde and, to a lesser degree, by acetaldehyde which is the second smallest aldehyde after formaldehyde (R = CH_3). However, none of the alcohols were able to trigger P_{frmRA} induction. These outcomes are consistent with results obtained by fluorescence anisotropy (Figure 5. 2).

5.3 Mechanism of formaldehyde responsiveness by FrmR

5.3.1 *Proposed formaldehyde binding mechanism*

It has been suggested that the mechanism used by *E.coli* FrmR to bind formaldehyde and trigger

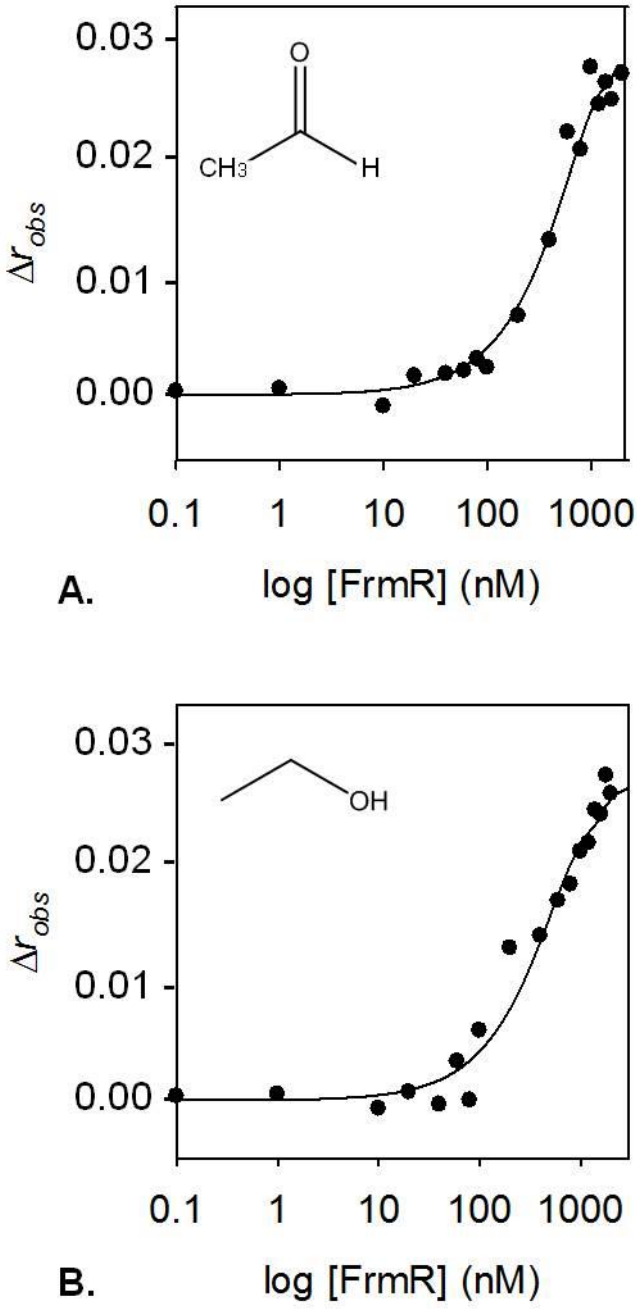


Figure 5. 2 *frmRA*Pro (10 nM) was anaerobically titrated with FrmR in the presence of 5 mM EDTA and in the presence of 20 µM of (**A**) acetaldehyde and (**B**) ethanol. DNA-binding was monitored by fluorescence anisotropy. Data were fit to a model describing a 2:1 FrmR tetramer (non-dissociable):DNA stoichiometry.

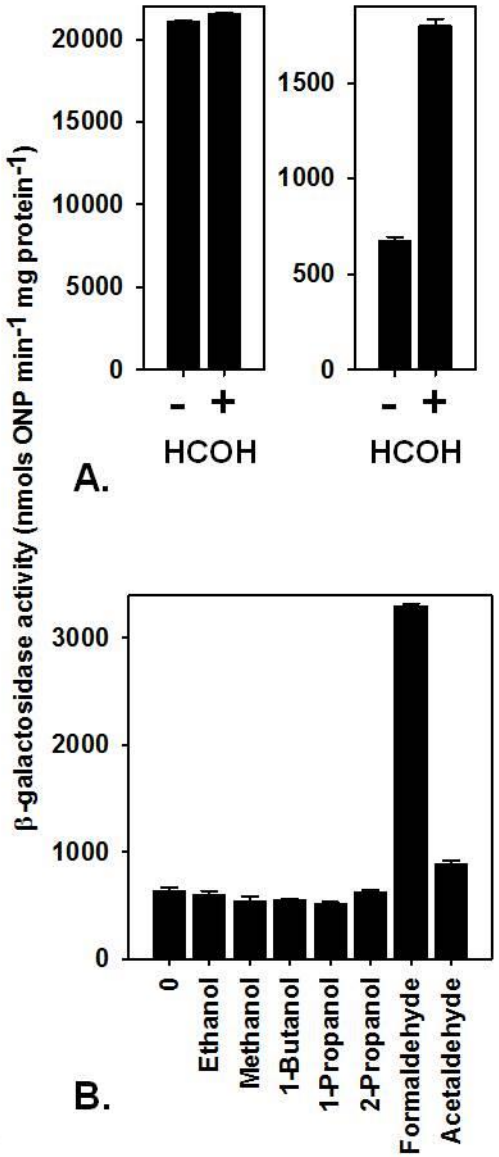


Figure 5. 3 A *In vivo* expression from P_{frmRA}-*frmR::lacZ.* Representative (n=3) *β-galactosidase* activity following growth of SL1344Δ*frmR*, harbouring either P_{frmRA}*::lacZ* or P_{frmRA}-*frmR::lacZ* construct, in M9 minimal media in the absence or presence of minimum non-inhibitory concentrations of formaldehyde. **B** *In vivo* expression from P_{frmRA}-*frmR::lacZ* Representative (n>3) β-galactosidase activity following growth of SL1344Δ*frmR*, harbouring a P_{frmRA}-*frmR::lacZ* construct, in M9 minimal media in the absence or presence of minimum non-inhibitory concentrations of ethanol, methanol, 1-butanol, 1-propanol, 2-propanol, formaldehyde and acetaldehyde. These data were obtained by Dr. Deenah Osman.

the conformational change that causes the release of its promoter region involves a cysteine (Higgins & Giedroc 2014; Law 2012). Cysteine residues are involved in sensing in other RcnR/CsoR family members (Iwig & Chivers 2009; Foster *et al.* 2012; Ma *et al.* 2009a). Moreover, reactions of thiols with carbonyl compounds are well known (Jencks 1969; Kallen 1971; Lienhard & Jencks 1966). In particular, the reaction of cysteine residues with formaldehyde can be divided in two steps. First, the sulphur atom on the thiol group attacks the carbonyl group of formaldehyde producing a hemithioacetal intermediate, followed by removal of one molecule of water and ring closure due to the attack of the amino group from cysteine backbone to yield a thiazolidinecarboxylic acid (Schubert 1935; Ratner & Clarke 1937).

Aqueous formaldehyde is present in its hydrated form, methanediol ($CH_2(OH)_2$); however, carboxyl compounds are reactive toward nucleophilic attack in their unhydrated forms (Lienhard & Jencks 1966; Sander & Jencks 1968; Bell & Evans 1966; Bell 1966; Kallen & Jencks 1966). Formaldehyde dehydration is catalyzed by hydroxide ions above pH = 7.0 (Le Henaff 1960) and was shown to not be a rate determining step in the subsequent reactions with cysteine residues (Kallen 1971).

A conserved residue, Cys35, was identified as the cysteine that reacts with formaldehyde in *E. coli* FrmR (Law 2012), and coincides with the only cysteine residue (located on helix 2, monomer 1) found in *Salmonella* FrmR. Figure 5. 4 outlines the proposed reaction mechanism between *Salmonella* FrmR and formaldehyde when this is present in low concentrations. In order to generate the reactive thiolate group (deprotonated form of thiol) on Cys35 a nearby residue is needed to act as a Lewis base and remove the proton (H^+) from the sulfhydryl group. This residue, herein called amino acid X, has not been identified but good candidates could be His60 or Asp63 since both have a carboxylate group deprotonated at pH values close to 7.0 (pK_r^{His} = 6.04 and pK_r^{Asp} = 3.90 in aqueous solution at 20 $^{\circ}$C). The now negatively charged sulphur atom of the Cys35 thiol group can give a nucleophilic addition on the formaldehyde yielding a tetrahedral intermediate. Subsequent abstraction of a proton from amino acid X (restoring the initial deprotonated form) yields an alcohol group. It has been postulated that a second residue from the N-terminal region of a different monomer is involved in this mechanism using its amino group (Higgins & Giedroc 2014). The amino terminus region is also required for metal-sensing by *E. coli* RcnR (Iwig *et al.* 2008; Higgins *et al.* 2012b). We here suggest that this residue could be Pro2 (on helix 1, monomer 2) which retains a basic pyrrolidine ring.

When a sample of purified FrmR was examined by quantitative amino acid analysis, the result was consistent with the presence of only three methionine residues instead of four, as the theoretical sequence would predict. Results from the Amino Acid Analysis performed on the protein are shown in the Appendix. If the number of moles per ml of Met, 513, is divided by the number of moles per ml of Tyr (since FrmR possesses only one tyrosine), 181, the resulting

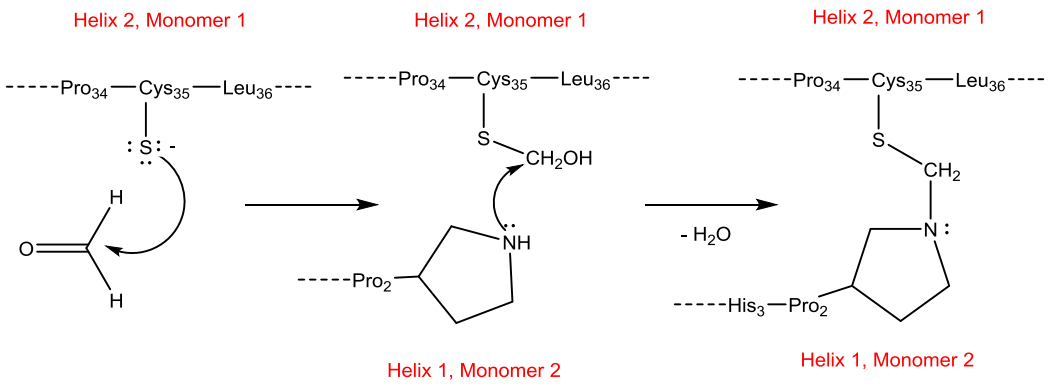


Figure 5. 4 Proposed mechanism of action of formaldehyde crosslinking FrmR. The reaction takes place in several steps. An amino acid (possibly His60 or Asp63, herein called amino acid X) close to Cys35 (helix 2, monomer 1) deprotonates the sulphur atom of the Cys35 thiol group, placing a negative charge on sulphur and making the thiol group more nucleophilic. Subsequent nucleophilic addition on the formaldehyde by the lone-pair electrons of the thiol on the Cys35 yields a tetrahedral intermediate. The basic intermediate abstracts a proton (H^+) from amino acid X to yield an alcohol group and regenerate the deprotonated form of amino acid X. Pro2 on helix 1, monomer 2 is the first residue of the N-terminal region and possesses a pirrolydine ring. Pyrrolidine is a strong base and it is among the most basic simple amines in nature (its conjugate acid has pKa = 11.27) (Hall 1957). The nitrogen on this residue attacks the positively polarized carbon. The alcohol group abstracts the proton from the now positively charged nitrogen and acts as a good living group (H_2O) to yield the final product.

value, 2.83, can be rounded up to 3. The analysis has been repeated twice on different preparations of FrmR, confirming the detection of only three methionine residues. This outcome was interpreted as evidence of the cleavage of Met1, a co-translational process predicted to occur in the ~ 80 % of the totality of proteins (Waller 1963; Matheson *et al.* 1975; Brown 1970; Frottin *et al.* 2006). Moreover, the cleavage of Met1 was confirmed by LC-MS analysis, performed by our collaborators Dr. Huggins and Dr. Chen (Procter and Gamble Mason Business Centre, Cincinnati, Ohio), on FrmR and E64HFrmR samples following digestion with trypsin (0.5 mg ml^{-1}) to obtain specific transitions (Figure 8. 3, Appendix). The presence of both PHSPEDK (where Met1 has been cleaved) and MPHSPEDK peptides was detected, but the qualitative nature of the analysis did not inform of their proportions. However, the presence of MPHSPEDK can be explained by considering that purified proteins are routinely overexpressed up to ≥ 5 mg/ml and the activity of the methionine aminopeptidase, the enzyme designated to cleave the methionine, would be hence limited by the large amount of available substrate. Since the Amino Acid Analysis previously discussed did not detect a fourth methionine, it is plausible to assume that a substantial proportion of the purified protein has undergone the cleavage.

This result is in accord with what was observed for *E. coli* FrmR by Law (Law 2012). The main FrmR peak detected by time-of-flight mass-spectrometry (TOF-MS) (ionization source: electrospray) has a molecular mass of 10186 Da instead of 10318 Da (the predicted FrmR molecular mass) (Law 2012). The difference between the predicted and the experimental mass is 132 Da which coincides with the cleavage of the methionine (131 Da).

In this scenario Pro2 is the first residue of the N-terminal region and it possesses a secondary amine group not involved in peptide bonding. According to the mechanism proposed here, the nitrogen on this residue attacks the electrophilic carbon on the tetrahedral intermediate and the alcohol group abstracts the proton from the now positively charged nitrogen creating a good leaving group (H_2O) to yield the final product. The specific crosslink would connect monomer 1 to monomer 2 of tetrameric FrmR causing a conformational rearrangement that produces a new protein assembly with a lower DNA-binding affinity.

In order to further explore these hypotheses, DNA-binding affinities of mutant FrmR proteins in the absence and in presence of stoichiometric concentration of formaldehyde were analysed.

5.4 Probing the formaldehyde-sensing mechanism of FrmR

In order to probe the proposed mechanism and identify residues necessary for formaldehyde-sensing, two single-point mutants were produced and tested by fluorescence anisotropy. The mechanism suggests that Cys35 and Pro2 contribute to the reaction with formaldehyde. Hence, Cys35 was mutated to an alanine, which is unable to interact with formaldehyde, whereas Pro2 was substituted with a serine, which is located in the same position in RcnR and CsoR

homologues (formaldehyde-unresponsive) (Figure 3. 2). Production and purification of C35AFrmR was discussed in Chapter 3 (Section 3.8.1).

5.4.1 *C35AFrmR:DNA K_{DNA} is unaffected by formaldehyde*

Titration of *frmRA*Pro with C35AFrmR was previously shown in Chapter 3 (Section 3.8.3), resulting in K_{DNA} = 158.2 ± 17.6 nM (Figure 3. 28A). Titration of *frmRA*Pro with C35AFrmR in the presence of 20 µM of formaldehyde does not affect DNA binding as confirmed by the unaltered $K_{C35AFrmR \cdot DNA}$ = 151.4 ± 2.2 nM (Figure 5. 5A, Appendix, Table 8. 2). Addition of 50 µM and 100 µM of formaldehyde to the reaction buffer resulted in a tighter DNA binding affinity ($K_{DNA}^{C35AFrmR}$ = 128.6 ± 5.6 nM and $K_{DNA}^{C35AFrmR}$ = 96.7 ± 4.5 nM, respectively) (Figure 5. 5B-C, Appendix, Table 8. 2) which might be explained by potential non-specific contacts, between protein and DNA operator, mediated by formaldehyde. Moreover the almost unaltered DNA-binding affinity of C35AFrmR variant suggests that the cysteine residue is the first to react, as suggested by the proposed mechanism (Figure 5. 4). Figure 5. 5D summaries these outcomes showing that DNA binding curves in the absence and presence of increasing concentrations of formaldehyde almost completely overlay. Coupling free energy values were $\Delta G_c^{C35AFrmR}$ -0.02 ± 0.01 kcal mol^{-1} (20 µM formaldehyde), $\Delta G_c^{C35AFrmR}$ -0.15 ± 0.02 kcal mol^{-1} (50 µM formaldehyde), and $\Delta G_c^{C35AFrmR}$ -0.29 ± 0.05 kcal mol^{-1} (100 µM formaldehyde).

5.4.2 *Production and purification of P2SFrmR*

Pro2→Ser mutation was inserted by site-directed mutagenesis of pETfrmR and was confirmed by sequence analysis. The production and purification steps (Figure 5. 6) followed the same protocol used for wild-type FrmR (Chapter 2) and protein was moved into the anaerobic glove-box and buffer exchanged with N_2-purged, chelex-treated buffer. Analysis of oxidized sulphur content and metal contamination by DTNB assay and ICP-MS respectively were performed to confirm reduced thiol content > 85 % and metal contamination < 5 %.

5.4.3 *P2SFrmR:DNA interactions in the presence of various formaldehyde concentrations by fluorescence anisotropy*

Fluorescence anisotropy was employed to test the role of Pro2 in formaldehyde binding, as previously shown with C35AFrmR (Section 5.4.1). As shown by Law, Pro2 is strictly conserved among predicted bacterial FrmR proteins located in operons with genes encoding deduced formaldehyde dehydrogenases (Law 2012). As in *Salmonella* FrmR, the N-terminus methionine is cleaved from *E. coli* FrmR, resulting in Pro2 being the first residue (Law 2012). If the lone electron pair of the Pro2 amino group was involved in the formation of a peptide bond with Met1 (or any other amino acid residue) it would not be available to coordinate formaldehyde. Fluorescence anisotropy analyses on the apo-P2SFrmR (Figure 5. 7A) reported a DNA-binding affinity weaker than wild-type but comparable to C35AFrmR ($K_{DNA}^{P2SFrmR}$ = 146.45 ±16.37 nM)

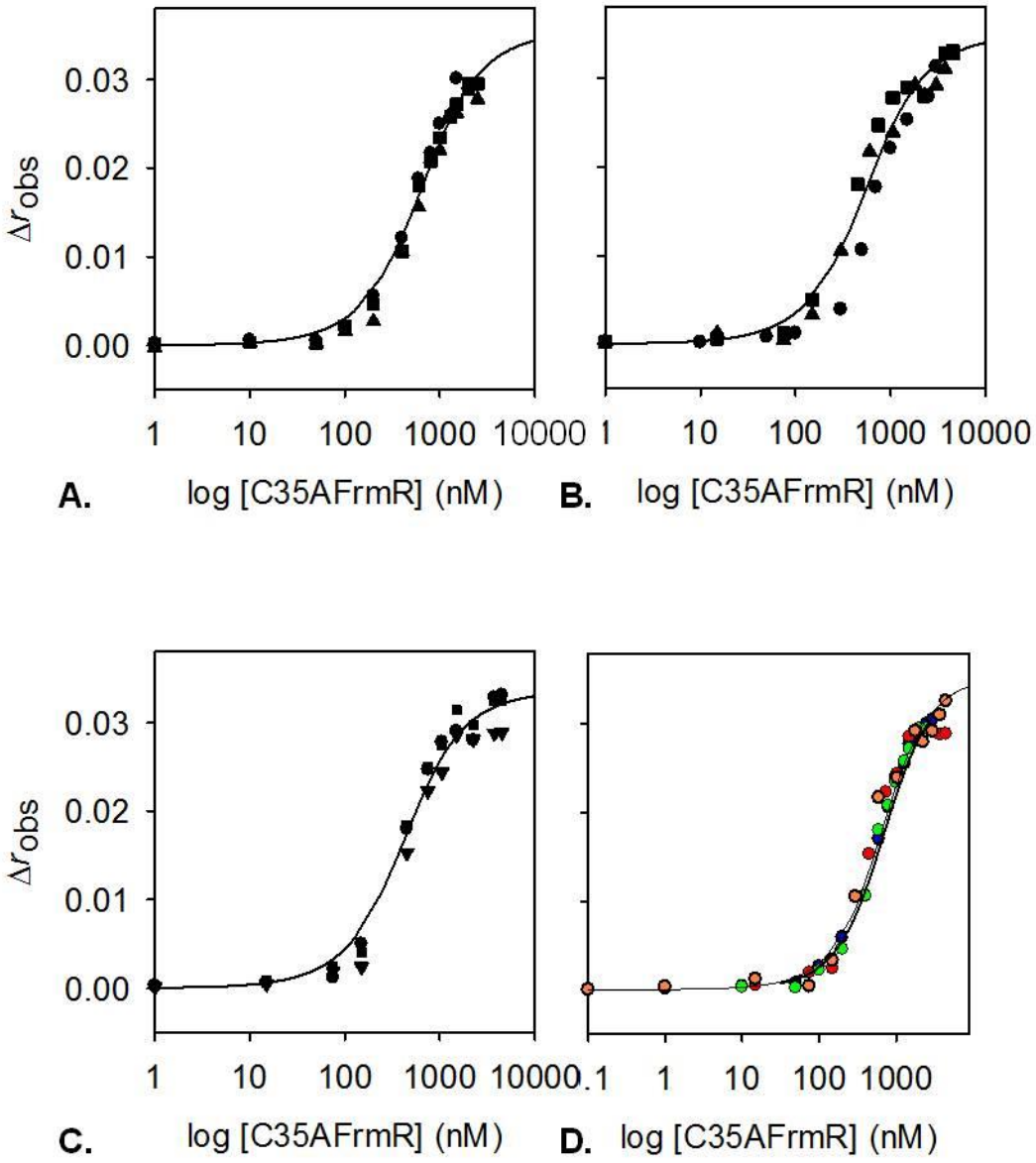


Figure 5. 5 *frmRA*Pro (10 nM) was anaerobically titrated with C35AFrmR in the presence of 5 mM EDTA plus 20 µM (**A**), 50 µM (**B**) and 100 µM (**C**) of formaldehyde. Figure **D** shows the comparison of one representative data set from C35AFrmR:DNA associations in the absence (blue circles), or presence of 20 µM (green circles), 50 µM (orange circles) and 100 µM (red circles) of formaldehyde. DNA binding was monitored by fluorescence anisotropy. Solid line represents simulated curves produced from the average K_{DNA} determined across the experiment replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 C35AFrmR tetramer (non-dissociable):DNA stoichiometry.

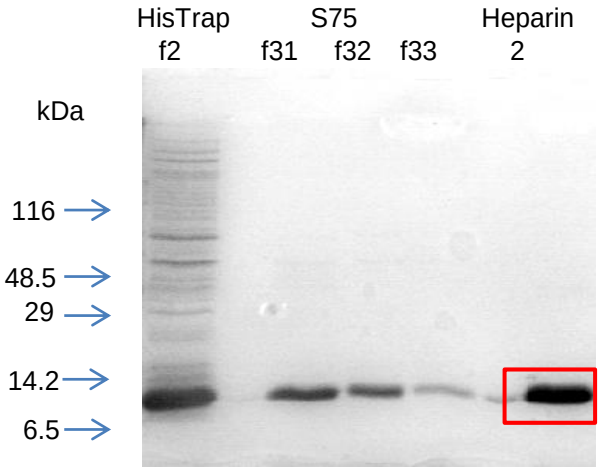


Figure 5. 6 Purification of P2SFrmR by nickel affinity chromatography, size exclusion and heparin affinity chromatography. SDS-PAGE analysis (18 % w/v acryl-bis) showing fractions containing P2SFrmR from each purification step. "HisTrap f2" refers to the second fraction eluted from a 5 ml Ni(II)-affinity column with buffer A containing 0.3 M imidazole. Fractions "S75 f31-f33" were eluted from a Superdex S75 column loaded with fraction "HisTrap f2". Fraction "Heparin 2" was eluted at 0.5 M NaCl from two 1 ml Heparin columns linked together loaded with size exclusion fractions f31-f32 combined. Presence of only one band at around 10 kDa confirms the purity of the protein.

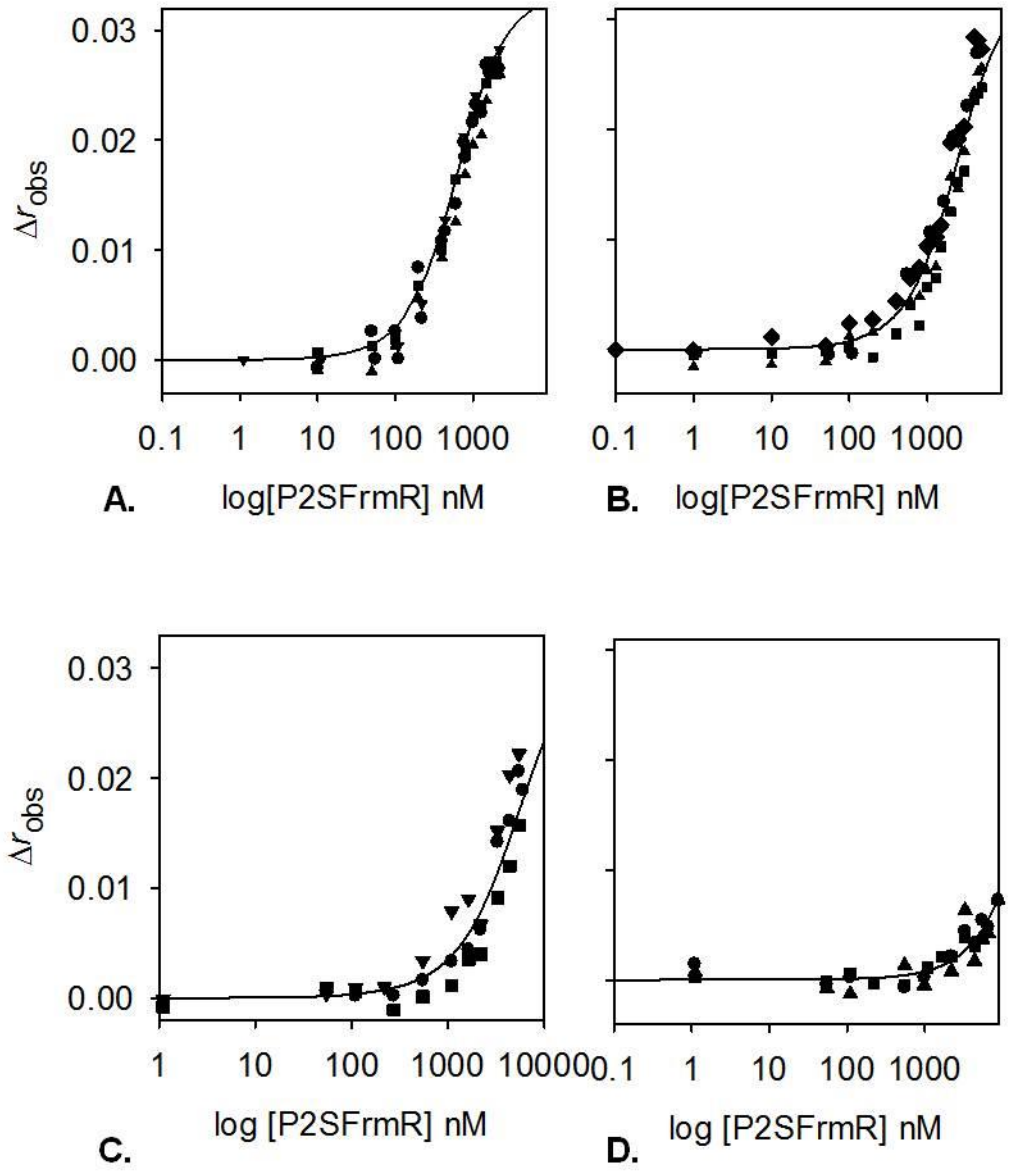


Figure 5. 7 *frmRA*Pro (10 nM) was anaerobically titrated with P2SFrmR in the presence of 5 mM EDTA in presence of 0 µM (**A**), 20 µM (**B**), 50 µM (**C**) and 100 µM (**D**) of formaldehyde. DNA binding was monitored by fluorescence anisotropy. Solid line represents simulated curves produced from the average K_{DNA} determined across the experiment replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 P2SFrmR tetramer (non-dissociable):DNA stoichiometry.

suggesting that these mutations affect FrmR DNA-binding to a modest degree. The same experiment was conducted in the presence of various concentrations of formaldehyde (20 µM – 50 µM – 100 µM, see Figure 5. 7 B-C-D respectively) producing the following values: $K_{DNA}^{HCOH\text{-}P2SFrmR}$ = 586.4 ± 130.3 nM, $\Delta G_c^{HCOH\text{-}P2SFrmR}$ = 0.81 ± 0.13 kcal mol^{-1} (20 µM formaldehyde), $K_{DNA}^{HCOH\text{-}P2SFrmR}$ = 1349.1 ± 468.7 nM, $\Delta G_c^{HCOH\text{-}P2SFrmR}$ = 1.29 ± 0.22 kcal mol^{-1} (50 µM formaldehyde) and $K_{DNA}^{HCOH\text{-}P2SFrmR}$ = 5757.7 ± 462.6 nM, $\Delta G_c^{HCOH\text{-}P2SFrmR}$ = 2.18 ± 0.07 kcal mol^{-1} (100 µM formaldehyde). As shown in Figure 5. 8, which reviews and compares the P2SFrmR:DNA association curves in the presence and absence of formaldehyde, the protein is still capable of interacting with formaldehyde but retains a tight DNA-binding affinity at a HCOH concentration of 20 µM. On the contrary, this concentration was sufficient to almost completely abolish any interaction with DNA in FrmR (Figure 5. 1B).

Comparison of the ΔG_c values calculated for P2SFrmR and FrmR (Table 8. 2, Appendix) indicates that a concentration of formaldehyde of 100 µM is needed to achieve a response in P2SFrmR comparable to that shown by FrmR when treated with 20 µM of formaldehyde. However, inclusion of 100 µM of formaldehyde in association experiments using C35AFrmR with *frmRA*Pro did not affect DNA-binding. These outcomes suggest that, unlike C35AFrmR, P2SFrmR is still capable of interacting with formaldehyde, presumably with a different mechanism. Formaldehyde may bind Cys35 yielding the tetrahedral intermediate, and a nearby residue may replace the amino group of the Pro2 (e.g. amino group of His3 imidazole). Alternatively, the modification on Cys35 alone may be sufficient to impair the ability of P2SFrmR to bind DNA with the same affinity of apo-protein.

5.5 *In vitro* analysis of the RcnR:*rcnRA* promoter interaction

In order to establish whether or not the range of [HCOH] used in the previous sections lead to non-specific formaldehyde contacts within protein and/or with DNA, *Salmonella* RcnR was produced and DNA binding properties were investigated under the same reaction conditions.

5.5.1 *Recombinant Salmonella RcnR production and purification*

Salmonella RcnR was overexpressed as outlined in Chapter 2 from single colonies of *E. coli* BL21 (DE3) containing pETRcnR. The protein was purified using a three-step protocol which involved 5 ml Heparin-affinity column, HiLoad 16/60 Superdex S75 GE Healthcare column and 5 ml HiTrap SP column (Chapter 2). SDS-PAGEs 18 % w/v were run after each step to determine the fractions containing RcnR and to confirm the purity of the protein after the last step (Figure 5. 9A-C). Pure protein from HiTrap SP column step was concentrated and buffer exchanged inside the anaerobic glove-box using N_2-purged, chelex-treated buffers. RcnR was quantified by measuring absorbance at 280 nm using the experimental extinction coefficient

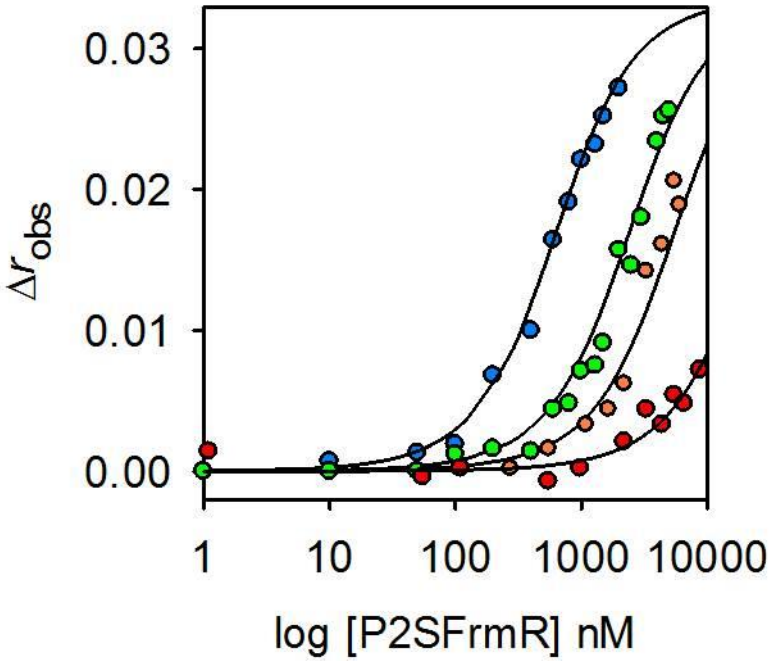


Figure 5. 8 Comparison of one representative data from P2SFrmR:DNA associations in the absence (blue circles), or in the presence of 20 µM (green circles), 50 µM (orange circles) and 100 µM (red circles) of formaldehyde, as previously shown in Figure 5.A-D. DNA binding was monitored by fluorescence anisotropy. Solid lines represent simulated curves produced from the average K_{DNA} determined across the experiment replicates shown. Data were fit to a model describing a 2:1 P2SFrmR tetramer (non-dissociable):DNA stoichiometry.

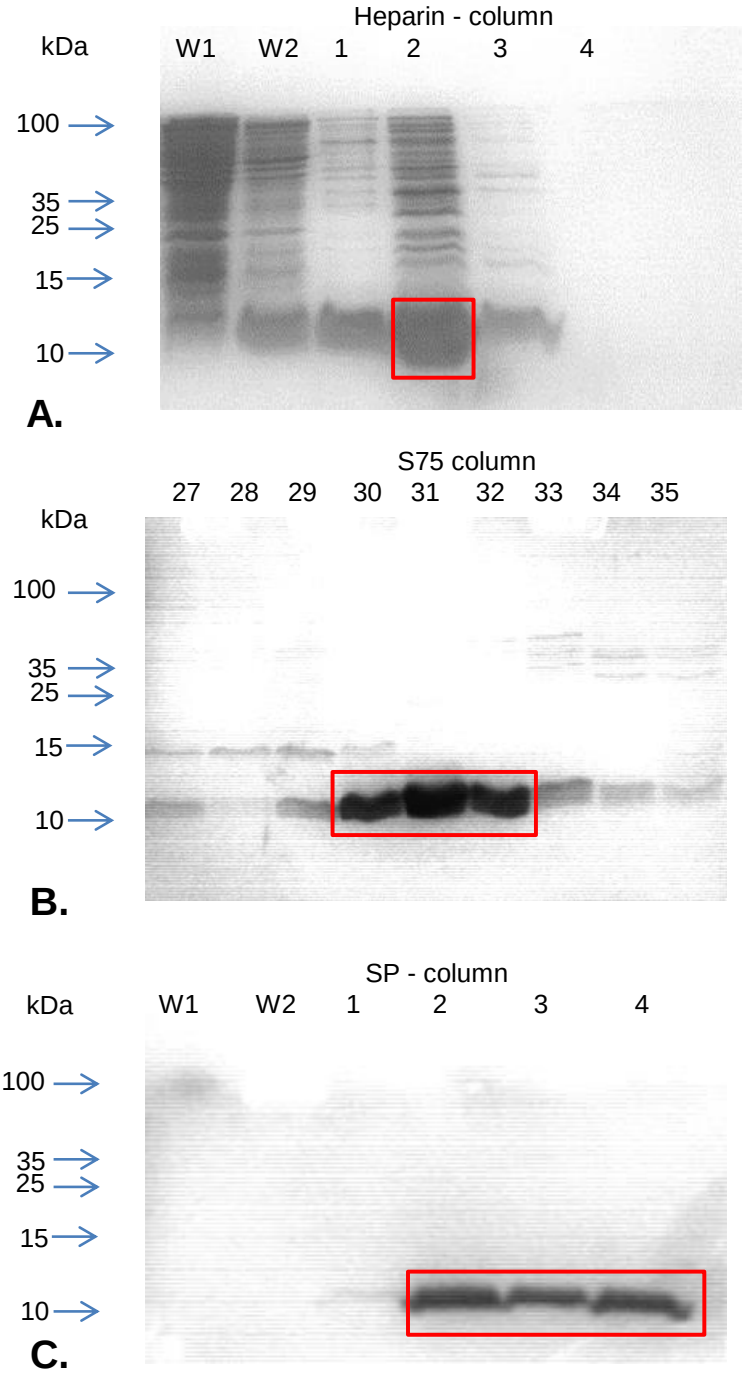


Figure 5. 9 Purification of RcnR by heparin affinity chromatography, size exclusion and HiTrap SP- chromatography. A. SDS-PAGE analysis (18 % w/v acryl-bis) of fractions eluted from a 5ml Heparin-affinity column with buffer B800 containing 800 mM NaCl. Fractions W1 and W2 contain material not bound to the column and flow through during wash step with 25 ml buffer B300 (300 mM NaCl) respectively. Fraction 2 was routinely found to contain RcnR in the highest concentration. B. SDS-PAGE analysis (18 % w/v acryl-bis) of fractions 27-35 eluted from a Superdex S75 column loaded with fraction 2 (5.5 ml) from the previous purification step (A.) C. SDS-PAGE analysis (18 % w/v acryl bis) of fractions eluted at 0.3 M NaCl from a 5 ml SP column with size exclusion fractions enriched for RcnR.

(ε_{280} = 2422 $M^{-1}cm^{-1}$) determined by quantitative amino acid analysis (Table 8.5, Figure 8. 2). Amino acid analysis on the purified protein also revealed the absence of one methionine, consistent with the loss of the N-terminal Met1 residue, as seen for *E. coli* RcnR (Iwig *et al.* 2008). Reduced cysteine content (> 85% free thiol) and metal contamination (< 5%) were determined by DTNB assay and ICP-MS. Protein was stored at 4 ºC for up to one month although a few RcnR preps showed signs of protein precipitation after the last buffer exchange step inside the glove-box and were therefore discarded.

5.5.2 *PrcnRA unlabelled oligonucleotides production and competitive EMSA to study RcnR:DNA interaction*

In order to predict and test for binding to the promoters of gene(s) regulated by *Salmonella* RcnR, bioinformatics comparison with the *E. coli* RcnR operator region and EMSA screening were adopted in order to locate promoter elements that might represent RcnR binding sites.

The *Salmonella* genome upstream of the *rcnR* gene was searched for promoter regions with similarity to the DNA binding site of *E. coli* RcnR with which *Salmonella* RcnR shares an identity of 92.2 % (Figure 5. 10, Table 5. 1). RcnR recognizes the pseudo palindromic, type 1, sequence TACT-G_6-N-AGTA, two of which are located in the *E. coli rcnA-rcnR* intergenic region (Iwig & Chivers 2009). This DNA recognition takes place through multiple contacts in the minor groove and non-specific interactions with the flanking DNA regions promoting DNA wrapping (Section 1.4.4).

The region including 100 bp immediately upstream of the *rcnR* translational start codon was therefore searched and two sequence (herein called site 1 and site 2) presenting features similar to those recognized by *E. coli* RcnR (type 1 category, containing one G/C tract flanked by AT-rich inverted repeats) were identified as shown in Figure 5. 11. As with *E. coli* RcnR the distance between the two centred A/T-rich tracts is 9 bp for each separate RcnR binding site. We initially explored the interaction between *Salmonella* RcnR and this region by electrophoretic mobility shift assay (EMSA). The nucleotide sequence containing the start of the *rcnR* coding region and the region immediately upstream was amplified by PCR, ligated to pGEM-T and the newly constructed plasmid was propagated in *E. coli* DH5α. Primers were designed homologous to the two regions underlined in Figure 5. 11A. The DNA probe (177bp, designated *rcnRA*ProEM) was then amplified by PCR. As a control for non-specific protein-DNA binding in EMSA assay, a non-specific DNA probe (141 bp, designated pGEMCon3) was PCR amplified from re-circularised pGEM plasmid using primers homologous to either side of the multi-cloning site (Table 2.1). Figure 5. 11B shows how, upon addition of ~ 3 µM RcnR to the EMSA binding reaction a first shifted *rcnRA*ProEM band is visible whereas after 8 µM RcnR the entire population of *rcnRA*ProEM has been retarded. The presence of two retarded bands may represent the formation of multimeric complexes and has been observed in EMSA

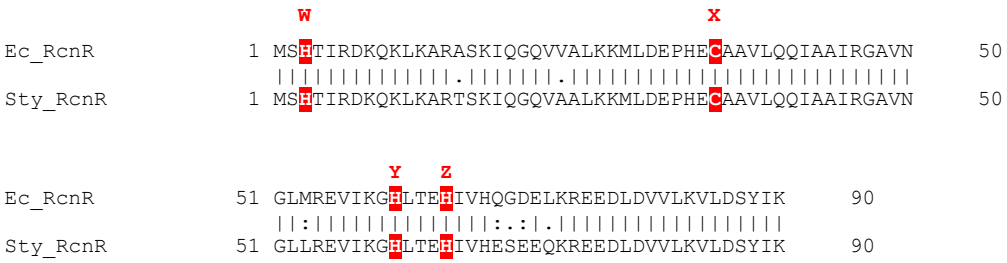
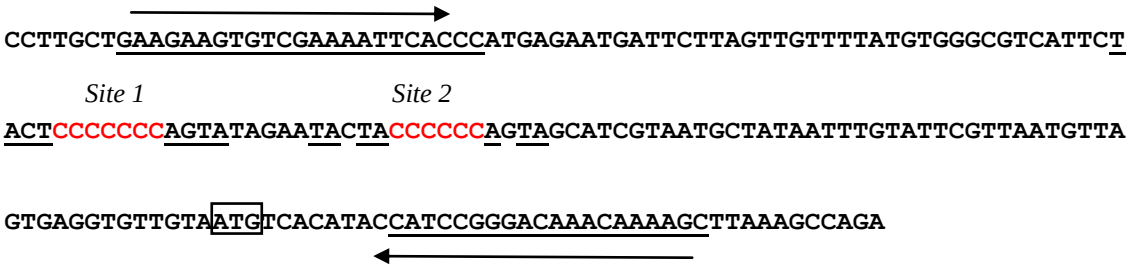


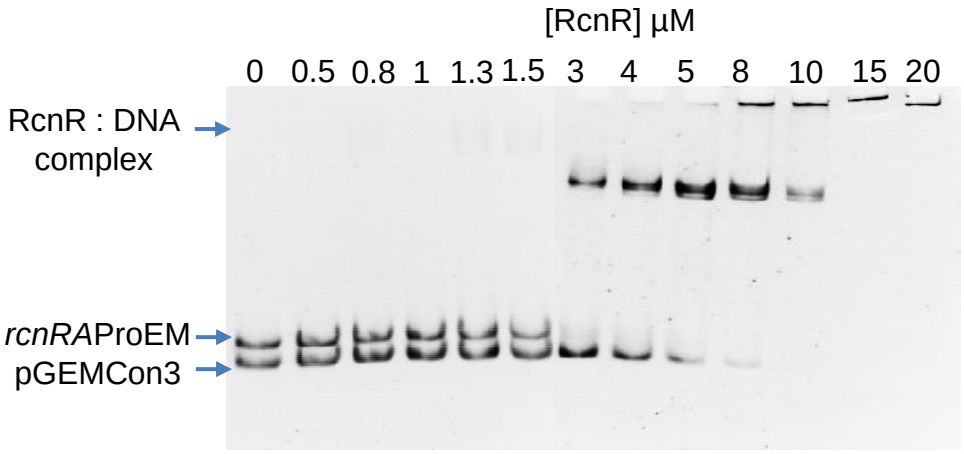
Figure 5. 10 Alignments of *S. typhimurium* RcnR (Sty_RcnR) with *E. coli* RcnR (Ec_RcnR) by Clustalw2. The residues at positions of the WXYZ fingerprint are highlighted.

E. coli RcnR	**Similarity (%)**	**Identity (%)**
S. typhimurium RcnR	95.6	92.2

Table 5. 1 Similarity and identity of amino acid sequence of *S. typhimurium* RcnR compared to *E. coli* RcnR using EMBOSS NEEDLE.



A.



B.

Figure 5. 11 A Nucleotide sequence of the start of the coding region and upstream region of *rcnRA*ProEM. The potential RcnR binding sites (containing one of the two C/G tracts) are highlighted in red and the translational start codon is highlighted in the box. Primers (Table 2.1) were designed homologous to the two regions underlined, and used to amplify this region from *Salmonella* genomic DNA producing the probe named *rcnRA*ProEM used in EMSA analysis. **B** Competitive EMSA assessing RcnR binding to *rcnRA*ProEM. Competitive EMSA with 177 bp *rcnRA*ProEM (100 nM) and 141 bp non-specific DNA probe pGEMCon3 (100 nM) incubated with 0-20 µM RcnR.

analyses using other RcnR/CsoR family members such as *M. tuberculosis* (Liu *et al.* 2007), *L. monocytogenes* CsoR (Corbett *et al.* 2011) and *Synechocystis* InrS (Foster *et al.* 2012). Alternatively the lower mass retarded band may indicate the specific interaction between RcnR and *rcnRA*ProEM while the higher mass retarded may represent the non-specific interaction of RcnR with the control fragment, consistent with the pGEMCon3 band fading approximately at the same [RcnR]. In fact, as the concentration of RcnR is further increased, a decrease in the intensity of the free pGEMCon3 (after ~ 5 µM RcnR) starts to be visible indicating an affinity of RcnR, although weak, for the control fragment. The apparent affinity of RcnR for *rcnRA*ProEM is $K_{DNA,app}$ ~ 1.5 µM, which is significantly weaker than that observed for *E. coli* RcnR by Iwig and collaborators ($K_{DNA\text{-}site1}$ = 126 nM and $K_{DNA\text{-}site2}$ = 174 nM) (Iwig & Chivers 2009). Since this may be due to the non-equilibrium nature of the EMSA assay, a quantitative approach intended to estimate RcnR DNA-affinity was performed at equilibrium by fluorescence anisotropy and presented in Sections 5.5.4 and 5.5.5.

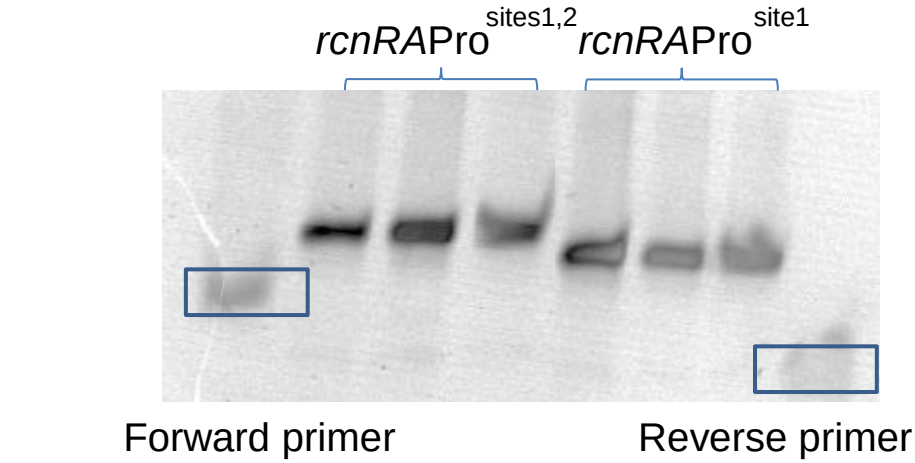
5.5.3 *RcnR labelled oligonucleotides production*

Production of a HEX-labelled double stranded, oligonucleotides containing the RcnR recognition region for use in fluorescence anisotropy was performed as described for FrmR in Chapter 3, section 3.6.3. Two DNA sequences were produced, *rcnRA*Pro^{sites1,2} and *rcnRA*Pro^{site1}, the first is composed of 35 nucleotides and contains the two identified RcnR recognition sites and flanking nucleotides whereas the second (28 bp) contains only RcnR recognition site 1 and flanking nucleotides (Figure 5. 12). Successful annealing of HEX-labelled oligonucleotides with their unlabelled reverse complements was confirmed by native PAGE analysis as shown in Figure 5. 12.

5.5.4 *RcnR:DNA stoichiometry explored by fluorescence anisotropy*

It has been shown that the DNA binding stoichiometry in *E. coli* RcnR is one RcnR tetramer per DNA recognition site (Iwig & Chivers 2009). Although *E. coli* RcnR is able to singularly interact with the two TACT-G_6-N-AGTA motifs almost with the same affinity, Iwig and Chivers demonstrated that the combination of site 1 and 2 is important for observing the extended footprint and DNA wrapping (Iwig & Chivers 2009). In order to determine the stoichiometry of *Salmonella* RcnR:DNA complex, increasing amounts of RcnR were added to concentrated samples of *rcnRA*Pro^{sites1,2} and *rcnRA*Pro^{site1} (2.5 µM) (Figure 5. 13A-B, respectively) and variation in anisotropy (r_{obs}) upon each protein addition was monitored by a modified Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies) as described in Chapter 2.

Since only 9 bp separates the centers of *E. coli* RcnR AT-rich sequences, resulting in a rotation of DNA ~ 324 º which places the two regions on the same face (Iwig & Chivers 2009), the protein is able to simultaneously interact with the two inverted repeat motifs flanking the G-rich



site 1 site 2

5'-[HEX]TACTCCCCCCCAGTATAGAATACTACCCCCCAGTA

site 1

5'-[HEX]CATTCTACTCCCCCCCAGTATAGAATAC

Figure 5. 12 Production of double stranded, 5'-[HEX]-labelled, oligonucleotides containing both RcnR recognition sites (*rcnRA*Pro^{sites1,2}) or only site 1 (*rcnRA*Pro^{site1}) and flanking nucleotides for use in fluorescence anisotropy assays. The HEX-labelled primers were annealed with their unlabelled reverse complements. The native PAGE shown here confirms the successful annealing of the oligonucleotide pairs to produce the double stranded *rcnRA*Pro^{sites1,2} (first three lanes) and *rcnRA*Pro^{site1} (last three lanes). Forward and reverse primer samples shown in the gel refers to *rcnRA*Pro^{sites1,2} oligonucleotides. Sample concentrations: 10 µM.

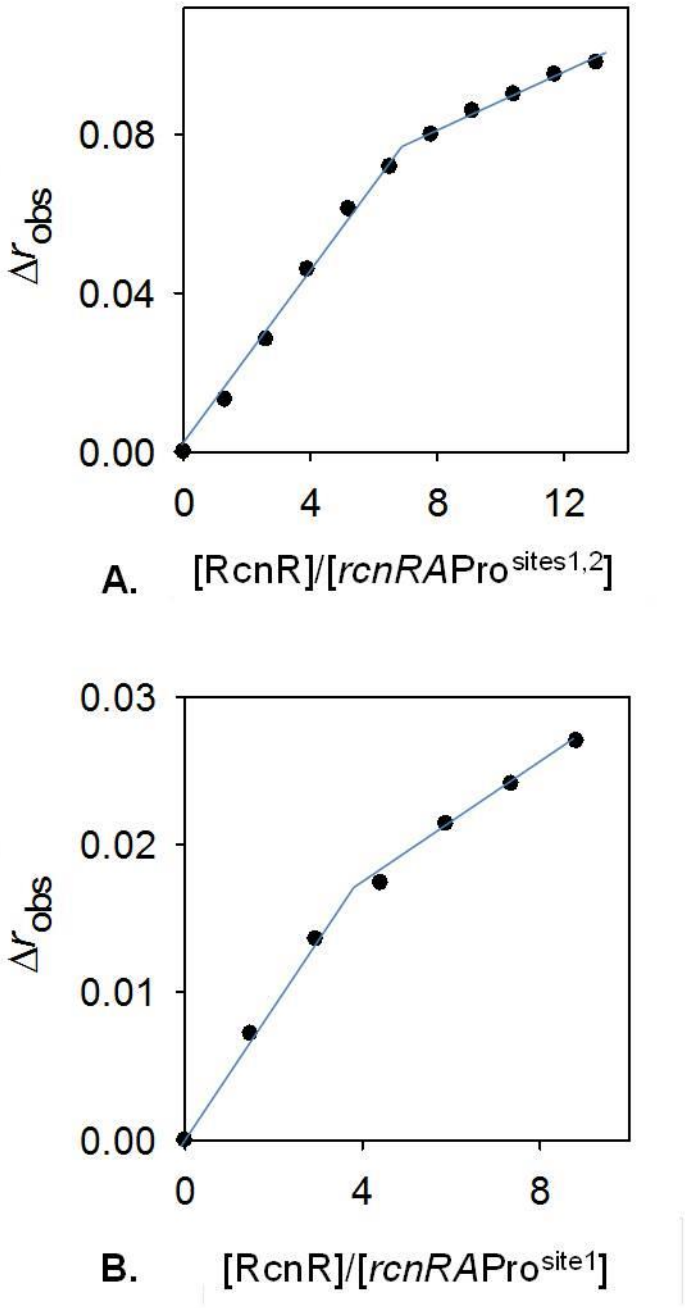


Figure 5. 13 A Anisotropy change upon titration of *rcnRA*Pro^{sites1,2} (2.5μM) with RcnR (2.5 μM monomer with each addition) (n = 1). A point of inflection is observed after addition of ~ 8 equivalents of protein monomer indicating a protein:DNA stoichiometry of 8:1 (two tetramers per *rcnRA*Pro^{sites1,2}). **B** Anisotropy change upon titration of *rcnRA*Pro^{site1} (2.5μM) with RcnR (2.5 μM monomer with each addition) (n = 1). A point of inflection is observed after addition of ~ 4 equivalents of protein monomer indicating a protein:DNA stoichiometry of 4:1 (one tetramer per *rcnRA*Pro^{site1}). The overall response is lower than half the response obtained using DNA containing the two tracts rich in C/Gs therefore it remains unclear if the binding of only one RcnR tetramer to the DNA needs both halves of the inverted repeat of the recognition sequence.

tract resulting in a 1:1 RcnR:DNA stoichiometry for each recognition site (previously illustrated in Figure 3. 17). An analogous result was observed with *Salmonella* RcnR where an inflection after 8 molar equivalents of protein (consistent with two RcnR tetramers per DNA region) was detected with *rcnRA*Pro^{sites1,2} (Figure 5. 13A), which contain both recognition site 1 and 2, and after 4 when *rcnRA*Pro^{site1} (containing only site 2) was used (Figure 5. 13B). In the experiment with *rcnRA*Pro^{site1} the overall response (Δr_{obs} at the inflection point) appears smaller than the expected value which was theoretically predicted to be ~ half the Δr_{obs} obtained when *rcnRA*Pro^{sites1,2} was used. This could indicate that the response is strictly specific to the length and composition of the DNA sequence.

5.5.5 *RcnR:rcnRAPro interaction explored by fluorescence anisotropy*

Titrations of fluorescently labelled *rcnRA*Pro^{sites1,2} (10 nM) with RcnR performed anaerobically in the presence of 5 mM EDTA, produced binding isotherms shown in Figure 5. 14A (symbol shapes represent individual experiments). Since it has been established by ITC that *E. coli* RcnR is able to bind essentially with the same affinity for each single site ($K_{DNA\text{-}site1}$ = 126 nM and $K_{DNA\text{-}site2}$ = 174 nM) (Iwig & Chivers 2009), data were fit to a model describing binding of two RcnR tetramer to one DNA molecule (containing two G/C sites), making the assumption that the two tetramers bind DNA with equal affinity and that the relationship between protein binding and r_{obs} is linear. $K_{rcnRAPro\text{-}sites1,2}{}^{RcnR}$ was found to be 152 ± 8 nM which is very close to the values calculated for *E. coli* RcnR (Iwig & Chivers 2009).

Anaerobic titration of RcnR with *rcnRA*Pro^{site1} (10 nM), which contains only site 1, yielded a weaker DNA binding affinity ($K_{rcnRAPro\text{-}site1}{}^{RcnR}$ = 1562 nM) (Figure 5. 14B) suggesting a positive cooperativity, which was seen in *E. coli* RcnR foot printing experiments involving large DNA duplex (~ 300 bp) but not for short fragments (Iwig & Chivers 2009).

5.5.6 *Analysis of the specificity of RcnR:rcnRAPro interaction*

To probe whether *Salmonella* RcnR-DNA contact is restricted at its consensus operator site, we tested if RcnR is capable of binding to *frmRA*Pro with affinities sufficiently tight to mediate effective repression. In fact, *frmRA*Pro belongs to type 1 site and the *rcnRA*Pro^{sites1,2} contains a tandem operator site consisting of two type 1 sites. *frmRA*Pro (10 nM) was anaerobically titrated with RcnR in the presence of 5 mM EDTA and change in anisotropy monitored by fluorescence anisotropy, showing an overall response (final Δr_{obs}) consistent with what has already been observed for *frmRA*Pro (Chapter 3 sections 3.6.5). $K_{frmRAPro}{}^{RcnR}$ was calculated to be 2.0 µM using Dynafit software with a model describing a 1:1 RcnR tetramer:*frmRA*Pro ratio (red line in Figure 5. 15) and 745.8 nM considering two RcnR tetramers:*frmRA*Pro (blue line in Figure 5. 15). The difference between the two fitted curves is very small, however the latter model

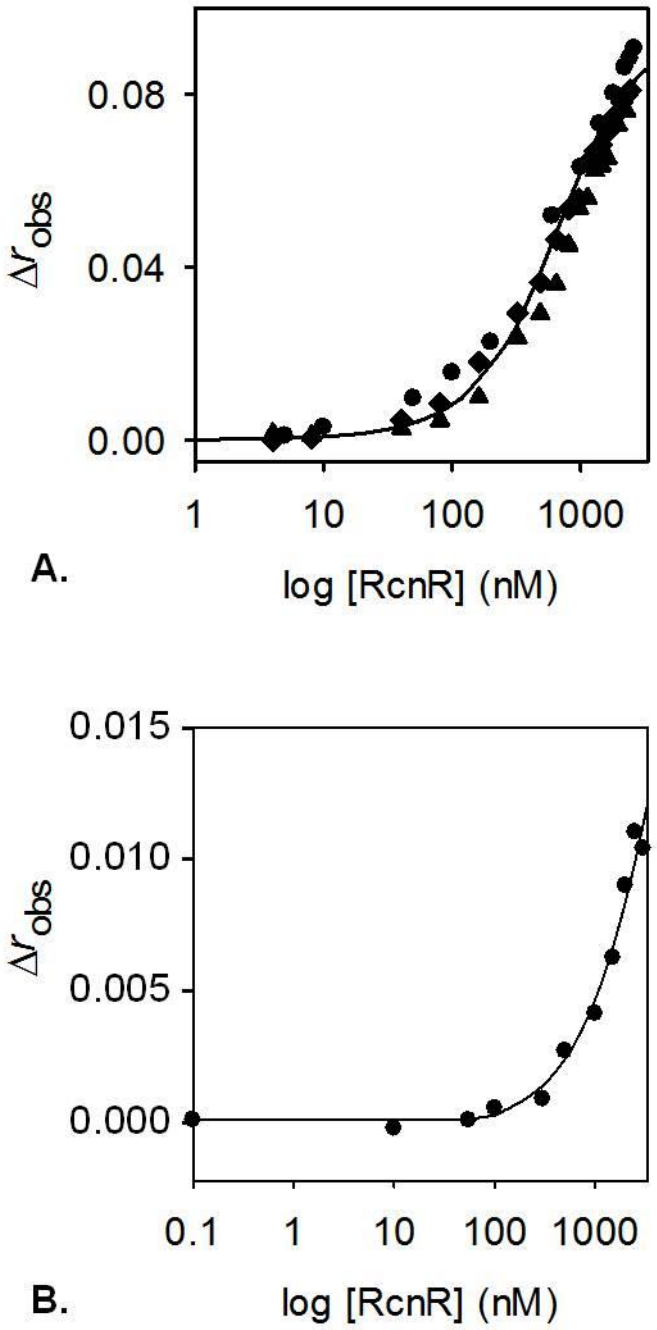


Figure 5. 14 *rcnRA*Pro^{sites1,2} (10 nM) (containing two C/G tracts (**A**)) and *rcnRA*Pro^{site1} (containing only one C/G tract (**B**)) were anaerobically titrated with RcnR in the presence of 5 mM EDTA. DNA binding was monitored by fluorescence anisotropy. Symbol shapes represent individual experiments. Solid lines represent simulated curves (in A the average K_{DNA} has been determined across the experiment replicates shown). Data were fit to a model describing (A) a 2:1 RcnR tetramer (non-dissociable):DNA stoichiometry), and (B) a 1:1 RcnR tetramer (non-dissociable):DNA stoichiometry).

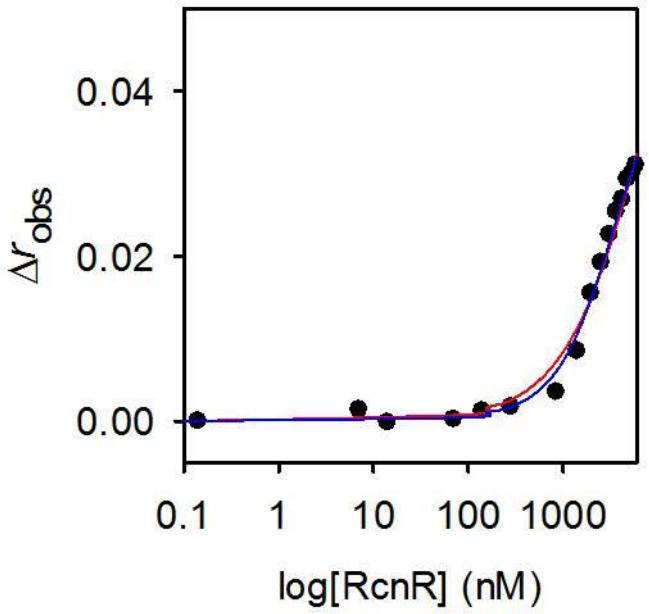


Figure 5. 15 *frmRA*Pro (10 nM) was anaerobically titrated with RcnR in the presence of 5 mM EDTA. DNA binding was monitored by fluorescence anisotropy. Final Δr_{obs} is consistent with what was previously observed for FrmR: *frmRA*Pro association. Solid lines represent simulated curves. Data were fit to a model describing a 2:1 RcnR tetramer (non-dissociable):DNA stoichiometry (blue line), and a 1:1 RcnR tetramer (non-dissociable):DNA stoichiometry (red line).

appears to better describe the experimental data, suggesting that RcnR is capable of binding this DNA region with a 2:1 stoichiometry, although with low affinity. In fact, both values are outside the range of DNA binding affinities commonly found in transcriptional regulators (~30 - 200 nM) (Ma *et al.* 2009, Iwig & Chivers 2009, Foster *et al.* 2012) therefore it is unlikely that cross regulation occurs.

5.5.7 *Effect of Co(II)- and Ni(II)-binding on formation of RcnR:rcnRAPro interaction*

As discussed in Chapter 1, *E. coli* RcnR responds to cobalt and nickel *in vivo*. To explore *Salmonella* RcnR *in vitro*, FA was employed to analyse the effect of Ni(II) and Co(II) on DNA binding. RcnR has upper limits on the K_D of Ni(II)- and Co(II)-binding of of 25 nM and 5 nM, respectively (Iwig *et al.* 2008). Assuming a comparable affinity for Ni(II)- and Co(II)-binding in *Salmonella* RcnR, titration of DNA was performed with RcnR pre-incubated with [Ni(II)] or [Co(II)] = 5 µM, which is likely to saturate the metal binding sites. DNA-binding reactions were monitored by fluorescence anisotropy. No increase in measured r_{obs} values was observed when the association experiment was performed in the presence of Ni(II), indicating an extremely weak binding of Ni(II)-RcnR to DNA under these conditions (Figure 5. 16A). $K_{rcnRAProsites1,2}^{Ni(II)RcnR}$ = 5911.7 ± 1287.0 nM was calculated using a model describing a 2:1 RcnR tetramer:DNA ratio. Figure 5. 16B shows comparison with apo-RcnR association binding curve already presented in Section 5.5.5.

The allosteric coupling free energy was calculated, using the thermodynamic equation described in Section 3.6.2, to be $\Delta Gc^{Ni(II)RcnR}$ = 2.17 ± 0.15 kcal mol^{-1}. An analogous result was obtained when the same experiment was conducted with Co(II) (Figure 5. 17A-B), giving $K_{rcnRAPro-sites1,2}^{Co(II)RcnR}$ = 14950 ± 1837 nM and $\Delta Gc^{Co(II)RcnR}$ = 2.72 ± 0.13 kcal mol^{-1}. These findings are consistent with Ni(II) and Co(II) functioning as allosteric effectors of *Salmonella* RcnR, as in *E. coli* RcnR (Iwig *et al.* 2008).

RcnR can now be used as a control to explore the specificity of the FrmR response to formaldehyde (Section 5.3.1).

5.6 Assessing *Salmonella* RcnR responsiveness to formaldehyde *in vitro*

5.6.1 *RcnR:DNA interaction by fluorescence anisotropy in the presence and absence of various concentrations of formaldehyde*

Analyses of the effects of formaldehyde at 20 µM, 50 µM, and 100 µM on DNA-binding was carried out by fluorescence anisotropy in the presence of a large excess of EDTA (5 mM)

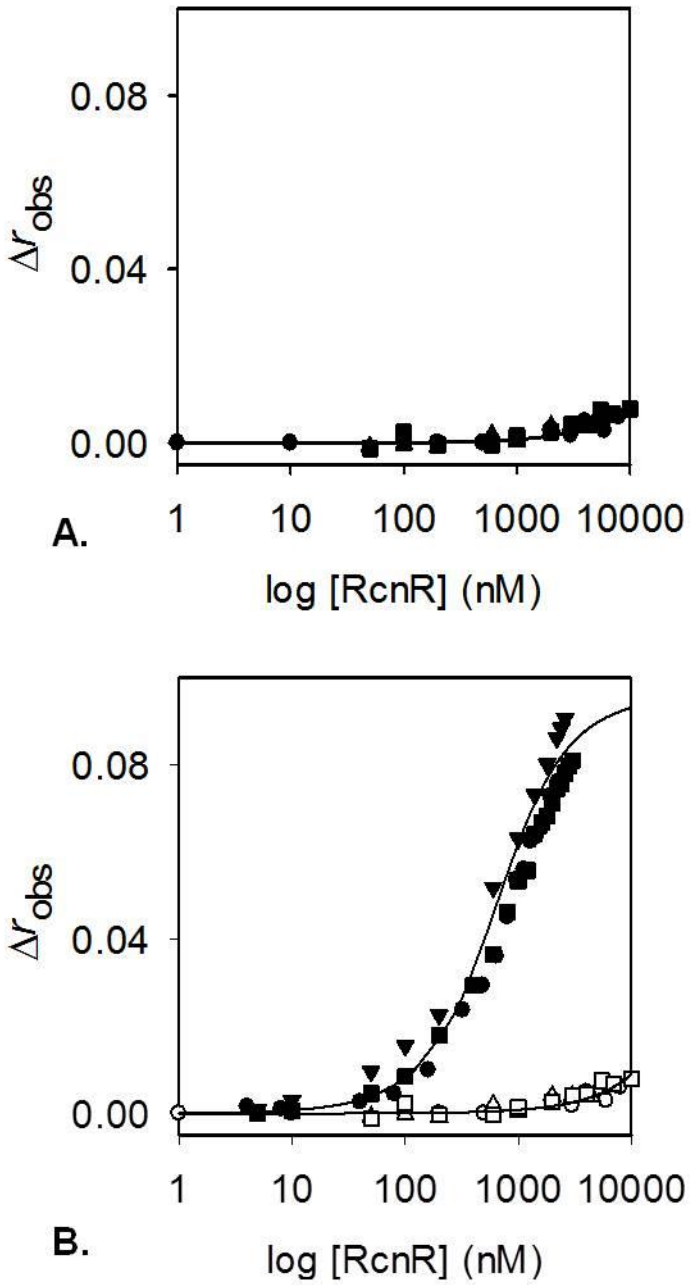


Figure 5. 16 A *rcnRA*Pro^{sites1,2} (10 nM) was anaerobically titrated with RcnR in the presence of of 5 µM $NiCl_2$ in the reaction buffer. RcnR samples were also incubated with 1.2 molar equivalent of $NiCl_2$. **B** Comparison between the data shown in A and the same experiment carried out in presence of EDTA and in absence of metal (previously shown in Figure 5. 15A). Solid lines represent simulated curves produced from the average K_{DNA} determined across the experiment replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 RcnR tetramer (non-dissociable):DNA stoichiometry.

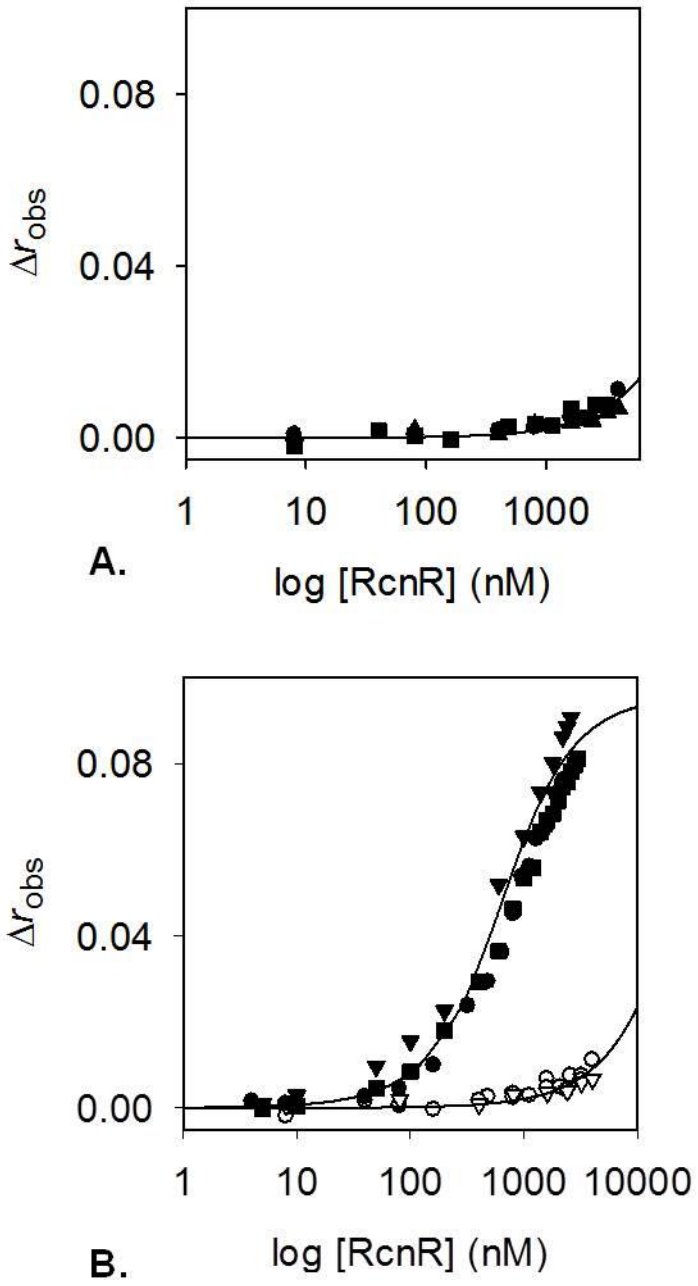


Figure 5. 17 Titration of *rcnRA*Pro^{sites1,2} with apo-RcnR and Co(II)-RcnR. **A** *rcnRA*Pro^{sites1,2} (10 nM) was anaerobically titrated with RcnR in the presence of of 5 μM $CoCl_2$ in the reaction buffer. RcnR samples were also incubated with 1.2 molar equivalent of $CoCl_2$. **B** Comparison between the data shown in A and the same experiment carried out in the presence of EDTA and in the absence of metal (previously shown in Figure 5. A). Solid lines represent simulated curves produced from the average K_{DNA} determined across the experimental replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 RcnR tetramer (non-dissociable):DNA stoichiometry.

(Figure 5. 18A-B-C). The estimated dissociation constant for protein binding to DNA ($K_{rcnRAPro\text{-}sites1,2}$) in the presence of 20 µM of formaldehyde is 157.6 ± 8.2 nM, (obtained for replicate titrations as routine) and shows that HCOH-RcnR interacts with *rcnRA*Pro essentially identically to apo-RcnR. Binding parameters for HCOH-RcnR when the concentration of formaldehyde is 50 µM ($K_{rcnRAPro\text{-}sites1,2}$ = 366.7 ± 110.3 nM, ΔG_C = 0.51 ± 0.18 kcal mol^{-1}) and 100 µM ($K_{rcnRAPro\text{-}sites1,2}$ = 774.9 ± 285.2 nM, ΔG_C = 0.94 ± 0.21 kcal mol^{-1}) describe a much weaker interaction between RcnR and formaldehyde compared with what was obtained for FrmR in the same conditions (Section 5.2.1). These results support the hypothesis that formaldehyde detection by FrmR is specific, with FrmR being especially reactive with formaldehyde relative to RcnR. In fact, 20 µM of formaldehyde does not prevent RcnR from binding to DNA whereas FrmR binding is almost completely disrupted in the same conditions. Also, when the experiment was carried out with the highest formaldehyde concentration used in this work (100 µM), we measured a ΔG_c for RcnR still smaller than the value obtained for FrmR at a [HCOH] ten-times lower (Table 8. 2, Appendix).

5.7 RcnR Ser2 → Pro: Generation of a formaldehyde sensor *in vitro* by a single point mutation

As the experiments described in the previous sections of this Chapter seem to confirm that Cys35 and Pro2 are involved in specific formaldehyde-binding (as shown in Section 5.4) and in order to further probe the predicted mechanism an attempt was made to create a FrmR-like formaldehyde binding site in *Salmonella* RcnR. RcnR already possesses a cysteine residue in position 35 whereas a serine, and not a proline, is located at the start of the N-terminal region. As previously observed for FrmR (Section 3.3.1), only two methionine residues (instead of three, as predicted by the theoretical sequence) were detected by Amino Acid Analysis (Table 8. 5, Appendix), suggesting that Met1 has been cleaved. In this scenario Ser2 is the actual first residue. In the next sections Ser2→Pro mutation will be described and S2PRcnR:*rcnRA*Pro^{sites1,2} interaction characterised in the absence and in presence of formaldehyde.

5.7.1 *Production and purification of recombinant S2PRcnR*

The serine residue in RcnR (Ser2), adjacent to the conserved W position of the CsoR/RcnR WXYZ motif (Figure 5. 10), was replaced with proline by site directed mutagenesis of pETRcnR to create the hypothesized FrmR-like formaldehyde binding site. The site directed mutant was purified following the protocol designed for wild-type RcnR (Chapter 2) (Figure 5. 19). Assays to determine free thiol content and metal contamination were performed as routine (Chapter 2).

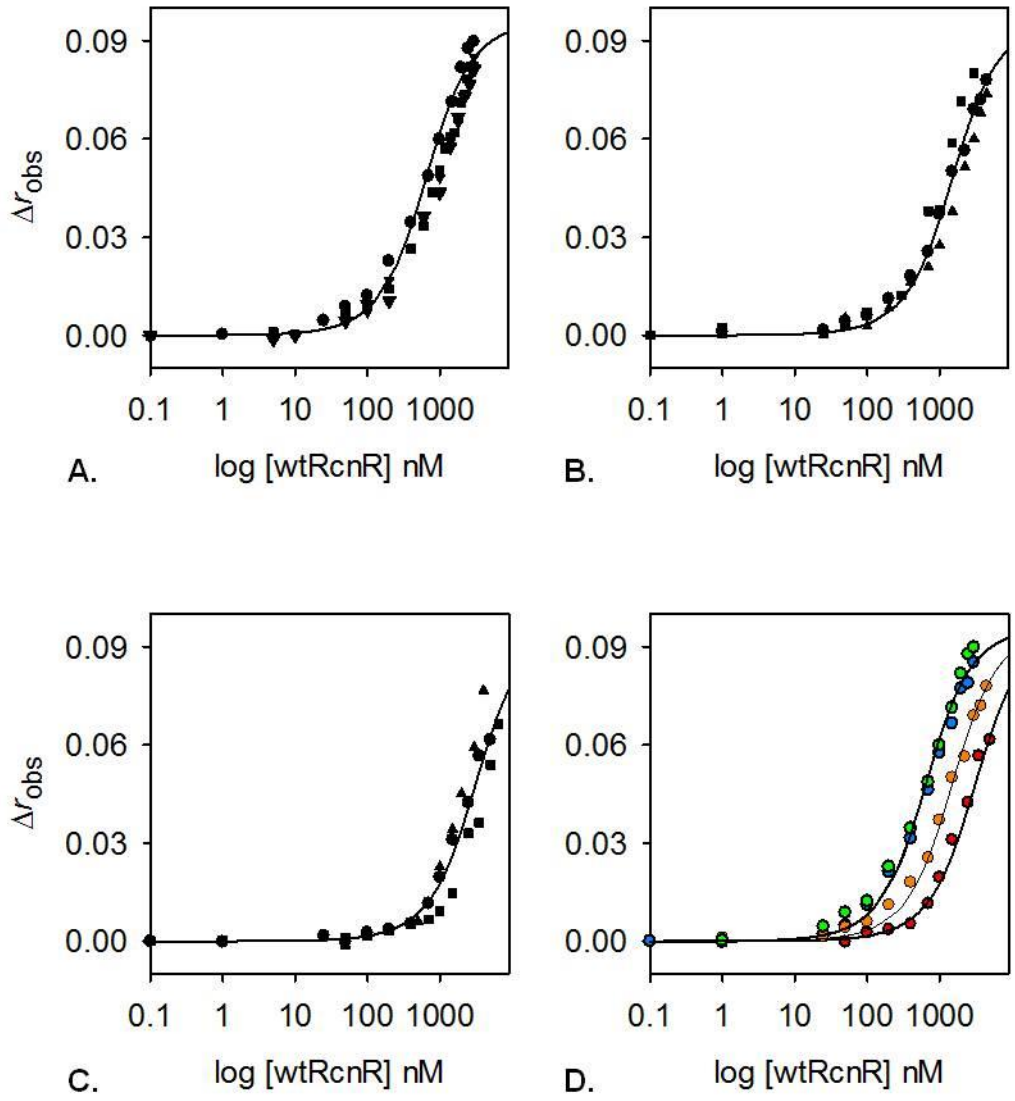


Figure 5. 18 *rcnRA*Pro^{sites1,2} (10 nM) was anaerobically titrated with RcnR in the presence of 5 mM EDTA plus 20 µM (**A**), 50 µM (**B**) and 100 µM (**C**) of formaldehyde. Figure **D** shows the comparison of data from the same RcnR:DNA association in the absence (blue circles), or presence of 20 µM (green circles), 50 µM (orange circles) and 100 µM (red circles) of formaldehyde. DNA binding was monitored by fluorescence anisotropy. Solid line represents simulated curves produced from the average K_{DNA} determined across the experimental replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 RcnR tetramer (non-dissociable):DNA stoichiometry.

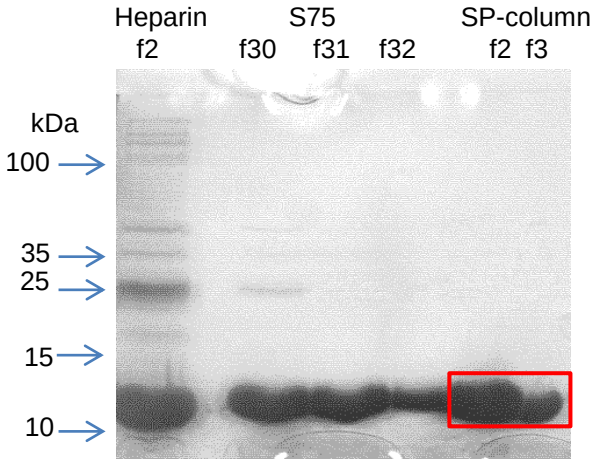


Figure 5. 19 SDS-PAGE analysis (18 % w/v acryl-bis) showing purification steps of S2PRcnR by nickel affinity chromatography, size exclusion and heparin affinity chromatography. "Heparin f2" represents the second fraction eluted from a 5 ml Heparin-affinity column with buffer B800 (Table 8. 1, Appendix) containing 800 mM NaCl. "S75 f30-f32" represents fractions 30-32 eluted from a Superdex S75 column loaded with fraction 2 (5.5 ml) from the previous purification step. "SP-column f2-f3" represents fractions eluted at 0.3 M NaCl from a 5 ml SP column with size exclusion fractions enriched for RcnR (f2 elution volume = 10 ml, f3 elution volume = 5 ml).

5.7.2 *S2PRcnR:DNA interaction by fluorescence anisotropy in presence and in absence of formaldehyde.*

To ensure that the substitution does not affect DNA binding affinity of RcnR, the ability of S2PRcnR to bind *rcnRA*Pro^{sites1,2} was assessed by fluorescence anisotropy (Figure 5. 20A). The DNA-binding affinity of the RcnR mutant was found to be $K_{rcnRAPro\text{-}sites1,2}{}^{S2PRcnR}$ = 155 ± 10 nM, which is identical to the value calculated for wild-type RcnR ($K_{rcnRAPro\text{-}sites1,2}{}^{RcnR}$ = 152 ± 8 nM, Section 5.5.5). This result indicates that S2PRcnR retains the ability to bind DNA with comparable affinity to wild-type. Fluorescence anisotropy was also used to monitor the interaction of S2PRcnR with DNA in the presence of [HCOH] = 20 µM, 50 µM and 100 µM (Figure 5. 20B-C and D, respectively). The binding parameters show a weakening in DNA-binding affinity even after only [HCOH] = 20 µM ($K_{rcnRAPro\text{-}sites1,2}{}^{S2PRcnR}$ = 696 ± 141 nM, ΔG_C = 0.88 ± 0.14 kcal mol^{-1}) which was not observed for RcnR and, as formaldehyde concentration increases, the differences with wild-type protein becomes more evident ([HCOH] = 50 µM, $K_{rcnRAPro\text{-}sites1,2}{}^{S2PRcnR}$ = 1781 ± 471 nM, ΔG_C = 1.43 ± 0.18 kcal mol^{-1}; [HCOH] = 100 µM, $K_{rcnRAPro\text{-}sites1,2}{}^{S2PRcnR}$ = 2620 ± 327 nM, ΔG_C = 1.67 ± 0.09 kcal mol^{-1}) (Table 8. 2, Appendix) (Figure 5. 21).

Although these outcomes provide additional support for the hypothesis that Pro2, in addition to Cys35, may be involved in formaldehyde sensing in FrmR, the smaller (compared to FrmR), ΔG_C values obtained with the engineered formaldehyde-binding RcnR protein suggest the possibility that the formaldehyde-mediated mechanism of allosteric negative regulation may differ in the two proteins and may involve other residues. Furthermore, while in FrmR the presence of Cys35 only (in P2SFrmR mutant) was sufficient to substantially impair DNA binding *in vitro*, in wild-type RcnR (which inherently lacks Pro2 but possesses Cys35) the effect of HCOH on DNA binding is noticeably smaller (Figures 5. 7 - 5. 18). The origin of these differences is not yet known and production of RcnR X-ray crystal structure could elucidate why the allosteric inhibition of DNA binding by formaldehyde-binding is not as effective as in FrmR.

5.8 Concluding remarks

Salmonella FrmR is a formaldehyde-sensing transcription factor belonging to the RcnR/CsoR family. In this chapter the detection of formaldehyde by FrmR has been explored and a putative reaction mechanism has been presented and analysed (see Sections 7.5.1, 7.5.2 and 7.5.3 for further discussion).

This mechanism has been directly tested by site-directed mutagenesis producing two FrmR mutants (C35AFrmR and P2SFrmR) each lacking one of the two residues thought to be involved in the selective detection of formaldehyde at low concentrations. Substitution of Cys35

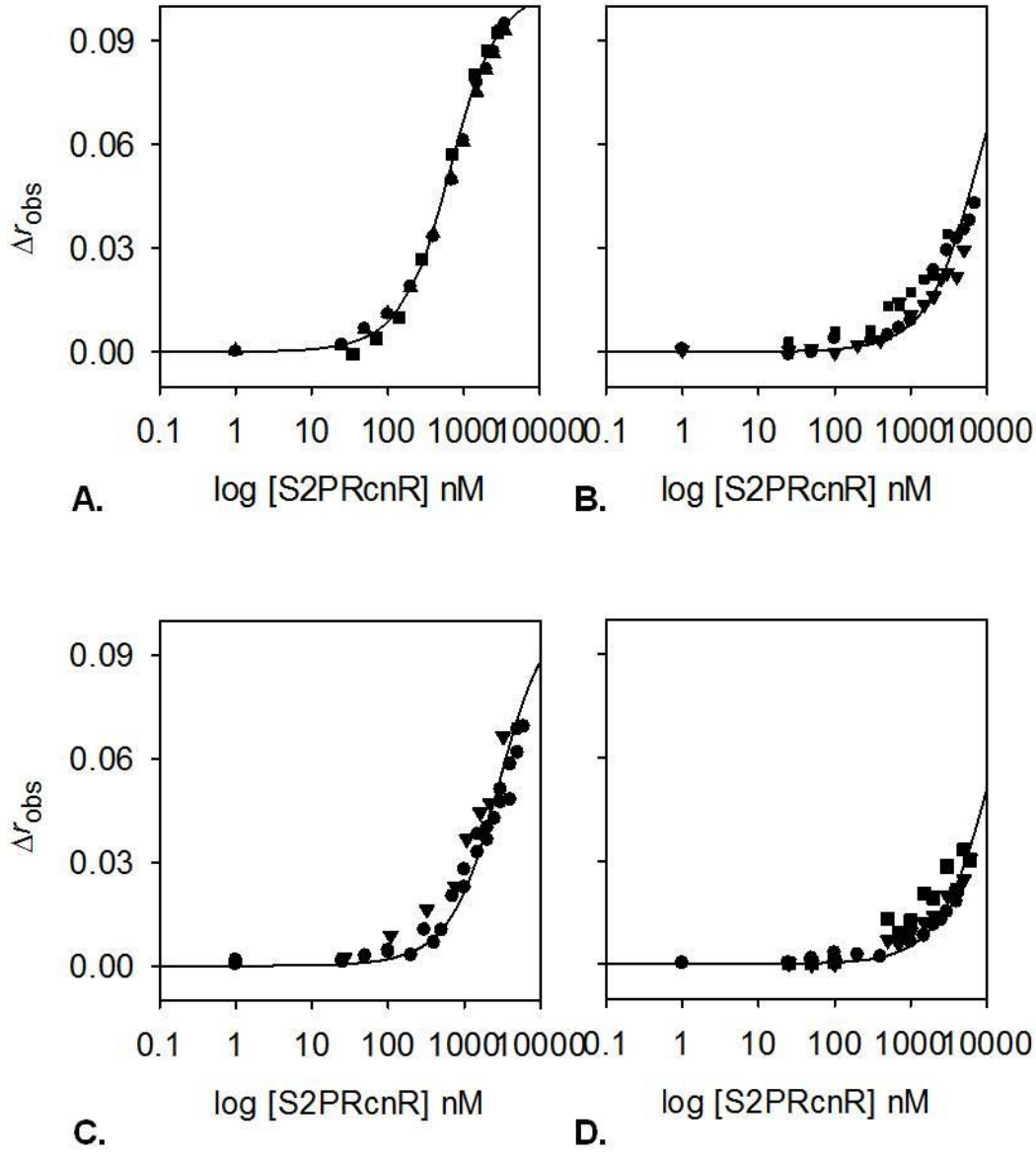


Figure 5. 20 *rcnRA*Pro^{sites1,2} (10 nM) was anaerobically titrated with S2PRcnR in the presence of 5 mM EDTA and in the presence of 0 µM (**A**), 20 µM (**B**), 50 µM (**C**) and 100 µM (**D**) of formaldehyde. DNA binding was monitored by fluorescence anisotropy. Solid line represents simulated curves produced from the average K_{DNA} determined across the experiment replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 S2PRcnR tetramer (non-dissociable):DNA stoichiometry.

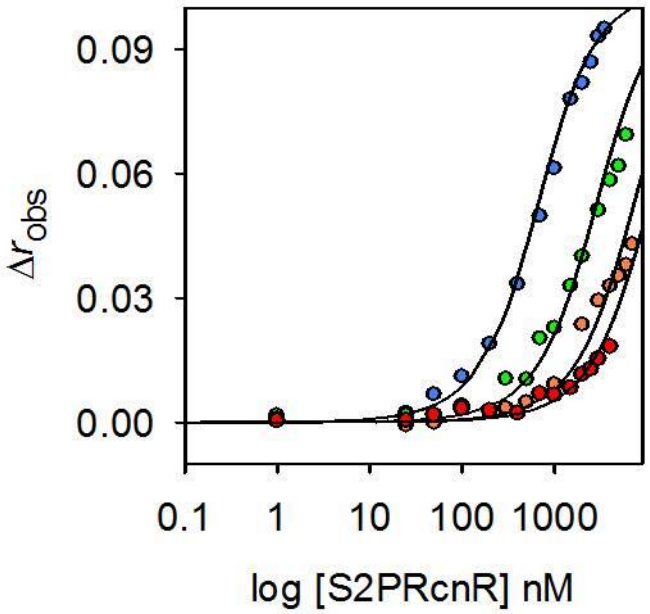


Figure 5. 21 Comparison of data from the S2PRcnR:DNA association in the absence (blue circles), or presence of 20 µM (green circles), 50 µM (orange circles) and 100 µM (red circles) of formaldehyde as previously shown in Figure 5. 21A-C. DNA-binding was monitored by fluorescence anisotropy. Solid line represents simulated curves produced from the average K_{DNA} determined across the experimental replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 S2PRcnR tetramer (non-dissociable):DNA stoichiometry.

(α2' helix) with Ala completely abolishes the protein's ability to sense formaldehyde suggesting that this residue is involved in the first step of the reaction with the aldehyde, probably yielding a thiohemiacetal intermediate. Pro2 (α1 helix) appears to be less fundamental as formaldehyde is still able to elicit an allosteric response in P2SFrmR, although to a lesser extent and at a higher concentration of formaldehyde. This result suggests that once Cys35 binds formaldehyde, other residues could intervene and interact with the Cys-formaldehyde adduct, inducing a perturbation in the protein conformation not as efficient as the Cys35-formaldehyde-Pro2 cross-link hypothesized in wild-type FrmR. Another explanation could be that the steric hindrance of the Cys-formaldehyde adduct may be sufficient to mis-align the two helices involved in DNA-binding (see Section 7.5.1 for further discussion).

In order to test the specificity of the formaldehyde detection by FrmR, *Salmonella* RcnR was identified and produced. Analysis of the interaction with its promoter operator region was investigated by EMSA and fluorescence anisotropy. Effects of low concentration formaldehyde along with cognate metals Ni(II) and Co(II) on DNA-binding were tested by fluorescence anisotropy, confirming that the formaldehyde concentration must be increased by 10-fold in order to detect a similar magnitude of K_{DNA} weakening as observed in FrmR (Section 7.6.1).

Furthermore, we applied our proposed mechanism to successfully create a novel *in vitro* formaldehyde-sensor by substituting the residue in position 2 of *Salmonella* RcnR with a Pro, (RcnR already possesses a Cys residue in position 35) (Section 7.7).

Chapter 6

Crystal structure determination of E64HFrmR

Acknowledgements are given to Dr. Ehmke Pohl, Dr. Morten Grøftehauge and Mr. Ian Edwards for their help and assistance given in the experiments described in this section. Data refinement and model building are currently being performed by Dr. Ehmke Pohl.

6.1 Aims and objectives

This chapter describes attempts to generate high resolution structures of the purified apo and Zn(II) forms of FrmR and E64HFrmR (from Chapter 3 and 4).

If successful, the production of a detailed molecular structure of the apo-form will provide information needed to elucidate the atomic details of protein interaction with formaldehyde and DNA operator promoter, whereas the structure of the Zn(II)-form will clarify which ligands participate in the zinc(II)-binding site. Furthermore, an aim is to detect the difference, if any, in the structure of FrmR and E64H variant as suggested by the biophysics assays findings discussed in Chapter 4 and the conformational change mediated by effector- (Zn(II)) binding.

6.2 Crystallization background

6.2.1 *Crystal assembly*

Crystallization is a purification technique and a separation process which occurs when a set of randomly arranged molecules are organized in a three-dimensional arrangement which is orderly, highly reproducible and very accurate. The selectivity of the process by which a molecule in the fluid is incorporated in the crystal, arises from the fact that every specific surface of the crystal recognizes the molecule and leads it to be arranged in the position that will have in the final structure. The final structure will be the one with a 3D layout where the bulk is more stabilized by intermolecular interactions (i.e. with the minimum lattice energy).

Equation 4 shows the fundamental equation for the free energy of crystallization:

$$\Delta G_c = \Delta H_c - T\,\Delta S_c$$

Equation 4

The thermodynamic driving force of the process is the increase of entropy of the solvent ($\Delta S_s > 0$) due to the release of water molecules which were ordered around both hydrophobic and polar residues. This factor must overcompensate other destabilizing contributions to ΔG_s such as the loss of entropy of the protein ($-25 < \Delta S_p < -75$ cal mol^{-1} K^{-1}) due to the reduction of degrees of freedom (during the formation of the crystal lattice, the protein loses translational and rotational freedom and limits conformational freedom). The contribution by enthalpy ΔH_p is favourable ($\Delta H_p > 0$) but negligible because the few contact points formed within the crystal are mediated by weak interactions. Therefore the thermodynamic equation (Equation 5) can be re-written taking in account the contributions to the system composed by protein and solvent:

$$\Delta G_c = \Delta H_c - T\,(\Delta S_p + \Delta S_s)$$

Equation 5

However, achievement of a negative ΔG_c is not the only requirement for crystal formation. Crystallization rate depends on the concentration of the molecule (here the protein). The concentration range in which the process occurs has a lower limit which is defined by the equilibrium composition of the saturated solution $c_{eq}(T)$. The solubility curve in Figure 6. 1 shows how the composition of a saturated solution varies with temperature *T*, which is only one of the many parameters involved in nucleation and crystal growth. The kinetic driving force of the self-assembly of molecules into a crystal is therefore constituted by the supersaturation. Supersaturation is a dimensionless quantity indicated with σ and corresponds to the difference between the chemical potential of a molecule at steady state and at supersaturation state. In the supersaturated solution, local and labile aggregates of molecules will form by thermal agitation (Brownian motion). This state is metastable and only if the radius of these aggregates exceeds a threshold value (critical radius) the formation of nuclei will occur (Figure 6. 1). The existence of a critical radius for the size of stable aggregates is linked to an energy barrier which opposes the nucleation. In fact the creation of a nucleus is accompanied by the formation of stabilizing interactions in the "bulk" but also by the creation of a surface at the interface between the solution (liquid phase) and the crystal (solid phase). The creation of this surface requires energy, because the molecules are in a state of tension not being able to satisfy all the interactions available.

6.2.2 *Crystallization techniques*

In order to crystallize biological molecules, factors such as temperature, pH, dielectric constant, presence of metal ions and/or additives, must be considered, in addition to the nature of the precipitant agent of choice. A common strategy involves a two-step approach: the use of commercial matrix sparse screens to coarse test a large number of crystallization and solubility conditions (with the aid of a crystallization robot), followed by optimization of specific crystallization conditions varying specific components of the crystallization cocktail (Jancarik & Kim 1991; Dessau & Modis 2011).

Spontaneous crystallization of a molecule occurs when the process is favored thermodynamically, therefore when the free energy of the system reaches a minimum. Since a biological macromolecule is located at a minimum of energy when it is fully solvated (and this condition must remain valid even in the crystal), a crystal protein contains a large number of water molecules. The difficulty of crystallizing a protein therefore consists in finding the chemical-physical conditions that allow the transition from solution to solid phase while the macromolecule remains solvated and consequently can assemble in a 3D ordered arrangement instead of an amorphous aggregate.

Since often it depends on the rate at which the saturation is reached, the strategy is to bring the system very slowly toward a minimum of solubility so as to obtain crystals suitable for X-ray

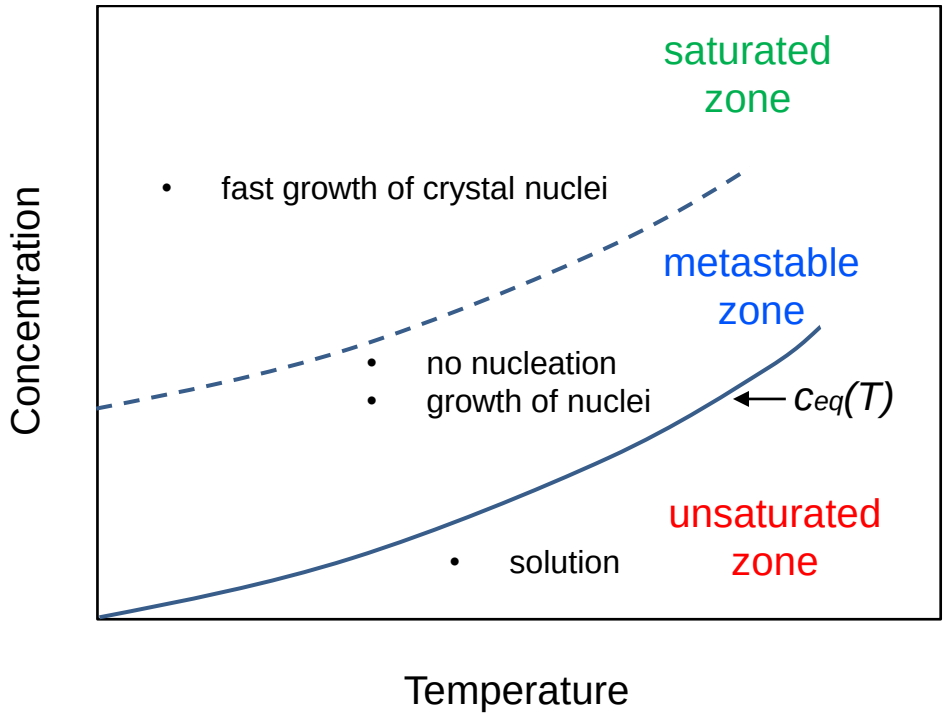


Figure 6. 1 Schematic representation of the solubility curve. If the solution has a concentration of solute below c_{eq} it is undersaturated and any crystal introduced in the solution will melt. A solution over the upper curve is supersaturated, and the content of dissolved solute is higher than c_{eq}. Crystals can nucleate and grow only from supersaturated solutions. Between these conditions there is a zone of supersaturation "metastable", in which nucleation does not take place and only preformed crystals will grow.

diffraction (Rupp 2009). There are various techniques to obtain supersaturation of the protein solution but microdialysis and vapor diffusion methods are the most common. In this work we used vapor diffusion methods, in particular sitting-drop for the initial screening of crystallization conditions and hanging-drop for the optimization. Vapor diffusion techniques exploit a system where, in a closed system, two solutions of the same substance, but at different concentrations, are allowed to equilibrate.

In the sitting-drop method (Figure 6. 2) a micro drop (200 nL) of the protein of interest is mixed with the precipitant (200 nL) and placed in a microbridge inside a transparent container (e.g. 96-well plate) which already contains ~ 100 µL of precipitant solution ("reservoir"). The concentration of precipitant equilibrates through the vapor phase and if the supersaturation condition is reached crystal growth can occur (Rupp 2009).

In the hanging-drop method (Figure 6. 2) a micro drop of protein solution (0.5 – 1 µL) is mixed with an analogous volume of precipitant and placed on a microscope coverslip. The coverslip is then suspended over a well containing 0.5 mL of precipitant solution (Rupp 2009).

6.2.3 *Diffraction experiment theory and principles*

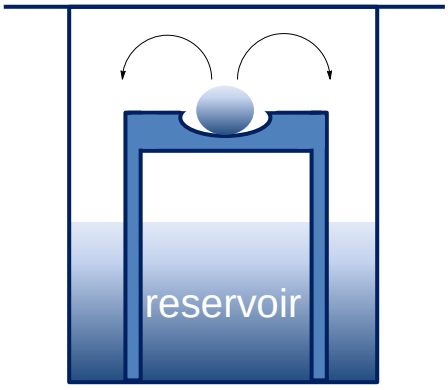
X-ray diffraction happens when high energy electromagnetic radiation interacts with the electrons of the atoms in the crystal. During a single-crystal diffraction experiment the X-ray beam encounters the single crystal of the material of interest. At this point the photons will either travel through the object (probability of ~ 98 %) or interact with the obstacle (~ 2 %). In the latter scenario, the electrons in the crystal will start to oscillate coherently within the photon's coherence length and will emit virtual waves which interfere constructively or destructively depending on their directions, following Equation 6 known as Bragg's Law (Bragg 1913; Rupp 2009):

$$n\lambda = 2d\sin\theta$$

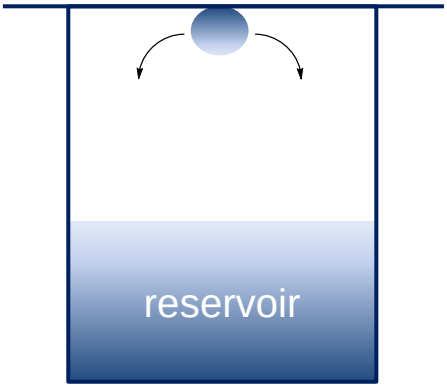
Equation 6

where n is an integer, λ is the wavelength of the photons, d the interplanar distance, and θ the scattering angle (see Figure 6. 3). W.H and W.L Bragg observed and explained for the first time the relationship between the angle of reflection to the X-ray and several parallel planes (due to the periodic nature of a crystal, this can be described as the sum of a series of equally spaced planes).

Only certain angles of incidence will produce in-phase diffracted waves and the probability of finding the scattered photon in a particular direction will be proportional to the amplitude of the resulting wave in that direction. The constructive and destructive events of a single photon will be recorded by a detector, such as CCDs (charge-coupled devices) (Walter *et al.* 1995) or pixel



Sitting drop



Hanging drop

Figure 6. 2 Schematic representation of sitting-drop and hanging-drop vapour diffusion techniques.

array detectors (Broennimann *et al.* 2006), and the sum of the diffraction events of all the photons will produce a pattern of spots scattered around the main incident beam (diffraction pattern). Bragg defined the relationship between the diffraction pattern and the crystal structure. Diffraction can be seen as the reflection of the primary beam by a set of parallel planes passing through the elementary cells of the crystal. Once the diffraction pattern has been recorded, Fourier methods are used to reconstruct the electron density of the molecule.

The electron density $\rho(x,y,z)$ of the crystal sample is a periodic function and therefore can be represented by a Fourier series (see Equation 7).

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} F_{hkl}\, e^{-2\pi i(hx+ky+lz)}$$

Equation 7

h,*k* and *l* are the indices of reflection *hkl* and identify a set of equivalent, parallel planes, whereas F_{hkl} is the structure factor describing the reflection (Rhodes 1993).

6.2.4 *Data processing and phase problem*

In order to produce a model for protein structure, data collection is followed by "indexing" (assignment of a set of integers *h k l* to each reflection and identification of a cell with sides *a, b, c* and angles *α, β, γ*), determination of potential space-group (as a result of crystal architecture) and "integration" (prediction of a diffraction pattern, measurement of the spot intensities and summing of the spots belonging to the same reflection). "Scaling" and "merging" steps consist in merging equivalent measurements and in putting all the reflections on a common scale. Subsequent determination of the space group and conversion of the intensities into structure factor amplitudes (|F|) provides information about twinning, B-factors, and many other factors.

Since the structure factor F_{hkl} is a wave equation (Section 6.2.3), it is defined by an amplitude, which is directly proportional to the intensity of the spot (I_{hkl}), and a phase, which cannot be directly obtained from the diffraction experiment. The phase is therefore the only additional parameter needed in order to obtain $\rho(x,y,z)$ and produce the 3D-structure of the desired molecule.

In order to solve the "phase problem", several common techniques, such as molecular replacement, heavy-atom method and anomalous scattering, can be adopted (Rhodes 1993). These methods will only provide an estimate of a limited number of phases but each of the thousands of reflections used in the Fourier series approximating the electron density has a phase. Further refinement using phase improvement and phase extension techniques is necessary to provide an interpretable electron density map (Rhodes 1993).

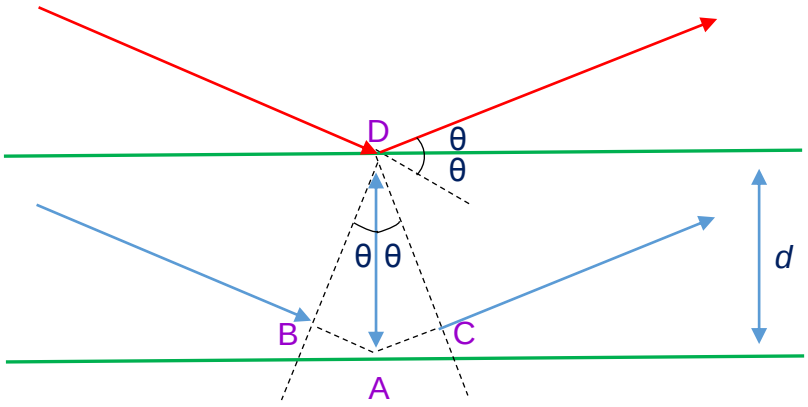


Figure 6. 3 The scheme shows the graphical interpretation of the Bragg's law. The blue and red arrows represent X-rays whereas the green lines represent planes of atoms. The interaction of the incident X-rays with the crystalline sample produces parallel diffracted rays with a reflection angle θ which is equal to the incident. The path difference between the two incident beams is BA + CA where BA = CA = $d\sin\theta$, consequently BA + CA = 2 $d\sin\theta$. Reflection will occur only if this "extra" distance corresponds to the wavelength of the X-ray.

6.3 Crystallization of apo- and zinc-forms of FrmR and E64HFrmR

An initial screening was performed using the sitting-drop vapor diffusion method. Several crystallization screens by Molecular Dimensions (Hampton Research) were set up in a 96-well crystallization plate format and the protein samples were dispensed with a Screen Maker robot (Innovadyne technologies). Crystallization kits (Hampton Research) adopted in this work are listed in Chapter 2 table 2.1.

Drops were inspected for appearance of crystals every day during the first week and every week afterwards. Both FrmR and E64HFrmR in the presence and absence of Zn(II) readily formed crystals in several conditions. In particular JCSG-*plus*™ (Page *et al.* 2003; Newman *et al.* 2005a-b; McPherson *et al.* 2001) (Table 2.2, Section 2.7.1) and PACT *premier*™ (Newman *et al.* 2005a-b) (Table 2.2, Section 2.7.1) kits were the most successful screens, providing hits where various crystal forms were found. Although the crystal size was often too small to be used in a diffraction experiment, these preliminary tests allowed identification of successful crystallization agents and pH range, and determination of the crystallization potential of the protein of interest in the presence of different ions.

The hanging-drop technique was adopted applying a systematic approach, performing grid screens of ammonium sulphate or PEG (polyethylene glycol) with various molecular masses and exploring the effect of small variation of pH and [NaCl] on nucleation and crystal growth (as described in Chapter 2). Larger and improved crystals were obtained within a few days although complete growth required ~ one week. A comprehensive list of successful crystallization conditions used at this stage can be found in Chapter 2 table 2.2.

Attempts to crystallize *E. coli* FrmR were reported in the Ph.D. dissertation by J. Law, although none of these were successful probably due to oxidation sensitivity of the protein (Law 2012). Crystals of C36AFrmR mutant (where Cys36 is the only cysteine residue present in *E coli* FrmR and corresponds to Cys35 in *Salmonella* FrmR) were obtained and diffracted with an adequate resolution (2.3 Å) although, no resulting structure was deposited in the Protein Data Bank (PDB) (Law 2012). Although the similarity between the two homologues is significant (74 %, Table 3.1), we did not encounter the same difficulties and various crystal forms of apo and Zn(II)-loaded FrmR and E64HFrmR were obtained (the most representative are shown in Figure 6. 4 and Figure 6. 5).

6.4 Diffraction data collection on FrmR and E64HFrmR crystals

Protein crystals, soaked into the reservoir solution, were cryomounted in small loops, tightly sealed in order to prevent evaporation of the solvent and disintegration of the crystal (Teng

1990). During the diffraction experiment the crystal was slowly rotated so as to have all the reflections in diffraction conditions. Data collection occurred at cryogenic temperatures (77 K) since high X-ray radiation elicits formation of free radicals which can destroy crystal contacts (Rupp 2009; Henderson 1990; Meents *et al.* 2010). Addition of a cryoprotectant to the sample prevents formation of ice in the crystals (Petsko 1975; Teng 1990; Hope *et al.* 1989).

Diffraction experiments were carried out by Dr. Ehmke Pohl on apo and zinc-forms of FrmR and E64HFrmR at the Diamond Light Source synchrotron facility using beamline ID: IO3. The optimal wavelength, 0.97625 Å, was selected by a monochromator and directed at the crystal. The images were recorded through 180^{o} with an individual oscillation angle of 0.5^{o} using an imaging plate. Most of the various crystal forms obtained during the course of this work diffracted, although the presence of a sub-optimal cryo-protectant caused low resolution and/or poor completeness in the collected data sets. However, a few data sets presented higher resolution and the best diffraction pattern in the set (relative to apo-E64HFrmR) is shown in Figure 6. 6. The plate-like apo-E64HFrmR crystals diffracted to a resolution of 2.19 Å. The data obtained from the diffraction experiment are listed in Table 6.3.

6.5 Data processing of apo-E64HFrmR diffraction data

Due to time constrain, diffraction data processing and refinement was carried out by Dr. Ehmke Pohl, therefore it will not be discussed in detail. Solution of the phase problem (Section 6.2.4) through molecular replacement was first attempted using Phaser (McCoy *et al.* 2007). This method exploits the use of a known structure of a protein with a great degree of similarity (generally > 30 %) with the protein of interest. It is based on the assumption that proteins with similar sequences tend to fold in a similar conformation (Anfinsen 1973). The atomic model for the known structure is rotated and translated in the unit cell of the unknown structure until the best fit between the calculated data (from the replaced model) and the observed (experimental) is reached (Rupp 2009). The unknown phases can then be estimated and a preliminary model can be built. The next step involves model refinement, where the differences between the amplitudes, observed and calculated by the model, are minimized thus to optimize the geometry of the model. Cycles of refinement and reconstruction are repeated until convergence between the calculated and the target model is reached. The phasing models for E64HFrmR were the two CsoR structures (PDB ID: 2HH7 from *Mycobacterium tuberculosis* and PDB ID: 3AAI from *Thermus thermophilus HB8*) and *Synechocystis* InrS structure which has been recently solved by colleagues in my research group (Foster *et al.*, in preparation) (Table 6.2). Unfortunately, it was not possible to obtain a plausible molecular replacement solution.

ARCIMBOLDO, a program designed to solve the phase problem for difficult macromolecule structures at medium resolution (~ 2 Å) (Rodriguez *et al.* 2009; Rodriguez 2012), was then implemented and the phase was successfully determined. Its approach combines the use of

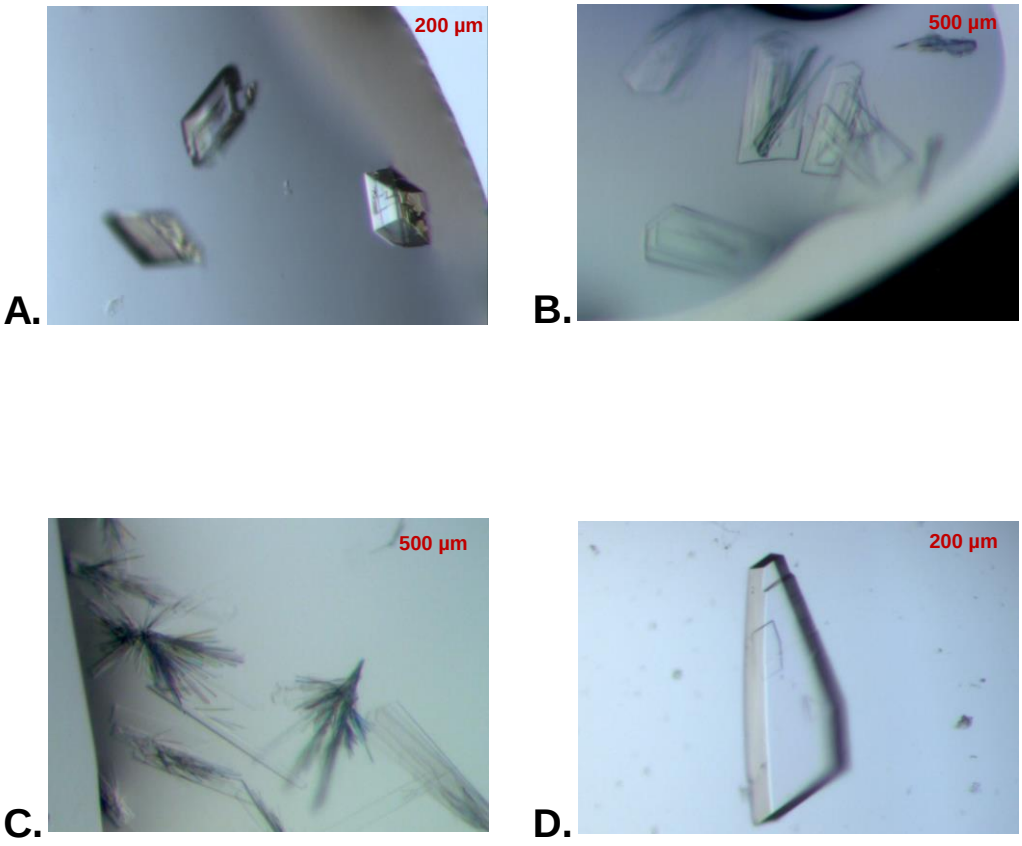


Figure 6. 4 Pictures taken on E64HFrmR crystals formed in **A** screen 6. (ammonium sulphate 1.8M, pH 5.0): perfectly formed cubic single crystals in hanging drop; **B** screen 4. (PEG2000 MME 23 %, pH 4.6): thin plates in hanging drop; **C** screen 8. (PEG1500 15 %, pH 5.1): cluster of thin needles in hanging drop; **D** screen 1. (PEG4000 17 %, pH 6.0): single irregular plate-like crystal in hanging drop. Screen compositions are described in Table 2.3, Chapter 2.

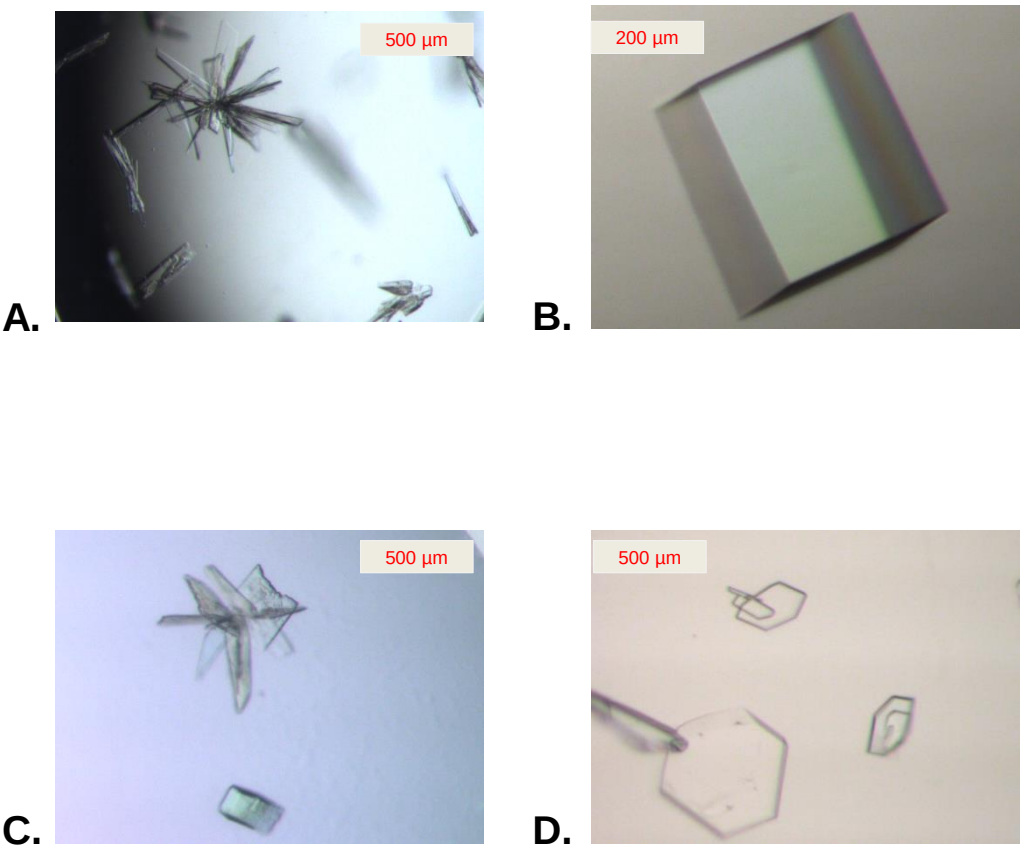


Figure 6. 5 Pictures taken on FrmR crystals formed in **A** well A4 of screen 2. (PEG3350 17 %, pH 4.9): cluster of thin needles in hanging drop; **B** screen 1. (PEG4000 17 % pH 6.5): perfectly formed cubic single crystal in hanging drop; **C** well C1 of JCSG-*plus*™: irregular dendritic growth of crystals in sitting drop, and **D** screen 1. (PEG4000 19 %, pH 6.5): thin plates in hanging drop. Screen compositions are described in Table 2.3, Chapter 2.

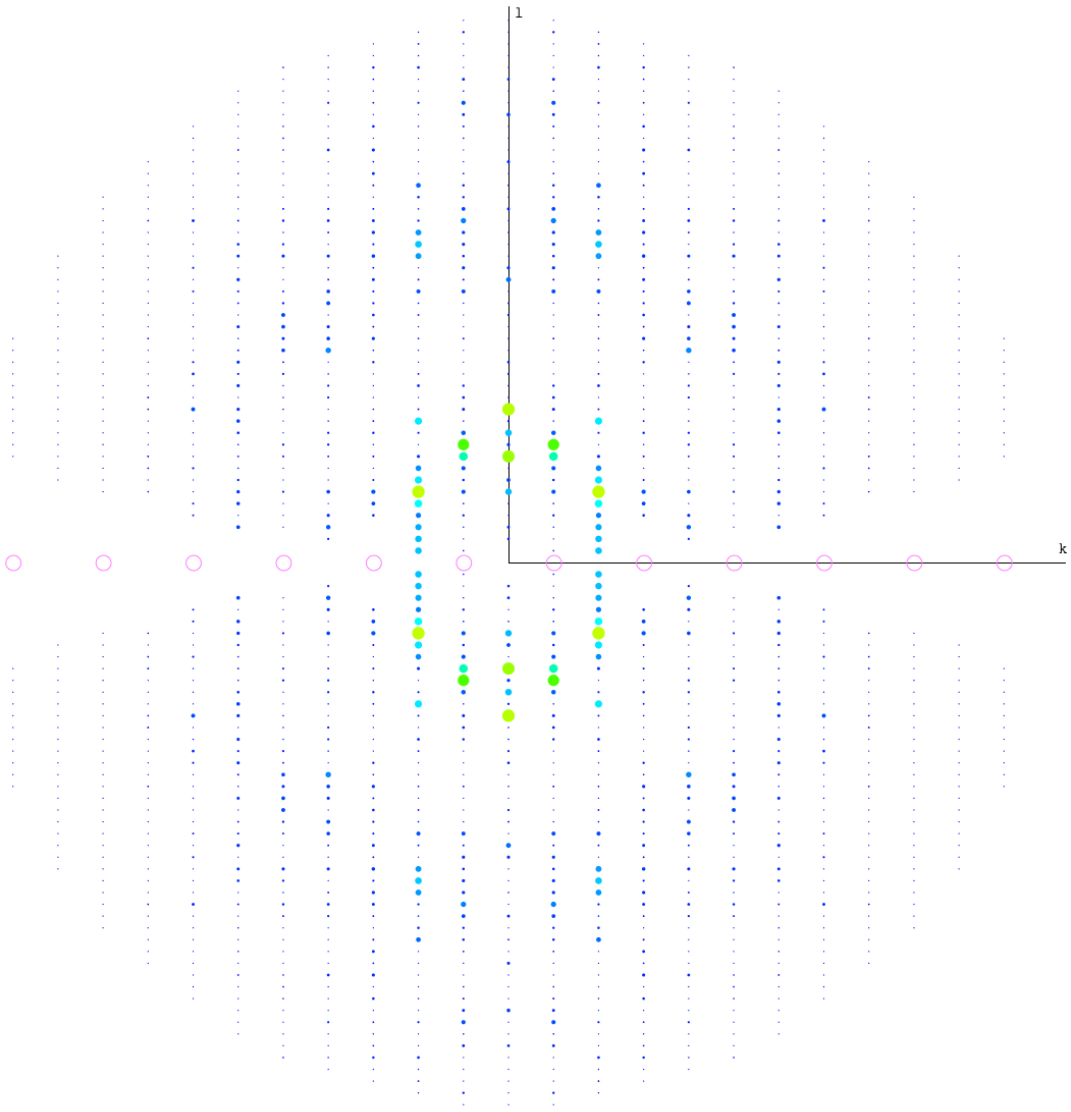


Figure 6. 6 Diffraction pattern (h = 0, k, l) showing integrated data obtained from the plate-like apo-E64HFrmR crystal using the Phenix software (Adams *et al*. 2002). The intensity of each dot is informative of the intensity of each reflection and the color coding refers to the signal-to-noise ratio I/σ(I). The empty red circles (0, k, 0) are the systematic extinct reflections.

Phaser, to determine the accurate location of small fragments, and SHELXE (Sheldrick 2008, 2010) to perform electron density modification and autotracing in order to expand the small fragment structures to a substantial part of the macromolecule structure (Rodriguez 2012).

At the time of writing, a partial structure has been determined and preliminary model refinement is still ongoing; any communication can be directed to Dr. Ehmke Pohl.

6.6 Concluding remarks

In order to further investigate and validate the mechanistic hypotheses proposed in the previous chapters of this work, a 3D model of FrmR proteins was pursued by X-ray crystallography. Several crystals of FrmR and E64HFrmR were obtained in the presence or absence of $ZnCl_2$ and analysed at the Diamond Light Source synchrotron facility consenting collection of various data-sets. In particular, one of the tested apo-E64HFrmR crystals was suitably diffracting (2.19 Å). Data processing and model building have been performed by Dr. Ehmke Pohl. At the time of Ph.D. thesis submission the 3D electron density model is still undergoing refinement, however, preliminary outcomes will be shown in Chapter 7 in order to aid the interpretation of results presented in previous sections of this work.

PROCESSING PROGRAMS	
iMosflm	Indexing and integration of images
SCALA	Scaling and merging
Pointless	Space group determination
Truncate	Conversion of intensities to structure factors, Generation of a 5 % R-free dataset
Phaser	Phase determination
Arcimboldo	Phase determination

Table 6. 1 List and description of programmes adopted during processing of diffraction data-sets collected at Diamond Light Source facility.

***Salmonella* FrmR**	**IDENTITY**	**SIMILARITY**
***Mycobacterium tuberculosis* CsoR** **PDB ID: 2HH7**	21.9 %	35.2 %
***Thermus thermophilus* CsoR** **PDB ID: 3AAI**	28.3 %	47.5 %
***Synechocystis* InrS** **PDB ID: 5FMN**	13.6 %	27.8 %

Table 6. 2 Identity and similarity values obtained comparing *Salmonella* FrmR sequence with those of proteins adopted in MR process by use of EBLOSUM62 matrix (EMBOSS NEEDLE). At the time of writing InrS structure (ID: 5FMN) is under embargo until paper publication.

	E64HFrmR
Space group	P 1 2_1 1
Unite cell dimensions	a = 68.79 Å b = 25.68 Å c = 100.50 Å α = 90.00 ° β = 103.11 ° γ = 90.00 °
Wavelength	0.97625 Å
Resolution limits	2.19 – 29.70 Å
Number of observations/ Unique reflections	124698 / 19983
Completeness (%)	98.5
Mean I/σI	12.2
R_{merge}	0.081
Multiplicity	6.3

Table 6. 3 Data collection statistics for apo-E64HFrmR high resolution synchrotron data set. The data analysis has been performed by Dr. Ehmke Pohl using iMosflm (Leslie *et al.* 2002; Battye *et al.* 2011).

Chapter 7

Discussion and future work

7.1 Summary of results

Prior to the start of this work, *stm1628* was hypothesised to encode a metal-sensing transcription factor due to the high degree of similarity (64.1 %, Table 3.1) with the *E. coli* Ni(II)/Co(II)-sensing RcnR, member of the RcnR/CsoR family of metalloregulators.

Due to its genomic location, STM1628 was supposed to regulate *stm1627* (located right upstream *stm1628,* Figure 3. 1A), predicted to encode a putative Zn(II)-requiring class III alcohol dehydrogenase. Alcohol dehydrogenases can be found in animals, yeasts, plants and bacteria (Branden *et al.* 1975) and are classified depending on the presence or absence of metal ions, typically zinc or iron. In zinc-requiring alcohol dehydrogenases, Zn(II) ions play both catalytic and structural roles (Jörnvall *et al.* 1987). Since Zn(II) was presumably necessary for STM1627 activity, this transition metal appeared a plausible candidate as STM1628 effector, able to trigger an allosteric response in the transcription regulator.

In the light of these observations, STM1628 metal-binding properties were investigated *in vitro*. Initially, protein quantification of purified samples was carried out by applying the Lambert-Beer equation (Section 3. 3), using FrmR predicted extinction coefficient, $\varepsilon^{theoretical}$ 1490 M^{-1} cm^{-1}. This value was obtained using the ProtParam tool (Gasteiger *et al.* 2005), available at the Expasy Proteomics Server (Gasteiger *et al.* 2003). Only when characterization experiments did not show an exact stoichiometric ratio between the protein and various metals, a protein sample was analysed by Amino Acid Analysis (AltaBioscience Ltd) to determine the experimental extinction coefficient, which resulted to be $\varepsilon^{experimental}$ 1951 M^{-1} cm^{-1}, allowing rigorous quantification of FrmR samples. The protein was found to bind various transition metals, such as Zn(II), Cu(I), Co(II) and Ni(II), adopting a number of geometries and involving a mercapto group from the only cysteine residue (Cys35) (Figure 7. 4).

Metal-binding affinities were measured by competition with chelators of known affinities, revealing sub-nanomolar and femtomolar values, respectively, of $K_{Zn(II)}^{FrmR}$ and $K_{Cu(I)}^{FrmR}$ for the tightest sites (Figures 3. 1-3. 13). Fluorescence anisotropy studies revealed that Zn(II) and Cu(I) were capable of disrupting STM1628:DNA interaction, contributing to the formulation of a hypothetical mechanism where STM1628 acts as a metallo-regulatory protein, presumably detecting zinc.

Twelve months in the project our research group got approval to establish a class 2 laboratory, allowing us to study *Salmonella* STM1628' behavior *in vivo*. *B-galactosidase* assays carried out by Dr. Osman on *Salmonella* Δ*stm1628* cells with $P_{stm1628\text{-}stm1627}$-*stm1628* fused to *lacZ* showed that only exposure to maximum non-inhibitory concentrations (MNIC) of formaldehyde alleviates repression by *Salmonella* STM1628 but not exposure to various metals (Mn(II), Fe(III), Co(II), Zn(II), Cu(II), Ni(II), Se(IV) and Te(IV)) (Figures 4. 1A-B).

These studies supported STM1627 identification by bioinformatics analyses as the formaldehyde dehydrogenase FrmA (STM1627 shares 90.6 % of sequence similarity with *E. coli* FrmA, Figure 7. 1), and STM1628 as a *Salmonella* homologue of *E. coli* formaldehyde-sensing transcriptional factor FrmR.

In light of these findings, and being aware of the ability of Zn(II) not only to bind to STM1628 (now called FrmR) with sub-nanomolar affinity, but also to elicit the negative allosteric regulation of *stm1628-stm1627* (now called *frmRA*), we decided to investigate if there may be a relationship between Zn(II)-binding and formaldehyde sensing. A model where Zn(II) directly coordinates formaldehyde, further increasing the polarity of the carbonyl group due to its Lewis acid character, was then proposed (inset Figure 7. 2) and tested by time-course fluorescence anisotropy. As shown in Figure 7. 2A-B, disruption of the pre-formed FrmR:DNA complex due to formaldehyde, occurred faster when 0.2 molar equivalents of zinc was added to the reaction buffer. The same experiment was repeated in the absence of formaldehyde and adding $ZnCl_2$ (0.2 molar equivalent) or EDTA (5 mM) Figure 7. 2C-D, showing a small degree of dissociation of the protein-DNA complex in both cases.

It was not clear if the finding shown in Figure 7. 2B was due to the formation of a zinc-formaldehyde adduct with enhanced electrophilicity, or if we were simply additively monitoring the distinct effects that zinc and formaldehyde have on FrmR-DNA complexes. The hypothesis of Zn(II) aiding the nucleophilic attack on the carbon by enhancing the polarization of the carbonyl group was later discarded due to lack of supporting evidence.

The discovery that the formaldehyde-sensor FrmR is capable of binding metals *in vitro* but fails to do so in the cell, where only formaldehyde is detected, confirms that the relative (and not the absolute) effector preference of a sensor determines which effector species will be detected *in vivo*. FrmR cannot access Zn(II) or Cu(I) *in vivo* because *Salmonella* Zn(II)- and Cu(I)-sensors (ZntR, ZuR for Zn(II), and CueR for Cu(I)) maintain the metal level below a lower set point (Section 4.6.1; Osman *et al*. 2015).

In the attempt to construct an RcnR-like metal-binding site in FrmR by single-point mutation (Glu64 → His), we created a sensor of cellular zinc and cobalt (Section 4.3). FrmR and E64HFrmR properties were consequently compared in order to directly test hypotheses about the mechanisms determining which metals are detected by a given sensor in cells, previously introduced in Section 1.3. These results will be further discussed in Section 7.3.

In addition, we investigated the ability of FrmR to detect cellular formaldehyde, proposing a reaction mechanism which was tested by site-directed mutagenesis *in vitro*. Mutagenesis and sequence conservation studies already present in literature suggested a model where formaldehyde interacts with Cys35 and/or an amino acid residue in the N-terminus region (Higgins & Giedroc 2014). In this work we suggest that Pro2 in the N-terminus region can be a

```
Sty_STM1627        1 MKSRAAVAFGPGQPLKIVEIDVAPPKKGEVLVKITHTGVCHTDAFTLSGD     50
                     |||||||||.||:||:|||||||||||||||:|:||||||||||||||||
Ecoli_FrmA         1 MKSRAAVAFAPGKPLEIVEIDVAPPKKGEVLIKVTHTGVCHTDAFTLSGD     50

Sty_STM1627       51 DPEGVFPAVLGHEGGGVVVEVGEGVTSLKPGDHVIPLYTAECGECKFCKS    100
                     |||||||.||||||.||||||||||:|||||||||||||||||:||:|
Ecoli_FrmA        51 DPEGVFPVVLGHEGAGVVVEVGEGVTSVKPGDHVIPLYTAECGECEFCRS    100

Sty_STM1627      101 GKTNLCQAVRATQGKGLMPDGTTRFSYNGEPVYHYMGTSTFSEYTVCAEI    150
                     ||||||.|||.||||||||||||||||:|:|||||.||||||||.||:
Ecoli_FrmA       101 GKTNLCVAVRETQGKGLMPDGTTRFSYNGQPLYHYMGCSTFSEYTVVAEV    150

Sty_STM1627      151 SLAKVNPQAPLDKVCLLGCGVTTGIGAVHNTAKVKAGDTVAVFGLGGIGL    200
                     ||||:||:|..:.|||||||||||||||||:.||:||||||||.|||
Ecoli_FrmA       151 SLAKINPEANHEHVCLLGCGVTTGIGAVHNTAKVQPGDSVAVFGLGAIGL    200

Sty_STM1627      201 AVIQGAVQAKAGRILAVDTNPEKFTLAGEMGATDFINPNDYDKPIQDVIV    250
                     ||:|||.|||||||:|:||||:||.||...||||.||||||||||:||::
Ecoli_FrmA       201 AVVQGARQAKAGRIIAIDTNPKKFDLARRFGATDCINPNDYDKPIKDVLL    250

Sty_STM1627      251 ELTDGGVDFSFECIGNVNVMRAALECCHKGWGESIIIGVAGAGQEIKTRP    300
                     ::...|:|.:|||||||||||||||..|:|||:|:||||.|||||.|||
Ecoli_FrmA       251 DINKWGIDHTFECIGNVNVMRAALESAHRGWGQSVIIGVAVAGQEISTRP    300

Sty_STM1627      301 FQLVTGRVWRGSAFGGVKGRTQLPGMVEDAMNGKIRLDPFITHRLPLEQI    350
                     |||||||||:||||||||||:||||||||||.|.|.|:||:||.:.|::|
Ecoli_FrmA       301 FQLVTGRVWKGSAFGGVKGRSQLPGMVEDAMKGDIDLEPFVTHTMSLDEI    350

Sty_STM1627      351 NDAFELMHQGKSIRTVIHFGDN    372
                     ||||:|||:||||||||.:
Ecoli_FrmA       351 NDAFDLMHEGKSIRTVIRY---    369
```

Figure 7. 1 Alignments of *S. Typhimurium* STM1627 (Sty_STM1627) with *E. coli* FrmA (Ecoli_FrmA) by Clustalw2.

S. typhimurium STM1627	**Similarity (%)**	**Identity (%)**
E. coli FrmA	90.6	80.1

Table 7. 1 Similarity and identity of amino acid sequence of *E. coli* FrmA compared to *S. Typhimurium* STM1627 using EMBOSS NEEDLE.

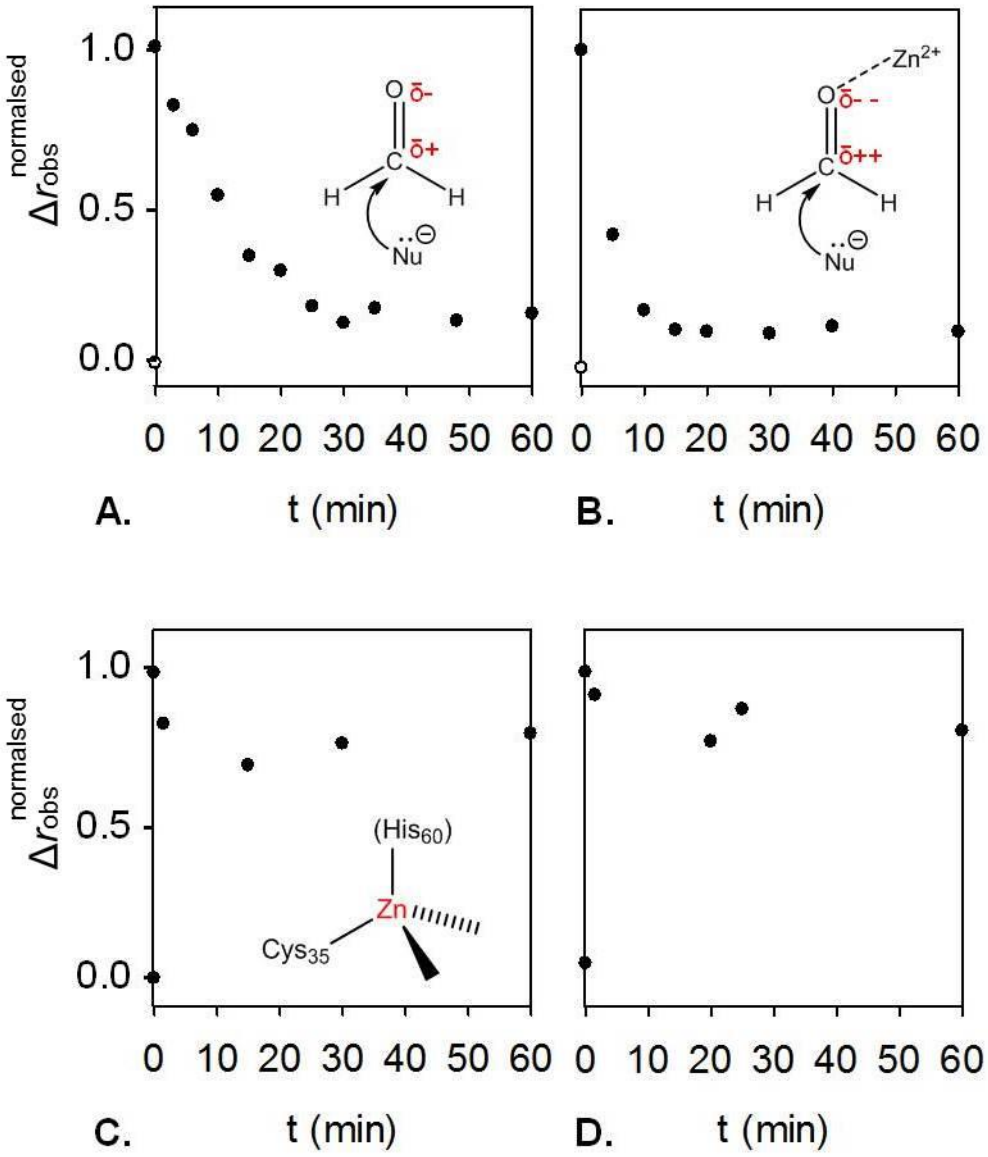


Figure 7. 2 Dissociation of pre-formed FrmR:DNA complexes in the presence of HCOH and/or Zn(II) was monitored by fluorescence anisotropy (TIME drive software). *frmRA*Pro (10 nM) was pre-incubated with FrmR (2 µM) before addition of HCOH (20 µM) (**A**), HCOH (20 µM) and Zn(II) (0.4 µM) (**B**), Zn(II) (0.4 µM) (**C**), or nothing (**D**). Experiments shown in A and D were conducted in the presence of 5 mM EDTA. Experiments were performed anaerobically at pH 7.0. The total anisotropy change is normalised to the apo-FrmR value (Δr_{obs}^{max}). Inset in Figure B shows Zn(II) enhancing the electrophilic character of the carbonyl C, aiding attack from a nucleophilic species (Nu^-). The charge of the metal ion is highly concentrated because of its small ionic radius (0.65 Å) (Frausto Da Silva & Williams 2001), and, in addition to a high ionization potential, it makes Zn(II) a strong Lewis acid.

possible site of formaldehyde modification. Furthermore, the gene encoding for *Salmonella* RcnR was identified and the protein, belonging to the same family of transcriptional regulators of FrmR, was used to test the specificity of FrmR formaldehyde responsiveness. By a single point-mutation (Ser2 → Pro) we successfully switched the Ni(II)/Co(II) sensor RcnR into a formaldehyde sensor *in vitro*, further endorsing our proposed mechanism.

Investigation of FrmR structure was pursued by producing crystals of apo- and Zn(II)-FrmR and E64HFrmR forms, which were then analysed at the Diamond Light Source. The best dataset (apo-E64HFrmR, resolution 2.19 Å) is currently being processed by Dr. Ehmke Pohl (Durham University) in order to produce a 3D-model.

In the following sections, selected topics from this work will be further discussed and developed.

7.2 Formaldehyde-sensing *Salmonella* FrmR is capable of binding metals

Zn(II) and Co(II) share the same binding site and are coordinated by FrmR with a tetrahedral geometry and a 1:1 stoichiometry (Figures 3. 4 - 3. 5 - 7. 4A). FrmR also binds at least one Cu(I) ion per protomer adopting a trigonal planar geometry and one Ni(II) ion with a square planar geometry (Figures 3. 8 - 3. 9- 7. 4 A).

E. coli RcnR, which shares a high degree of similarity with *Salmonella* FrmR, uses the primary amine group of the amino terminus region and the side chains of His3, Cys35, His60 and His64 (*E. coli* RcnR notation) to detect Co(II) (Iwig *et al*. 2008). The same residues are used to coordinate Ni(II), with the exception of His60. (Iwig *et al*. 2008). These residues (WXYZ motif) constitute plausible candidates for metal-binding ligands in FrmR and were therefore tested by site-directed mutagenesis (Section 3.8).

Cys35 (residue X) is the only cysteine in the FrmR sequence and spectral features attributable to LMCTs transitions between a thiol group and the metal ion have been observed with all the spectrally-active tested metals (Sections 3.4.1-3.4.2). Mutation of Cys35 to Ala abolished Zn(II)-binding ability as monitored using size-exclusion chromatography and fluorescence anisotropy (Figures 3. 27A-B, 3. 28B-C) confirming that the side chain of this residue acts as a zinc ligand in FrmR.

His60 (residue Y) was also tested via a His60 → Leu mutation. H60LFrmR appears to possess a diminished Zn(II)-binding affinity compared to that of wild type as it is able to retain 1 molar equivalent of Zn(II) when applied to size-exclusion column only if 20 µM of Zn(II) is present in the elution buffer (Figure 3. 30). Quenching of the intrinsic fluorescence of Tyr89 in H60LFrmR upon zinc titration reports of a 1:1 protein:metal stoichiometry and of an unaltered

tyrosine environment, suggesting that His60 may have been replaced by a nearby ligand (e.g. Glu58) in the metal coordination. Alternatively, His60 may not be a coordinating ligand in wild type FrmR, but its substitution with a Leu may cause a local rearrangement determining a weakened zinc affinity. In fact, unlike the hydrophobic amino acid Leu, which is usually buried in the protein core, His has a side-chain with a p*K*a (~ 6.5) close to the physiological pH, hence it may prefer a more hydrophobic or hydrophilic (solvent-exposed) environment depending on the local pH (Betts & Russell 2003).

Position Z of FrmR is occupied by a glutamate (Glu64), which was mutated to a His during the course of this work in order to investigate other protein features (see Section 7.4). The resulting E64HFrmR variant showed a tighter $K_{Zn(II)}$ of the tightest site (Appendix, Table 8. 3), suggesting that residue Z is in the correct position to coordinate the metal ion, in accord to what was observed for *E. coli* RcnR (Iwig *et al*. 2008). Metal-binding properties of E64HFrmR mutant will be extensively discussed in Section 7.4. However, there is not sufficient evidence to indicate that Glu64 contributes to Zn(II)-binding in wild type FrmR.

A possible role in metal-binding of the N-terminus region (amino terminal and residues W) was not explored, although probable since His3 coordinates cognate metals in *E. coli* RcnR, as discussed in Section 1.2.1.1 (Iwig *et al*. 2008; Higgins *et al.* 2012). Investigation by site-directed mutagenesis and X-ray crystallography analyses would be advisable. In fact, the allosteric rearrangement observed upon formaldehyde-binding (discussed in Section 7.5) appears to be triggered by an inter-subunit cross-link between Cys35 and Pro2 (the first residue of the N-terminus region). An intriguing hypothesis may relate to FrmR, which shares an evolutionary clade with *E. coli* Ni(II)/Co(II)-sensing RcnR (Figure 7. 6), modifying its function in the cell while evolving from a metal- to a formaldehyde-sensing protein exploiting a similar allosteric mechanism upon binding of the effector. This may have occurred as a consequence of the spontaneous mutation of His64 (residue Z, present in *E. coli* RcnR) to a Glu, since during the course of this work we proved that E64HFrmR mutant is capable of detecting Zn(II) and Co(II) *in vivo* (see discussion in Section 7.4).

Figure 7. 3 shows a preliminary model of the electron density of apo-E64HFrmR metal-binding site produced by Dr. Ehmke Pohl from crystals produced in this project. Putative Zn(II)-binding ligands (His3, His60, His64, and Cys35) are labelled. These residues are in the correct position to coordinate a metal ion, upon a small spatial rearrangement. It is possible to identify the imidazole of His3 (helix α1) pointing toward the residues from helices α2 (His60, His64) and α2'(Cys35), endorsing the hypothesis that a residue from the N-terminus region, presumably His3, may be involved in metal-binding.

Figure 7. 4B shows the coordination geometries and the ligands used by some of RcnR/CsoR sensors to bind their cognate metals. Complexes of Zn(II)- and Cu(I) have been included in the set since these ions are able to elicit an allosteric response in FrmR *in vitro.*

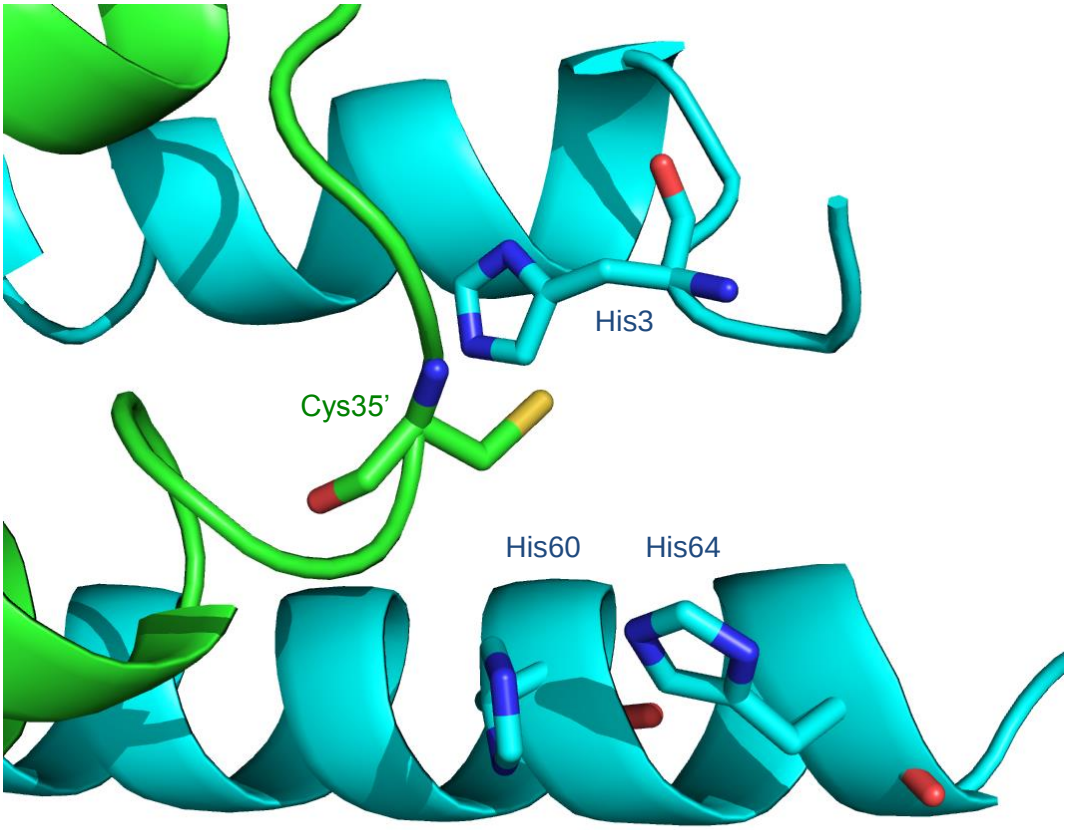


Figure 7. 3 Putative E64HFrmR metal binding site.

(Å)	His3	Cys35	His60	Asp63	His64
His3	/	3.1	6.9	3.3	7.6
Cys35	3.1	/	3.7	3.5	5.6
His60	6.9	3.7	/	6.0	2.8
Asp63	3.3	3.5	6.0	/	5.8
His64	7.6	5.6	2.8	5.8	/

Table 7. 2 The table lists the shortest approximate distances (in Å) between the side chain atoms of E64HFrmR putative metal ligands shown in Figure 7. 3. Asp63 has been added to the set as it is located at an appropriate distance from the other ligands to potentially allow participation to metal-binding, upon small rearrangement.

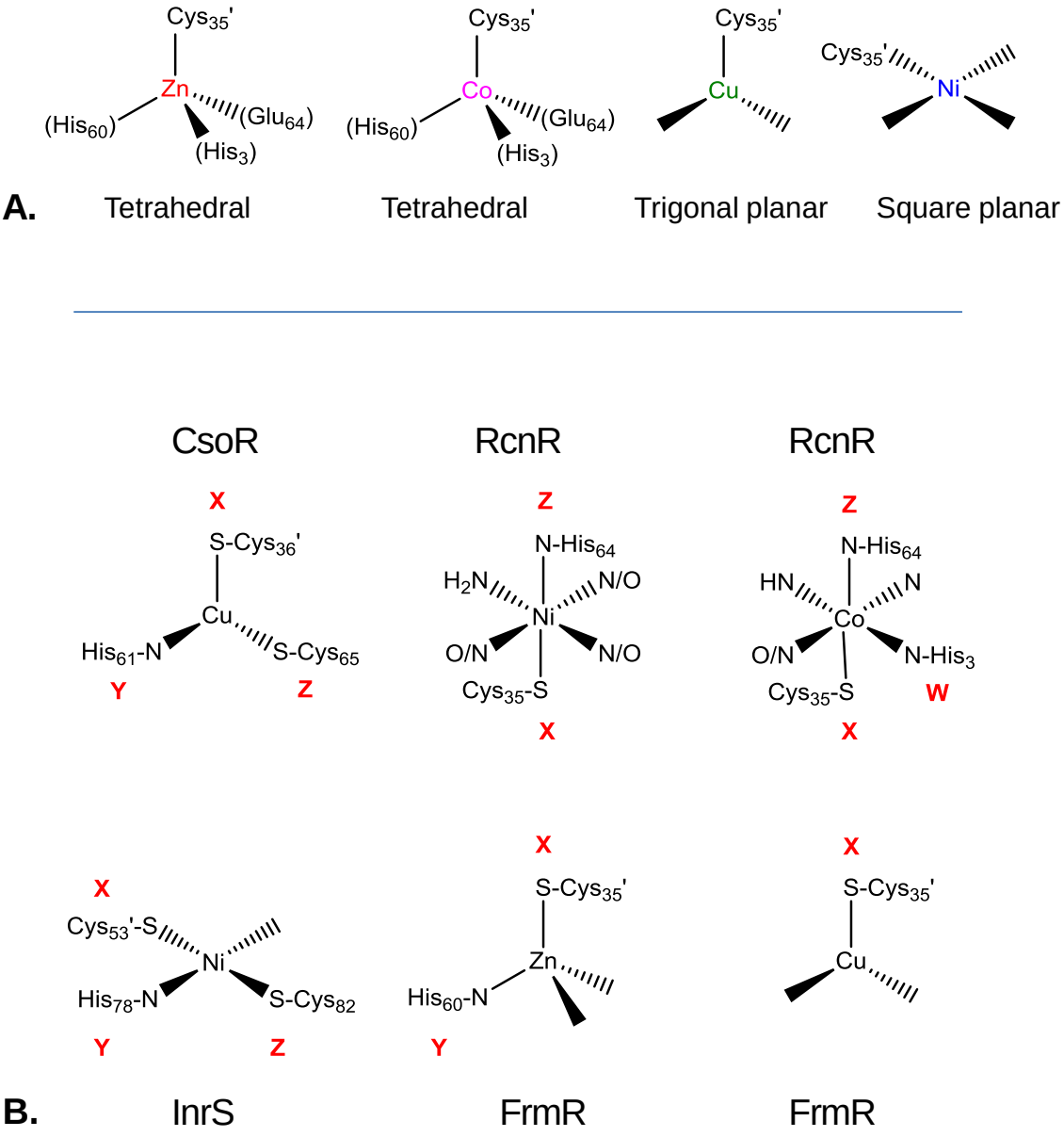


Figure 7. 4 A FrmR can adopt several coordination geometries depending on the nature of the metal ion. The His60 residue is in brackets as its role in metal coordination has not been univocally demonstrated. **B** Coordination geometries adopted by RcnR/CsoR family members while binding their cognate metals. FrmR complexes of Zn(II) and Cu(I) have been included in the scheme since these metals are able to elicit the allosteric mechanism *in vitro*. Letters in red indicate the position occupied by specific amino acid residues in the W-X-Y-Z motif.

7.3 *Salmonella* FrmR allosteric regulation of transcription

7.3.1 *Salmonella FrmR (and RcnR) bind(s) type-1 operator promoter regions*

The palindromic motif recognized by FrmR, ATAGTATAC$_6$TATAGTAT, is located in the region upstream *frmRA* (Figure 3. 1).

As previously discussed in Chapter 1, transcriptional factors belonging to the RcnR/CsoR family have been grouped by Iwig and Chivers according to the recognized promoter region (Iwig *et al.* 2009; Tan *et al*. 2014). *E. coli* RcnR protein binds a tandem of type 1 DNA regions, characterized by a single G/C tract (3-8 bp) flanked by inverted repeats rich in A/T (Figure 7. 5A) (Iwig *et al*. 2009). CsoR homologues recognize sequences with two shorter G/C tracts separated by 2-4 bp and A/T rich inverted repeats outside the G/C tracts (Figure 7. 5B) (Tan *et al*. 2014). The region recognized by FrmR comprises T/A tracts slightly longer than those observed in *E. coli* RcnR and CsoR homologues and, with the exception of one nucleotide (G), contains only thymines and adenosines. Fluorescence anisotropy analyses revealed that two FrmR tetramers interact with this sequence (Figure 7. 5C). The spacing between the centres of the inverted repeats (approximately coincident with the only G base) is 14 bp, while in the region recognized by *B. subtilis* CsoR is 13 bp (Ma *et al*. 2009a) and for *E. coli* RcnR is 9 bp (Iwig *et al*. 2009). Considering a rotation per bp ~ 34.3^o in the T/A tracts (B-DNA) and ~ 32.7^o in the G/C tract (A-DNA), we can assume that the T/A tracts are placed on opposite sides in FrmR and CsoR recognition sites and on the same side in RcnR recognition site (Figure 7. 5A-C). Modification of one of the T/A rich tracts (ATAGTATA → GTTCAACA) resulted in the loss of a tight DNA-binding affinity, suggesting that these regions play a determinant role in DNA recognition by FrmR.

Like *E. coli* RcnR, also the *Salmonella* homologue interacts with two TACT-N-G$_6$-AGTA type 1 sites, which are located in the region within *rcnA* and *rcnR* (Section 5.5). However, unlike the homologue in *E. coli*, this protein binds DNA very weakly when only one site is available (Figure 5.14B), suggesting positive cooperativity between the two RcnR tetramers, which was previously observed in *E. coli* RcnR foot printing experiments involving large DNA duplex (~ 300 bp) but not for short fragments (Iwig *et al*. 2009).

7.3.2 *Zn(II) and Cu(I) negatively regulate in vitro binding of frmRA promoter by Salmonella FrmR*

Zn(II) and Cu(I) were shown by fluorescence anisotropy to impair FrmR binding to DNA (Section 3.6.6). Although these metals adopt different geometries when coordinated to the protein, they share the Cys35 ligand, suggesting that this residue may be involved in the allosteric mechanism. This result is perhaps not surprising as FrmR and *E. coli* metal-sensor

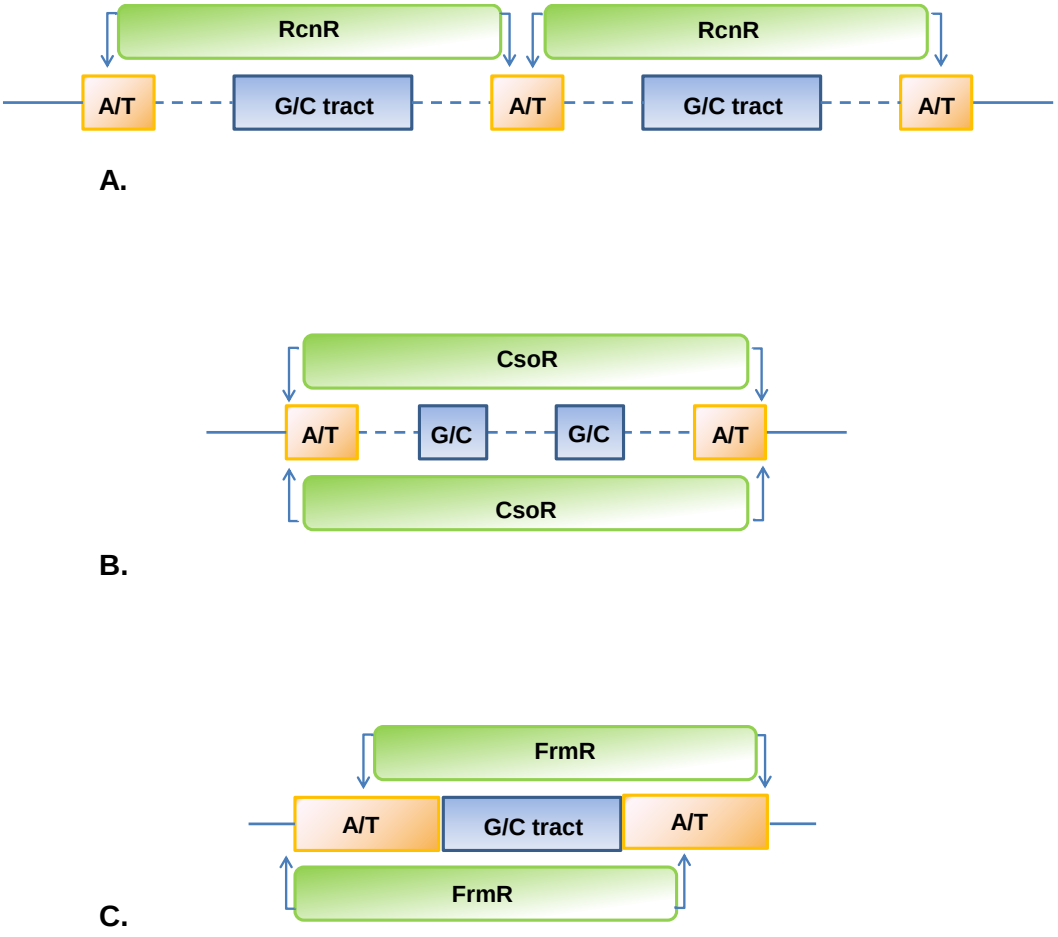
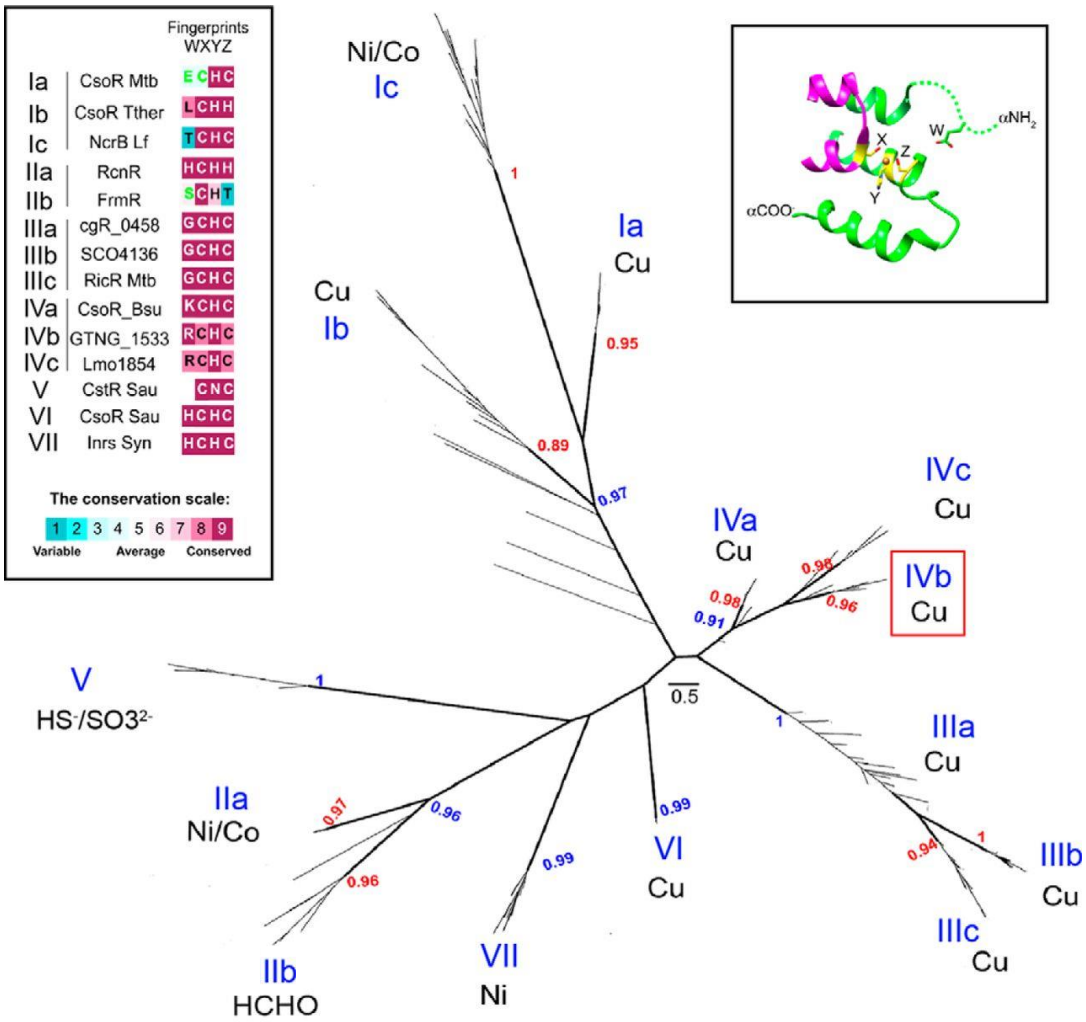


Figure 7. 5 Proposed scheme of RcnR homologues (**A**) *S. lividans* CsoR (**B**) and *S. typhimurium* FrmR (**C**) binding to their respective operators. G/C tract is depicted in blue and sites rich in T/A in yellow.

RcnR have 64.1 % of sequence similarity (Table 3. 1), and share a common evolutionary origin as shown by Chang and colleagues (Chang *et al.* 2014). A phylogenetic analysis on DUF_{156} (RcnR/CsoR) proteins, analyzed the characterized proteins belonging to this family plus the uncharacterized homologues, associating the sequences to clades (Figure 7. 6). ConSurf (Glaser *et al*. 2003) was used to estimate the evolutionary variation rates of specific amino acids. Seven distinct clades were identified within the RcnR/CsoR family, with CsoR proteins distributed over four. *E. coli* FrmR and RcnR proteins belong to the same clade (II) and share two of the four ligands of the WXYZ motif (Figure 7. 6- 3. 2). It was also hypothesized that proteins in the same clade may share residues involved in DNA-binding or in ion-pairing interactions (Chang *et al.* 2014).

The mechanism that drives the disassembly of the protein:DNA complex upon metal-binding was explored in RcnR/CsoR family members. For instance, Cu(I) ion is detected by *Geobacillus thermodenitrificans* (Gt) CsoR by coordinating Cys79 and His75 from one protomer and Cys50` from a different protomer (Chang *et al.* 2014). In the Cu(I)-bound form, a kink in the α2 helix (between Ala76 and His78) moves Cys79 closer to His75, allowing the coordination of Cu(I). Folding of one CsoR dimer over the other one in the tetramer upon Cu(I)-binding was described as resembling a "bow-tie" (Chang *et al.* 2014). Cu(I)-binding to GtCsoR remodels the protein conformation, perturbing a large number of residues, including those directly involved in DNA binding (such as Lys101 and Arg65), causing the release of the DNA. His61 plays a fundamental role in *M. tuberculosis* (Mtb) CsoR, where it is involved in both Cu(I)-coordination (through Nδ1) and in the propagation of the conformational change elicited by Cu(I) binding (through Nε2) (Liu *et al.* 2007; Ma *et al.* 2009b). Examination of MtbCsoR structure revealed a hydrogen-bond network amongst His61 (donor), Tyr35 and Glu81 (acceptors) (Liu *et al.* 2007; Eiamphungporn *et al.* 2009). When His60 is mutated to Ala, the DNA-binding affinity is drastically impaired, although CsoR maintains the original folding (Liu *et al.* 2007). In order to further explore the specific role of Nε2 in the stabilization of the Cu(I) bound-form of CsoR, His61 was substituted *in vitro* with Nε2-methyl-histidine (MeH) or (thiazolyl)-L-alanine (Thz) amino acids (Ma *et al.* 2009b).

Both these CsoR variants retain wild-type' Cu(I)- binding affinity and coordination geometry, but Cu(I) coordination is no longer coupled with DNA release ($\Delta Gc^{MeH\text{-}CsoR} = 0.0 \pm 0.4$ kcal mol^{-1} and $\Delta Gc^{Thz\text{-}CsoR} = 0.5 \pm 0.2$ kcal mol^{-1}). Moreover, site-directed mutagenesis was adopted to create Tyr35Phe and several Glu81 mutants, showing impairment of the propagation of the allosteric signal. Since double mutant Y35F/E81QCsoR presented a decrease in Δ*G*c that approximately corresponds to the sum of the decrease for the single mutants, it was suggested that Glu81 and Tyr35 do not interact with each other, but simply accept the hydrogen-bond from His61 (Ma *et al.* 2009b).



Chang *et al.* (2014)

jbc

Figure 7. 6 Phylogenetic tree and group-specific amino acid signatures of functionally characterised DUF_{156} (RcnR/CsoR) proteins. This research was originally published in Journal of Biological Chemistry. Chang F. M. J., Coyne H. J., Cubillas C., Vinuesa P., Fang X., Ma Z., Ma D., Helmann J. D., Garcia-de los Santos A., Wang Y. X., Dann C. E. & Giedroc D. P., Cu(I)-mediated allosteric switching in a copper-sensing operon repressor (CsoR). ***J. Biol. Chem***. 2014; 289:19204-19217. © the American Society for Biochemistry and Molecular Biology.

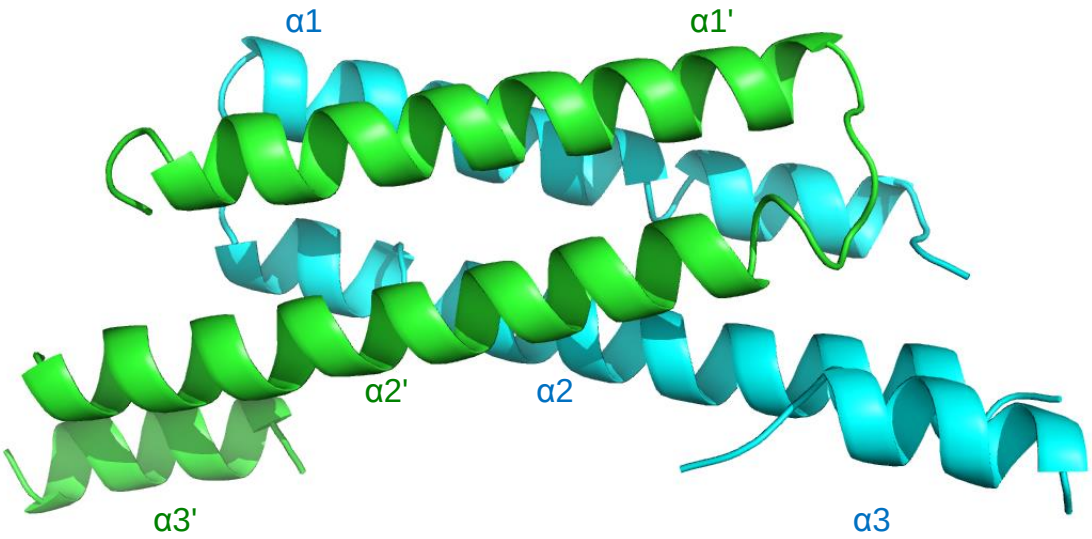
Similarly, a histidine residue plays an analogous role in CzrA, member of the ArsR family, where Nδ1 and Nε2 are involved in Zn(II) coordination and in the propagation of the allosteric signal, respectively (Arunkumar *et al.* 2009).

Synechocystis InrS possesses a similar hydrogen-bond network involving His78 and Glu98 (Foster *et al.* 2014a). However, this interaction was shown to be not absolutely required for allostery, presumably due to an additional connection between the α2 and α3 helices, the salt bridge between Arg85 and Glu95 (Foster *et al.* 2014a) (Figure 3. 23B).

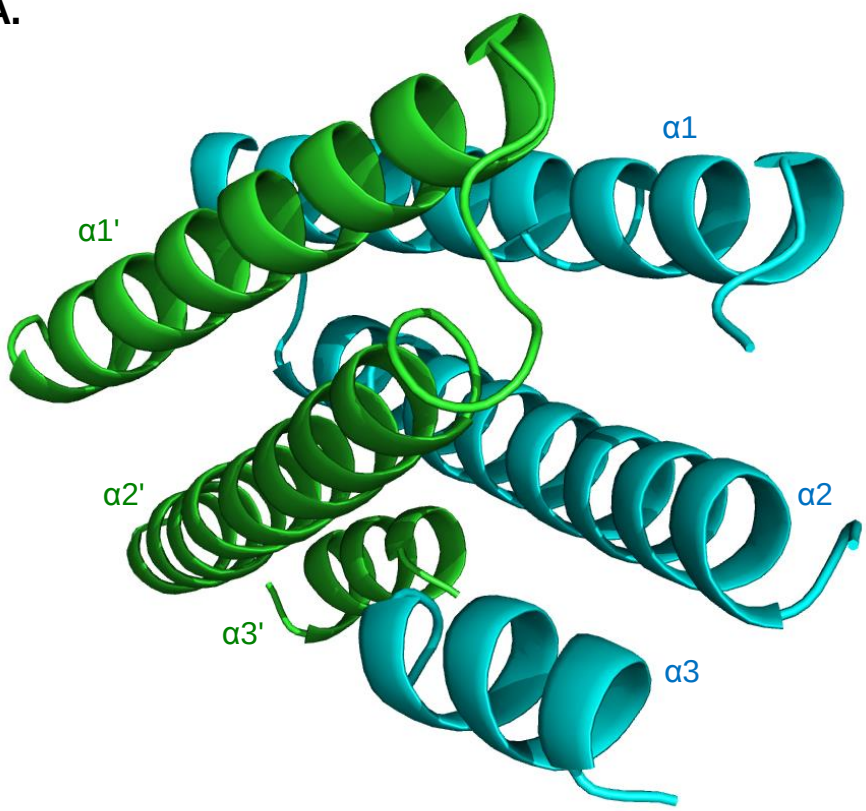
In light of these studies, we investigated the role of two conserved residues in *Salmonella* FrmR, His60 (Y) and Glu81 (B), in DNA-binding in order to study the propagation of the allosteric mechanism. Based on the proximity of the two helices a connection between His60 (helix α2) and Glu81 (helix α3) is plausible. FrmR also lacks the Tyr residue in position A of the secondary coordination shell, as InrS (Foster *et al.* 2014a). H60L and E81AFrmR mutants were therefore produced to test the effects of these mutations on the DNA-binding affinity by fluorescence anisotropy. As shown in Figure 3. 25A-B, in the absence of the only identified connection between helices α2 and α3 mediated by Y and B residues, K_{DNA} is significantly impaired (Table 8.2, Appendix). Therefore, it was not possible to gather evidence of the connection between metal-binding and DNA-binding as both mutations H60→L and E81→A already affected the K_{DNA}.

Since His60 is believed to be a putative ligand in FrmR (Section 7.2), upon metal binding it presumably undergoes a local rearrangement in order to orientate its side chain toward the metal ion. This rearrangement may cause the disruption of the hydrogen bond with Glu81, resulting in a reciprocal misalignment of helices α2 and α3. A set of basic and polar residues involved in DNA-binding in *E. coli* FrmR have been identified by bioinformatics analyses and tested *in vivo* (Law 2012). These residues (Arg14, Arg46, Gly47, Lys91) are conserved in *Salmonella* FrmR (Figure 3. 2) suggesting that they play a similar role. The location of these residues (Arg46 and Gly47 on helix α2, Lys91 on α3) may explain why the reciprocal orientation of helices α2 and α3 may be important in preserving a conformation with a tight DNA-binding affinity since a local rearrangement of His60 would cause an overall modification of helices α2 and α3 spatial position, as can be deduced by the preliminary X-ray 3D model in Figure 7. 7A-B.

Least-square superposition methodology was adopted to compare apo-E64HFrmR preliminary structure with that of *Streptomyces lividans* (Sli) apo-CsoR (PDB ID: 4ADZ) (Dwarakanath *et al*. 2012), revealing that the two structures differ for the position of helix α3 (Figures 7. 8A-B). Further investigation of His60 and Glu81 positions in FrmR shows that, unlike in CsoR, the distance (> 7 Å) between the two residues is not compatible with a chemical reaction. If even in the absence of an effector the two helices are not linked via this connection, other reasons must be found to understand the weakened K_{DNA} observed in H60LFrmR and E81AFrmR mutants. It is possible that the reciprocal disposition of the two helices is not vital in FrmR and that an



A.



B.

Figure 7. 7 A-B show different orientations of E64HFrmR dimer (chains A,B depicted in cyan/green, respectively). The second dimer (chains C,D) is omitted for clarity. The three helices per chain are labelled.

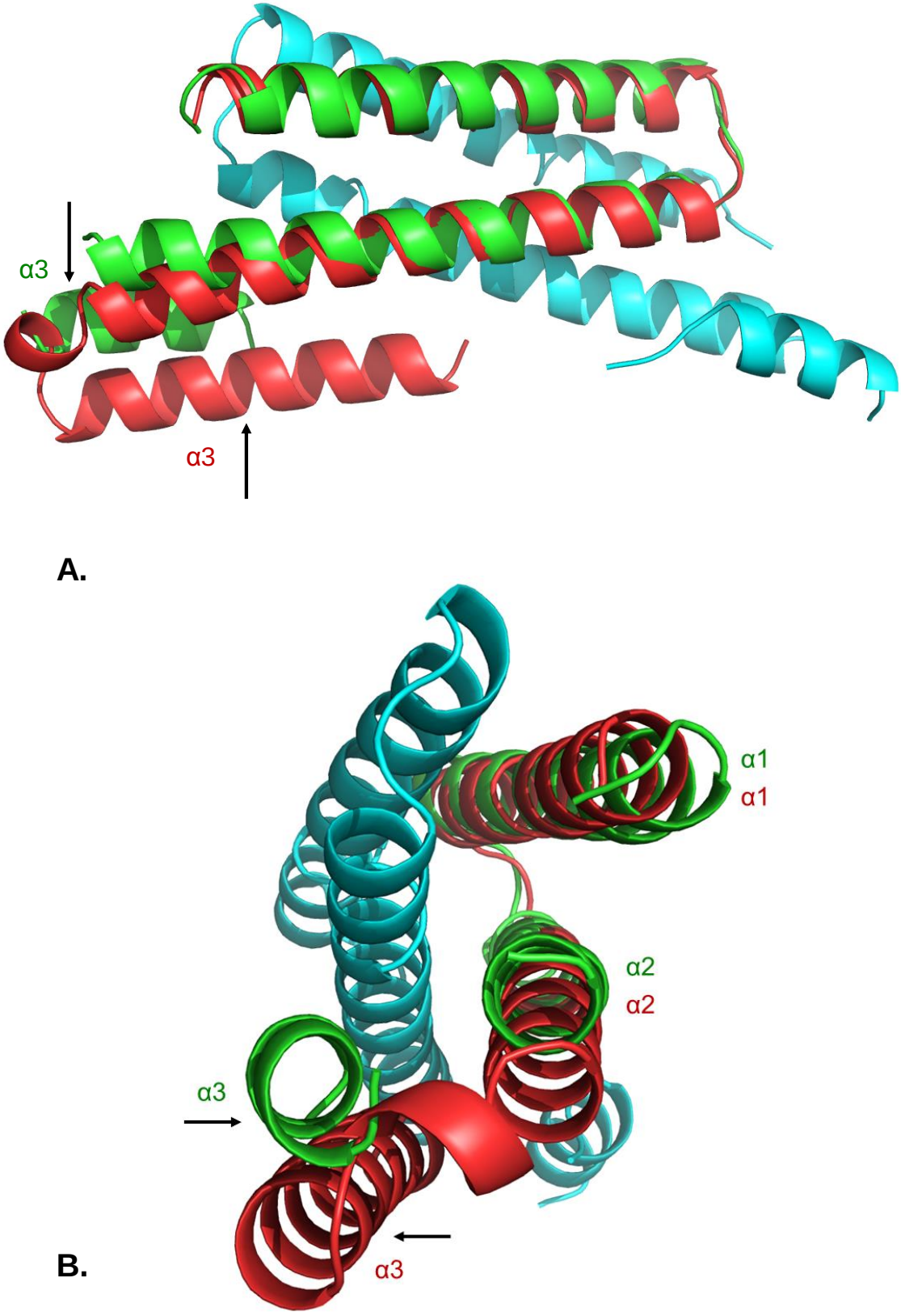


Figure 7. 8 Least-squares superposition of E64HFrmR (from this work) with apo-CsoR from *Streptomyces lividans* (PDB ID: 4ADZ) produced by Dr. Ehmke Pohl. **A-B** show different orientations. Only one chain of CsoR is shown (in red) to illustrate how the first two helices fit very well, whereas the third helix occupies a very different position. The A,B dimer of E64HFrmR is shown in cyan/green. The second dimer (C,D) is omitted for clarity but shows the same difference.

impaired DNA-binding may result from the substitution of two charged residues (Glu is negatively charged, whereas His is positively charged) with neutral residues such as Ala and Leu, causing an alteration of the charged patch that recognize DNA double helices.

Figures 7. 9A-B show the least-square superposition of apoE64HFrmR preliminary structure with the Cu(I)-bound form of CsoR from *Mycobacterium tuberculosis* (PDB ID: 2HH7) (Liu *et al.* 2007). Superposing helices α1 of the two proteins resulted in an overall body motion where each E64HFrnR molecule moves with respect to each other (~ 5 Å movement between chains A and C). Interestingly, the position of E64HFrmR' helix α3 resembles more that of helix α3 in the Cu(I)-bound form of MtbCsoR (Figure 7. 9B) rather than that of the apo-SliCsoR (Figure 7. 8B).

7.4 One single point mutation allows *Salmonella* FrmR to sense Zn(II) and Co(II) *in vivo*

During the course of this work we showed that substitution of one amino acid created a metal-sensor from the formaldehyde-responsive, DNA-binding, transcriptional de-repressor, FrmR (Figure 4. 2B). The substitution (Glu64 → His), was originally designed as part of sub-project intended to enhance the Ni(II)- and Co(II)-affinity of FrmR. In fact, FrmR shares with *E. coli* RcnR and CsoR homologues some of the ligands used by these proteins in Ni(II)/Co(II) and Cu(I) recognition, although with weaker affinity (Table 8.3, Appendix). FrmR' WYXZ fingerprint (H-C-H-E, Section 3.2) includes a glutamate in position Z, which is not present in either *E. coli* RcnR (Z = His) nor in the majority of CsoR homologues (Z = Cys). Hence, by substituting the Glu64 with a His, it was our intention to create an RcnR-like metal binding site. An analogous substitution, aimed at creating a CsoR-like site (Glu64 → Cys), was planned although not pursued due to the interesting outcomes achieved with E64HFrmR.

In contrast with our assumptions, E64HFrmR variant did not appear to possess an improved Ni(II)-binding affinity as suggested by the results presented in Section 4.4.3, where it is shown that the mutant is not capable of retaining Ni(II) ions when resolved on a size exclusion column (Figure 4. 8B). Measurement of the Ni(II)-binding affinity was therefore abandoned as presumably it is very weak.

Co(II)-binding affinity of E64HFrmR was shown to be enhanced by ~ 10-fold (K_1 = 2.56 ± 0.4 x 10^{-7}, Table 8.3, Appendix), however it is still significantly weak if compared to *E. coli* RcnR K_{app} < 5 nM (Iwig *et al.* 2008).

In an attempt to fully characterize E64HFrmR in order to understand if any other property may have been altered by the substitution, *in vivo* studies were carried out on the mutant (Section 4.2.2, Figure 4. 2B). Surprisingly, E64HFrmR is now able to detect cellular Zn(II) and Co(II).

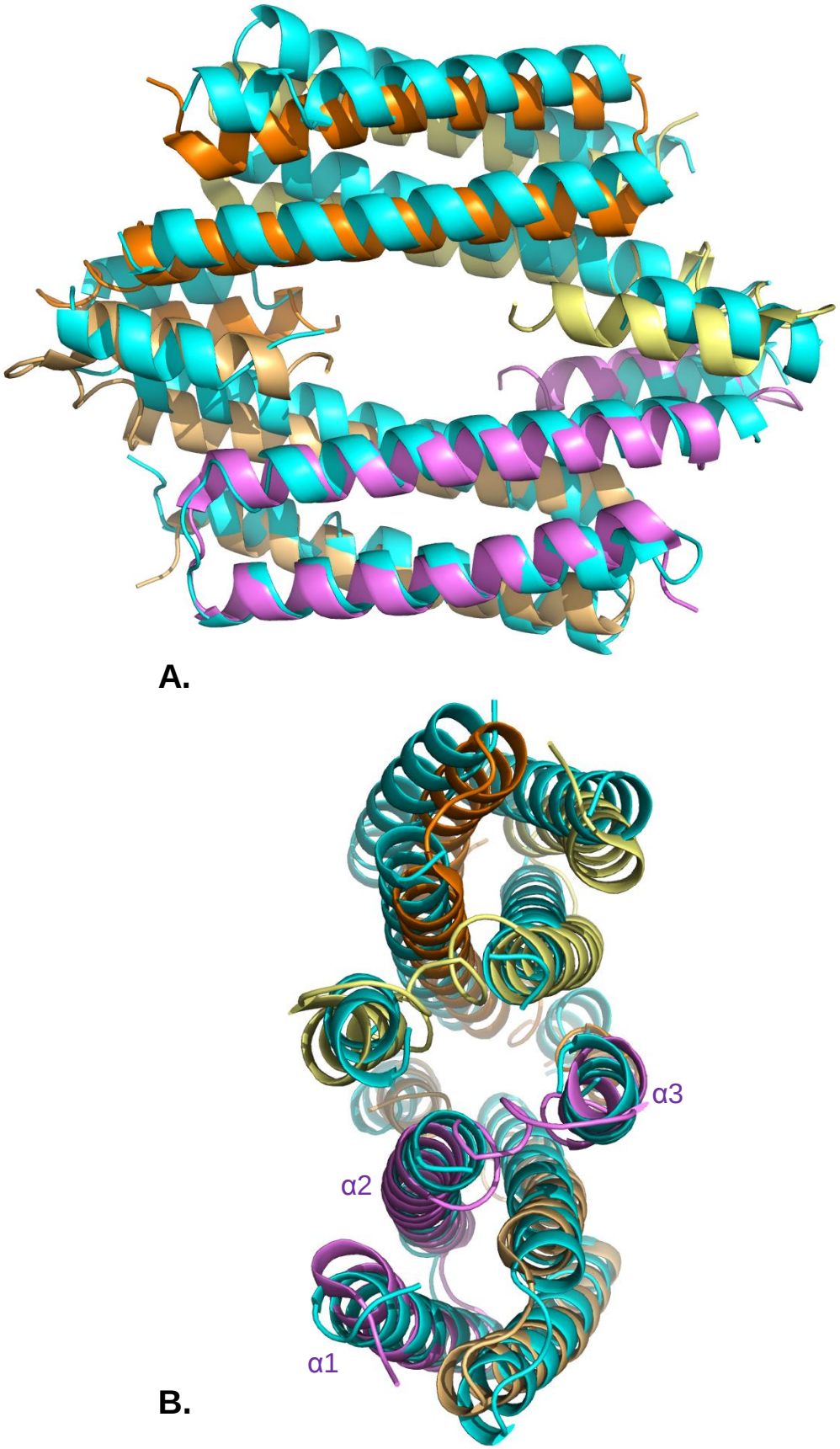


Figure 7. 9 Least-squares superposition of E64HFrmR (from this work) with Cu(I)-bound form of CsoR from *Mycobacterium tuberculosis* (PDB ID: 2HH7) produced by Dr. Ehmke Pohl. E64HFrmR tetramer is shown in cyan, whereas CsoR, which has been deposited in the PDB as a monomer, was converted to a tetramer (chain A in pink, chain B in pale orange, chain C in orange, chain D in yellow) using the crystallographic symmetry with the program COOT (Emsley & Cowtan 2004; Emsley *et al.* 2010**). A-B** show different orientations.

In 2012 Higgins and collaborators created a Zn(II) site in *E. coli* RcnR through a His3 → Glu mutation (Higgins *et al.* 2012). Zn(II) was able to elicit modifications in the conformation of H3ERcnR by binding to the same positions as Ni(II) and Co(II) would do in the wild-type. However, the result described in this work does not have a precedent as it is the first time that a non metal-sensing transcriptional regulator is converted in a metal-sensor.

The following sections will further analyse Zn(II)- and Co(II)-recognition by E64HFrmR. The inability of FrmR to detect these metals *in vivo* will also be exploited in order to recognize the determinants behind E64HFrmR' Zn(II) and Co(II) gain of function.

As previously mentioned in Section 7.2, the residue W (His3) points toward the other putative metal ligands (Cys35', His60 and His64) even if the model describes the apo-form of the protein (Figure 7. 3). Since E64HFrmR possesses a weakened K_{DNA} and a tightened $K_{Zn(II)}$ when compared to wild-type, it is possible to suppose that His3 is somehow constrained, maybe by interacting with His64, in a location more favorable to bind the metal, even in the absence of the latter. In fact, a histidine has a bulkier side chain and could be in a more suitable position to interact with the other residues in the metal-binding pocket. This would cause a local rearrangement of helix α1, similar to that induced by the effector-binding. Hence, E64HFrmR may have a tighter affinity for metals, as less energy is needed in order to rearrange the ligands to form a pocket, and a weakened affinity for DNA, as the protein conformation is already partially reorganized toward a metal-bound form. These are merely speculations that need to be tested by X-ray crystallography on apo-FrmR, and Zn(II)-bound forms of both FrmR and E64HFrmR.

7.4.1 *FrmR and E64HFrmR cannot compete for metal-binding with Salmonella cognate metal sensors*

As introduced in Section 1.3.1, relative effector-binding affinity is an important parameter which may determine which effector is detected by which sensor. *Salmonella* FrmR is able to coordinate various metals (Zn(II), Co(II), Ni(II) and Cu(I)) (Section 3.3) and some of them (Zn(II) and Cu(I)) have been shown to elicit the conformational change coupled with DNA release (Section 3.6.6).

Salmonella possesses two zinc sensors ZntR (MerR-family) and Zur (Fur-family), a copper sensor, CueR (MerR-family), and the Ni(II)/Co(II) sensor RcnR (RcnR/CsoR family). Since it is possible that FrmR is never able to access Zn(II), Cu(I) or Co(II) ions because the corresponding cognate metal sensor maintains the cellular level below FrmR metal-binding affinity, whereas E64HFrmR may have an enhanced Zn(II)-binding affinity that allows this protein to compete with zinc sensors. In order to test these hypotheses, cognate metal affinities of *Salmonella* ZntR, RcnR and CueR were measured by competition assays performed by Dr. Osman and published in recent work (Osman *et al*. 2015). Findings will be briefly summarised

in this section because comparison with values calculated for FrmR and E64H mutant with the same assays may elucidate why E64HFrmR, but not wild-type FrmR, responds to metals *in vivo*.

Salmonella ZntR binds two Zn(II) ions per dimer (K_{Zn1} = 3.20 ± 0.73 x 10^{-12} M, K_{Zn2} = 2.68 ± 0.73 x 10^{-11} M, Osman *et al*. 2015). *E. coli* ZntR has been structurally characterised with a binuclear Zn(II)-site across the dimer interface giving a total Zn(II) binding stoichiometry of 2:1 in contrast to that observed for *Salmonella* ZntR (1:1, Zn(II):ZntR monomer) (Changela *et al.* 2003; Osman *et al.* 2015).

In common with other Fur-family members, *Salmonella* Zur exists as dimers. Moreover it has at least three exchangeable sites per dimer which are accessible to both Co(II) and Zn(II) (Osman *et al*. 2015). *E. coli* Zur has been structurally characterised with four Zn(II) ions per dimer, assigned to two structural sites and two sensory sites (whose residue are conserved in *Salmonella* Zur) (McGuire *et al*. 2013). Six Zn(II)-binding sites have been observed in other Zur-family members and *E. coli* Zur purifies with three molar equivalents of Zn(II) per dimer, one of which is lost during crystallisation (McGuire *et al*. 2013; Foster *et al*. 2014; Ma *et al*. 2009). Affinities for Zn(II) binding to *Salmonella* Zur (K_{Zn1-2} = 6.36 ± 0.41 x 10^{-13} M, K_{Zn3} = 8.04 ± 2.92 x 10^{-11} M, K_{Zn4} > 5 x 10^{-7} M) were determined by Dr. Osman considering Zn(II) binding to three sites (K_{Zn1-2} and K_{Zn3}) on a Zur dimer (with the structural site already filled) (Osman *et al*. 2015). Since *Salmonella* ZntR has a weaker zinc affinity than Zur, competition for zinc binding was conducted only with ZntR.

Salmonella RcnR, like the *E. coli* homologue, exists as a tetramer containing one Co(II) site per monomer (Iwig *et al*. 2008; Osman *et al*. 2015). Two of the four sensory sites per RcnR tetramer have a tighter affinity, which has been measured by Dr. Osman to be K_{Co1-2} = 5.06 ± 0.86 x 10^{-10} M, where only lower and upper range for the third site could be determined (3 x 10^{-5}< K_{Co3} < 10^{-7} M) (Osman *et al*. 2015).

Salmonella CueR was previously shown to outcompete a ten-fold molar excess of BCS (Osman *et al*. 2013) therefore a 100-fold and then a 75-fold excess of BCS was used to estimate K_{Cu1} = 3.25 ± 0.66 x 10^{-19} M (Osman *et al*. 2015).

These findings show that the tightest exchangeable sites of the endogenous metal-sensors are tighter for their cognate metals than either FrmR or E64HFrmR in every case (Table 8.3, Appendix) (Osman *et al*. 2015). However, the difference in $K_{Zn(II)}$ between E64HFrmR and cognate Zn(II) sensors is the smallest. To directly investigate whether metal-binding affinities of FrmR and E64HFrmR are too weak to allow the proteins to access metal ions in cells we carried out pair-wise competitions between FrmR and *Salmonella* cognate metal-sensors ZntR, RcnR and CueR (Section 4.6.1), and E64HFrmR and *Salmonella* ZntR (Section 4.6.2). Results show that in each competition both metal-loaded FrmR proteins failed to retain the metal, confirming

that their metal-binding affinities are not tight enough to allow competition with cognate metal sensors (Figure 7. 10, Table 7. 3). However, it must be noted that the competition experiment between E64HFrmR and ZntR was performed only once and experimental replicates are needed to confirm, or otherwise, the result.

This outcome explains why FrmR does not detect any cellular metal, as shown in Figure 4. 1A (Section 4.2.1) and supports the belief that the behavior of a group of metal sensors is determined by a combination of parameters, which allow the best (thermodynamically or kinetically favored) sensor to detect the cognate metal. *In vivo* Zn(II)-sensing by E64HFrmR (Figure 4. 2B, Section 4.2.2) cannot solely be explained by Zn(II) affinity. In fact, tightening of $K_{Zn(II)}^{E64HFrmR}$ by ~ one order of magnitude compared to $K_{Zn(II)}^{FrmR}$ appears insufficient to allow competition with *Salmonella* ZntR, as shown by preliminary experiments, hence other parameters must be involved.

7.4.2 *Combination of relative affinity and relative-allostery explains Zn(II) gain of function*

In chapter 4 the biochemical properties of FrmR and the E64H mutant were compared in order to identify how such a small modification can have a great impact on the protein behavior. As summarized in Tables 8.2 and 8.3, E64HFrmR possesses a tighter Zn(II) affinity (variation of ~ 10 fold). Although this difference would not be enough to explain how FrmR mutant can now detect cellular zinc, in combination with a ~ 10 fold weakened apo-K_{DNA}, it enhances the zinc-sensitivity by a sufficient degree to match the endogenous zinc-sensors in *Salmonella*, ZntR and ZuR. This result is counterintuitive as a smaller apo-K_{DNA} results in Zn(II) being less allosterically effective upon binding to E64HFrmR, since the associated ΔG_C is smaller (Table 8.2). As previously discussed in Section 1.3.2, an opposite result (a greater ΔG_C) might be expected for a de-repressor sensor with an enhanced sensitivity. However, a weaker apo-K_{DNA} implies that a greater number of apo-protein molecules will be in the unbound-form, and a smaller concentration of Zn(II) will be needed to dissociate the fewer protein molecules bound to the operator region (Figure 7. 11).

The previous findings help to explain the higher base line observed in *β*-galactosidase assays carried out in Δ*frmR* cells containing P_{frmRA}-*frmRE64H* when compared to the same assay performed in Δ*frmR* cells containing P_{frmRA}-*frmR* (Figures 4. 1A-4. 2B). Since the DNA occupancy of E64HFrmR is lower than FrmR, at a fixed concentration of zinc ions, the repression of *frmR* by E64HFrmR is also smaller.

In a recent paper by our group, (Osman *et al.* 2015), we also suggest that since the allosteric coupling scheme reported by Grossoehme and Giedroc (2012) (Figure 3. 21) refers to a closed system, each constant would have a reciprocal effect on the opposite one. In fact:

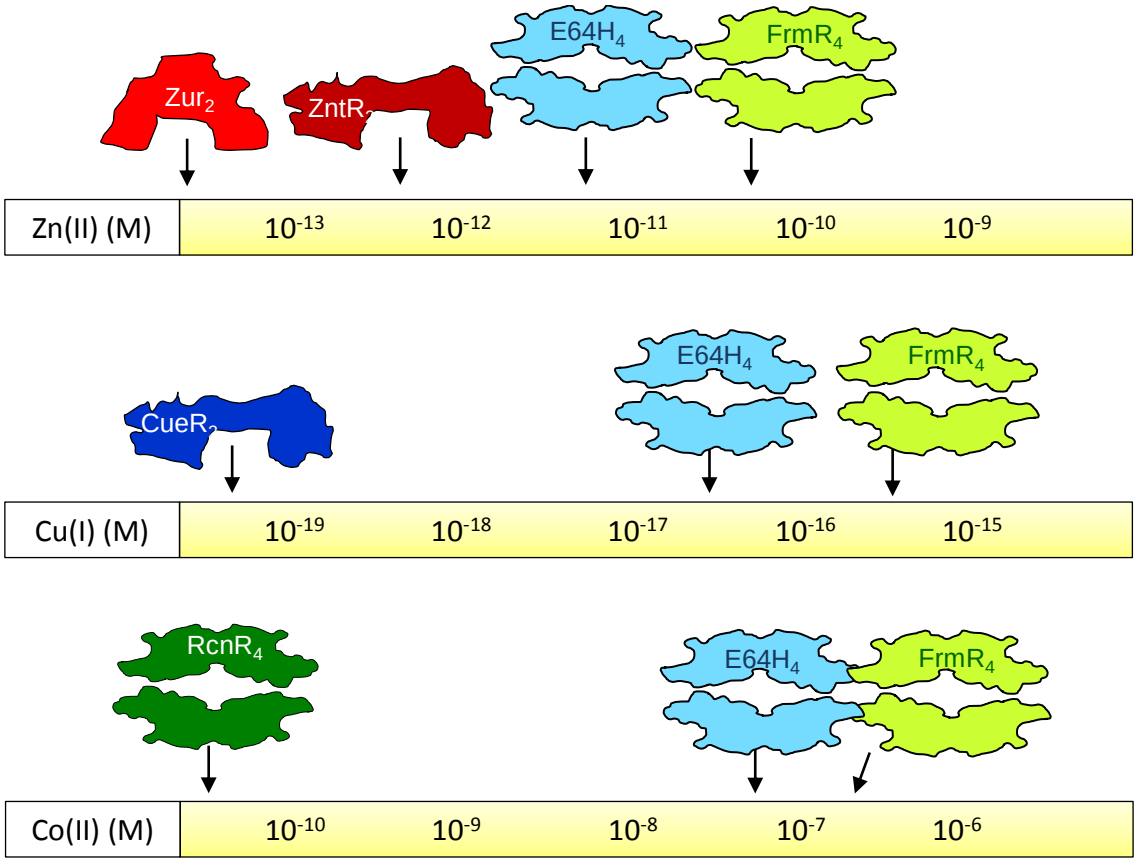


Figure 7. 10 Scheme showing metal affinities of *Salmonella* Zn(II), Co(II) and Cu(I) sensors plus FrmR and E64HFrmR. Values regarding ZuR, ZntR, RcnR, and CueR have been obtained from Osman *et al*. 2015 and are summarised in Table 7.3.

Sensor	Metal	K_{metal} (M)
ZuR	Zn(II)	$K_{1\text{-}2} = 6.36 \pm \text{x } 10^{-13}$; $K_3 = 8.04 \pm 2.92 \text{ x } 10^{-11}$; $K_4 \geq 5 \text{ x } 10^{-7}$
ZntR	Zn(II)	$K_1 = 3.2 \pm 0.73 \text{ x } 10^{-12}$; $K_2 = 2.68 \pm 0.73 \text{ x } 10^{-11}$
RcnR	Co(II)	$K_{1\text{-}2} = 5.06 \pm 0.86 \text{ x } 10^{-10}$; $3 \text{ x } 10^{-5} \geq K_3 \geq 10^{-7}$
CueR	Cu(I)	$K_1 = 3.25 \pm 0.66\ 8 \text{ x } 10^{-19}$

Table 7. 3 Metal affinities of ZuR, ZntR, RcnR and CueR determined by Osman and colleagues (adapted from Osman *et al.* 2015).

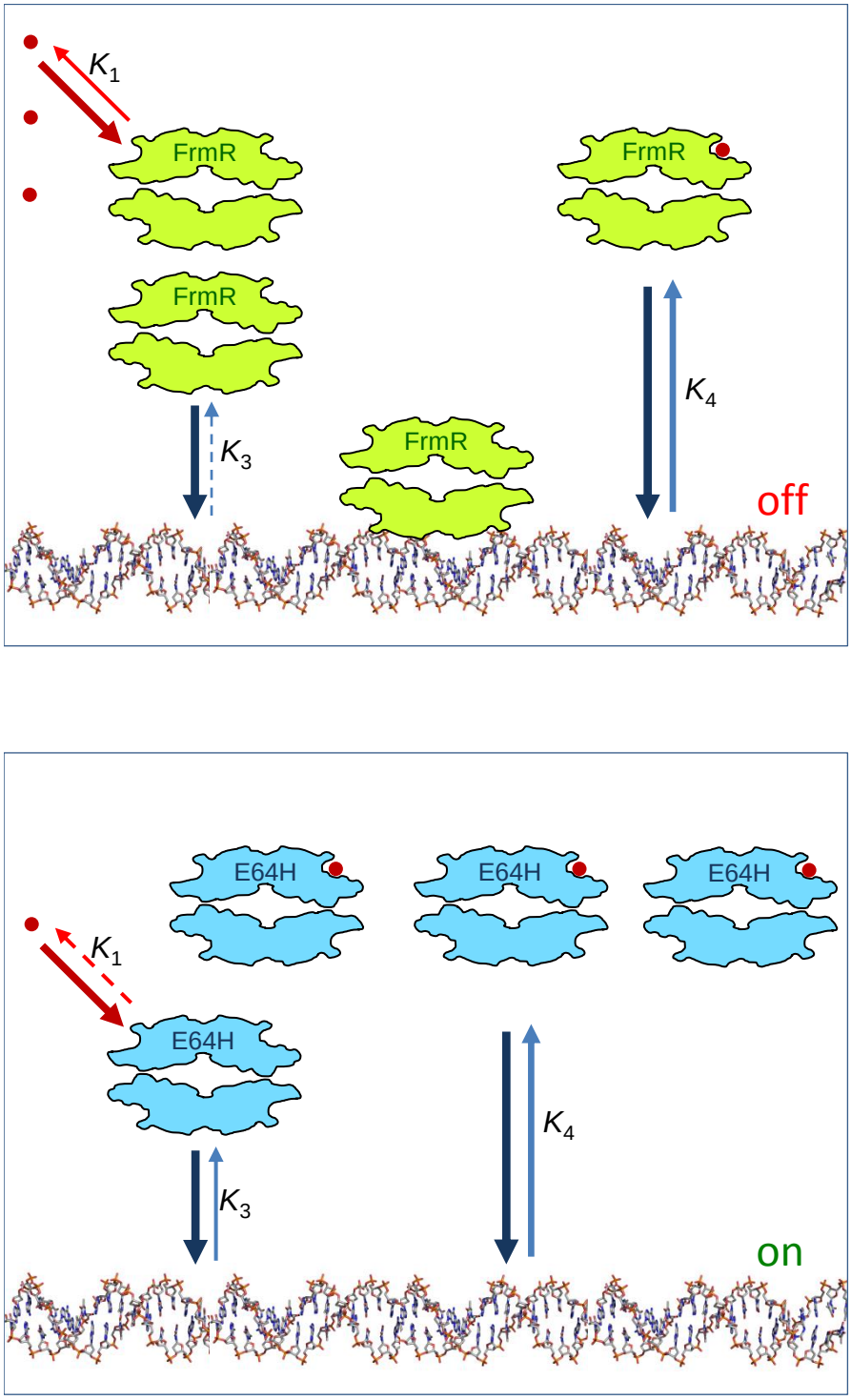


Figure 7. 11 Scheme representing the effects that the weaker K_{DNA} and tighter $K_{Zn(II)}$ of E64HFrmR (bottom box), compared to wild-type FrmR (upper box), have in the cell. Zinc ions are represented by red circles. Arrows width represents the different off-rates for each described process (the greater the width, the faster the rate). On-rates are likely diffusion-limited and, hence, faster than off-rates. Since E64HFrmR has a tighter $K_{Zn(II)}^{apo\text{-}protein}$ (K_1), and a weaker $K_{DNA}^{apo\text{-}protein}$ (K_3), compared to wild-type FrmR, at a given concentration of Zn(II) DNA-occupancy of E64HFrmR is lower. Expression of *frmRA* regulated by E64HFrmR is therefore permitted at lower [Zn(II)].

$$Kc = \frac{K4}{K3} = \frac{K2}{K1}$$

Equation 8

and:

$$K2 = \frac{K1\,K4}{K3}$$

The tightening of K_1 ($K_{Zn(II)}^{protein}$) and the weakening of K_3 ($K_{DNA}^{protein}$) in E64HFrmR reflects on K_2 ($K_{Zn(II)}^{DNA\cdot protein}$), which was calculated to be 1.9×10^{-10} M knowing that K_4 is $K_{DNA}^{Zn(II)}\cdot$protein. For a rigorous analysis of FrmR and E64HFrmR ability as de-repressors, we should compare $K_{Zn(II)}^{DNA\cdot protein}$ (K_2) values, since the gene transcription is activated upon binding of the effector (here Zn(II)) to the DNA-bound protein). Thus, K_2 was calculated also for wild-type FrmR (5.3×10^{-9} M), showing a greater differentiation between the two protein species (Osman *et al.* 2015). Therefore, in E64HFrmR the contribution toward Zn(II)-sensitivity is achieved by a tighter $K_{Zn(II)}$ and, surprisingly, a weaker K_{DNA}. It must be noticed that this outcome is valid under the assumption that the described system is closed. It may be not absolutely true if metal ions are in exchange with the buffer, or if dissociation of the DNA-bound form is not the regulatory step.

7.4.3 *Glutathione enhanced metal-detection by E64HFrmR may explain Co(II)-sensing*

While a tighter K_{metal} and a weaker K_{DNA} may explain the gain of zinc-sensing by E64HFrmR, as discussed in Section 7.4.2, this is not sufficient to understand why the FrmR variant is now able to detect cellular Co(II) as $K_{Co(II)}^{E64HFrmR}$ is very weak compared to that of *Salmonella* Ni(II)/Co(II) sensor RcnR. The possibility that an additional factor may be involved in cobalt delivery to the protein was therefore investigated. In Section 1.3.3 we discussed the factors behind cobalt-sensing by *Synechocystis* PCC 6803 CoaR, which possesses a weaker Co(II) affinity than other metal-sensors from the same organism (ZiaR and Zur). Co(II) is also able to trigger the allosteric mechanism that leads to the activation of transcription in ZiaR and Zur, however CoaR has evolved to selectively respond to cobalt. Co(II) appeared to be channeled directly to its precorrin isomerase-like domain (Patterson *et al.* 2013).

A hypothesis to explain why a variant of the formaldehyde-sensor FrmR should have a preferential access to Co(II), may refer to the common origin with Co(II)-sensing RcnR and to a consequent interaction with a Co(II) donor. Although FrmR has, somehow, cobalt ions channeled to its binding sites, the affinity for the metal is still too weak to allow detection. However, E64HFrmR has a $K_{Co(II)}$ enhanced enough to allow the response to the metal (Table 8. 3, Appendix). A plausible candidate as a cobalt-donor may be glutathione and/or its adducts as

FrmA (the formaldehyde dehydrogenase regulated by FrmR) acts on the substrates S-(hydroxymethyl)glutathione and S-nitrosoglutathione (Osman *et al.* 2015; Gutheil *et al.* 1992; Liu *et al.* 2001).

Glutathione is the most abundant low molecular weight molecule carrying a thiol functional group in the cell and is widespread amongst bacteria, animals and plants. Its intracellular concentration is estimated to be in the mM range (0.5 mM < [GSH] < 10 mM) (Maher 2005; Halliwell & Gutteridge 1989; Lushchak 2012). It is a tripeptide (L-glutamyl-L-cysteinyl-glycine) characterized by an unusual bond between the amino group of the cysteine residue and the Cγ of the glutamate residue instead of the Cα, as typically observed in proteins. This feature protects it from the intracellular peptidases action which usually cleaves the Cα-NH bond.

The primary biological function of glutathione is as an antioxidant, reacting with ROS/RNS (reactive oxygen and nitrogen species) or during peroxidase-catalysed reactions (Lushchak 2012; Halliwell & Gutteridge 1989). Moreover, it acts as a cysteine reserve and in the storage and transport of nitric oxide (NO), and many others (Lushchak 2012; Halliwell & Gutteridge 1989). For example, it is involved in the detoxification of endogenously produced toxic substances, such as formaldehyde (as discussed in Chapter 1) and methylglyoxal (Inagi *et al.* 2010; Kalapos 1999; Yadav *et al.* 2008), by reducing the reactivity of these molecules and, in some cases, aiding the transport across the cellular membrane (Wang & Ballatori 1998).

Since glutathione can coordinate various metals, including copper, zinc, cadmium, mercury, lead and gold (Ballatori 1994), it also plays a major role in metal homeostasis by aiding transfer of cations from one ligand to another, transporting cations through the membranes, acting as co-factor in redox processes and as cysteine source (Wang & Ballatori 1998; Lushchak 2012).

Within this tripeptide it is possible to identify six potential sites for metal coordination, with cysteine thiol being the most reactive group in the set (Figure 7. 12). Glutathione' metal coordination is a spontaneous process and some of the most stable metal complexes are achieved with divalent cations adopting a 2:1 GSH:metal stoichiometry (Lushchak 2012).

Biological roles for glutathione in iron and copper metabolism have been studied in detail. Glutathione is involved in Cu(II) reduction to Cu(I), in its mobilization from storage sites and in the insertion of Cu(I) during the biosynthesis of copper-containing proteins, such as Cu,Zn-superoxide dismutase (Cu,Zn-SOD), (Lushchak 2012; Freedman *et al*. 1989, Ciriolo *et al*. 1990; Brouwer & Brouwer-Hoexum 1992; Suzuki *et al*. 1989). Moreover, glutathione is believed to play a similar role in the metabolism of iron. Its contribution to the maturation of extra mitochondrial Fe,S-proteins in *S. cerevisiae* has been documented by Sipos and collaborators (Sipos *et al*. 2002).

A putative role for glutathione in Co(II)- and Zn(II)-sensing by FrmR and E64HFrmR was investigated by *β-galactosidase* assays by Dr. Deenah Osman and published in a recent work

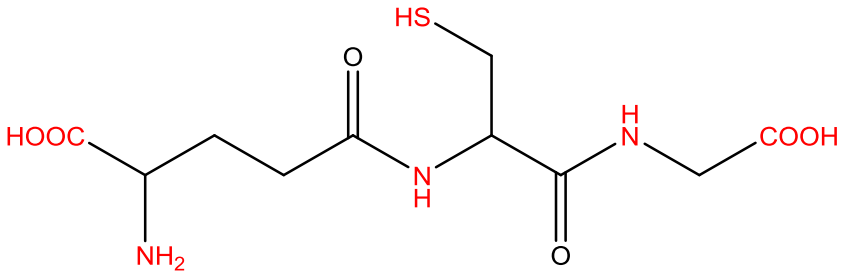


Figure 7. 12 Glutathione possesses at least six possible sites (highlighted in red) to coordinate metals. Shankar and colleagues showed that zwitterionic glutathione forms tri- and tetradentate complexes, highly stabilizing the metal cation (Shankar *et al*. 2012).

(Osman *et al*. 2015). Cells lacking *gshA*, the glutathione synthetic enzyme, after 2 h of exposure to Zn(II) and Co(II), showed a significantly decreased E64HFrmR- mediated expression compared to that obtained with constructs including *gshA* (Osman *et al*. 2015).

The effect that the presence of glutathione has on ZntR and RcnR-mediated gene-expression, upon incubation with Zn(II) and Co(II) respectively, was also tested. Regulation controlled by RcnR showed a dependency for glutathione similar to that observed with E64HFrmR (Osman *et al*. 2015), suggesting that glutathione assists cobalt-sensing by both proteins. However, ZntR-mediated expression in Δ*gshA* cells upon Zn(II) incubation did not differ from wild-type (Osman *et al*. 2015). Therefore glutathione helps the detection of Co(II) by both E64HFrmR and RcnR, while the effect on Zn(II)-sensing seems more controversial (glutathione competes with ZntR whereas assists E64HFrmR).

In light of these findings, here we further explored metal detection by E64HFrmR *in vitro*. Since FrmR and E64HFrmR possess a weak cobalt affinity (Table 8.3, Appendix), it was possible to assess only the contribution of glutathione to Zn(II) binding under the experimental conditions described in Section 4.7.1. The experimental asset involves evaluating the rate of zinc extraction by the protein from mag-fura-2 in the presence or absence of glutathione, by inspection of curve slope (Figure 4. 19A-B). However, it was not possible to detect any significant difference suggesting that glutathione is involved in zinc delivery to FrmR or E64HFrmR in this experimental setup. It appears that these proteins receive Zn(II) directly from the chelator. The reason for the different behaviour observed *in vitro* and *in vivo* in zinc detection remains unclear. It may be ascribed to a suboptimal glutathione/zinc ratio adopted for the *in vitro* competition. Glutathione could readily react with a small amount of zinc, competing with mag-fura-2 and FrmR/E64HFrmR, and form a stable zinc complex failing to deliver the metal to the protein. This hypothesis may be supported by the different intensity of the spectral feature at 366 nm at the end of the competition experiment whether this was conducted in the presence or absence of glutathione. Support to this thesis is also provided by a recent study by Ma and colleagues that demonstrates that low molecular weight thiols and bacillithiols can compete for zinc binding with zinc-sensors (Ma *et al*. 2014) (see Section 7.8).

7.4.4 *Detection of two distinct E64HFrmR conformations*

During the course of this work two distinct conformations of E64HFrmR were identified. An additional absorbance spectral feature at around ~ 300 nm and a different DNA-binding affinity of the apo-state (Table 8.2, Appendix) allow differentiation between the two protein forms.

In order to investigate the nature of this spectral feature, UV-vis absorbance spectra of both E64HFrmR variants were recorded under aerobic conditions or upon treatment with formaldehyde and glutathione (Figure 8. 4BC, Appendix). However these treatments did not induce any change in the absorbance spectra suggesting that the feature at ~ 300 nm is not due

to oxidation of the cysteine residue caused by inter/intra- subunit cross-linking or formation of adduct with formaldehyde or glutathione.

The two preparations showing a UV-vis spectrum indistinguishable from wild-type, were produced amongst the very first batches of E64HFrmR. The first experiments carried out to assess metal- and DNA-binding properties were then carried out with both types of preparations since, initially, we did not discern between the two. The experiments performed with E64HFrmR lacking the spectral feature at 300 nm (E64HFrmR*****), include fluorescence Zn(II)-titration (Figure 8. 5, Appendix), and protein:DNA interaction studied in the absence and presence of Zn(II) by fluorescence anisotropy (Figure 8. 9AB, Appendix). Interestingly, $K_{DNA}^{E64HFrmR*}$ and $K_{DNA}^{Zn(II)\text{-}E64HFrmR*}$ were very close to the corresponding values obtained with FrmR. Moreover, Zn(II) titration of E64HFrmR***** quenches Tyr89 intrinsic fluorescence (Figure 8. 5, Appendix) whereas the same experiment performed with E64HFrmR shows an increase of the fluorescence intensity at 304 nm (Figure 4. 6A).

These outcomes may suggest that the mutant displaying the extra spectral feature may have a different spatial disposition compared to E64HFrmR***** and, possibly, to FrmR.

Although the nature of this additional spectral feature remains unclear, understanding the thermodynamics of protein folding may help to elucidate the origin of the two stable conformations of E64HFrmR. To acquire its native three-dimensional structure a protein undergoes a folding process which involves the formation of protein domains (α-helices, β-sheets, loops, etc.) from a one-dimensional linear chain of amino acids (Branden & Tooze 2001). In a simplistic model every amino acid residue in a denatured protein can assume three different positions (α, β, L in the Ramachandran diagram) relative to each other according to its steric hindrance (limited freedom of rotation of the two dihedral angles, ψ and φ). A polypeptide chain of 150 residues may therefore undertake $3^{150} = 10^{68}$ different configurations and need ~ 10^{48} years (~ 10^{-13} sec for each conformational transition) to explore all of them (Branden & Tooze 2001). This could not be compatible with cellular life and since the average folding process, both *in vitro* and *in vivo*, lasts 0.1 – 1000 seconds, the transition from a denatured form to the native state must be somehow facilitated.

The unfolded protein assumes a filamentous state which is adopts a globular state by interaction with water molecules from the solvent. At this stage the protein can adopt various high energy forms and identification of secondary structure elements is not yet possible. Under appropriate conditions the protein will undergo the folding reaction which can involve the formation of intermediates such as the molten globule, "MG" (Creighton 1990; Richards 1991). During this stage secondary structure elements will form. Moreover hydrophilic amino acid residues will become solvent exposed whereas the hydrophobic residues will be hidden creating a hydrophobic core. Further rearrangements are needed to reach the folded conformation. The "Energy landscape theory" is often adopted to model the protein's potential surface in order to

explain the thermodynamics of protein stability during the $U \Leftrightarrow N$ (unfolded to native state) process (Onuchic *et al.* 1997). According to this theory the protein thermodynamic landscape resembles a rugged funnel. (Bryngelson *et al.* 1995; Leopold *et al.* 1992; Onuchic *et al*. 1995). In the early stages the unfolded protein is guided through a multiplicity of folding routes and as the protein proceeds in the folding the number of available conformations decreases until only the lowest energy state is available. However, in the late stages the protein may be trapped in partially unfolded conformations and need the aid of specialised proteins, called chaperonins, which assist the folding process providing favorable conditions.

It is possible to hypothesise that the single point mutation introduced in FrmR may have had severe effects on the pathways toward the native structure undertaken by the protein. In this perspective E64HFrmR could possess at least two local minima on its free energy landscape each with a distinct structure although very close in energy.

7.5 Specific formaldehyde-sensing by *Salmonella* FrmR

7.5.1 *Cys35 and Pro2 play an essential role in formaldehyde detection*

The mechanism used by *E. coli* FrmR to detect formaldehyde is unknown, however it has been suggested that a cysteine residue may react with the aldehyde to yield a S-formyl adduct, which will then react with an amine from the N-terminal region producing a thiazoline-like adduct (Higgins & Giedroc 2014). Since during the course of this work we showed that *Salmonella* FrmR is capable of binding metals, presumably using most of the residues in the WXYZ motif (Chapter 3), a plausible mechanism for formaldehyde-sensing, which includes residues involved in or nearby the metal-binding site, was delineated. Pro2 is the first residue in the amino acid chain (Section 5.3.1), hence it possesses a secondary amine group not involved in peptide bond. The pyrrolidine ring on the proline is a strong base and among the most basic simple amines in nature (its conjugate acid has pK_a = 11.27) (Hall 1957), making this residue a good candidate for the nucleophile species involved in the second step of the hypothesized mechanism.

The mechanism has been confirmed by testing FrmR mutants, each lacking one of these residues (C35AFrmR and P2SFrmR), *in vitro* by fluorescence anisotropy (Section 5.4). These experiments allowed also the determination of the order of nucleophilic additions to formaldehyde. Since formaldehyde detection induces a significantly larger allosteric response in P2SFrmR than in C35AFrmR (Figures 5. 5 - 5. 7), it is possible to assume that the first step in formaldehyde sensing involves the nucleophilic attack of the mercapto group from Cys35 to yield the tetrameric intermediate, as previously suggested for *E. coli* FrmR (Higgins & Giedroc 2014). Whether, in P2SFrmR, the amino group belonging to a different residue (e.g. the imidazole of His3) could replace Pro2 pyrrolidine ring activity, or simply the modification on

Cys35 is enough to partially impair the ability of the protein to bind DNA with the same affinity of the wild-type is not yet clear.

The mechanism of sulphur nucleophile addition on formaldehyde has been extensively described in Section 1.5.1.1, however the reaction cannot proceed after the formation of the hemithioacetal species due to the absence of a second sulphur nucleophile nearby (here cysteine residue). Instead, the reaction proceeds with the nucleophilic attack of the secondary amino group on Pro2. As a result, helix α2 from one monomer is covalently linked to the N-terminal region of helix α1 from a different monomer, perturbing the local conformation of FrmR. The cross-link induced by formaldehyde presumably results in a weakening of K_{DNA} in a similar fashion to that induced by metal-binding (Section 7.3.2). The HCOH-bound FrmR species then releases its operator promoter region, allowing the RNA polymerase to bind and start the transcription of the *frmRA* operon.

Figure 7. 13 shows the electron density map of E64HFrmR were Cys35 (helix α2) and Pro2 (helix α1) are represented in ball-and-stick model. Pro2 is located at the start of the N-terminus and, along with some of the residues at the C-terminus, did not show any electron density. This residue was therefore manually added to the model by Dr. E. Pohl using COOT (Emsley & Cowtan 2004; Emsley *et al*. 2010). Cys35 and Pro2 are located at a distance in the range of chemical reactions (~ 4.6 Å). Based on the proximity of the helices, the formaldehyde-mediated link between Cys35 and Pro2 could be possible. In order to determine an approximate distance between the two residues, it was necessary to add an alanine as the first residue of the N-terminus, to minimize the Pro2 position. If Pro2 is the actual first residue of the chain, as previously suggested, it will possess more conformational freedom.

7.5.2 *frmRA expression is not induced by other aldehydes or alcohols*

FrmR detection of a number of alcohols and acetaldehyde was tested *in vivo* (by Dr. D. Osman) and *in vitro* to assess selectivity toward formaldehyde-sensing (Figures 5. 2-5. 3B). None of the compounds used in this work elicited a response in FrmR. Alcohol binding sites often consists of a loop region with hydrogen bond acceptors, for example lysines, usually located at the end of the N-terminal region (Bukiya *et al*. 2014; Dwyer & Bradley 2000). Alcohol binding can alter the local structure of the protein or displace solvent molecules inducing an allosteric response. FrmR' N-terminus presents two lysine residues (Lys8 and Lys9) that could form hydrogen bonds with alcohol molecules, although this interaction is unlikely to perturb FrmR conformation and, therefore, DNA-binding affinity.

Not even addition of acetaldehyde induces an allosteric response in FrmR (Figure 5. 2A). Acetaldehyde (CH_3CHO) differs from formaldehyde for the presence of an alkyl group. As a consequence of the electro-donating character of this substituent, the carbonyl group is less polarised and therefore less reactive toward nucleophilic additions. Moreover, since

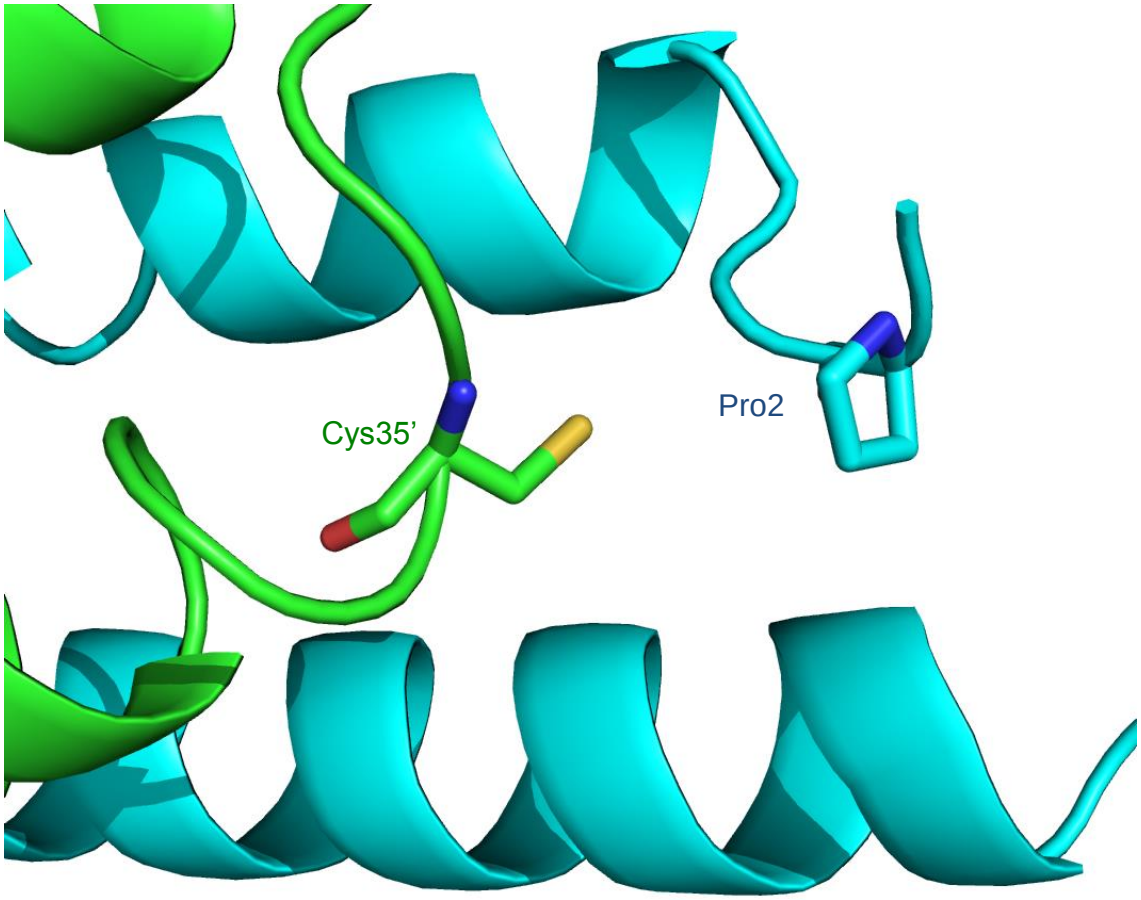


Figure 7. 13 Putative formaldehyde-binding site consisting of Cys35 (helix α2, monomer 1, in green) and Pro2 (helix α1, monomer 2, in cyan). The shortest distance between the side chains atoms is ~ 4.6 Å.

acetaldehyde is slightly bulkier than formaldehyde, the steric hindrance may prevent penetration of this aldehyde in the binding-site. In this case, it may be possible that acetaldehyde would still react with Pro2, which being located on the N-terminus loop is more flexible, yielding an enamine (Figure 7. 14A). Also in this scenario little, or no, perturbation of FrmR conformation would be expected and K_{DNA} would remain unaltered (Table 8.2, Appendix).

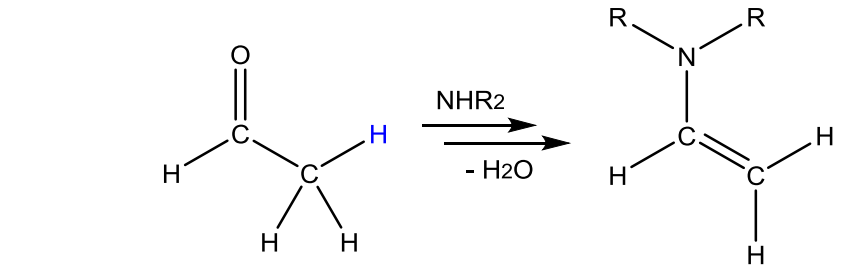
However, once the X-ray structure of one of the FrmR protein variants became available, it was possible to study the surface of the effector-binding site (Figure 8. 10, Appendix) which appeared enough accessible to allow interaction of Cys35 with substrates more sterically hindered than formaldehyde. It may be possible that this residue does not react with acetaldehyde because of inductive effects or that the Cys35-Pro2 cross-link mediated by bulkier molecules may not lead to a conformational rearrangement of helices α2 and α3 resembling that induced by formaldehyde.

7.5.3 *Glutathione-formaldehyde adducts*

During the first step of the glutathione-dependent formaldehyde detoxification pathway, glutathione is required to form an adduct with formaldehyde, S-hydroxymethylglutathione, which is then oxidised to S-formylglutathione by the formaldehyde dehydrogenase (FDH) FrmA.

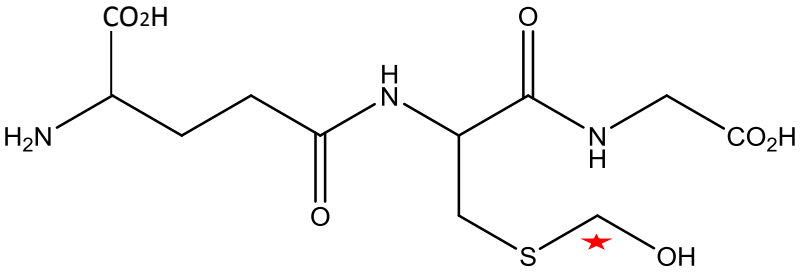
Sanghani and collaborators (Sanghani *et al.* 2000) observed a number of formaldehyde-glutathione adducts depending on the concentration of the two species and the pH (Figure 7. 14B-C-D). At pH = 7.5 (close to intracellular pH), if [GSH] equals or exceeds [HCOH] the predominant adducts will be S-hydroxymethylglutathione (Figure 7. 14B) (Sanghani *et al.* 2000). However, if formaldehyde is the most abundant species two additional adducts will form (Figure 7. 14C-D). Moreover, pH plays a crucial role as the ionization states of mercapto and amino groups determine whether formation of additional adducts will occur. Since these adducts lack a free primary alcohol group, they cannot act as substrates for the formaldehyde dehydrogenase (FDH) causing either inhibition of enzyme reaction or decrease of S-formylglutathione available (Sanghani *et al.* 2000). Nevertheless, cellular [HCOH] and [GSH] have been measured to be ~0.1 mM and ~4-5 mM, respectively (Heck *et al.* 1985; Osman *et al.* 2015) therefore the predominant species will be S-hydroxymethylglutathione.

As previously discussed in Section 7.5.2, sterically hindered substrates may be unable to trigger the conformational change which results in a weaker DNA-affinity. In this scenario, unlike FrmA, FrmR would react directly with formaldehyde instead of S-hydroxymethylglutathione. In fact, treatment of FrmR with acetaldehyde showed only a little response *in vivo* (Figure 5. 3B) and no response *in vitro* (Figure 5. 2A). Similar results have been obtained by Law with bulkier aldehydes, as furaldehyde and tribromoacetaldehyde (Law 2012). A selection performed on the basis of the effector size would be consistent with the hypothesis that the FrmR binding site

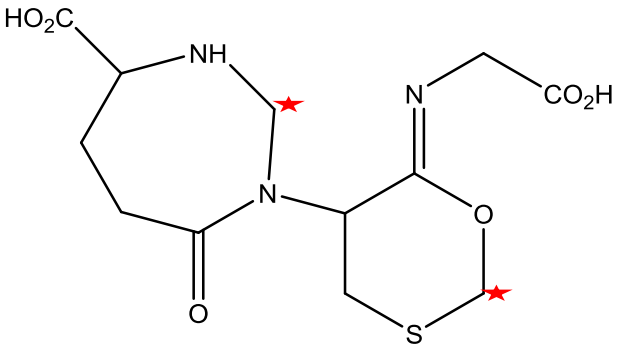


A.

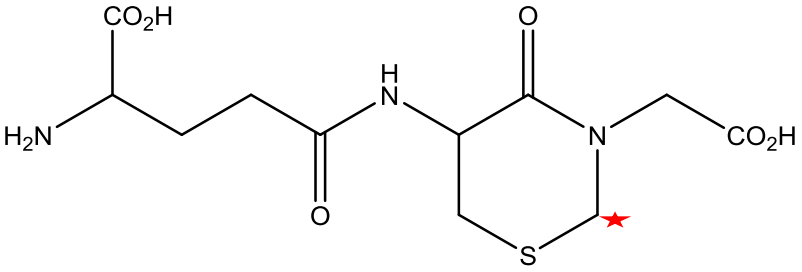
S-hydroxymethyl-glutathione



B.



C.



D.

Figure 7. 14 A Acetaldehyde reaction with secondary amine (Pro2) yields an enamine. **B-C-D** show adducts produced by glutathione reaction with formaldehyde at different pH values and initial [GSH] and [HCOH]. This scheme has been adapted from Sanghani *et al.* 2000 research paper. The symbol marks the carbon originated from formaldehyde.

evolved from one originally established to accept small species like metal ions. The same selection seems to not occur in the *B. subtilis* AdhR formaldehyde sensor, member of the MerR/NmlR family of transcriptional regulators (Section 1.5.5). Huyen and colleagues showed that, not only formaldehyde, but also methylglyoxal can trigger the transcriptional response of AdhR (Huyen *et al*. 2009). A cysteinyl residue, conserved amongst AdhR homologues (Cys52), is believed to be involved in the formaldehyde-binding mechanism, probably by being alkylated by electrophilic carbonyl groups (Huyen *et al*. 2009). This modification would be sufficient to elicit an allosteric response and activate transcription of the regulated genes. Moreover, S-hydroxymethylglutathione is less reactive than formaldehyde towards nucleophilic additions as both carbonyl groups are bounded to electron-donator groups (NHR_2) which decrease the polarization of the C=O.

However, even if FrmR might sense only formaldehyde, the S-hydroxymethylglutathione synthesis occurs spontaneously, resulting in a decreased pool of free formaldehyde that could be detected by the protein. Since the equilibrium constant for the glutathione reaction with formaldehyde is in the mM range (*K*eq = 1.77 mM) (Sanghani *et al.* 2000), this could represent the limiting factor for the evolution of the formaldehyde detection of FrmR.

7.6 *Salmonella* RcnR regulates expression of *rcnRA* operon

7.6.1 *Ni(II) and Co(II), but not formaldehyde, negatively regulate in vitro binding of rcnRAPro by Salmonella RcnR*

In order to test the specificity of the expression of the formaldehyde responsive mechanism of FrmR, *Salmonella rcnR* was identified and the encoded protein, RcnR, was produced. In a recent paper from our group it was shown that *Salmonella* RcnR binds Ni(II) and Co(II), and that it has $K_{Co(II)1-2}$= 5.06 ± 0.86 × 10^{-10} and 3 × $10^{-5} \geq K_{Co(II)3} \geq 10^{-7}$ (Osman *et al.* 2015).

As previously discussed in Section 7.3.1, *Salmonella* RcnR interacts with two TACT-N-G_6-AGTA type 1 sites located in the region within *rcnA* and *rcnR* (Section 5.5). DNA-binding affinity was first determined by EMSA, giving a $K_{DNA,app}$ of ~ 1.5 µM using an amplified region that contains both type 1 sites. However, this value departs considerably from the K_{DNA} values determined for *E. coli* RcnR ($K_{DNA\text{-}site1}$ = 126 nM and $K_{DNA\text{-}site2}$ = 174 nM) (Iwig *et al*. 2009b), therefore we decided to use fluorescence anisotropy as this technique works under equilibrium conditions. Under these experimental conditions $K_{rcnRAPro\text{-}sites1,2}^{RcnR}$ was calculated to be 152 ± 8 nM. The same methodology was also used to investigate *in vitro* Ni(II)- and Co(II)-sensing by RcnR, confirming that the two metals induce a large allosteric modification (Figures 5. 16- 5. 17, Table 8.2, Appendix).

Salmonella RcnR was then tested for *in vitro* formaldehyde detection by fluorescence anisotropy by carrying out association experiments between the protein and the labelled

rcnRAPro-sites1,2 DNA fragment in the presence of various concentrations of formaldehyde (Figure 5. 18), as previously done with FrmR proteins and *frmRAPro* (Figure 5. 1). A concentration of 100 μM of formaldehyde was needed in order to observe a $\Delta G_C = 0.94 \pm 0.21$ kcal mol^{-1}, which is still much lower than the value obtained with FrmR at a five-fold lower [HCOH] (Table 8.2, Appendix). This result supports the hypothesis of a specific formaldehyde-sensing mechanism in FrmR.

7.7 Single point mutation enhances the capacity to detect low concentration of formaldehyde *in vitro* by *Salmonella* RcnR

Cys35 and Pro2 are the putative formaldehyde-binding residues in *Salmonella* FrmR (Section 7.5.1). *Salmonella* RcnR conserves a Cys at the residue 35, however it possesses a Ser in position 2. In order to further test the mechanism hypothesized for FrmR, a single-point mutation Ser2→Pro was generated in RcnR, to create a FrmR-like pair of putative formaldehyde ligands (Cys35, Pro2). Since the cleavage of Met1 has been established by Amino Acid Analysis carried out on wild-type RcnR (Table 8. 5, Appendix), the inserted Pro2 in the S2PRcnR variant is the first residue of the N-terminus α1-helix, and hence its amino group is not involved in peptide bond with other residues.

The ability of S2PRcnR mutant to detect formaldehyde *in vitro* was tested by fluorescence anisotropy showing an encouraging shift in the DNA-binding curve after only 20 μM of formaldehyde when compared to the same experiment carried out in the absence of formaldehyde (Figure 5. 20AB, Table 8.2 Appendix). These findings support a decisive role for Pro2 and therefore sustain the hypothesis delineated in Section 5.3.1, which sees Pro2 from one protein monomer cross-linked to Cys35 from a second monomer, eliciting an overall spatial modification which impairs the ability of the protein to bind DNA. It is plausible that this structural change may resemble the theorized allosteric modification induced by metal-binding. In fact, Cys35 would be exploited in both events, as well as a residue from the N-terminus region (Pro2 for formaldehyde, probably His3 for metals).

7.8 Future work

The *Salmonella* formaldehyde sensor FrmR has been used as a probe to investigate how metal selectivity is achieved by the cell. The creation of a Zn(II)/Co(II) sensor *in vivo* by site-directed mutagenesis represents an important step in understanding the determinants involved in metal-sensing in RcnR/CsoR family members. Moreover, the work presented here highlights an interesting mechanism of selective formaldehyde-sensing. However, many questions still remain unanswered and opportunities for further investigation of FrmR' metal- and formaldehyde-sensing mechanisms have now arisen.

At the time of writing, no DNA co-crystal structure of RcnR/CsoR family members is available, and only a small number of X-ray structures of DNA-bound proteins belonging to different metal-sensor families have been deposited in the PDB (Figure 7. 15). In addition, Arunkumar and colleagues studied the DNA-bound structure of the ArsR-family member CzrA by ^{1}H-^{15}N HSQC NMR spectroscopy (Arunkumar *et al.* 2009) (Figure 7. 15).

Little is known about how FrmR recognizes the *frmRA* promoter. In the course of this work we showed that two protein tetramers presumably bind to opposite faces of the recognition site (Figure 7. 5). We also investigated the effect of the presence of Zn(II), Cu(I), and formaldehyde on the DNA:protein association. As previously discussed in Section 1.4.2, DNA-binding often implicates shape variations which may either involve a few bp or have a greater effect. Switching the protein to its holo-form may stabilize the protein:DNA complex in an energetically distinct state, as recently shown for *E. coli* CueR (Philips *et al.* 2015). X-ray crystallography may be applied to study the FrmR:DNA complex with 1:1 and 2:1 stoichiometry, in the presence or absence of metals or formaldehyde, in order to investigate if FrmR binding on the *frmRA* promoter induces a local modification in the DNA shape prior to/upon effector binding. Moreover, small angle X-ray scattering (SAXS) analyses may be carried out on FrmR-DNA complexes at the Diamond Light Source facility to determine if holo-FrmR (de-repressor-form) undergoes a similar compression to that observed by Chang and colleagues with GtCsoR (Chang *et al*. 2014), previously described in Section 7.3.2.

The rate limiting step by which *frmRA* transcription is activated in the presence of formaldehyde- (or metal-, in the case of the E64H variant) may involve binding of the effector by on-DNA apo-FrmR to encourage release of the holo-protein from the DNA. By analogy with a recent work of P. Chen and collaborators (albeit in their case on the mechanism of de-activation by sensors), this DNA-dissociation process might be assisted by other protein molecules present in the cytoplasm or weakly bound to nearby unspecific DNA regions. Chen and collaborators used *in vivo* high resolution live cell imaging methods to explore deactivation of transcription by MerR-family members CueR and ZntR when the concentration of metal ions is restored to non-toxic levels (Joshi *et al.* 2012; Chen P. *et al.* 2013; Chen T. Y. *et al*. 2015).

Transcription factors belonging to the MerR-family bind to a symmetric operator promoter region in both their apo- and holo-forms with a tight affinity (Brown *et al*. 2003; O'Halloran *et al.* 1989; Newberry & Brennan 2004; Outten *et al*. 1999; Frantz & O'Halloran 1990). In the absence of metal stress they induce a local distortion in the DNA conformation which prevents the RNAp to bind to the -10 and -35 elements, hence blocking the expression of the corresponding metal resistance genes. In the presence of toxic concentrations of metal, the metallated-protein elicits a conformational change in the DNA shape allowing the correct interaction of RNAp with -10 and -35 elements which results in the transcription activation (Brown *et al*. 2003; O'Halloran *et al.* 1989; Newberry & Brennan 2004; Outten *et al*. 1999;

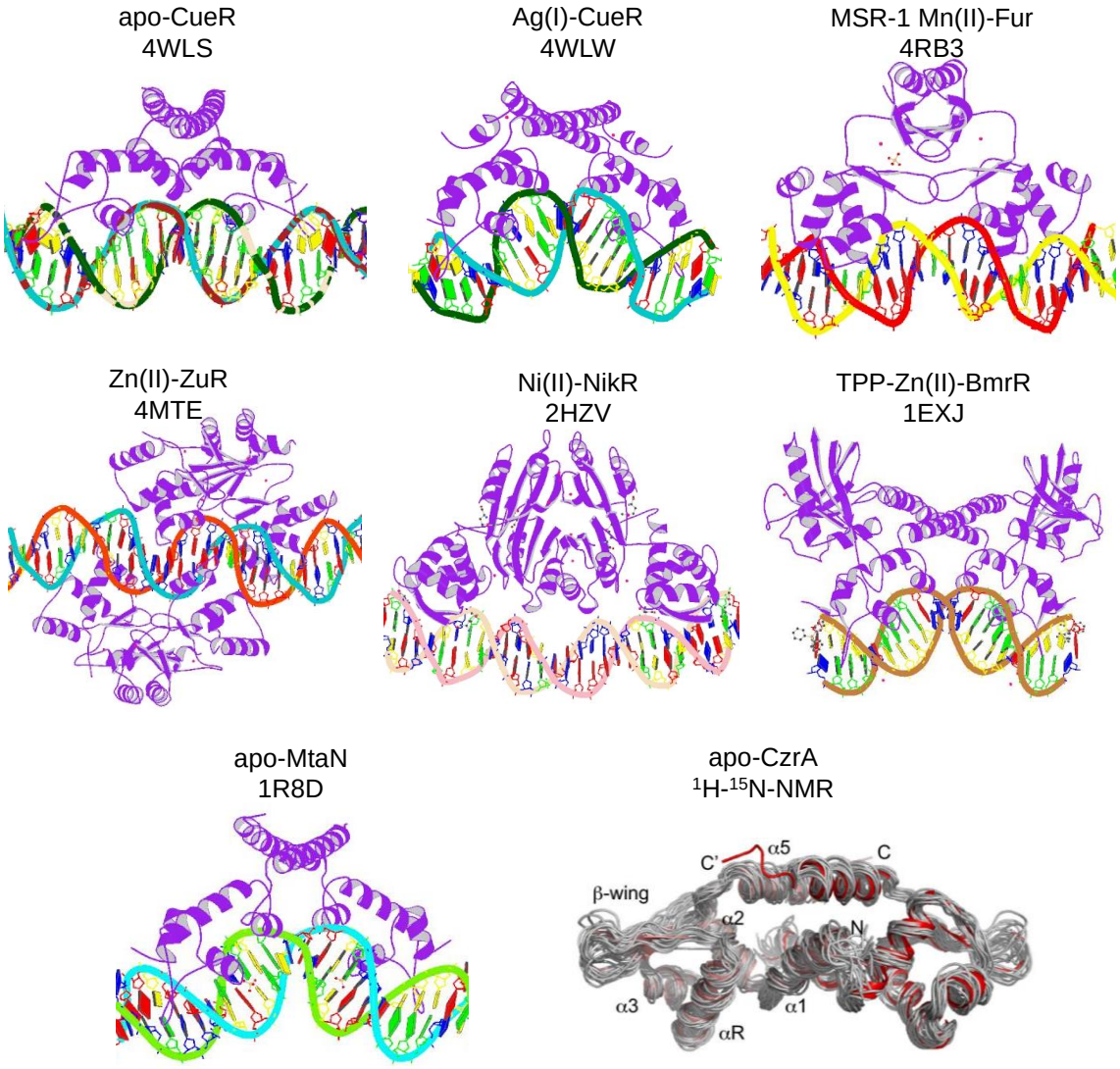


Figure 7. 15 X-ray crystal structures of DNA-bound form of metal-transcriptional regulators available in the PDB at the time of writing (Philips *et al.* 2015; Gilston *et al.* 2014; Schreiter *et al.* 2006; Heldwein & Brennan 2001; Newberry & Brennan 2004). The authors choose different colour codes to depict the DNA components as indicative of the specific modifications induced by protein-binding. The ^{1}H-^{15}N HSQC NMR structure of the DNA-bound form of apo-CzrA (Arunkumar *et al.* 2009) was added for completeness.

Frantz & O'Halloran 1990).

Once the metal stress is relieved, the cell must adopt a strategy to turn off the transcription of metal resistance genes since these are no longer needed. Two distinct pathways, the assisted protein dissociation and the direct protein substitution, were postulated to describe the unbinding of the transcription factor from the DNA (Joshi *et al.* 2012). Both processes appear to involve the formation of a ternary complex where the incumbent holo-protein partially releases one half of the symmetric DNA recognition site, allowing the incoming apo-protein to contact the second half (Chen P. *et al.* 2013). In the assisted dissociation pathway the process proceeds with the dissociation of both proteins whereas the direct protein substitution sees the release of only one of the two proteins and therefore a 50 % of probability of the apo-protein (repressor) now being bound to the DNA (Joshi *et al.* 2012). These mechanisms have been investigated in real time *in vivo* using stroboscopic single molecule tracking (SMT) showing that the kinetics of both processes depends on the protein concentration which is inversely proportional to the residence time τ on the chromosome (Chen T. Y. *et al*. 2015). The number of direct substitution and assisted dissociation events were measured by fluorescence resonance energy transfer (FRET) showing that the former pathway is more probable at high protein concentration whereas the latter is likely to occur when the concentration of protein is low (Joshi *et al.* 2012). Moreover, it was determined that direct substitution mainly governs the DNA recognition of holo-CueR, whereas apo-CueR interacts with DNA predominantly via assisted dissociation (Joshi *et al.* 2012).

It would be of great interest to investigate if mechanistic pathways similar to those observed for CueR and ZntR (Chen T. Y. *et al*. 2015) may occur with FrmR, or if unbinding from the *frmRA* operator (here resulting in the derepression of transcription) follows spontaneous routes. Notably, this becomes more likely if on-DNA FrmR proteins do have a weaker $K_{Zn(II)}$ than off-DNA molecules, and this may similarly be true for K_{HCOH}. In the presence of inducer(s), off-DNA apo-FrmRs would then be more likely to bind the effector but the resultant off-DNA holo-FrmRs might then aid disassembly of apoFrmR-*frmRA*Pro adduct through ternary complexes as observed in CueR and ZntR. The pathway could proceed with the release of both proteins from the DNA (assisted dissociation) or with the 50 % of probability to have the substitution of apo-FrmR with the holo-form (direct substitution). Holo-FrmR would then spontaneously dissociate as it has a poor affinity for the *frmRA* operator. In both cases the transcription would be activated.

The determination of the copy number of protein molecules in the cell appears to be a key parameter in order to further understand transcriptional regulation. The cellular abundances of apo-FrmR in logarithmic cultures growth in M9 minimal medium was measured to be 135 ± 17 tetramer cell^{-1} (Osman *et al.* 2015). However, it must be noted that this value refers to FrmR expressed from the pET29a vector and not from the chromosome. Using an estimate of protein

abundance approximately 10-fold lower and a cell volume of ~ 1 fL, the concentration of FrmR is then calculated to be ~ 20 nM. The same methodology can be applied to measure the protein abundance at several time points during the cell growth. Use of stroboscopic single molecule tracking (SMT) to measure the *in vivo* residence time of apo- and holo-forms of FrmR molecules on the DNA would then allow determination of any dependency on protein concentration as observed in MerR family-members.

It would be interesting to determine what proportion of FrmR molecules in their activator form (formaldehyde-bound) versus those in the repressor state (apo-form) is necessary to allow formaldehyde-sensing. If formaldehyde-binding induces a substantial structural rearrangement in FrmR, it might be possible to use FRET to visualize the holo-proteins in the cell. Moreover, calculating the copy number of E64HFrmR and *Salmonella* cognate metal sensors RcnR, ZntR, and ZuR expressed from the chromosome may provide insights to determine at which extent this parameter affects metal-selectivity. In fact, if the apo-form of a certain metal sensor is greatly in excess compared to the others, this would be able to access the metal with a higher probability.

Structural studies are needed to further explore some aspects of the mechanisms adopted by FrmR and to investigate the nature of the spectral feature at 300 nm observed with the majority of E64HFrmR preparations. In this regard, mass spectrometry analyses using hard ionization techniques may be adopted on E64HFrmR, E64HFrmR* and FrmR. Furthermore, X-ray diffraction may investigate the potential discrepancy in their 3D spatial disposition. In particular, in Section 7.3.2 we discussed the location of the helix α3 in E64HFrmR which seems to occupy a position approximately analogous to that of helix α3 in the copper-bound GtCsoR structure (Figure 7. 9). One hypothesis sees the enhanced K_{met} (and, possibly, the weakened K_{DNA}) observed in E64HFrmR (Table 8. 3, Appendix) as a consequence of a different spatial disposition of the three helices, in particular of helix α3 (compared to that of apo-wild type), which resembles more the effector-binding state of RcnR/CsoR family members. Since preliminary experiments carried out on E64HFrmR* showed a K_{DNA} comparable to that of FrmR (Table 8. 3, Appendix), it is possible that these proteins share a similar 3D conformation. Production and superposition of apo- and holo-forms of FrmR, E64HFrmR, and E64HFrmR* X-ray structures would be ideal.

These studies may also provide important information regarding the ligands involved in metal-coordination and the propagation of the allosteric mechanism connecting the effector-binding event with a weakened K_{DNA}. Moreover, by producing the formaldehyde-bound structure of FrmR it would be possible to determine if the formaldehyde-mediated, inter-subunit cross-link postulated in this work induces a quaternary structure similar to that achieved by the conformational rearrangement elicited by metal-coordination.

Formaldehyde-detection by FrmR may also be investigated by mass-spectrometry to study the cross-linked FrmR species after exposure to appropriate formaldehyde concentrations. Moreover, fluorescence anisotropy can be adopted to confirm, or otherwise, that adducts of formaldehyde with other molecules present in high concentration in the cytoplasm (such as glutathione) may not be able to disrupt the protein interactions with the promoter region, as predicted in this work (Section 7.5.4).

Another interesting aspect disclosed by the work carried out on FrmR and E64HFrmR is the mechanism of Co(II)-sensing. Although the Co(II)-binding affinity of E64HFrmR is significantly weaker than compared to that of *Salmonella* Ni(II)/Co(II) sensor RcnR ($K_{Co(II)1\text{-}2}^{RcnR}$ 5.06 ± 0.86 x 10^{-10} M, $K_{Co(II)1}^{E64HFrmR}$ = 2.56 ± 0.4 x 10^{-7}M; Osman *et al.* 2015, Table 8. 3, Appendix), E64HFrmR is still capable of detection of cellular cobalt. The possibility that glutathione may be a key contributor to Co(II)-sensing by E64HFrmR has been investigated *in vivo* by Dr. Deenah Osman (Section 7.4.4) revealing a glutathione-dependency of Co(II)- and Zn(II)-sensing by the protein. The hypothesis is that Co(II)-bound glutathione molecules interact with the protein, transferring the metal ion directly to the binding site. It is plausible to assume that the glutathione may, at least partially, penetrate the pocket to deliver the Co(II). However, the physiological purpose of a specific interaction between glutathione and FrmR remains enigmatic. As glutathione also appears to aid the cobalt-sensing by the *Salmonella* cobalt sensor RcnR (Osman *et al.* 2015), it was hypothesized that RcnR has evolved to receive Co(II) ions from glutathione and that the co-evolved FrmR may have retained the same ability. Otherwise, it is possible that FrmR is able to detect glutathione-formaldehyde adducts although it appears unlikely, as previously discussed in Section 7.5.4. Glutathione may also be necessary to maintain the reduced state of FrmR and RcnR, both thiol containing sensors, thus explaining the *in vivo* dependency of metal-sensing.

Recently the role of another low molecular weight (LMW) thiol, bacillithiol (BSH), has been investigated in *B. subtilis* revealing a contrasting behavior to that observed in *Salmonella* with glutathione (Ma *et al.* 2014; Osman *et al.* 2015). *B. subtilis* does not produce glutathione but possesses a concentration of bacillithiol that varies from ~ 1 mM in mid-log growth phase to ~ 5 mM in stationary growth phase (Sharma *et al.* 2013). Compared to that of glutathione, the thiol on bacillithiol is more acidic, and therefore more reactive (Sharma *et al.* 2013). Bacillithiol can form both 2:1 and 1:1 BSH:Zn(II) complexes with a tight affinity ($\beta_{Zn(II)}$ 1.9 x 10^{-12} M^{-2}) whereas glutathione has a $K_{Zn(II)}^{GSH}$ 3 x 10^{4} M^{-1} for the 1:1 complex (Ma *et al.* 2014).

A higher sensitivity to thiophilic metals, such as Cd(II), Cu(II) and Zn(II), has been detected in cells lacking bacillithiol (Rajkarnikar *et al.* 2013; Ma *et al.* 2014). Moreover, Ma and colleagues observed that bacillithiol acts as a competitor for Zn(II)-binding with the Zn(II)-sensor BsuCzrA (StmB/ArsR family). A similar hypothesis was also inferred to explain the slightly

enhanced induction of *copZA*, regulated by BsuCsoR, in the bacillithiol null strain (Ma *et al.* 2014).

In order to discern if the effect that glutathione has on *Salmonella* FrmR and RcnR metal-sensing is to be ascribed to channeling or to the maintenance of the reduced state of these sensors, the *in vivo* sensitivity to metal stress may be measured in the presence of bacillithiol in a glutathione null *Salmonella* strain. In this scenario, the cell would still possess an efficient reductant whereas any interaction between the protein and glutathione would be absent. Preservation of Co(II)-sensing, even upon substitution of glutathione with bacillithiol, would mean that no specific channeling occurs in the wild-type strain. On the contrary, an inhibited Co(II)-detection may be explained by supposing that *Salmonella* metal sensors have evolved from a common ancestor to receive the metal by ligand exchange reaction with glutathione. However, the insertion of the bacillithiol machinery in the *Salmonella* strain may be complex. The bacillithiol biosynthetic pathway involves BshA (a glycosyl transferase), a BshB paralog (a deacetylase) and BshC (a cysteine ligase) (Gaballa *et al.* 2010; Parsonage *et al.* 2010). As an alternative, metal-sensing by *Salmonella* FrmR and RcnR may be studied in wild-type *B. subtilis* (which inherently lacks the glutathione machinery), or Zn(II)-sensing by BsuCzrA may be investigated in a bacillithiol null *B. subtilis* strain with an inserted glutathione biosynthetic pathway (Gsh1, γ-glutamylcysteine synthetase, and Gsh2, glutathione synthetase) (White *et al.* 2003) or in a wild-type *Salmonella* strain.

Further understanding of the formaldehyde- and metal-sensing mechanisms adopted by FrmR proteins will aid a deeper comprehension of the transcriptional control accomplished by bacterial transcription factors. During the course of this work, many thermodynamic parameters have been measured (Tables 8.2 - 8.3, Appendix) and used to produce transcriptional regulation models. However, the extent to which these thermodynamic predictions are truly descriptive of what occurs in the cell remains to be evaluated. What is the contribution of kinetic factors and associative processes?

Chapter 8

Appendix

Buffer name	Composition
Buffer A	50 mM sodium phosphate pH 7.4, 300 mM NaCl, 5 mM DTT, 1 mM PMSF, 10 mM imidazole
Buffer B	50 mM sodium phosphate pH 7.4, 300 mM NaCl, 5 mM DTT, 300 mM imidazole
Buffer E	1.2 M NaCl, 40 mM Hepes pH 7.8, 8 mM EDTA, 8 mM DTT
B100	100 mM NaCl, 10 mM Hepes pH 7.0, 10 mM EDTA, 10 mM DTT
B200	200 mM NaCl, 10 mM Hepes pH 7.0, 10 mM EDTA, 10 mM DTT
B300	300 mM NaCl, 10 mM Hepes pH 7.0, 10 mM EDTA, 10 mM DTT
B400	400 mM NaCl, 10 mM Hepes pH 7.0, 10 mM EDTA , 10 mM DTT
B500	500 mM NaCl, 10 mM Hepes pH 7.0, 10 mM EDTA, 10 mM DTT
B800	800 mM NaCl, 10 mM Hepes pH 7.0, 10 mM EDTA, 10 mM DTT
B1000	1000 mM NaCl, 10 mM Hepes pH 7.0, 10 mM EDTA, 10 mM DTT
C100 (chelex-treated, N_2-purged)	80 mM KCl, 20 mM NaCl, 10 mM Hepes pH 7.0
C300 (chelex-treated, N_2-purged)	240 mM KCl, 60 mM NaCl, 10 mM Hepes pH 7.0
C500 (chelex-treated, N_2-purged)	400 mM KCl, 100 mM NaCl, 10 mM Hepes pH 7.0
C1000 (chelex-treated, N_2-purged)	800 mM KCl, 200 mM NaCl, 10 mM Hepes pH 7.0

Table 8. 1 List of buffers used during the course of this work.

Protein	**Effector**	**K_{DNA} (M)**	**ΔG_c (kcal mol^{-1})**
FrmR	Apo	$9.94 \pm 0.30 \times 10^{-8}$	/
	Zn(II)	$3.11 \pm 0.40 \times 10^{-6}$	2.03 ± 0.08
	Cu(I)	$6.54 \pm 1.30 \times 10^{-7}$	1.10 ± 0.10
	10 µM HCOH	$6.37 \pm 0.16 \times 10^{-7}$	1.10 ± 0.02
	20 µM HCOH	$7.09 \pm 0.60 \times 10^{-6}$	2.52 ± 0.05
	20 µM EtOH	9.21×10^{-8}	/
	20 µM C_2H_4O	8.56×10^{-8}	/
E64HFrmR	Apo	$4.26 \pm 0.4 \times 10^{-7}$	/
	Zn(II)	$3.51 \pm 0.7 \times 10^{-6}$	1.24 ± 0.16
E64HFrmR*	Apo	$1.06 \pm 0.7 \times 10^{-7}$	/
	Zn(II)	$2.75 \pm 0.4 \times 10^{-6}$	1.92 ± 0.09
C35AFrmR	Apo	$1.58 \pm 0.18 \times 10^{-7}$	/
	Zn(II)	$1.47 \pm 0.54 \times 10^{-7}$	-0.07 ± 0.21
	20 µM HCOH	$1.51 \pm 0.22 \times 10^{-7}$	-0.02 ± 0.01
	50 µM HCOH	$1.29 \pm 0.56 \times 10^{-7}$	-0.15 ± 0.02
	100 µM HCOH	$9.7 \pm 0.45 \times 10^{-8}$	-0.29 ± 0.05
H60LFrmR	Apo	1.92×10^{-6}	/
E81AFrmR	Apo	8.71×10^{-7}	/
P2SFrmR	Apo	$1.47 \pm 0.16 \times 10^{-7}$	/
	20 µM HCOH	$5.86 \pm 1.30 \times 10^{-7}$	0.81 ± 0.13
	50 µM HCOH	$1.35 \pm 0.47 \times 10^{-6}$	1.29 ± 0.22
	100 µM HCOH	$5.76 \pm 0.46 \times 10^{-6}$	2.18 ± 0.07
RcnR on *rcnRA*Pro^{site1}	Apo	1.56×10^{-6}	/

RcnR on *rcnRA*Pro^{sites1,2}	Apo	$1.52 \pm 0.80 \times 10^{-7}$	/
	Ni(II)	$5.91 \pm 1.29 \times 10^{-6}$	2.17 ± 0.15
	Co(II)	$1.50 \pm 0.18 \times 10^{-5}$	2.72 ± 0.15
	20 µM HCOH	$1.58 \pm 0.08 \times 10^{-7}$	0.03 ± 0.10
	50 µM HCOH	$3.67 \pm 1.10 \times 10^{-7}$	0.51 ± 0.18
	100 µM HCOH	$7.75 \pm 2.85 \times 10^{-7}$	0.94 ± 0.21
RcnR on *frmRAPro*	Apo	7.46×10^{-7} – 2.00×10^{-6}	/
S2PRcnR	Apo	$1.55 \pm 0.10 \times 10^{-7}$	/
	20 µM HCOH	$6.96 \pm 1.41 \times 10^{-7}$	0.88 ± 0.14
	50 µM HCOH	$1.78 \pm 0.47 \times 10^{-6}$	1.43 ± 0.18
	100 µM HCOH	$2.62 \pm 0.33 \times 10^{-6}$	1.67 ± 0.09

Table 8. 2 DNA-binding affinities and allosteric coupling free energies of proteins used during the course of this work.

Sensor	Metal	K_{metal} (M)
FrmR	Zn(II)	$K_{1-3} = 1.7 \pm 0.7 \times 10^{-10}$
	Co(II)	$K_{1-4} = 7.59 \pm 0.4 \times 10^{-6}$
	Cu(I)	$K_{1-2} = 4.9 \pm 1.6 \times 10^{-15}$; $K_{3-4} = 1.72 \pm 0.7 \times 10^{-12}$; $K_{5-8} \geq 8 \times 10^{-11}$
E64HFrmR	Zn(II)	$K_{1-3} = 2.33 \pm 0.3 \times 10^{-11}$
	Co(II)	$K_1 = 2.56 \pm 0.4 \times 10^{-7}$; $K_{1-4} \leq 10^{-6}$
	Cu(I)	$K_{1-2} \sim 5 \times 10^{-16}$; $K_{3-4} = 7.29 \pm 1.29 \times 10^{-15}$; $K_{5-6} = 5.6 \pm 2.0 \times 10^{-12}$; $K_{7-8} \geq 4 \times 10^{-10}$

Table 8. 3 Metal-binding affinities of FrmR and E64HFrmR.

DynaFit Scripts

Dynafit script used to determine $K_{Co(II)}$ in competition experiment between FrmR and BisTris

```
[task]

data = equilibria
task = fit

[components]

; C = Co(II)

; P = protein Cu(I) sites

; B = BisTris

[mechanism]

     C + P <==> CP     :    Keq1   dissociation
     C + B <==> CB     :    Keq2   dissociation


[concentrations](µM)

P = [protein monomer]x 4
B = [BisTris]


[constants]

Keq1 = hypothesised value ?, Keq2 = K_Co(II)^BisTris

[responses]

CP = final ΔAbs336?


[equilibria]


variable C

offset = auto
```

Dynafit script used to determine $K_{Cu(I)}$ in competition experiment between FrmR and BCA

```
[task]

data = equilibria
task = fit


[components]

; C = Cu(I)

; B = BCA

; W = 1-2 protein Cu(I) sites

; X = 3-4 protein Cu(I) sites

; Y = 5-8 protein Cu(I) sites


[mechanism]

     C + B + B <==> CB2      :    Keq1   association
     C + W <==> CW           :    Keq2   association
     C + X <==> CX           :    Keq3   association
     C + Y <==> CY           :    Keq4   association

[concentrations](M)

B = [BCA]
W = [protein monomer]
X = [protein monomer]
Y = [protein monomer]X 2

[constants]

Keq1 = K_Cu(I)^BCA, Keq2 = hypothesised value?,  Keq3 = hypothesised
value?,  Keq4 = hypothesised value?

[response]

CB2 = final ΔAbs562 ?

[equilibria]

variable C
offset = auto
```

Dynafit script used to determine $K_{Cu(I)}$ in competition experiment between E64HFrmR and BCA

```
[task]

   data = equilibria
   task = fit

[components]

; C = Cu(I)

; B = BCA

; W = 1-2 protein Cu(I) sites

; X = 3-4 protein Cu(I) sites

; Y = 5-6 protein Cu(I) sites

; Z = 7-8 protein Cu(I) sites


[mechanism]

     C + B + B <==> CB2     :    Keq1   association
     C + W <==> CW          :    Keq2   association
     C + X <==> CX          :    Keq3   association
     C + Y <==> CY          :    Keq4   association
     C + Z <==> CZ          :    Keq5   association

[concentrations](M)
B = [BCA]
W = [protein monomer]
X = [protein monomer]
Y = [protein monomer]
Z = [protein monomer]

[constants]

Keq1 = K_Cu(I)^BCA, Keq2 = hypothesised value ?,  Keq3 = hypothesised
value?,  Keq4 = hypothesised value?, Keq5 = hypothesised value?

[response]

 CB2 = final ΔAbs562 ?

[equilibria]

 variable C
 offset = auto
```

Dynafit script used to determine $K_{Zn(II)}$ in competition experiment between FrmR proteins and mag-fura-2

```
[task]

data = equilibria
task = fit

[components]

; Z = Zn(II)
; M = Mag-fura2
; P = 1-3 protein Zn(II) sites


[mechanism]

      Z + M  <==> ZM     :    Keq1   dissociation
      Z + P  <==> ZP     :    Keq2   dissociation
      Z + ZP <==> ZZP     :    Keq3   dissociation
      Z + ZZP <==> ZZZP     :    Keq4   dissociation


 [concentrations](M)

M = [mag-fura-2]
P = [protein monomer]


[constants]

Keq1 = K_Zn(II)^mag-fura-2, Keq2 = hypothesised value?, Keq3 =
hypothesised value?, Keq4 = hypothesised value?


[Response]

ZM = - final ΔAbs366?


[equilibria]

variable Z
offset = auto
```

Dynafit script used to determine $K_{Zn(II)}$ in competition experiment between FrmR proteins and quin-2

```
[task]

data = equilibria
task = fit

[components]

; Z = Zn(II)
; Q = Quin-2
; P = 1-3 protein Zn(II) sites


[mechanism]

      Z + Q  <==> ZQ     :    Keq1   dissociation
      Z + P  <==> ZP     :    Keq2   dissociation
      Z + ZP <==> ZZP     :    Keq3   dissociation
      Z + ZZP <==> ZZZP     :    Keq4   dissociation

 [concentrations](M)

Q = [quin-2]
P = [protein monomer]

 [constants]

Keq1 = K_Zn(II)^quin-2, Keq2 = 0.002?, Keq3 = 0.002?, Keq4 = 0.002?

[Response]

ZQ = - final ΔAbs265?


[equilibria]

variable Z
offset = auto
```

Dynafit script used to determine $K_{DNA}^{apo\text{-}protein}$ of FrmR and RcnR proteins by fluorescence anisotropy (FA)

```
[task]

data = equilibria
task = fit

[components]

; P = protein monomer
; D = DNA


[mechanism]

     P + P + P + P <==> P4     :    Keq1   dissociation
     P4 + D <==> (P4)D     :    Keq2   dissociation
     P4 + (P4)D <==> (P4)2D     :    Keq2   dissociation


[concentrations](nM)

D = 10

[constants]

Keq1 = 0.00000000001
Keq2 = supposed K_DNA^apo-protein?


[Response]

(P4)D = (final Δr_obs^apo)/2

(P4)2D = final Δr_obs^apo?

[equilibria]
variable P
offset = auto
```

Ref.: Liu T. *et al.* 2008

Dynafit script used to determine $K_{DNA}^{effector\text{-}protein}$ of FrmR and RcnR proteins by fluorescence anisotropy (FA)

```
[task]

data = equilibria
task = fit

[components]

; P = protein monomer
; D = DNA


[mechanism]

      P + P + P + P <==> P4     :    Keq1   dissociation
       P4 + D <==> (P4)D     :    Keq2   dissociation
       P4 + (P4)D <==> (P4)2D     :    Keq2   dissociation


[constants]

Keq1 = 0.00000000001
Keq2 = supposed K_DNA^effector-protein?

[concentrations](nM)

D = 10

[Response]

(P4)D = (mean final Δr_obs^apo)/2

(P4)2D = mean final Δr_obs^apo

[equilibria]
variable P
offset = auto
```

Ref.: Liu T. *et al.* 2008

AA	n.mole/ml	ug/ml	mg/ml
Cysteic acid	-	-	-
Hydroxyproline	-	-	-
Aspartic acid	784	90.3	0.0903
Threonine	559	56.5	0.0565
Serine	967	84.2	0.0842
Glutamic acid	2720	351	0.351
Proline	783	76.0	0.0760
Glycine	1190	67.9	0.0679
Alanine	1770	126	0.126
Cysteine, est	*190*	*19.6*	*0.0196*
Valine	1110	110	0.110
Methionine	513	67.3	0.0673
Isoleucine	1150	130	0.130
Leucine	2110	238	0.238
Tyrosine	181	29.6	0.0296
Phenylalanine	-	-	-
Histidine	650	89.2	0.0892
Tryptophan	-	-	-
Lysine	763	97.9	0.0979
Arginine	1830	286	0.286
Totals	17300	1920	1.92

Table 8. 1 Sample (n = 2) of Amino Acid Analysis after hydrolysis (24 hours at 110 °C) carried out on FrmR (AltaBioscience Ltd). Documentation provided by the company states: "Asn and Gln are completely converted to Asp and Glu during the acid hydrolysis of the protein. The values for Thr and Ser have been corrected for hydrolysis losses of 5% and 10% respectively. Trp usually suffers complete loss during acid hydrolysis and is not normally quantified. Cys is usually observed as cystine and its recovery is variable using standard hydrolysis conditions. The values for His are sometimes affected by co-eluting compounds from the sample. The reported values have been rounded off to either 2 or 3 significant figures, depending on peak size".

AA	n.mole/ml	ug/ml	mg/ml
Cysteic acid	-	-	-
Hydroxyproline	-	-	-
Aspartic acid	478	55	0.055
Threonine	210	21.2	0.0212
Serine	254	22.2	0.0222
Glutamic acid	844	109	0.109
Proline	74.9	7.27	0.00727
Glycine	279	15.9	0.0159
Alanine	524	37.3	0.0373
Cysteine, est	*28.8*	*6.39*	*0.00639*
Valine	367	36.3	0.0363
Methionine	124	16.3	0.0163
Isoleucine	530	60	0.06
Leucine	603	68.2	0.0682
Tyrosine	58.7	9.58	0.00958
Phenylalanine	-	-	-
Histidine	406	55.8	0.0558
Tryptophan	-	-	-
Lysine	605	77.6	0.0776
Arginine	387	60.5	0.0605
Totals	5770	658	0.658

Table 8. 2 Amino acid analysis after hydrolysis (24 hours at 110 °C) carried out on RcnR (AltaBioscience Ltd). Documentation provided by the company states: "Asn and Gln are completely converted to Asp and Glu during the acid hydrolysis of the protein. The values for Thr and Ser have been corrected for hydrolysis losses of 5% and 10% respectively. Trp usually suffers complete loss during acid hydrolysis and is not normally quantified. Cys is usually observed as cystine and its recovery is variable using standard hydrolysis conditions, therefore the value related to cysteic acid has been obtained from separate analysis. The values for His are sometimes affected by co-eluting compounds from the sample. The reported values have been rounded off to either 2 or 3 significant figures, depending on peak size".

Calculations of protein experimental extinction coefficient

Accurate extinction coefficient was calculated by using the experimental protein concentration, obtained from Amino Acid Analysis (Table 8.4 – 8.5), and the A_{280} recorded on the same sample adopted in the analysis (Figures 8. 1 - 8. 2), in the Lambert-Beer equation (Equation 2).

FrmR:

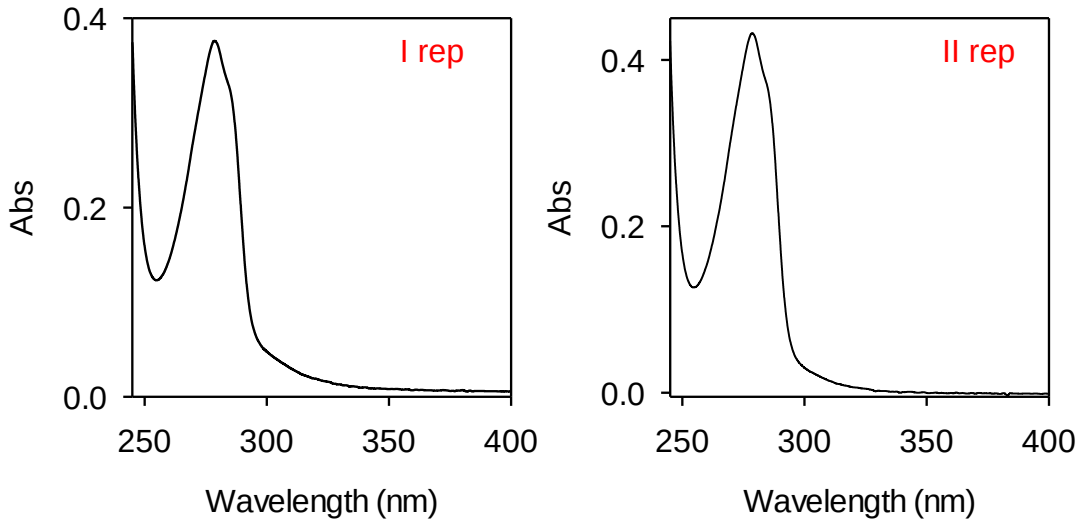


Figure 8. 1 UV-visible spectra of the apo-FrmR samples analysed for Amino Acid Analysis by AltaBioscience Ltd, recorded in buffer C500 at pH 7.0.

I rep

c = 1.92 mg/ml, equivalent to 188.5 µM (MW^{FrmR} 10185.8 Da)

A_{280} = 0.371075

ε_{280} = 0.371075/(1cm x 0.0001885 M) = 1968.1 $M^{-1}cm^{-1}$

II rep

c = 2.24 mg/ml, equivalent to 219.9 µM (MW^{FrmR} 10185.8 Da)

A_{280} = 0.425267

ε_{280} = 0.425267/(1cm * 0.0002199 M) = 1933.9 $M^{-1}cm^{-1}$

FrmR extinction coefficient used in this work, 1951 $M^{-1}cm^{-1}$, is the average of the two values calculated here.

Calculations of protein experimental extinction coefficient

RcnR:

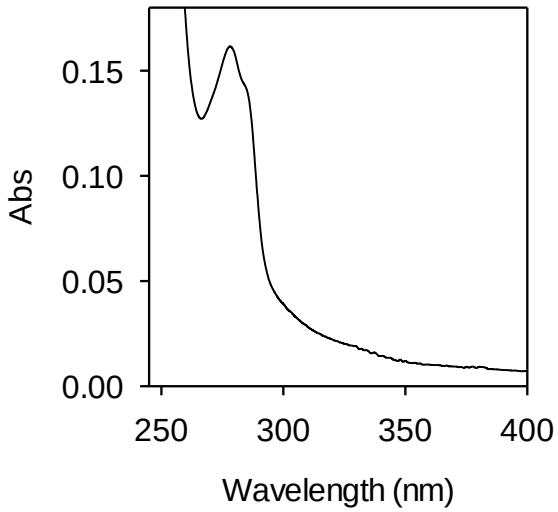


Figure 8. 2 UV-visible spectrum of the apo-RcnR sample analysed for Amino Acid Analysis by AltaBioscience Ltd, recorded in buffer C500 at pH 7.0.

c = 0.658 mg/ml, equivalent to 65.3 µM (MW^{RcnR} 10207.9 Da)

A_{280} = 0.158183

ε_{280} = 0.158183/(1 cm * 0.0000653 M) = 2422.4 $M^{-1}cm^{-1}$

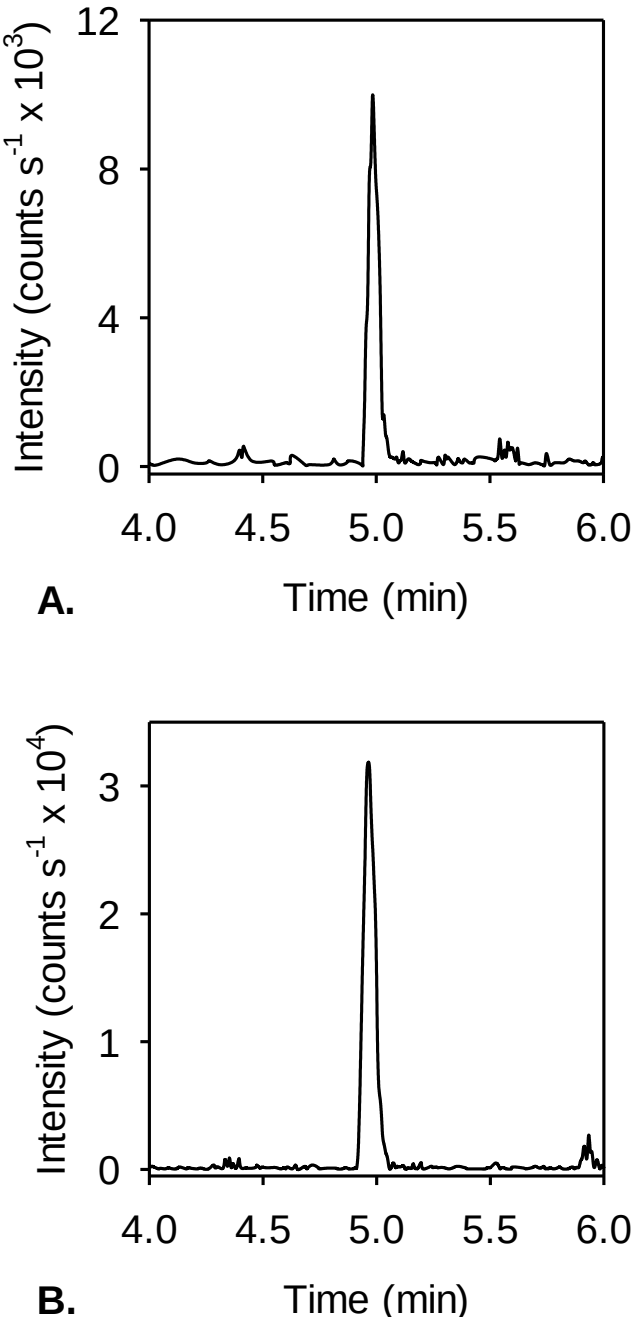


Figure 8. 3 Extracted LC-MS chromatograms of ion transitions (405.193-488.235 Da) monitored for the FrmR (**A**) and E64HFrmR (**B**) PHSPEDK peptide detected in trypsin digested protein samples. These analyses were performed by Dr. Chen and Dr. Huggins (Procter and Gamble Mason Business Centre, Cincinnati, Ohio; Section 2.4.10).

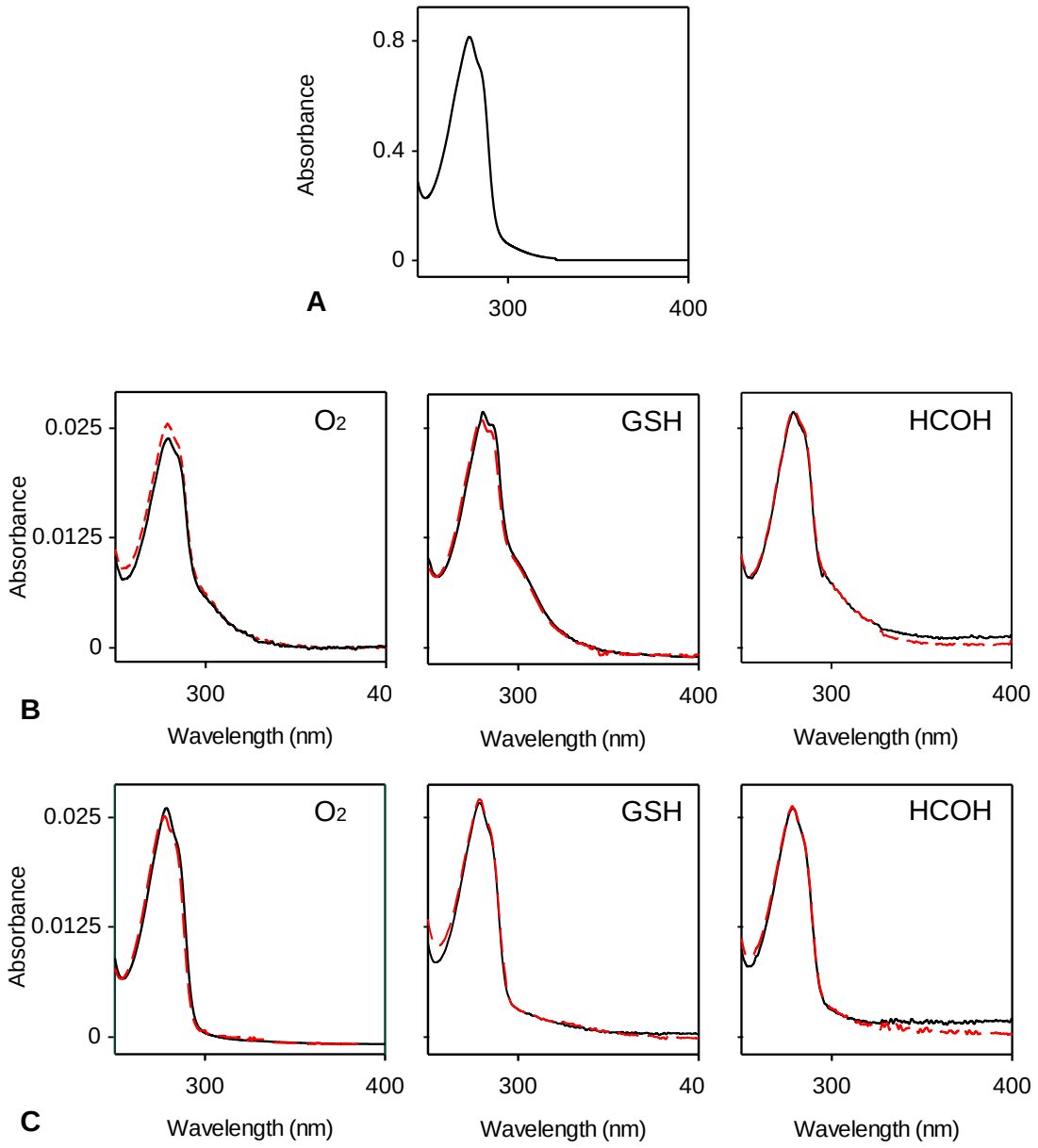


Figure 8. 4 A UV-visible spectrum of apo-FrmR (412 µM), recorded anaerobically in buffer C500 at pH 7.0. **B** UV-visible spectra of E64HFrmR (~ 12 µM), recorded anaerobically (solid black line) and again aerobically (left panel, red line), or upon addition of 250 µM glutathione (center panel, red line), or 250 µM formaldehyde (right panel, red line) after 16 h incubation. Buffer C500, pH 7.0. **C** UV-visible spectra of E64HFrmR* (~ 12 µM), recorded anaerobically (solid black line) and again aerobically (left panel, red line), or upon addition of 250 µM glutathione (center panel, red line), or 250 µM formaldehyde (right panel, red line) after 16 h incubation. Buffer C500, pH 7.0.

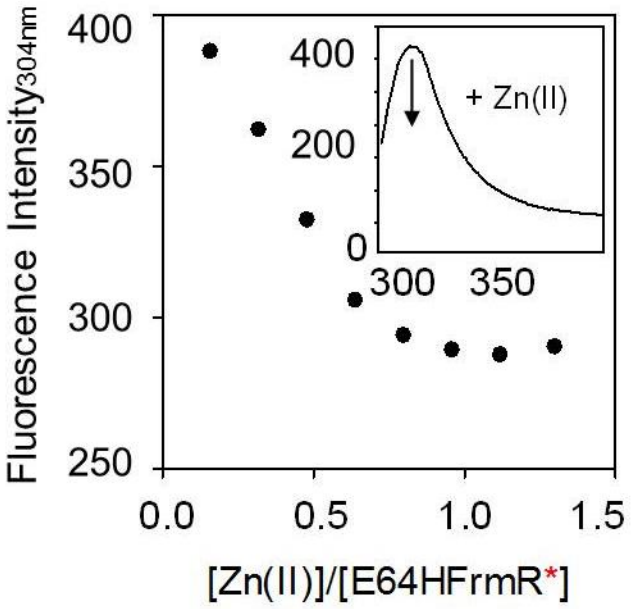


Figure 8. 5 Fluorescence emission at 304 nm (λ_{ex} = 280 nm) of E64HFrmR***** (10.21 µM) and following titration with $ZnCl_2$. Fluorescence is quenched upon addition of ~ 1 molar equivalent of zinc.

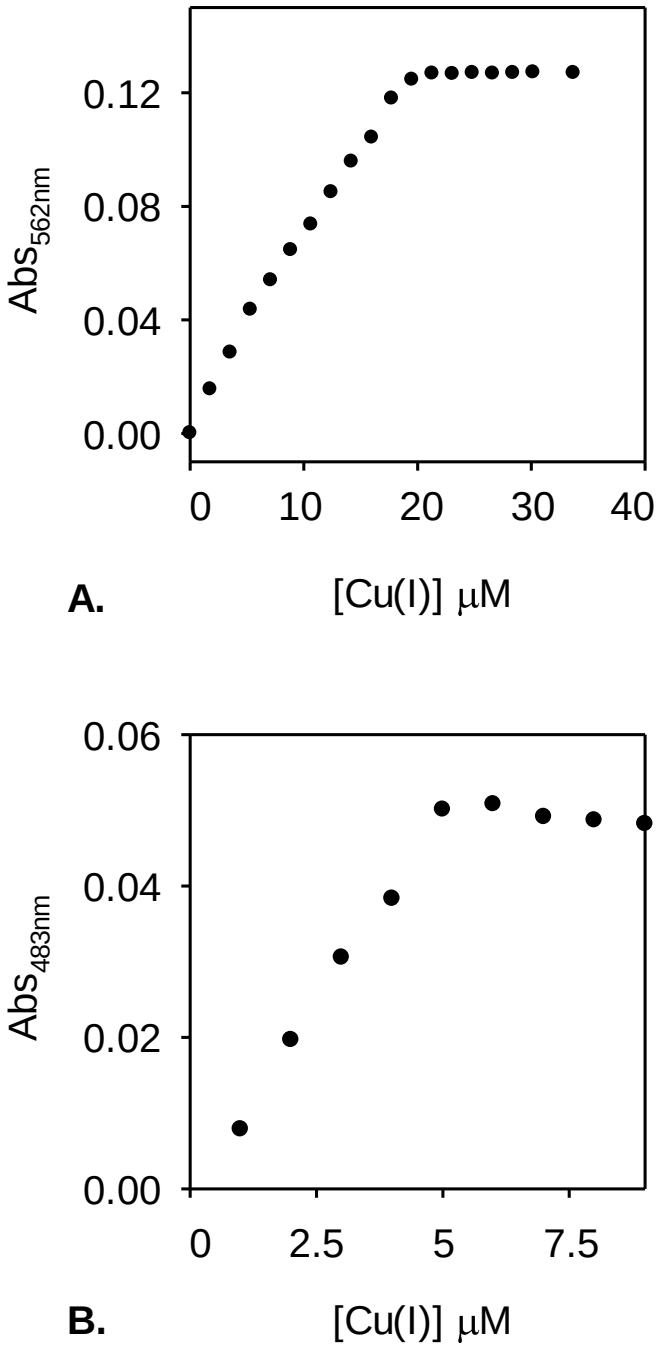


Figure 8. 6 A Sample of control titration of BCA (40 μM) with CuCl [> 95% Cu(I)] monitored anaerobically by UV-vis spectrophotometer. System conditions: buffer C500, pH7.0. **B** Sample of control titration of BCS (10 μM) with CuCl [> 95% Cu(I)] monitored anaerobically by UV-vis spectrophotometer. System conditions: buffer C500, pH 7.0.

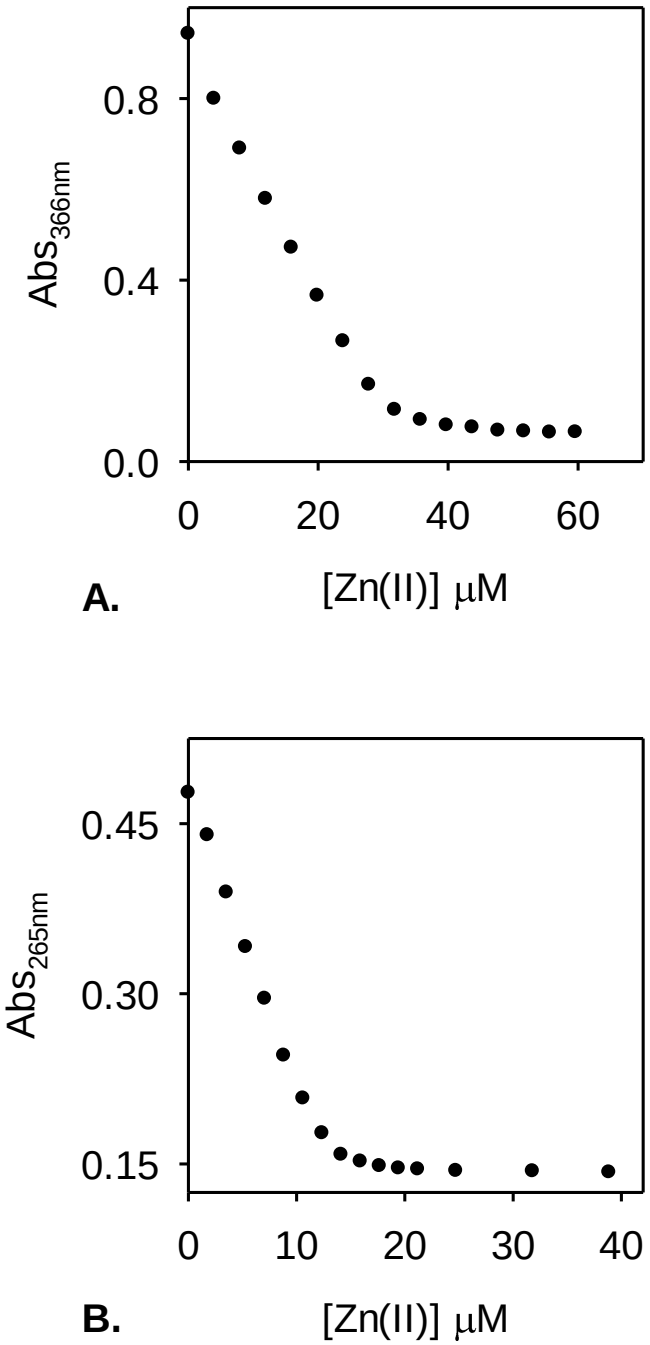


Figure 8. 7 A Sample of control titration of mag-fura-2 (42.5 μM) with $ZnCl_2$ monitored anaerobically by UV-vis spectrophotometer. System conditions: buffer C500, pH7.0. **B** Sample of control titration of quin-2 (13.4 μM) with $ZnCl_2$ monitored anaerobically by UV-vis spectrophotometer. System conditions: buffer C500, pH 7.0.

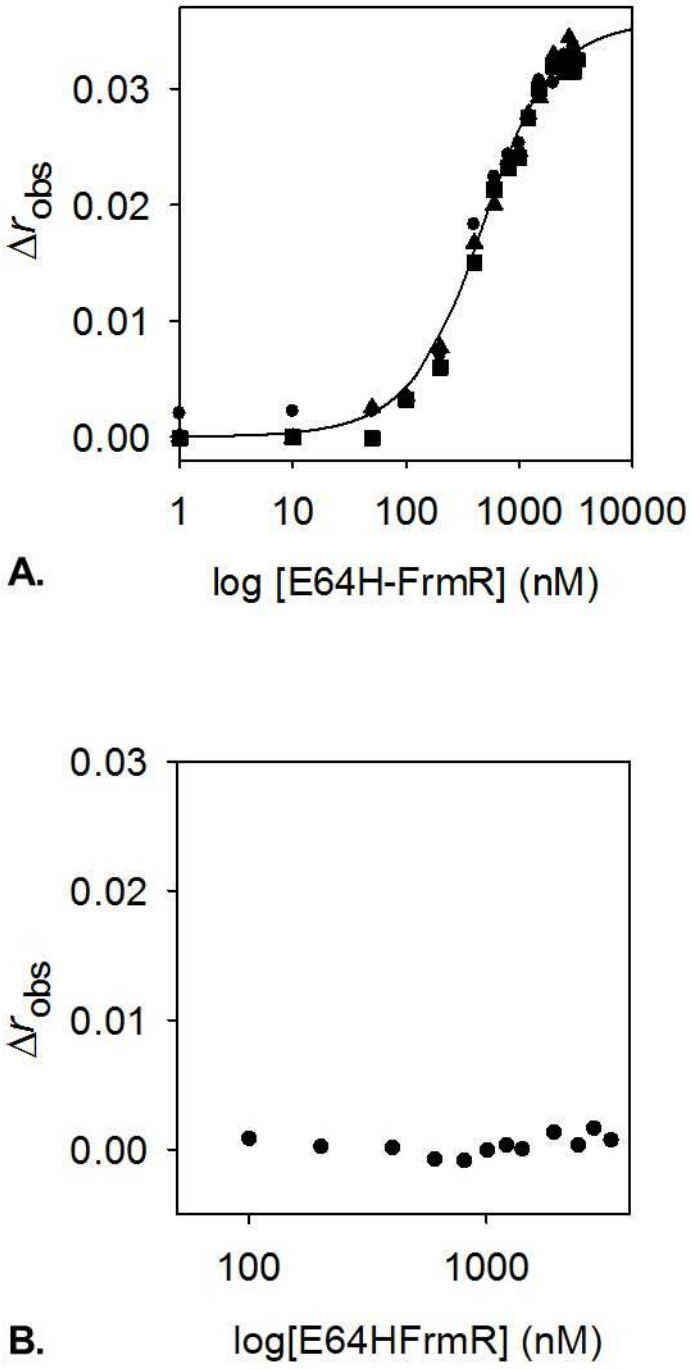


Figure 8. 8 A *frmRA*Pro (10 nM) was anaerobically titrated with a purified stock of E64HFrmR not displaying the shoulder at Abs ~ 300 nm (herein called E64HFrmR*) in the presence of 5 mM EDTA. DNA binding was monitored by fluorescence anisotropy. Solid line represents simulated curves produced from the average K_{DNA} determined across the experiment replicates shown. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 E64HFrmR* tetramer (non-dissociable):DNA stoichiometry. **B** *frmRA*Pro-T/A-mod (10 nM) was anaerobically titrated with E64HFrmR* (protein prep not displaying the shoulder at Abs ~ 300 nm) in the presence of 5 mM EDTA. Experiment was carried out by fluorescence anisotropy.

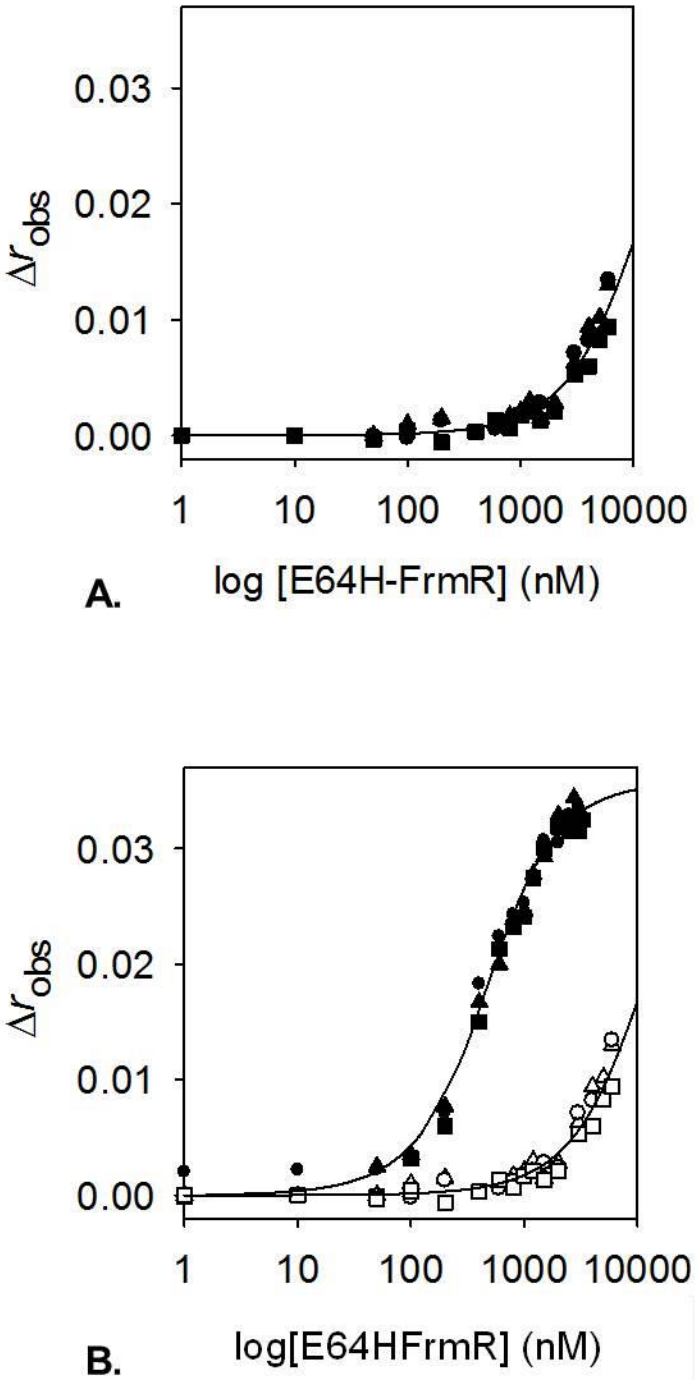


Figure 8. 9 A *frmRA*Pro (10 nM) was anaerobically titrated with E64HFrmR* in the presence of 5 µM ZnCl. The protein was incubated with 1.2 molar equivalents of $ZnCl_2$ and EDTA was omitted. **B** Comparison of the anisotropy change upon titration of *frmRA*Pro with apo-E64HFrmR* (filled circles) and Zn(II)-E64HFrmR* (empty circles). DNA binding was monitored by fluorescence anisotropy. Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 E64HFrmR* tetramer (non-dissociable):DNA stoichiometry. Solid line represents simulated curves produced from the average K_{DNA} determined across the experiment replicates shown.

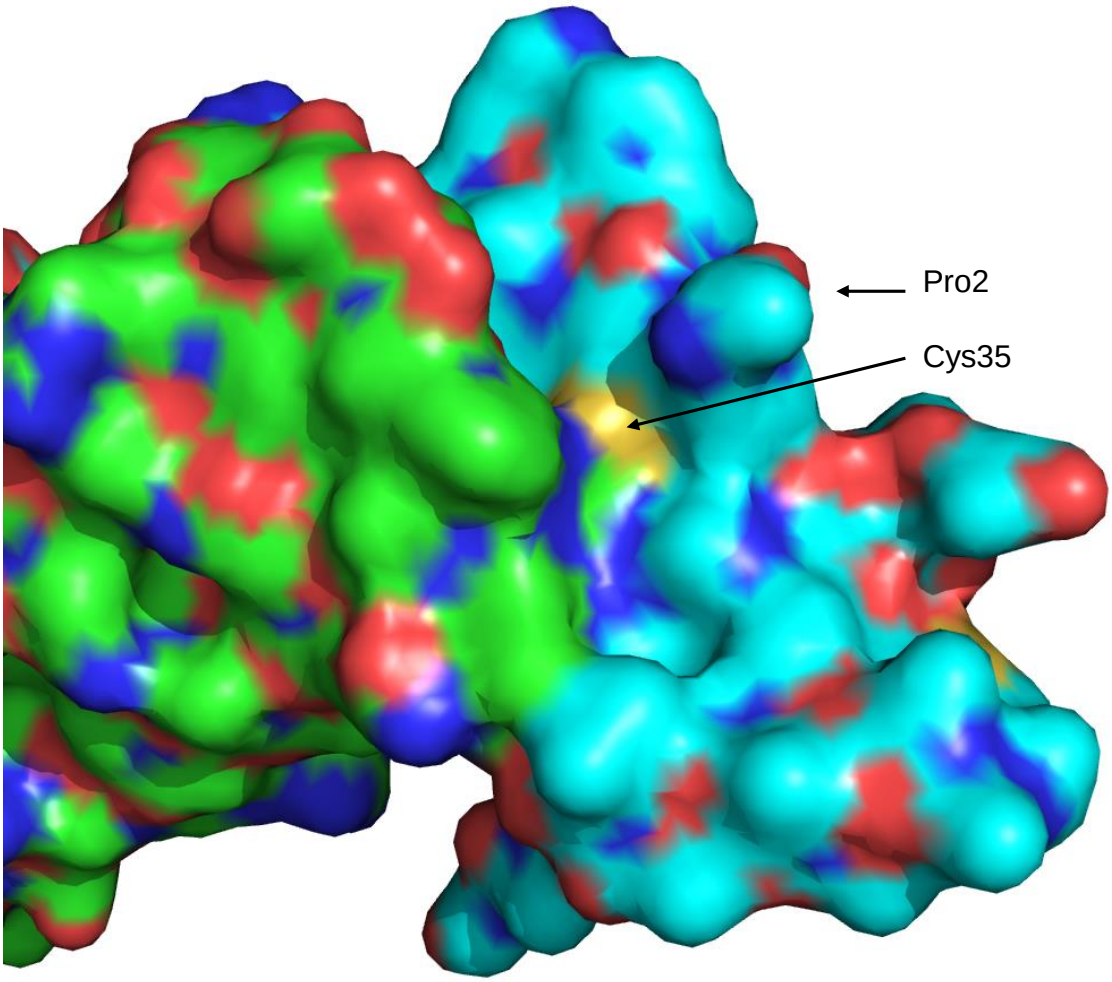


Figure 8. 10 Surface of effector-binding site in apo-E64HFrmR crystal structure. Color-coding refers to different chain (chain A in green, chain B in cyan) and atoms (oxygen in red, nitrogen in blue and sulfur in yellow). This is not an electrostatic representation as this would not allow distinction between oxygen and sulphur atoms

Chapter 9

References

Abraham L. O., Li Y. & Zamble D. B., The metal- and DNA-binding activities of *Helicobacter pylori* NikR. *J. Inorg. Biochem*. 100:1005-1014 (2006).

Adams P. D., Grosse-Kunstleve R. W., Hung L. W., Ioerger T. R., McCoy A. J., Moriarty N. W., Read R. J., Sacchettini J. C., Sauter N. K. & Terwilliger T. C., PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D. Biol. Crystallogr*. 58(11):1948-54 (2002).

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

Alberts B., Johnson A., Lewis J., Raff M., Roberts K. & Walter P., Molecular Biology of the Cell. 4th edition. *Garland Science,* New York, 11 (2002)

Altschul S., Gish W., Miller W., Myers E. & Lipman D., Basic local alignment search tool. *Molecular Biology* 215(3):403–410 (1990).

Ammendola S., Cerasi M. & Battistoni A., Deregulation of transition metals homeostasis is a key feature of cadmium toxicity in *Salmonella. Biometals.* 27(4):703-14 (2014).

An Y. J., Ahn B. E., Han A. R., Kim H. M., Chung K. M., Shin J. H., Cho Y. B., Roe J. H. & Cha S. S., Structural basis for the specialization of Nur, a nickel-specific Fur homolog, in metal sensing and DNA recognition. *Nucleic Acids Res.* 37(10):3442-51 (2009).

Anderson L. A., Palmer T., Price N. C., Bornemann S., Boxer D. H. & Pau R. N., Characterisation of the molybdenum-responsive ModE regulatory protein and its binding to the promoter region of the *modABCD* (molybdenum transport) operon of *Escherichia coli*. *Eur J Biochem.* 246(1):119-26 (1997).

Andoy N. M., Sarkar S. K., Wang Q., Panda D., Benítez J. J., Kalininskiy A. & Chen P., Single-molecule study of metalloregulator CueR-DNA interactions using engineered Holliday junctions. *Biophys J*. 97(3):844-52 (2009).

Anfinsen C. B., Principles that govern the folding of protein chains. *Science* 181(4096):223-230 (1973).

Angeletti B., Waldron K. J., Freeman K. B., Bawagan H., Hussain I., Miller C. C., Lau K. F., Tennant M. E., Dennison C., Robinson N. J. & Dingwall C., BACE1 cytoplasmic domain interacts with the copper chaperone for superoxide dismutase-1 and binds copper. *J Biol Chem* 280(18):17930-7 (2005).

Anjem A. & Imlay J. A., Mononuclear iron enzymes are primary targets of hydrogen peroxide stress. *J. Biol. Chem.* 287, 15544–15556 (2012).

Anjem A., Varghese, S. & Imlay J. A., Manganese import is a key element of the OxyR response to hydrogen peroxide in *Escherichia coli*. *Mol. Microbiol.* 72, 844–858 (2009)

Anthony C. The Biochemistry of Methylotrophs. *Academic Press*, New York (1982).

Arnott S., Chandrasekaran R., Birdsall D. L., Leslie A. G. & Ratliff R. L. Left-handed DNA helices. *Nature* 283(5749):743-5 (1980).

Arunkumar A. I., Campanello G. C. & Giedroc D. P., Solution structure of a paradigm ArsR family zinc sensor in the DNA-bound state. *Proc Natl Acad Sci U S A.* 106(43):18177-82 (2009).

Attwood M. M. & Quayle J. R. Formaldehyde as a central intermediary metabolite of methylotrophic metabolism, in: "Proceedings of the 4th International Symposiumon Microbial Growthon C1 compounds" (Crawford R. L. & Hanson R. S., eds.), *American society for microbiology*, Washington DC 315-323 (1984).

Bagg A. & Neilands J. B., Mapping of a mutation affecting regulation of iron uptake systems in *Escherichia coli* K-12 *J Bacteriol.* 161(1):450-3 (1985)

Bagg A. & Neilands J. B., Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochemistry*. 26(17):5471-7 (1987).

Baker J., Sengupta M., Jayaswal R. K., Morrissey J. A., The *Staphylococcus aureus* CsoR regulates both chromosomal and plasmid-encoded copper resistance mechanisms. *Environ Microbiol.*13(9):2495-507 (2011).

Ballatori N., Glutathione mercaptides as transport forms of metals. *Adv Pharmacol.* 27:271-98 (1994).

Barber R. D. & Donohue T. J., Pathways for transcriptional activation of a glutathione-dependent formaldehyde dehydrogenase gene. *J Mol Biol*. 280:775–784 (1998).

Barry A. N., Otoikhian A., Bhatt S., Shinde U., Tsivkovskii Ri, Blackburn N. J. & Lutsenko S., The luminal loop Met672-Pro707 of copper-transporting ATPase ATP7A binds metals and facilitates copper release from the intramembrane sites. *J. Biol. Chem.* 286:26585-26594 (2011).

Basham B., Schroth G. P. & Ho P. S. An A-DNA triplet code: thermodynamic rules for predicting A- and B-DNA *Proc Natl Acad Sci U S A.* 92(14):6464-8 (1995).

Battye T. G. G., Kontogiannis L., Johnson O., Powell H. R. & Leslie A. G. W., iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta crystallographica Section D, Biol crystal* 67:271-81 (2011).

Beer A., Bestimmung der Absorption des rothen Lichts in farbigen Flüssigkeiten [Determination of the absorption of red light in colored liquids], *Annalen der Physik und Chemie*, 86:78–88 (1852).

Bell R. P., The reversible hydration of carbonyl compounds. *Advan. Phys. Org. Chem*. 4:1-29 (1966).

Bell R. P., & Evans J. E., Kinetics of the dehydration of methylene glycol in aqueous solution. *Proc. Roy. Soc.*, Series A 291:297-323 (1966).

Betts M. J. & Russell R. B., Amino acid properties and consequences of substitutions; in "Bioinformatics for geneticists" (Barnes M. R. & Gray I. C., eds.) *John Wiley & Sons, Ltd*. Chapter 14 (2003).

Binet M. R. & Poole R. K., Cd(II), Pb(II) and Zn(II) ions regulate expression of the metal-transporting P-type ATPase ZntA in *Escherichia coli*. *FEBS Lett.* 473(1):67-70 (2000).

Black C. B. & Cowan J.A., Magnesium-dependent enzymes in nucleic acid biochemistry, in "The Biological Chemistry of Magnesium". (Cowan J. A. ed) *New York: VCH* (1995).

Blaha, D., Arous, S., Bleriot, C., Dorel, C., Mandrand-Berthelot, M. A. & Rodrigue, A., The *Escherichia coli* metallo-regulator RcnR represses *rcnA* and *rcnR* transcription through binding on a shared operator site: Insights into regulatory specificity towards nickel and cobalt. *Biochimie,* 93(3), 434-439 (2011).

Bloom S. L. & Zamble D. B. Metal-Selective DNA-Binding Response of *Escherichia coli* NikR. *Biochemistry* 43:10029-10038 (2004).

Bragg W. L., The diffraction of short electromagnetic waves by a crystal. *Proceedings of the Cambridge Philosophical Society* (17):43-57 (1913).

Branden C. I., Jornvall H., Eklund H. & Furungren B., Alcohol dehydrogenase, in "The Enzymes" (Boyer P. D., ed.) *Academic Press*, New York. 104-190 (1975).

Branden C. & Tooze J., *Intro. to protein structure, Zanichelli (*Branden C. & Tooze J., eds) 2nd edition (2001).

Bryngelson J. D., Onuchic J. N., Socci N. D. & Wolynes P. G., Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins.* 21(3):167-95 (1995).

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

Brodolin K., Protein-DNA crosslinking with formaldehyde in vitro, in DNA-protein interactions: A practical approach (Travers A. & Buckle M., eds), *Oxford University Press Inc.*, New York. Chapter 10 (2010).

Broennimann C. H., Eikenberry E. F., Henrich B., Horisberger R., Hülsen G., Pohl E., Schmitt B., Schulze-Briese C., Suzuki M., Tomizaki T., Toyokawa H. & Wagner A. *J. Synchrotron Rad.*13:120-130 (2006).

Brouwer M. & Brouwer-Hoexum T., Glutathione-mediated transfer of copper(I) into American lobster apohemocyanin. *Biochemistry* 31(16):4096-102 (1992).

Brown J. L., The N-terminal region of soluble proteins from procaryotes and eucaryotes. *Biochim. Biophys. Acta* 221:480– 488 (1970).

Brown N. L., Stoyanov J. V., Kidd S. P. & Hobman, J. L. The MerR family of transcriptional regulators. *FEMS Microbiology Reviews.* 27:145-63 (2003).

Bostian K. A. & Betts G. F., Kinetics and reaction mechanism of potassium-activated aldehyde dehydrogenase from *Saccharomyces cerevisiae*. *Biochem J*. 173(3):787-798 (1978).

Boyer P. D., Lardy H. A. & Phillips P. H., The role of potassium in muscle phosphorylations. *J Biol Chem* 146:673–681 (1942).

Boyer P., Pyruvate Kinase, *Enzyme* 6, 95 (1962).

Boyer E., Bergevin I., Malo D., Gros P. & Cellier M. F. M., Acquisition of Mn(II) in addition to Fe(II) is required for full virulence of Salmonella enterica serovar typhimurium. Infect. Immun. 70(11):6032-42 (2002).

Bukiva A. N., Kuntamallappanavar G., Edwards J., Singh A. K., Shivakumar B. & Dopico A. M., An anlcohol-sensing site in the calcium- and voltage-gated, larged conductance potassium (BK) channel. *Proc Natl Acad Sci U S A.* 111(25):9313-8 (2014).

Busenlehner L. S., Apuy J. L., Giedroc D. P., Characterization of a metalloregulatory bismuth(III) site in *Staphylococcus aureus* pI258 CadC repressor. *J Biol Inorg Chem.* 7(4-5):551-9 (2002a).

Busenlehner L. S., Weng T., C., Penner-Hahn J. E., Giedroc D. P., Elucidation of primary (alpha(3)N) and vestigial (alpha(5)) heavy metal-binding sites in *Staphylococcus aureus* pI258 CadC: evolutionary implications for metal ion selectivity of ArsR/SmtB metal sensor proteins. *J Mol Biol.* 319(3):685-701. (2002b).

Busenlehner L. S., Pennella M. A. & Giedroc D. P., The SmtB/ArsR family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance. *FEMS Microbiol Rev.* 27(2-3):131-43 (2003).

Butcher J., Sarvan S., Brunzelle J. S., Couture J. F. & Stintzi A., Structure and regulon of *Campylobacter jejuni* ferric uptake regulator Fur define apo-Fur regulation *Proc Natl Acad Sci U S A* 109(25):10047-52 (2012).

Calderwood S. B. & Mekalanos J. J., Iron regulation of Shiga-like toxin expression in *Escherichia coli* is mediated by the fur locus. *J Bacteriol.* 169(10):4759-64 (1987).

Calderwood S. B. & Mekalanos J. J., Confirmation of the Fur operator site by insertion of a synthetic oligonucleotide into an operon fusion plasmid. *J Bacteriol.* 170(2):1015-7 (1988).

Canfield D. E., Glazer A. N. & Falkowski P. G. The evolution and future of Earth's nitrogen cycle *Science*. 330(6001):192-6. doi: 10.1126/science.1186120. (2010).

Carlier P., Hannachi H. & Mouvier G., The chemistry of carbonyl compounds in the atmosphere. *Atmos. Environ*. 20: 2079-2099 (1986).

Carrington P. E., Chivers P. T., Al-Mjeni F., Sauer R. T. & Maroney M. J., Nickel coordination is regulated by the DNA-bound state of NikR. *Nat Struct Biol*., 10(2):126-30 (2003).

Cavet J. S., Meng W., Pennella M. A., Appelhoff R. J., Giedroc D. P. & Robinson N. J., A nickel-cobalt-sensing ArsR-SmtB family repressor. Contributions of cytosol and effector binding sites to metal selectivity. *J. Biol. Chem*., 277(41):38441-38448 (2002).

Chang F. M. J., Coyne H. J., Cubillas C., Vinuesa P., Fang X., Ma Z., Ma D., Helmann J. D., García-de los Santos A., Wang Y. X., Dann C. E. & Giedroc D. P., Cu(I)-mediated allosteric switching in a copper-sensing operon repressor (CsoR) *J Biol Chem.* 289(27):19204-17 (2014).

Changela A., Chen K., Xue Y., Holschen J., Outten C. E., O'Halloran T. V. & Mondragon A., Molecular basis of metal-ion selectivity and zeptomolar sensitivity by CueR. *Science* 301:1383-1387 (2003).

Chaplin A. K., Tan B. G., Vijgenboom E. & Worrall J. A., Copper trafficking in the CsoR regulon of Streptomyces lividans. *Metallomics* 7(1):145-55 (2015).

Checa S. K., Espariz M., Audero M. E., Botta P. E., Spinelli S. V. & Soncini F. C. Bacterial sensing of and resistance to gold salts. *Mol Microbiol.* 63(5):1307-18 (2007).

Chen X., Chu M. & Giedroc D. P., Spectroscopic characterization of Co(II)-, Ni(II)-, and Cd(II)- substituted wild-type and non-native retroviral-type zinc finger peptides. *J Biol Chem* 5:93-101 (2000).

Chen P. R. & He C. Selective recognition of metal ions by metalloregulatory proteins. *Curr Opin Chem Biol.* 12(2):214-21 (2008).

Chen P., Keller A. M., Joshi C. P., Martell D. J., Andov N. M., Benitez J. J., Chen T. Y., Santiago A. G. & Yang F., Single-molecule dynamics and mechanisms of metalloregulators and metallochaperones. *Biochemistry* 52:7170-7183 (2013).

Chen T. Y., Santiago A. G., Jung W., Krzeminski L., Yang Feng, Martell D. J., Helmann J. D. & Chen Peng, Concentration- and chromosome-organization-dependent regulator unbinding from DNA for transcription regulation in living cells. *Nat. Commun*. 6:7445 (2015).

Chistoserdova L., Gomelsky L., Vorholt J. A., Gomelsky M., Tsygankov Y. D. & Lidstrom M. E., Analysis of two formaldehyde oxidation pathways in *Methylobacillus flagellates* KT, a ribulose monophosphate cycle methylotroph. *Micobiology* 146(1):233-8 (2000).

Chistoserdova L., Kalyuzhnaya M. G. & Lidstrom M. E., The expanding world of methylotrophic metabolism *Annu Rev Microbiol.* 63:477-99 (2009).

Chivers P. T. & Sauer R. T., NikR is a ribbon-helix-helix DNA-binding protein. *Protein Sci.* 8:2494–2500 (1999).

Chivers P. T. & Sauer R. T., Regulation of high affinity nickel uptake in bacteria. Ni(II)-dependent interaction of NikR with wild-type and mutant operator sites. *J Biol Chem*. 275(26):19735-41 (2000).

Chivers P. T. & Sauer R. T., NikR Repressor: High-Affinity Nickel Binding to the C-Terminal Domain Regulates Binding to Operator DNA. *Chem. Biol.* 9:1141–1148 (2002).

Cho Y., Gorina S., Jeffrey P. D. & Pavletich N. P., Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science* 265(5170):346-55 (1994).

Choudhury S., Dance I. G., Guerney P. J. & Rae A. D., *Inorg Chim Acta* 70:227 (1983).

Ciccioli P., Brancaleoni E., Frattoni M., Cecinato A. & Brachetti A., Ubiquitous occurrence of semi-volatile carbonyl compounds in tropospheric samples and their possible source. *Atmos. Environ*. 27A:1891-1901 (1993).

Ciriolo M. R., Desideri A., Paci M. & Rotilio G., Reconstitution of Cu,Zn-superoxide dismutase by the Cu(I).glutathione complex. *J Biol Chem.* 265(19):11030-4 (1990).

Coburn B., Grassi G. A. & Finlay B. B., *Salmonella*, the host and disease; a brief review. *Immunol Cell Biol*. 85(2):112-8 (2007).

Coleman P. l., Iweibo I. & Weiner H., Role of zinc in horse liver alcohol dehydrogenase. Influence on structure and conformational changes. *Biochemistry* 11(6):1010-18 (1972).

Conrady D. G., Brescia C. C., Horii K., Weiss A. A., Hassett D. J. & Herr A. B., A zinc-dependent adhesion module is responsible for intercellular adhesion in staphylococcal biofilms. *Proc Natl Acad Sci U S A.* 9;105(49):19456-61(2008).

Corbett D., Schuler S., Glenn S., Andrew P. W., Cavet J. S. & Roberts I. S., The combined actions of the copper-responsive repressor CsoR and copper-metallochaperone CopZ modulate CopA-mediated copper efflux in the intracellular pathogen *Listeria monocytogenes*. *Mol Microbiol.* 81(2):457-72 (2011).

Cowan J. A., Introduction to the biological chemistry of magnesium, in "The Biological Chemistry of Magnesium". (Cowan J. A. ed) *New York: VCH* (1995).

Cowan J. A., Structural and catalytic chemistry of magnesium-dependent enzymes. *BioMetals* 15(3):225–235 (2002).

Crane-Robinson C., Dragan A. I., Privalov P. L. The extended arms of DNA-binding domains: a tale of tails *Trends Biochem Sci.* 31(10):547-52 (2006).

Creighton T. E., Review article on protein folding. *Biochem J.* 270:1-16 (1990).

Crothers D. M. & Shakked Z. DNA bending by adenine-thymine tracts. In "Oxford Handbook of Nucleic acid structure" (S. Neidle, ed.) *Oxford University Press*, Oxford 455-470 (1999).

Dainty S. J., Patterson C. J., Waldron K. J. & Robinson N. J., Interaction between cyanobacterial copper chaperone Atx1 and zinc homeostasis. *J. Biol. Inorg. Chem*. 15: 77-85 (2010).

De Lorenzo V., Wee S., Herrero M. & Neilands J. B. Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (fur) repressor. *J Bacteriol* 169:2624-2630 (1987).

De Lorenzo V., Herrero M., Giovanninini F. & Neilands J. B., Fur (ferric uptake regulation) protein and CAP (catabolite-activator protein) modulate transcription of fur gene in *Escherichia coli*. *Eur J Biochem.* 173(3):537-46 (1988).

Denby K. J., Rolfe M. D., Crick E., Sanguinetti G., Poole R. K. & Green J. Adaptation of anaerobic cultures of *Escherichia coli* K-12 in response to environmental trimethylamine-N-oxide. *Environ Microbiol.* (7):2477-91. (2015).

De Pina K., Desjardin V., Mandrand-Berthelot M. A., Giordano G. & Wu L. F. Isolation and characterization of the nikR gene encoding a nickel-responsive regulator in *Escherichia coli J Bacteriol.*181(2):670-4 (1999).

Dessau M. A. & Modis Y., Protein crystallization for X-ray crystallography. *J Vis Exp.* (47):2285 (2011).

Dian C., Vitale S., Leonard G. A., Bahlawane C., Fauquant C., Leduc D., Muller C., de Reuse H., Michaud-Soret I. & Terradot L., The structure of the *Helicobacter pylori* ferric uptake regulator Fur reveals three functional metal binding sites *Mol Microbiol.* 79(5):1260-75 (2011).

Di Cera E., A structural perspective on enzymes activated by monovalent cations. *J Biol Chem* 281:1305–1308 (2006).

Dijkhuizen, L., Levering P. R. & De Vries G. E., The physiology and biochemistry of aerobic methanol-utilizing gram negative and gram positive bacteria, in "Methane and methanol utilizers" (Murrell J. C. & Dalton, eds.), *Plenum Press*, NewYork (149-181) (1992).

Dudev T., Lin Y. L., Dudev M. & Lim C., First-second shell interactions in metal binding sites in proteins: a PDB survey and DFT/CDM calculations. *J Am Chem Soc* 12;125(10):3168-80 (2003).

Dudev T. & Lim C.,, Competition between protein ligands and cytoplasmic inorganic anions for the metal cation: a DFT/CDM study *J Am Chem Soc.* 128(32):10541-8. (2006).

Dudev T. & Lim C., Competition among metal ions for protein binding sites: determinants of metal ion selectivity in proteins. Chem. Rev. 114(1):538-556 (2014).

Dunn M. F. & Hutchison J. S., Roles of Zinc Ion and Reduced Coenzyme in the Formation of a Transient Chemical Intermediate During the Equine Liver Alcohol Dehydrogenase Catalyzed Reduction of an Aromatic Aldehyde. *Biochemistry* 12(24):4882-92 (1973).

Dwarakanath S., Chaplin A. K., Hough M. A., Rigali S., Vijgenboom E. & Worrall J. A. Response to copper stress in *Streptomyces lividans* extends beyond genes under direct control of a copper-sensitive operon repressor protein (CsoR). *J Biol Chem.* 18;287(21):17833-47 (2012).

Dwyer D. S. & Bradley R. J., Chemical properties of alcohols and their protein binding sites. Cell Mol life Sci. 57(2):265-75 (2000).

Eiamphungporn W., Soonsanga S., Lee J. W. & Helmann J. D., Oxidation of a single active site suffices for the functional inactivation of the dimeric *Bacillus subtilis* OhrR repressor *in vitro*. *Nucleic Acids Res*. 37(4):1174-81 (2009).

Eiberg H. & Mohr J., Identity of the polymorphisms for esterase D and S-formylglutathione hydrolase in red blood cells *Hum Genet.* 74(2):174-5 (1986).

Eicken C., Pennella M. A., Chen X., Koshlap K. M., VanZile M. L., Sacchettini J. C. & Giedroc D. P. A, Metal-ligand-mediated intersubunit allosteric switch in related SmtB/ArsR zinc sensor proteins. *J Mol Biol.* 333(4):683-95 (2003).

El Yazal J., Roe R. R. & Yuan-Ping Pang, Zinc's effect on proton transfer between imidazole and acetate predicted by ab initio calculations. *J Phys Chem B.* 104 (28)6662-6667 (2000).

Emsley P. & Cowtan K., Coot: model-building tools for molecular graphics. Acta Crystallographica D60, 2126-2132 (2004).

Emsley P., Lohkamp B., Scott W. G. & Cowtan K., Features and Development of Coot. *Acta Crystallographica D66*, 486-501 (2010).

Endo G., Silver S., CadC, the transcriptional regulatory protein of the cadmium resistance system of *Staphylococcus aureus* plasmid pI258. *J Bacteriol.* 177(15):4437-41 (1995).

Ernst J. F., Bennett R. L. & Rothfield L. I. Constitutive expression of the iron-enterochelin and ferrichrome uptake systems in a mutant strain of *Salmonella typhimurium*. *J Bacteriol.* 135(3):928-34 (1978).

Escolar L., De Lorenzo V., Pérez-Martín J. Metalloregulation in vitro of the aerobactin promoter of Escherichia coli by the Fur (ferric uptake regulation) protein. *Mol Microbiol.* 26(4):799-808 (1997).

Escolar L., Pérez-Martín J., De Lorenzo V. Coordinated repression *in vitro* of the divergent fepA-fes promoters of *Escherichia coli* by the iron uptake regulation (Fur) protein. *J Bacteriol.* 180(9):2579-82. (1998).

Espariz M., Checa S. K., Audero M. E., Pontel L. B. & Soncini F. C., Dissecting the *Salmonella* response to copper. *Microbiology* 153(9):2989-97 (2007).

Evans P., Scaling and assessment of data quality. *Acta crystallographica*. *Section D, Biol. crystal.* 62:72-82 (2006).

Farooqui A. A., Purification of enzymes by heparin-sepharose affinity chromatography. *J. Chrom A* 184(3):335-345 (1980).

Fenton H. J. H. *Oxidation of tartaric acid in presence of iron*. *J. Chem. Soc. Transactions 65*:899–910 *(1894).*

Ferenci T., Strøm T. & Quayle J. R., Purification and properties of 3-hexulose phosphate synthase and phospho-3-hexuloisomerase from *Methylococcus capsulatus. Biochem J.* 144(3):477-86 (1974).

Festa R. A., Jones M. B., Butler-Wu S., Sinsimer D., Gerads R., Bishai W. R., Peterson S. N. & Darwin K. H. A novel copper-responsive regulon in *Mycobacterium tuberculosis*. *Mol Microbiol.* 79(1):133-48 (2011).

Foster A. W., Characterisation of a cyanobacterial nickel sensor and a thermodynamic model of metal sensing. PhD thesis, Newcastle University (2012).

Foster A. W., Patterson C. J., Pernil R., Hess C. R. & Robinson N. J., Cytosolic Ni(II) sensor in cyanobacterium: nickel detection follows nickel affinity across four families of metal sensors *J Biol Chem.* 6;287(15):12142-51 (2012).

Foster A. W., Pernil R., Patterson C. J. & Robinson N. J., Metal specificity of cyanobacterial nickel-responsive repressor InrS: cells mantain zinc and copper below the detection threshold for InrS. *Mol. Microbiol*. 92(4):797-812 (2014a).

Foster A. W., Osman D. & Robinson N. J., Metal preferences and metallation. *J Biol Chem.* 289(41):28095-103 (2014b).

Frantz B., O'Halloran T. V., DNA distortion accompanies transcriptional activation by the metal-responsive gene-regulatory protein MerR. *Biochemistry* 29:4747-4751 (1990).

Frausto da Silva J. J. & Williams R. J. P., Zinc: Lewis acid catalysis and regulation, in "The biological chemistry of the elements: the inorganic chemistry of life", *OUP Oxford* (2001).

Freedman J. H., Ciriolo M. R. & Peisach J., The role of glutathione in copper metabolism and toxicity. *J Biol Chem.* 264(10):5598-605 (1989).

Freeman W. H., DNA can assume a variety of structural forms; in "Biochemistry". 5th edition (Berg J. M., Tymoczko J. L., Stryer L., eds.), *NCBI* New York (2002).

French S. & Wilson K. On the treatment of negative intensity observations. *Acta crystallographica Section A* 34:517-525 (1978).

Frottin F., Martinez A., Peynot P., Mitra S., Giglione C. & Meinnel T., The Proteomics of N-terminal Methionine Cleavage. *Molecular and Cellular Proteomics* 5:2336-2349 (2006).

Gaballa A., Newton G. L., Antelmann H., Parsonage D., Upton H., Rawat M., Claiborne A., Fahey R. C. & Helmann J. D., Biosynthesis and functions of bacillithiol, a major low-molecular weight thiol in Bacilli. *Proc Natl Acad Sci USA* 107(14):6482-6 (2010).

Garcia-Olalla C. & Garrido-Pertierra A., On the role of magnesium in the reaction of the pyruvate kinase from *Salmonella typhimurium. Z Naturforsch C.* 41(11-12):1018-22 (1986).

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

Garvie C. W. & Wolberger C. Recognition of specific DNA sequences *Mol Cell.* 8(5):937-46 (2001).

Gasteiger E., Gattiker A., Hoogland C., Ivanyi I., Appel R. D. & Bairoch A., ExPASy: the proteomic server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31:3784-3788 (2003).

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M. R., Appel R. D. & Bairoch A., *Protein Identification and Analysis Tools on the ExPASy Server* in "The Proteomics Protocols Handbook" (John M. Walker, ed), *Humana Press* 571-607 (2005).

Giedroc, D. P. & Arunkumar A.I. Metal sensor proteins: nature's metalloregulated allosteric switches. *Dalton Transactions* 3107-3120 (2007).

Gilston B. A., Wang S., Marcus M. D., Canalizo-Hernández M. A., Swindell E. P., Xue Y., Mondragón A. & O'Halloran T. V., Structural and mechanistic basis of zinc regulation across the *E. coli* Zur regulon. *PLoS Biol.* 12(11):e1001987 (2014).

Glaser F., Pupko T., Paz I., Bell R. E., Bechor-Shental D., Martz E., & Ben-Tal N., ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* 19:163-164 (2003).

Gold B., Deng H., Bryk R., Vargas D., Eliezer D., Roberts J., Jiang X. & Nathan C., Identification of a copper-binding metallothionein in pathogenic mycobacteria. *Nat Chem Biol*. 4(10):609-16 (2008).

Golynskiy M. V., Gunderson W. A. Hendrich M. P. & Cohen S. M., Metal binding studies and EPR spectroscopy of the manganese transport regulator MntR. *Biochemistry* 45:15359-15372 (2006).

Gonzalez C. F., Proudfoot M., Brown G., Korniyenko Y., Mori H., Savchenko A. V. & Yakunin A. F. Molecular basis of formaldehyde detoxification. Characterization of two S-formylglutathione hydrolases from *Escherichia coli*, FrmB and YeiG. *J Biol Chem.* 281(20):14514-22 (2006).

Grossoehme N. E. & Giedroc D. P., Energetics of allosteric negative coupling in the zinc sensor *S. aureus* CzrA. *J Am Chem Soc*. 131(49):17860-70 (2009).

Grossoehme N., Kehl-Fie T. E., Ma Z., Adams K. W., Cowart D. M., Scott R. A., Skaar E. P. & Giedroc D. P., Control of copper resistance and inorganic sulfur metabolism by paralogous regulators in *Staphylococcus aureus*. *J Biol Chem*. 286(15):13522-31 (2011).

Grossoehme N. E. & Giedroc D. P., Illuminating Allostery in Metal Sensing Transcriptional Regulators in "Spectroscopic Methods of Analysis: Methods and Protocols", Methods in Molecular Biology. (Wlodek M. & Bujalowski, eds.) *Springer Science + Business Media* New York. 165-190 (chapter 8) (2012).

Grunden A. M., Ray R. M., Rosentel J. K., Healy F. G. & Shanmugam K. T. Repression of the *Escherichia coli* modABCD (molybdate transport) operon by ModE. *J Bacteriol.* 178(3):735-44 (1996).

Guerra A. J. & Giedroc D. P., Metal site occupancy and allosteric switching in bacterial metal sensor proteins. *Arch. Biochem. Biophys*. 519:210-222 (2012).

Gutheil W. G., Holmquist B. & Vallee B. L. Purification, characterization, and partial sequence of the glutathione-dependent formaldehyde dehydrogenase from *Escherichia coli*: a class III alcohol dehydrogenase. *Biochemistry* 31(2):475-81 (1992).

Guzikevich-Guerstein G. & Shakked Z. A novel form of the DNA double helix imposed on the TATA-box by the TATA-binding protein *Nat Struct Biol.* 3(1):32-7 (1996).

Haenen, Guido R. M. M. & Aalt Bast. Glutathione Revisited: A Better Scavenger than Previously Thought. *Frontiers in Pharmacology* 5:260 (2014).

Hall H. K., Correlation of the Base Strengths of Amines. *Journal of the American Chemical Society* 79(20):5441 (1957).

Hall D. R., Gourley D. G., Duke E. M., Leonard G. A., Anderson L. A., Pau R. N., Boxer D. H. & Hunter W. N. Two crystal forms of ModE, the molybdate-dependent transcriptional regulator from *Escherichia coli*. *Acta Crystallogr D Biol Crystallogr.* 55(2):542-3 (1999).

Halliwell B. & Gutteridge J. M. C., Free Radicals in Biology and Medicine, *Clarendon Press Oxford* (1989).

Handler P., Bernheim M. L. C. and Klein J. R., The oxidative demethylation of sarcosine to glycine. *J Biol Chem* 138, 211 (1941).

Hantke K. Regulation of ferric iron transport in *Escherichia coli* K12: isolation of a constitutive mutant. *Mol Gen Genet.* 182(2):288-92 (1981).

Hantke K Mol Gen Genet. Cloning of the repressor protein gene of iron-regulated systems in *Escherichia coli* K12. *Mol Gen Genet.* 197(2):337-41 (1984). Haran T. E. & Mohanty U. The unique structure of A-tracts and intrinsic DNA bending *Q Rev Biophys.* 42(1):41-81 (2009).

Haran T. E. & Mohanty U., The unique structure of A-tracts and intrinsic DNA bending. *Q Rev Biophys* 42(1):41-81 (2009).

Harms N. & van Spanning R. J. C1 metabolism in Paracoccus denitrificans: genetics of *Paracoccus denitrificans*. *J Bioenerg Biomembr.* 23(2):187-210 (1991).

Harms N., Ras J., Reijnders W. N., van Spanning R. J., Stouthamer A. H. S-formylglutathione hydrolase of *Paracoccus denitrificans* is homologous to human esterase D: a universal pathway for formaldehyde detoxification? *J Bacteriol.* 178(21):6296-9 (1996).

Harms N., Reijnders W. N. M., Koning S. & van Spanning R. J. M., Two-component system that regulates methanol and formaldehyde oxidation in *Paracoccus denitrificans*. *J Bacteriol.* 183(2):664-670 (2001).

Harrison S. C. & Aggarwal A. K. DNA recognition by proteins with the helix-turn-helix motif *Annu Rev Biochem.* 59:933-69 (1990).

Harvie D. R., Andreini C., Cavallaro G., Meng W., Connolly B. A., Yoshida K., Fujita Y., Harwood C. R., Radford D. S., Tottey S., Cavet J. S. & Robinson N. J., Predicting metals sensed by ArsR-SmtB repressors: allosteric interference by a non-effector metal. *Mol Microbiol.* 59(4):1341-56 (2006).

Heck H. D., Casanova-Schmitz M., Dodd P. B., Schachter E. N., Witek T. J. & Tosun T., Formaldehyde (CH2O) concentrations in the blood of humans and Fischer-344 rats exposed to CH2O under controlled conditions *Am Ind Hyg Assoc J.* 46(1):1-3 (1985).

Heldwein E. E. & Brennan R. G., Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* 409:378-382 (2001).

Henderson R., Cryo-protection of protein crystals against radiation damage in electron and X-ray diffraction. *Proc. R. Soc. Lond.* B(241):6-8 (1990).

Herring C. D. & Blattner F. R., Global transcriptional effects of a suppressor tRNA and the inactivation of the regulator frmR. *J Bacteriol.*186(20):6714-20 (2004).

Hickman J. W., Witthuhn V. C. Jr, Dominguez M. & Donohue T. J., Positive and negative transcriptional regulators of glutathione-dependent formaldehyde metabolism. *J Bacteriol.* 186:7914–7925 (2004).

Higgins K. A., Carr C. E. & Maroney M. J., Specific metal recognition in nickel trafficking. *Biochemistry* 51:7816-7832 (2012a).

Higgins K. A., Chivers P. T. & Maroney M. J., Role of the N-terminus in determining metal-specific responses in the *E. coli* Ni- and Co-responsive metalloregulator, RcnR. *J Am Chem Soc*. 134(16):7081-93 (2012b).

Higgins K. A., Hu H. Q., Chivers P. T., Maroney M. J., Effects of select histidine to cysteine mutations on transcriptional regulation by *Escherichia coli* RcnR. *Biochemistry* 52(1):84-97 (2013).

Higgins K. A. & Giedroc D., Insights into protein allostery in the CsoR/RcnR family of transcriptional repressors. Chem Lett. 43(1):20-25 (2014).

Hildeman D. A., Mitchell T., Kappler J. & Marrack P., T cell apoptosis and reactive oxygen species. *J Clin Invest.* 111(5):575-81 (2003).

Hille R., The Mononuclear Molybdenum Enzymes. *Chem Rev.* 96(7):2757-2816 (1996).

Hizver J., Rozenberg H., Frolow F., Rabinovich D. & Shakked Z., DNA bending by an adenine--thymine tract and its role in gene regulation. *Proc Natl Acad Sci U S A* 98(15):8490-5 (2001).

Holm R. H., Kennepohl P. & Solomon E. I., Structural and functional aspects of metal sites in biology. *Chem. Rev*. 96:2239-2314 (1996).

Hong M., Fitzgerald M. X., Harper S., Luo C., Speicher D. W. & Marmorstein R. Structural basis for dimerization in DNA recognition by Gal4. *Structure* 16(7):1019-26 (2008).

Hope H., Frolow F., von Bohlen K., Makowski I., Kratky C., Halfon Y., Danz h., Webster P., Bartels K. S., Wittmann H. G. & Yonath A., Cryocrystallography of Ribosomal Particles. *Acta Cryst* B45, 190-199 (1989).

Hoskisson P. A. & Rigali S., Chapter 1: Variation in form and function the helix-turn-helix regulators of the GntR superfamily. *Adv Appl Microbiol.* 69:1-22 (2009).

Huckle J. W., Morby A. P., Turner J. S. & Robinson N. J., Isolation of a prokaryotic metallothionein locus and analysis of transcriptional control by trace metal ions. *Mol Microbiol.* 7(2):177-87 (1993).

Humbert M. V., Rasia R. M., Checa S. K. & Soncini F. C., Protein mixer that promote operator selectivity among paralog MerR monovalent metal-ion regulators. *J Biol Chem*. 288:20510-20519 (2013).

Huyen N. T. T., Eiamphungporn W., Mäder U., Liebeke M., Lalk M., Hecker M., Helmann J. D. & Antelmann H., Genome-wide responses to carbonyl electrophiles in *Bacillus subtilis*: control of the thiol-dependent formaldehyde dehydrogenase AdhA and cysteine proteinase YraA by the MerR-family regulator YraB (AdhR). *Mol Microbiol.* 71(4):876-94 (2009).

Ibáñez M. M., Checa S. K. & Soncini F. C., A single serine residue determines selectivity to monovalent metal ions in metalloregulators of the MerR family *J Bacteriol.* 197(9):1606-13 (2015).

Imlay J. A. & Linn S., DNA damage and oxygen radical toxicity. *Science* 240(4857):1302-9 (1988).

Inagi R., Kumagai T., Fujita T. & Nangaku M., The role of glyoxalase system in renal hypoxia. *Adv. Exp. Med. Biol.* 662:49–55 (2010).

Irving H. & Williams R. J. P., Order of stability of metal complexes. *Nature.* 746-747 (1948).

Irving H. & Williams R. J. P., The stability of transition-metal complexes. *J. Chem. Soc.* 3192-3210 (1953).

Iwig J. S., Rowe J. L. & Chivers P. T., Nickel homeostasis in *Escherichia coli* - the rcnR-rcnA efflux pathway and its linkage to NikR function. *Mol Microbiol.* 62(1):252-62 (2006).

Iwig J. S., Leitch S., Herbst R. W., Maroney M. J. & Chivers P. T., Ni(II) and Co(II) sensing by *Escherichia coli* RcnR. *J Am Chem Soc.*130(24):7592-606 (2008).

Iwig J. S. & Chivers P. T., DNA recognition and wrapping by *Escherichia coli* RcnR *J Mol Biol*. 393(2):514-26 (2009).

Jacob C., Giles G. I., Giles N. M. & Sies H., Sulfur and selenium: the role of oxidation state in protein structure and function. *Angew Chem* Int Ed Engl. 42(39):4742-58 (2003).

Jacquamet L., Traoré D. A., Ferrer J. L., Proux O., Testemale D., Hazemann J. L., Nazarenko E., El Ghazouani A., Caux-Thang C., Duarte V. & Latour J. M., Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding. *Mol Microbiol.* 73(1):20-31 (2009).

Jancarik J., Kim S. H., Sparse matrix sampling: a screening method for crystallization of proteins. *J Appl Cryst*. 23:409–411 (1991).

Janin J., Rodier F., Chakrabarti P. & Bahadur R. P., Macromolecular recognition in the Protein Data Bank *Acta Crystallogr D Biol Crystallogr.* 63(Pt 1):1-8 (2007).

Jauch R., Ng C. K., Narasimhan K., Kolatkar P. R., The crystal structure of the Sox4 HMG domain-DNA complex suggests a mechanism for positional interdependence in DNA recognition *Biochem J.* 443(1):39-47 (2012).

Jayaram B., Sharp K. A. & Honig B., The electrostatic potential of B-DNA *Biopolymers*28(5):975-93 (1989).

Jefferson J. R., Hunt J. B. & Ginsburg A., Characterization of indo-1 and quin-2 as spectroscopic probes for Zn^{2+}-protein interactions. *Anal. Biochem*. 187:328-336 (1990).

Jencks W. P., Catalysis in Chemistry and Enzymology. McGraw-Hill, New York 196-198 (1969).

Ji G. & Silver S., Reduction of arsenate to arsenite by the ArsC protein of the arsenic resistance operon of *Staphylococcus aureus* plasmid pI258. *Proc Natl Acad Sci U S A.* 89(20):9474-8 (1992).

Jörnvall H., Persson B. & Jeffery J., Characteristics of alcohol/polyol dehydrogenases. The zinc-containing long-chain alcohol dehydrogenases. *Eur J Biochem.* 167(2):195-201 (1987).

Joshi R., Passner J. M., Rohs R., Jain R., Sosinsky A., Crickmore M. A., Jacob V., Aggarwal A. K., Honig B. & Mann R. S., Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell* 131(3):530-43 (2007).

Joshi C. P., Panda D., Martell D. J., Andoy N. M., Chen T. Y., Gaballa A., Helmann J. D. & Chen P., Direct substitution and assisted dissociation pathways for turning off transcription by a MerR-family metalloregulator. *Proc Natl Acad Sci U S A.* 109(38):15121-6 (2012).

Kachmar F. & Boyer P. D., Kinetic analysis of enzyme reactions. II. The potassium activation and calcium inhibition of pyruvic phosphoferase. *J Biol Chem* 200:669–682 (1953).

Kalapos M. P., Methylglyoxal in living organisms: chemistry, biochemistry, toxicology and biological implications. *Toxicol Lett*. 110(3):145-75 (1999).

Kallen R. G. & Jencks W. P., The dissociation constants of tetrahydrofolic acid. *J Biol Chem.* 241(24):5845-5850 (1966).

Kallen R. G., The mechanism of reactions involving Schiff base intermediates. Thiazolidine formation from L-cysteine and formaldehyde. *J Am Chem Soc.*93(23):6236-48 (1971).

Kato N., Yurimoto H., Thauer R. K., The physiological role of the ribulose monophosphate pathway in bacteria and archaea. *Biosci Biotechnol Biochem.* 70(1):10-21 (2006).

Kelly S. M., Jess T. J. & Price N. C., How to study proteins by circular dichroism. *Biochim Biophys Acta* 1751(2):119-39 (2005).

Keren N., Aurora R. & Pakrasi H. B., Critical roles of bacterioferritins in iron storage and proliferation of cyanobacteria. Plant Physiol. 135(3):1666-73 (2004).

Kerns W. D., Pavkov K. L., Donofrio D. J., Gralla E. J., & Swenberg J. A., Carcinogenicity of Formaldehyde in Rats and Mice after Long-Term Inhalation Exposure1. *Cancer Research* 43: 4382-4392 (1983).

Khadem A. F., Wieczorek A. S., Pol A., Vuilleumier S., Harhangi H. R., Dunfield P. F., Kalyuzhnaya M. G., Murrell J. C., Francoijs K. J., Stunnenberg H. G., Stein L. Y., DiSpirito A. A., Semrau J. D., Lajus A., Medigue C., Klotz M. G., Jetten M. S. M. & Op den Camp H. J. M., Draft Genome Sequence of the Volcano-Inhabiting Thermoacidophilic Methanotroph *Methylacidiphilum fumariolicum* Strain SolV. *J Bacteriol.* 194:3729-3730 (2012).

Khan S., Brocklehurst K. R., Jones G. W. & Morby A. P., The functional analysis of directed amino-acid alterations in ZntR from *Escherichia coli. Biochem Biophys Res Commun* 299:438-445 (2002).

Kiefer L. L., Paterno S. A. & Fierke C. A. Hydrogen bond network in the metal binding site of carbonic anhydrase enhances zinc affinity and catalytic efficiency. *J Am Chem Soc* 117:6831-6837 (1995).

Kim J. L., Nikolov D. B. & Burley S. K., Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* 365(6446):520-7 (1993).

Kim Y., Geiger J. H., Hahn S. & Sigler P. B., Crystal structure of a yeast TBP/TATA-box complex. *Nature* 365(6446):512-20 (1993).

Kisker C., Schindelin H. & Rees D. C., Molybdenum-cofactor-containing enzymes: structure and mechanism. *Annu Rev Biochem.* 66:233-67 (1997).

Kosman D. J., Multicopper oxidases: a workshop on copper coordination chemistry, electron transfer, and metallophysiology. *J Biol Inorg Chem.* 15(1):15-28 (2010).

Kunkel G. R., Mehrabian M. & Martinson H. G., Contact-site cross-linking agents. *Mol Cell Biochem*. 34(1):3-13 (1981).

Kuppuraj G., Dudev M. & Lim C., Factors governing metal-ligand distances and coordination geometries of metal complexes. *J Phys Chem B.* 113(9):2952-60 (2009).

Kuzmic P., Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem*. 237:260–273 (1996).

Kwan C. Y. & Putney J. W. Jr., Uptake and intracellular sequestration of divalent cations in resting and methacholine-stimulated mouse lacrimal acinar cells. Dissociation by Sr(II) and Ba(II) of agonist-stimulated divalent cation entry from the refilling of the agonist-sensitive intracellular pool. *J Biol Chem* 265(2):678-84 (1990).

Lambert J. H., Photometria sive de mensura et gradibus luminis, colorum et umbrae [Photometry, or, On the measure and gradations of light, colors, and shade], *Eberhardt Klett* 391 (1760).

Lavery R. & Pullman B., The molecular electrostatic potential and steric accessibility of poly (dI.dC). Comparison with poly (dG.dC). *Nucleic Acids Res.* 9:7041–7051 (1981).

Law J. R., Molecular Basis of Bacterial Formaldehyde Sensing. PhD thesis, Manchester University (2012).

Lee J. W. & Helmann J. D., The PerR transcription factor senses H_2O_2 by metal-catalysed histidine oxidation. *Nature* 440(7082):363-7 (2006).

Lee Y. M. & Lim C., Factors controlling the reactivity of zinc finger cores. *J Am Chem Soc.* 133(22):8691-703 (2011).

Le Henaff P. *C. R., Acad. Sci Paris* 235:2706 (1961).

Leopold P. E., Montal M. & Onuchic J. N., Protein folding funnels: a kinetic approach to the sequence-structure relationship. *Proc Natl Acad Sci U S A.* 15;89(18):8721-5 (1992).

Lesburg C. A., Lloyd M. D., Cane D. E. & Christianson D. W., Crystallization and preliminary X-ray diffraction analysis of recombinant pentalenene synthase. *Protein Sci.* 4(11):2436-8 (1995). Leslie A., Powell H., Winter G., *et al*., Automation of the collection and processing of X-ray diffraction data – a generic approach. *Acta Crystallogr. D58*, 1924-1928.

Lever A. B. P., Inorganic electronic spectroscopy. *Elsevier*, Amsterdam, 2nd edition (1984).

Lewis M., Chang G., Horton N. C., Kercher M. A., Pace H. C., Schumacher M. A., Brennan R. G. & Lu P., Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271(5253):1247-54 (1996).

Li M. H., Jang J. H., Na H. K., Cha Y. N. & Surh Y. J., Carbon monoxide produced by heme oxygenase-1 in response to nitrosative stress induces expression of glutamate-cysteine ligase in PC12 cells via activation of phosphatidylinositol 3-kinase and Nrf2 signaling. *J Biol Chem.* 282(39):28577-86 (2007).

Li T., Jin Y., Vershon A. K. & Wolberger C., Crystal structure of the MATa1/MATalpha2 homeodomain heterodimer in complex with DNA containing an A-tract. *Nucleic Acids Res*. 26(24):5707-18 (1998).

Lienhard G. E. & Jencks W. P., Thiol addition to the carbonyl group. Equilibria and kinetics. *J Am Chem Soc.* 88(17):3982-94 (1966).

Lin C. S., Chao S. Y., Hammel M., Nix J. C., Tseng H. L., Tsou C. C., Fei C. H., Chiou H. S., Jeng U. S., Lin Y. S.,Chuang W. J., Wu J. J. & Wang S., Distinct structural features of the peroxide response regulator from group A *Streptococcus* drive DNA binding. *PLoS One* 9(2):e89027 (2014).

Lipari F., Dasch J. M. & Scruggs W. F., Aldehyde emissions from wood-burning fireplaces. *Environmental science and technology* 18:326–330 (1984).

Liu T., Ramesh A., Ma Z., Ward S. K., Zhang L., George G. N., Talaat A. M., Sacchettini J. C. & Giedroc D. P., CsoR is a novel *Mycobacterium tuberculosis* copper-sensing transcriptional regulator. *Nat Chem Biol.* 3(1):60-8 (2007).

Lisher J.P., Higgins K.A., Maroney M.J. & Giedroc D.P., Physical characterization of the manganese-sensing pneumococcal surface antigen repressor from *Streptococcus pneumonia*. *Biochem.* 52:7689-7701 (2013).

Lithgow J. K., Hayhurst E. J., Cohen G., Aharonowitz Y. & Foster S. J., Role of a cysteine synthase in *Staphylococcus aureus*. *J Bacteriol.* 186(6):1579-90 (2004).

Liu L., Hausladen A., Zeng M., Que l., Heitman J. & Stamler J. S., A metabolic enzyme for S-nitrosothiol conserved from bacteria to humans. *Nature* 410(6827):490-4 (2001).

Liu T., Ramesh A., Ma Z., Ward S. K., Zhang L., George G. N., Talaat A. M., Sacchettini J. C. & Giedroc D. P., CsoR is a novel *Mycobacterium tuberculosis* copper-sensing transcriptional regulator. *Nat Chem Biol*. 3(1):60-8 (2007).

Liu T., Chen X., Shokes J., Hemmingsen L., Scott R. A. & Giedroc D. P., A Cu(I)-sensing ArsR family metal sensor protein with a relaxed metal selectivity profile. *Biochemistry* 47(40):10564-10575 (2008).

Lloyd D. R., Phillips D. H. & Carmichael P. L. Generation of putative intrastrand cross-links and strand breaks in DNA by transition metal ion-mediated oxygen radical attack. *Chem Res Toxicol.* 10(4):393-400 (1997).

López-Maury L., García-Domínguez M., Florencio F. J. & Reyes J. C., A two-component signal transduction system involved in nickel sensing in the cyanobacterium *Synechocystis* sp. PCC 6803. *Mol Microbiol.* 43(1):247-56 (2002).

Lowy F. D., *Staphylococcus aureus* infections. *N Engl J Med.* 339(8):520-32 (1998).

Lu X. J., Shakked Z. & Olson W. K., A-form conformational motifs in ligand-bound DNA structures. *J Mol Biol.* 300(4):819-40 (2000).

Luebke J. L., Arnold R. J. & Giedroc D. P., Selenite and tellurite form mixed seleno- and tellurotrisulfides with CstR from *Staphylococcus aureus*. *Metallomics.* 5(4):335-42 (2013).

Luebke J. L., Shen J., Bruce K. E., Kehl-Fie T. E., Peng H., Skaar E. P. & Giedroc D. P., The CsoR-like sulfurtransferase repressor (CstR) is a persulfide sensor in Staphylococcus aureus. *Mol Microbiol*. 94(6):1343-60(2014).

Luscombe N. M., Austin S. E., Berman H. M. & Thornton J. M., An overview of the structures of protein-DNA complexes. *Genome Biol.* 1(1): reviews001.1-001.37 (2000).

Lushchak V. I., Glutathione homeostasis and functions: potential targets for medical interventions. *J. Amino Acids* 1-26 (2012).

Ma Z., Cowart D. M., Scott R. A. & Giedroc D. P., Molecular insights into the metal selectivity of the copper(I)-sensing repressor CsoR from *Bacillus subtilis*. *Biochemistry* 48(15):3325-34 (2009a).

Ma Z., Cowart D. M., Ward B. P., Arnold R. J., DiMarchi R. D., Zhang L., George G. N., Scott R.A. & Giedroc D.P., Unnatural amino acid substitution as a probe of the allosteric coupling pathway in a mycobacterial Cu(I) sensor. *J. Am. Chem. Soc.* 131:18044-18045 (2009b).

Ma Z., Gabriel S. E. & Helmann J. D., Sequential binding and sensing of zinc by *Bacillus subtilis* ZuR. *Nucleic Acids Res*. 39:9130-9138 (2011).

Ma Z., Chandrangsu P., Helmann T. C., Romsang A., Gaballa A. & Helmann J. D., Bacillithiol is a major buffer of the labile zinc pool in *Bacillus subtilis*. *Mol Microbiol.*94(4):756-70 (2014).

MacKenzie C. G., Formation of formaldehyde and formate in the bio-oxidation of the methyl group. *J Biol Chem.* 186(1):351-68 (1950).

Maher P., The effects of stress and aging on glutathione metabolism. *Ageing Research Reviews*, 4(2):288–314 (2005).

Majowicz S. E., Musto J., Scallan E., Angulo F. J., Kirk M., O'Brien S. J., Jones T. F., Fazil A., Hoekstra R. M., The global burden of nontyphoidal *Salmonella* gastroenteritis. *Clin Infect Dis* 50(6):882-9 (2010).

Marathias V. M. & Bolton P. H., Structures of the potassium-saturated, 2:1, and intermediate, 1:1, forms of a quadruplex DNA. *Nucleic Acids Res*. 28(9):1969–1977 (2000).

Marx C. J., Chistoserdova L. & Lidstrom M. E. Formaldehyde-detoxifying role of the tetrahydromethanopterin-linked pathway in *Methylobacterium extorquens* AM1 *J Bacteriol.* 185(24):7160-8 (2003).

Marx C. J., Miller J. A., Chistoserdova L. & Lidstrom M. E., Multiple formaldehyde oxidation/detoxification pathways in *Burkholderia fungorum* LB400. *J Bacteriol* 186(7):2173-2178 (2004).

Matheson A. T., Yaguchi M. & Visentin L. P., The conservation of amino acids in the N-terminal position of ribosomal and cytosol proteins from E*scherichia coli, Bacillus stearothermophilus*, and *Halobacterium cutirubrum. Can. J. Biochem*. 53:1323-1327 (1975).

McCoy A. J., Grosse-Kunstleve R. W., Adams P. D., Winn M. D., Storoni L. C. & Read R. J., *Phaser* crystallographic software. J. Appl. Cryst. 40:658-674 (2007).

McGuire A. M., Cuthbert B. J., Ma Z., Grauer-Gray K. D., Brunjes Brophy M., Spear K. A., Soonsanga S., Kliegman J. I., Griner S. L., Helmann J. D. & Glasfeld A., Roles of the A and C sites in the manganese-specific activation of MntR. *Biochemistry*. 52(4):701-13 (2013).

McMurry J., Fundamentals of Organic Chemistry, *Cengage Learning.* 7th Edition (2011).

McPherson, *Protein Science* (10):418-422 (2001).

Meents A., Gutmann S., Wagner A. & Schulze-Briese C., Origin and temperature dependence of radiation damage in biological samples at cryogenic temperatures. *PNAS*, 107(3):1094-1099 (2010).

Méjean V., Iobbi-Nivol C., Lepelletier M., Giordano G., Chippaux M. & Pascal M.C., TMAO anaerobic respiration in *Escherichia coli:* involvement of the tor operon. *Mol Microbiol* 11: 1169–1179 (1994).

Mendel R. R. & Kruse T., Cell biology of molybdenum in plants and humans. *Biochim Biophys Acta* 1823(9):1568-79 (2012).

Michoud V., Kukui A., Camredon M., Colomb A., Borbon A., Miet K., Aumont B., Beekmann M., Durand-Jolibois R., Perrier S., Zapf P., Siour G., Ait-Helal W., Locoge N., Sauvage S., Afif C., Gros V., Furger M., Ancellet G. & Doussin J. F., Radical budget analysis in a suburban European site during the MEGAPOLI summer field campaign. *Atmos Chem Phys* 12(24):11951–11974 (2012).

Min H., Shane B. & Stokstad E. L., Identification of 10-formyltetrahydrofolate dehydrogenase-hydrolase as a major folate binding protein in liver cytosol. *Biochim Biophys Acta* 967(3):348-53 (1988).

Mitsui R., Kusano Y., Yurimoto H., Sakai Y., Kato N. & Tanaka M. Formaldehyde fixation contributes to detoxification for growth of a nonmethylotroph, *Burkholderia cepacia* TM1, on vanillic acid *Appl Environ Microbiol*. 69(10):6128-32 (2003).

Mitsui R., Omori M., Kitazawa H. & Tanaka M., Formaldehyde-limited cultivation of a newly isolated methylotrophic bacterium, *Methylobacterium sp*. MF1: enzymatic analysis related to C1 metabolism. *J Biosci Bioeng.* 99(1):18-22 (2005).

Monks P. S., Gas-phase radical chemistry in the troposphere *Chem. Soc. Rev*. 34:376–395 (2005).

Monod J., Wyman J. & Changeux J., On the Nature of Allosteric Transitions: A Plausible Model. *J. Mol. Biol.* 12, 88-118 (1965).

Moore S. J. & Warren M. J., The anaerobic biosynthesis of vitamin B12. *Biochem Soc Trans.* 40(3):581-6 (2012).

Morby A. P., Turner J. S., Huckle J. W. & Robinson N. J., SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene *smtA*: identification of a Zn inhibited DNA-protein complex. *Nucleic Acids Res.* 21(4):921-5 (1993).

Moulis J. M., Holmquist B. & Vallee B. L., Hydrophobic anion activation of human liver χχ alcohol dehydrogenase. *Biochemistry* 30, 5743–5749 (1991).

Munishkina L. A. & Fink A. L., Fluorescence as a method to reveal structures and membrane-interactions of amyloidogenic proteins. *Biochimica et Biophysica* Acta (BBA)-Biomembranes 1768(8):1862-1885 (2007).

Munkelt D., Grass G. & Nies D. H., The chromosomally encoded cation diffusion facilitator proteins DmeF and FieF from *Wautersia metallidurans* CH34 are transporters of broad metal specificity. *J Bacteriol.*186(23):8036-43 (2004).

Murzin A. G., Brenner S. E., Hubbard T., Chothia C., SCOP: a structural classification of proteins database for the investigation of sequences and structures *J Mol Biol.* 247(4):536-40 (1995).

Nelson H. C., Finch J. T., Luisi B. F. & Klug A., The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature* 330(6145):221-6 (1987).

Newberry K. J. & Brennan R. G., The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N-terminus. *J Biol. Chem.* 279:20356-20362 (2004).

Newman J., Expanding screening space through the use of alternative reservoirs in vapor-diffusion experiments. *Acta Cryst.* D61, 490-493 (2005a).

Newman J. *et al.*, 2005., Towards rationalization of crystallisation screening for small- to medium-sized laboratories: the PACT/JCSG+ strategy. *Acta Cryst.* D61:1426-1431 (2005b).

Newton G. L., Rawat M., La Clair J. J., Jothivasan V. K., Budiarto T., Hamilton C. J., Claiborne A., Helmann J. D. & Fahey R. C., Bacillithiol is an antioxidant thiol produced in Bacilli. *Nature Chem. Biol.* 5:625-627 (2009).

Nizet V., Understanding how leading bacterial pathogens subvert innate immunity to reveal novel therapeutic targets. *J Allergy Clin Immunol.* 120(1):13-22 (2007).

Nobre L. S., Al-Shahrour F., Dopazo J. & Saraiva L. M., Exploring the antimicrobial action of a carbon monoxide-releasing compound through whole-genome transcription profiling of *Escherichia coli*. *Microbiology* 155(3):813-24 (2009).

Neilan ds J. B., Siderophores. *Arch Biochem Biophys.* 302(1):1-3 (1993).

Ogushi S., Ando M. & Tsuru D., Formaldehyde dehydrogenase from *Pseudomonas putida*: a zinc metalloenzyme. *J Biochem.* 96(5):1587-91 (1984).

O'Halloran T. V., Frantz B., Shin M. K., Ralston D. M., Wright J. G., The MerR heavy metal receptor mediates positive activation in a topologically novel transcription complex. *Cell* 56:119-129 (1989).

Olson W. K., Gorin A. A., Lu X. J., Hock L. M. & Zhurkin V. B., DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc Natl Acad Sci U S A.* 95(19):11163-8 (1998).

Onuchic J. N., Wolynes P. G., Luthey-Schulten Z. & Socci N. D., Toward an outline of the topography of a realistic protein-folding funnel. *Proc Natl Acad Sci U S A.* 92(8):3626-30 (1995).

Onuchic J. N., Luthey-Schulten Z. & Wolynes P. G., Theory of protein folding: the energy landscape perspective. *Annu Rev Phys Chem.* 48:545-600 (1997).

Oppenheimer N. J., Henehan G. T., Huete-Pérez J. A. & Ito K., *P. putida* formaldehyde dehydrogenase. An alcohol dehydrogenase masquerading as an aldehyde dehydrogenase. *Adv Exp Med Biol.* 414:417-23 (1997).

Oria-Hernandez J., Cabrera N., Perez-Montfort R. & Ramirez-Silva L., Pyruvate kinase revised: the activating effect of K^+. *J Biol Chem*. 280(45):37924-9 (2005).

Osman D., Waldron K. J., Denton H., Taylor C. M., Grant A. J., Mastroeni P., Robinson N. J. & Cavet J. S., Copper homeostasis in *Salmonella* is atypical and copper-CueP is a major periplasmic metal complex. *J Biol Chem.* 285(33):25259-68 (2010).

Osman D., Patterson C. J., Bailey K., Fisher K., Robinson N. J., Rigby S. E. & Cavet J. S., The copper supply pathway to a *Salmonella* Cu,Zn-superoxide dismutase (SodCII) involves P(1B)-type ATPase copper efflux and periplasmic CueP. *Mol Microbiol*. 87(3):466-77 (2013).

Osman D., Piergentili C., Chen J., Chakrabarti B., Foster A. W., Lurie-Luke E., Huggins T. G. & Robinson N. J., Generating a metal-responsive transcriptional regulator to test what confers metal sensing in cells. *J. Biol. Chem.* 290(32):19806-22 (2015).

Otwinowski Z., Schevitz R. W., Zhang R. G., Lawson C. L., Joachimiak A., Marmorstein R. Q., Luisi B. F. & Sigler P. B., Crystal structure of trp repressor/operator complex at atomic resolution. *Nature* 335(6193):837 (1988).

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

Owczarzy R., Moreira B. G., You Y., Behlke M. A. & Walder J. A., Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry* 47:5336-5353 (2008).

Page R., Grzechnik S. K., Canaves J. M., *et al.*, Shotgun crystallization strategy for structural genomics: an optimized two-tiered crystallization screen against the *Thermotoga maritima* proteome. *Acta Crystallogr*. D59(6):1028-1037 (2003).

Parsonage D., Newton G. L., Holder R. C., Wallace B. D., Paige C., Hamilton C. J., Dos Santos P. C., Redinbo M. R., Reid S. D. & Claiborne A., Characterization of the N-acetyl-α-D-glucosaminyl L-malate synthase and deacetylase functions for bacillithiol biosynthesis in *Bacillus anthracis. Biochem.* 49:8398-8414 (2010).

Patterson C. J., Pernil R., Dainty S. J., Chakrabarti B., Henry C. E., Money V. A., Foster A. W. & Robinson N. J., Co(II)-detection does not follow $K_{Co(II)}$ gradient: Channelling in Co(II)-sensing. *Metallomics* 5: 352-362 (2013).

Pavletich N. P. & Pabo C. O., Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A *Science* 252(5007):809-17 (1991).

Pecqueur L., D'Autréaux B., Dupuy J., Nicolet Y., Jacquamet L., Brutscher B., Michaud-Soret I. & Bersch B., Structural changes of *Escherichia coli* ferric uptake regulator during metal-dependent dimerization and activation explored by NMR and X-ray crystallography. *J Biol Chem.* 281(30):21286-95 (2006).

Pennella M. A., Shokes J. E., Cosper N. J., Scott R. A., & Giedroc D. P., Structural elements of metal selectivity in metal sensor proteins. *Proc Natl Acad Sci USA* 100(7), 3713-3718 (2003).

Pennella M. A., & Giedroc D. P., Structural determinants of metal selectivity in prokaryotic metal-responsive transcriptional regulators. *Biometals* 18:413-428 (2005).

Pennella M. A., Arunkumar A. I. & Giedroc D. P., Individual metal ligands play distinct functional roles in the zinc sensor *Staphylococcus aureus* CzrA. *J Mol Biol* 356(5):1124–36. (2006).

Pérez Audero M. E., Podoroska B. M., Ibáñez M. M., Cauerhff A., Checa S. K. & Soncini F. C., Target transcription binding sites differentiate two groups of MerR-monovalent metal ion sensors. *Mol Microbiol.* 78(4):853-65 (2010).

Peticolas W. L., Wang Y. & Thomas G. A., Some rules for predicting the base-sequence dependence of DNA conformation. *Proc Natl Acad Sci U S A.* 85(8):2579-83 (1988).

Petrarca P., Ammendola S., Pasquali P. & Battistoni A., The Zur-regulated ZinT protein is an auxiliary component of the high-affinity ZnuABC zinc transporter that facilitates metal recruitment during severe zinc shortage. *J Bacteriol.* 192(6):1553-64 (2010).

Petsko G. A., Protein crystallography at sub-zero temperatures: cryo-protective mother liquors for protein crystals. *J Mol Biol*. 96(3):381-92 (1975).

Philips S. J., Canalizo-Hernandez M., Yildirim I., Schatz G. C., Mondragon A.& O'Halloran T. V., Allosteric transcriptional regulation via changes in the overall topology of the core promoter. *Science* 349(6250):877-881 (2015).

Phillips C. M., Schreiter E. R., Guo Y., Wang S. C., Zamble D. B. & Drennan C. L., Structural Basis of the Metal Specificity for Nickel Regulatory Protein NikR. *Biochemistry* 47:1938–1946 (2008).

Phillips C. M., Schreiter E. R., Stultz C. M. & Drennan C. L., Structural Basis of Low-Affinity Nickel Binding to the Nickel-Responsive Transcription Factor NikR from *Escherichia coli. Biochemistry* 49(36): 7830–7838 (2010).

Pohl E., Haller J. C., Mijovilovich A., Meyer-Klaucke W., Garman E. & Vasil M. L., Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator. *Mol Microbiol.* 47(4):903-15 (2003).

Pomposiello P. J. & Demple B., Redox-operated genetic switches: the SoxR and OxyR transcription factors. *Trends Biotechnol.* 19:109–114 (2001).

Pontel L. B., Audero M. E., Espariz M., Checa S. K. & Soncini F. C., GolS controls the response to gold by the hierarchical induction of *Salmonella*-specific genes that include a CBA efflux-coding operon. *Mol Microbiol.* 66(3):814-25 (2007).

Potter A. J, Kidd S. P., McEwan A. G. & Paton J. C., The MerR/NmlR family transcription factor of *Streptococcus pneumoniae* responds to carbonyl stress and modulates hydrogen peroxide production. *J Bacteriol.* 192(15):4063-6 (2010).

Privalov P. L., Dragan A. I., Crane-Robinson C., Breslauer K. J., Remeta D. P. & Minetti C. A., What drives proteins into the major or minor grooves of DNA? *J Mol Biol.* 365(1):1-9 (2007).

Privalov P. L., Dragan A. I. & Crane-Robinson C., The cost of DNA bending *Trends Biochem Sci.* 34(9):464-70 (2009).

Quayle J. R. & Ferenci T., Evolutionary aspects of autotrophy *Microbiological reviews* 42:251-73 (1978).

Rajkarnikar A., Strankman A., Duran S., Vargas D., Roberts A. A., Barretto K., Upton H., Hamilton C. J.& Rawat M., Analysis of mutants disrupted in bacillithiol metabolism in *Staphylococcus aureus.* Biochem. Biophys Res Commun 436(2):128-33 (2013).

Rana S., Pozzi N., Pelc L. A. & Di Cera E., Redesigning allosteric activation in an enzyme. *Proc Natl Acad Sci U S A* 108(13):5221-5225 (2011).

Ratner S., Clarke H. T., The Action of Formaldehyde upon Cysteine. *J. Am. Chem. Soc.* 59 (1):200–206 (1937).

Raumann B. E., Rould M. A., Pabo C. O. & Sauer R. T., DNA recognition by beta-sheets in the Arc repressor-operator crystal structure *Nature* 367(6465):754-7 (1994).

Reinhart G. D., Quantitative analysis and interpretation of allosteric behavior *Methods Enzymol.* 380:187-203 (2004).

Reyes-Caballero H., Guerra A. J., Jacobsen F. E., Kazmierczak K. M., Cowart D., Koppolu U. M., Scott R. A., Winkler M. E. & Giedroc D. P., The metalloregulatory zinc site in *Streptococcus pneumonia* AdcR, a zinc-activated MarT family repressor. J. Mol. Biol. 403:197-216 (2010).

Rhee S., Martin R. G., Rosner J. L., Davies D. R., A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator *Proc Natl Acad Sci U S A* 95(18):10413-8 (1998).

Rhodes G., "Crystallography made crystal clear". *Academic Press*, London (1993).

Richards F. M., The protein folding problem. *Scientific American* 34-41 (1991).

Rice P., Longden I. & Bleasby A., EMBOSS; the European Molecular Biology Open Software Suite. *Trends in Genetics: TIG* 16(6):276-277 (2000).

Riddles P. W., Blakeley R. L. & Zerner B., Ellman's reagent: 5,5`-dithiobis(2-nitrobenzoic acid)-a re-examination. *Analytical Biochemistry* 94(1):75-81 (1979).

Rodríguez D. D., Grosse C., Himmel S., González C., de Ilarduya I. M., Becker S., Sheldrick G. M. & Usón I., Crystallographic *ab initio* protein structure solution below atomic resolution. *Nature Methods*, 6, 651–653 (2009).

Rodríguez, D. D., Sammito M., Meindl K., de Ilarduya I. M., Potratz M., Sheldrick G. M. & Usón I., Practical structure solution with ARCIMBOLDO. *Acta Cryst. D* 68:336–343 (2012).

Roe J. A., Shaw B. F. & Valentine J. S., Fundamentals of Coordination Chemistry, in "Biological Inorganic Chemistry, Structure and Reactivity" (Bertini I., Gray H. B., Stiefel E. I., Valentine J. S., eds.) *University Science Books* California 695-712 (2007).

Rohs R., West S. M., Liu P. & Honig B., Nuance in the double-helix and its role in protein-DNA recognition. *Curr Opin Struct Biol.* 19(2):171-7 (2009a).

Rohs R., West S. M., Sosinsky A., Liu P., Mann R. S. & Honig B., The role of DNA shape in protein-DNA recognition *Nature* 461(7268):1248-53 (2009b).

Rohs R., Jin X., West S. M., Joshi R., Honig B. & Mann R. S., Origins of specificity in protein-DNA recognition *Annu Rev Biochem.* 79:233-69 (2010).

Rowe J. L., Starnes G. L. & Chivers P. T., Complex transcriptional control links NikABCDE-dependent nickel transport with hydrogenase expression in *Escherichia coli. J Bacteriol.* 187(18):6317-23 (2005).

Rubio-Sanz L., Prieto R. I., Imperial J., Palacios J. M. & Brito B., Functional and Expression Analysis of the Metal-Inducible *dmeRF* System From *Rhizobium leguminosarum bv. viciae*. *Appl. Environ. Microbiol.* 79(20):6414-6422(2013).

Rupp B., "Biomolecular crystallography: principles, practice, and application to structural biology". 1st edition, *Garland Science* (2009).

Rutherford J.C., Cavet J. S. & Robinson N. J., Cobalt-dependent transcriptional switching by a dual-effector MerR-like protein regulates a cobalt-exporting variant CPx-type ATPase *J Biol Chem* 274:25827-25832 (1999).

Sakamoto K., Agari Y., Agari K., Kuramitsu S. & Shinkai A., Structural and functional characterization of the transcriptional repressor CsoR from *Thermus thermophilus* HB8. *Microbiology* 156(7):1993-2005 (2010).

Sambrook J. & Russell D. W., Molecular Biology: A laboratory manual. *CSHL Press* (1) (2001).

Sander E. G., Jencks W. P., Equilibria for additions to the carbonyl group. *J. Am. Chem. Soc.*, 90 (22):6154–6162 (1968).

Sanghani P. C., Stone C. L., Ray B. D., Pindel E. V., Hurley T. D. & Bosron W. F., Kinetic mechanism of human glutathione-dependent formaldehyde dehydrogenase. *Biochemistry* 39(35):10720-9 (2000).

Sanghani P. C., Bosron W. F., Hurley T. D., Human glutathione-dependent formaldehyde dehydrogenase. Structural changes associated with ternary complex formation. *Biochemistry* 41:15189–15194 (2002).

Sassine M., Picquet-Varrault B., Perraudin E., Chiappini L., Doussin J. F., George C., A new device for formaldehyde and total aldehydes real-time monitoring. *Environ Sci Pollut Res Int.* 21(2):1258-69 (2014).

Schelert J., Dixit V., Hoang V., Simbahan J., Drozda M. & Blum P., Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon *Sulfolobus solfataricus* by use of gene disruption. *J Bacteriol.* 186(2):427-37 (2004).

Scheller K. H., Abel T. H., Polanyi P. E., Wenk P. K., Fischer B. E. & Sigel H., Metal ion/buffer interactions. Stability of binary and ternary complexes containing 2-[bis(2-hydroxyethyl)amino]-2(hydroxymethyl)-1,3-propanediol (Bistris) and adenosine 5'-triphosphate (ATP). *Eur J Biochem.* 107(2):455-66 (1980).

Schreiter E. R., Sintchak M. D., Guo Y., Chivers P. T., Sauer R. T. & Drennan C. L., Crystal structure of the nickel-responsive transcription factor NikR *Nat Struct Biol.* 10(10):794-9 (2003).

Schreiter E. R., Wang S. C., Zamble D. B. & Drennan C. L., NikR-operator complex structure and the mechanism of repressor activation by metal ions. *Proc. Natl. Acad. Sci. U.S.A.* 103:13676–13681 (2006).

Schubert M. P., Combination of thiol acids with methylglyoxal. *J. Biol. Chem* 111:671-678 (1935).

Schumacher M. A., Choi K. Y., Zalkin H. & Brennan R. G., Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices *Science* 266(5186):763-70 (1994).

Schwartz T., Behlke J., Lowenhaupt K., Heinemann U. & Rich A., Structure of the DLM-1-Z-DNA complex reveals a conserved family of Z-DNA-binding proteins *Nat Struct Biol.* 8(9):761-5 (2001).

Seeman N. C., Rosenberg J. M. & Rich A., Sequence-specific recognition of double helical nucleic acids by proteins. *Proc Natl Acad Sci U S A* 73(3):804-8 (1976).

Shakked Z., Guerstein-Guzikevich G., Eisenstein M., Frolow F. & Rabinovich D., The conformation of the DNA double helix in the crystal is dependent on its environment. *Nature* 342(6248):456-60 (1989).

Shakked Z. & Rabinovich D., The effect of the base sequence on the fine structure of the DNA double helix. *Prog Biophys Mol Biol*. 47(3):159-95 (1986).

Shankar R., Kolandaivel P. & Senthil kumar L., Coordination and binding properties of zwitterionic glutathione with transition metal cations. *Inorg Chim Acta* 387:125-136 (2012).

Sharma S. V., Arbach M., Roberts A. A., Macdonald C. J., Groom M. & Hamilton C. J., Biophysical features of bacillithiol, the glutathione surrogate of *Bacillus subtilis* and other firmicutes. *Chembiochem*. 14(16):2160-8 (2013).

Sheldrick G. M., A short history of *SHELX*. *Acta cryst*. A64:112-122 (2008).

Sheldrick G. M., Experimental phasing with *SHELXC/D/E*: combining chain tracing with density modification. Acta Cryst. D Biol. Crystallogr. 66(4):479-85 (2010).

Shi X., Festa R. A., Ioerger T. R., Butler-Wu S., Sacchettini J. C., Darwin K. H. & Samanovic M. I., The copper-responsive RicR regulon contributes to *Mycobacterium tuberculosis* virulence. *MBio.* 18;5(1). pii: e00876-13 (2014).

Shin J. H., Jung H. J., An Y. J., Cho Y. B., Cha S. S. & Roe J. H., Graded expression of zinc-responsive genes through two regulatory zinc-binding sites in Zur. *Proc Natl Acad Sci U S A.* 108(12):5045-50 (2011).

Simons T.J.B., Measurement of free Zn^{2+}ion concentration with the fluorescent probe mag-fura-2 (furaptra). *J. Biochem. Biophys. Methods* 27:25–37 47 (1993).

Singh V. K., Xiong A., Usgaard T. R., Chakrabarti S., Deora R., Misra T. K. & Jayaswal R. K., ZntR is an autoregulatory protein and negatively regulates the chromosomal zinc resistance operon *znt* of *Staphylococcus aureus*. *Mol Microbiol.* 33(1):200-7 (1999).

Sipos K., Lange H., Fekete Z., Ullmann P., Lill R. & Kispal G., Maturation of cytosolic iron-sulfur proteins requires glutathione. *J Biol Chem.* 277(30):26944-9 (2002).

Sheikh M. A. & Taylor G. L., Crystal structure of the *Vibrio cholerae* ferric uptake regulator (Fur) reveals insights into metal co-ordination. *Mol Microbiol.* 72(5):1208-20 (2009).

Smaldone G. T. & Helmann J. D., CsoR regulates the copper efflux operon *copZA* in *Bacillus subtilis. Microbiology* 153(12):4123-8 (2007).

So A. K. & Espie G. S., Cloning, characterization and expression of carbonic anhydrase from the cyanobacterium *Synechocystis* PCC6803. *Plant Mol Biol.* 37(2):205-15 (1998).

Solomon M. J. & Varshavsky A., Formaldehyde-mediated DNA-protein crosslinking: a probe for in vivo chromatin structures. *Proc Natl Acad Sci USA* 82(19):6470-4 (1985).

Somers W. S., Phillips S. E., Crystal structure of the met repressor-operator complex at 2.8 A resolution reveals DNA recognition by beta-strands. *Nature* 359(6394):387-93 (1992).

Song Z. B., Xiao S. Q., You L., Wang S. S., Tan H., Li K. Z., Chen L. M., C1 metabolism and the Calvin cycle function simultaneously and independently during HCHO metabolism and

detoxification in *Arabidopsis thaliana* treated with HCHO solutions. *Plant Cell Environ.* 36(8):1490-506 (2013).

Soutourina O., Poupel O., Coppée J. Y., Danchin A., Msadek T. & Martin-Verstraete I., CymR, the master regulator of cysteine metabolism in *Staphylococcus aureus*, controls host sulphur source utilization and plays a role in biofilmformation. *Mol Microbiol.* 73(2):194-211 (2009).

Soutourina O., Dubrac S., Poupel O., Msadek T. & Martin-Verstraete I., The pleiotropic CymR regulator of *Staphylococcus aureus* plays an important role in virulence and stress response. *PLoS Pathog.* 6(5):e1000894 (2010).

Stojiljkovic I., Bäumler A. J. & Hantke K., Fur regulon in gram-negative bacteria. Identification and characterization of new iron-regulated *Escherichia coli* genes by a fur titration assay. *J Mol Biol* 236(2):531-45 (1994).

Stover N. A., Cavalcanti A. R., Li A. J., Richardson B. C. & Landweber L. F., Reciprocal fusions of two genes in the formaldehyde detoxification pathway in ciliates and diatoms. *Mol Biol Evol.* 22(7):1539-42 (2005).

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

Studholme D. J. & Pau R. N., A DNA element recognised by the molybdenum-responsive transcription factor ModE is conserved in Proteobacteria, green sulphur bacteria and Archaea. *BMC Microbiol.* 3:24 (2003).

Suelter C. H., Enzymes activated by monovalent cations. *Science* 168:789–795 (1970)

Suzuki K. T., Karasawa A. & Yamanaka K., Binding of copper to albumin and participation of cysteine *in vivo* and *in vitro*. *Arch Biochem Biophys*. 273(2):572-7 (1989).

Tahirov T. H., Inoue-Bungo T., Morii H., Fujikawa A., Sasaki M., Kimura K., Shiina M., Sato K., Kumasaka T., Yamamoto M., Ishii S. & Ogata K., Structural analyses of DNA recognition by the AML1/Runx-1 Runt domain and its allosteric control by CBFbeta *Cell* 104(5):755-67 (2001).

Tamilarasan R. & McMillin d. r., Ultraviolet absorbance increase on reduction of plastocyanin: conformational change or a new chromophore? *Biophys. J*. 50 (1986).

Tan B. G., Vijgenboom E. & Worrall J. A. R., Conformational and thermodynamic hallmarks of DNA operator site specificity in the copper sensitive operon repressor from *Streptomyces lividans*. *Nucleic Acids Research* 42(2):21326–1340 (2014).

Tang Q., Liu Y. P., Yan X. X. & Liang D. C., Structural and functional characterization of Cys4 zinc finger motif in the recombination mediator protein RecR. *DNA Repair (Amst)* 24:10-4 (2014).

Tateno M., Yamasaki K., Amano N., Kakinuma J., Koike H., Allen M. D. & Suzuki M., DNA recognition by beta-sheets. *Biopolymers* 44(4):335-59 (1997).

Teng T. Y., Mounting of crystals for macromolecular crystallography in a freestanding thin film. J. Appl. Cryst. 23:387-391 (1990).

Thompson J. D., Higgins D. G., Gibson T. J., CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*., 22:4673-4680 (1994).

Tolstorukov M. Y., Jernigan R. L. & Zhurkin V. B., Protein-DNA hydrophobic recognition in the minor groove is facilitated bysugar switching. *J Mol Biol*. 12;337(1):65-76 (2004).

Tottey S., Harvie D. R. & Robinson N. J., Understanding how cells allocate metals using metal sensors and metallochaperones. *Acc. Chem. Res.* 38: 775-783 (2005).

Traoré D. A, El Ghazouani A., Ilango S., Dupuy J., Jacquamet L., Ferrer J. L., Caux-Thang C., Duarte V. & Latour J. M., Crystal structure of the apo-PerR-Zn protein from *Bacillus subtilis*. *Mol Microbiol.* 61(5):1211-9 (2006).

Traoré D. A., El Ghazouani A., Jacquamet L., Borel F., Ferrer J. L., Lascoux D., Ravanat J. L., Jaquinod M., Blondin G., Caux-Thang C., Duarte V. & Latour J. M., Structural and functional characterization of 2-oxo-histidine in oxidized PerR protein. *Nat Chem Biol.* 5(1):53-9 (2009).

Travers A. A., DNA conformation and protein binding. *Annu Rev Biochem.* 58:427-52 (1989).

Travers A. A., Reading the minor groove. *Nat Struct Biol.* 2(8):615-8 (1995).

Trewick S. C., Henshaw T. F., Hausinger R. P., Lindahl T. & Sedgwick B., Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage *Nature* 419(6903):174-8 (2002).

Ünver H. & Hayvali Z., Synthesis, spectroscopic studies and structures of square-planar nickel(II) and copper(II) complexes derived from 2-{(Z)-[furan-2-ylmethyl]imino]methyl}-6-methoxyphenol. *Spectrochim Acta A Mol Biomol Spectrosc.* 75(2):782-8 (2010).

Valko M., Jomova K., Rhodes C. J., Kuča K. & Musílek K., Redox- and non-redox-metal-induced formation of free radicals and their role in human disease. *Arch Toxicol.* 90(1):1-37 (2016).

Van Ophem P. W. & Duine J. A., NAD- and co-substrate (GSH or factor)-dependent formaldehyde dehydrogenases from methylotrophic microorganisms act as a class III alcohol dehydrogenase. *FEMS Microbiol. Lett.* 116:87-94 (1994).

VanZile M. L., Cosper N. J., Scott R. A. & Giedroc D. P., The zinc metalloregulatory protein *Synechococcus* PCC7942 SmtB binds a single zinc ion per monomer wih high affinity i a tetrahedral coordination geometry. *Biochemistry* 39(38):118-29 (2000).

VanZile M. L., Chen X. & Giedroc D. P., Allosteric negative regulation of *smt* O/P binding of the zinc sensor, SmtB, by metal ions: a coupled equilibrium analysis. *Biochemistry* 41(31):9776-86 (2002a).

VanZile M. L., Chen X. & Giedroc D. P., Structural characterization of distinct α5N and α3 metal sites in the cyanobacterial zinc sensor SmtB. *Biochemistry* 41:9765-9775 (2002b).

Von Harsdorf R., Li P. F. & Dietz R., Signaling pathways in reactive oxygen species-induced cardiomyocyte apoptosis. *Circulation.* 99(22):2934-41 (1999).

Vorholt J. A., Marx C. J., Lidstrom M. E., Thauer R. K., Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol. *J. Bacteriol.* 182(23):6645-50 (2000).

Vorholt J. A., Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. *Arch Microbiol*. 178(4):239-49 (2002).

Waldron K. J. & Robinson N. J., How do bacterial cells ensure that metalloproteins get the correct metal? *Nat Rev Microbiol*. 7(1):25-35 (2009).

Waldron K. J., Rutherford J. C., Ford D. & Robinson N. J., Metalloproteins and metal sensing. *Nature*. 460(7257):823-30. (2009).

Waller J. P., The NH_2-terminal residue of the proteins from cell-free extract of *E. coli*. *J. Mol. Biol.* 7:483– 496 (1963).

Walter R. L., Thiel D. J., Barna S., Tate M. W., Wall M. E., Eikenberry E. F., Gruner S. M. & Ealick S. E., High-resolution Macromolecular Structure Determination using CCD Detectors and Synchrotron Radiation., *Structure* 3:835-844 (1995).

Wang A. H., Quigley G. J., Kolpak F. J., Crawford J. L., van Boom J. H., van der Marel G. & Rich A., Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* 282(5740):680-6 (1979).

Wang A. H., Gessner R. V., van der Marel G. A., van Boom J. H. & Rich A., Crystal structure of Z-DNA without an alternating purine-pyrimidine sequence. *Proc Natl Acad Sci U S A* 82(11): 3611–3615 (1985).

Wang S. C., Dias A. V., Bloom S. L. & Zamble D. B., Selectivity of Metal Binding and Metal-Induced Stability of *Escherichia coli* NikR. *Biochemistry* 43:10018-100281 (2004).

Wang S., Deng K., Zaremba S., Deng X., Lin C., Wang Q., Tortorello M. L. & Zhang W., Transcriptomic response of *Escherichia coli* O157:H7 to oxidative stress. *Appl. Environ. Microbiol.* 75(19):6110-23 (2009).

Wang W. & Ballatori N., Endogenous glutathione conjugates: occurrence and biological functions. *Pharmacol Rev.* 50(3):335-56 (1998).

Ward S. K., Abomoelak B., Hoye E. A., Steinberg H. & Talaat A. M., CtpV: a putative copper exporter required for full virulence of *Mycobacterium tuberculosis*. *Mol Microbiol.* 77(5):1096-110 (2010).

Watton S. P., Wright J. G., MacDonnell F. M., Bryson J. W., Sabat M. & O'Halloran T. V., Trigonal mercuric complex of an aliphatic thiolate: A spectroscopic and structural model for the receptor site in the Hg(II) biosensor MerR. *J. Am. Chem. Soc.* 112:2824-2826 (1990).

Weller M., Overton T., Rourke J., Armstrong F. & Atkins P., Inorganic Chemistry, 6th ed. *Oxford University*'s Press (2014).

Werner M. H., Huth J. R., Gronenborn A. M. & Clore G. M., Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell* 81(5):705-14 (1995).

White C. C., Viernes H., Krejsa C. M., Botta D. & Kavanagh T. J., Fluorescence-based microtiter plate assay for glutamate-cysteine ligase activity. *Analytical Biochem*. 318(2):175-180 (2003).

WHO IARC Monograph on the Evaluation of Carcinogenic Risks to Humans, Formaldehyde. 88:37-326 (2006).

Wilkins R. G., Kinetics and Mechanism of Reactions of Transition Metal Complexes. 2nd Thoroughly revised edition; *VCH Publishers*, Inc. New York. 23-24 (1991).

Wiethaus J., Wirsing A., Narberhaus F. & Masepohl B., Overlapping and specialized functions of the molybdenum-dependent regulators MopA and MopB in *Rhodobacter capsulatus*. *J Bacteriol.* 188(24):8441-51 (2006).

Wu J. & Rosen B. P., The ArsR protein is a trans-acting regulatory protein. *Mol Microbiol* 5(6):1331-6 (1991).

Xiao Z. & Wedd A. G., The challenges of determining metal-protein affinities. *Nat. Prod. Rep*. 27:768-789 (2010).

Xiong A. & Jayaswal R. K., Molecular characterization of a chromosomal determinant conferring resistance to zinc and cobalt ions in *Staphylococcus aureus*. *J Bacteriol.* 180(16):4024-9 (1998).

Xiong Y. & Sundaralingam M., Protein-nucleic acid interaction: major groove recognition determinants. *Encyclopedia of Life Sciences* (John Wiley & Sons, ed) (2001).

Xu C., Zhou T., Kuroda M. & Rosen B. P., Metalloid Resistance Mechanisms in Prokaryotes *J. Biochem. (Tokyo)* 123:16-23(1998).

Yadav S. K., Singla-Pareek S.L. & Sopory S. K., An overview on the role of methylglyoxal and glyoxalases in plants. *Drug Metabol Drug Interact*. 23(1-2):51-68 (2008).

Yam V. W. W., Molecular design of luminescent metal-based materials. *Pure Appl. Chem.* 73(3):543-548 (2001).

Yasueda H., Kawahara Y. & Sugimoto S.. *Bacillus subtilis yckG* and *yckF* encode two key enzymes of the ribulose monophosphate pathway used by methylotrophs, and *yckH* is required for their expression. *J Bacteriol.* 181(23):7154-60 (1999).

Yoon K. P., Misra T. K. & Silver S., Regulation of the cadA cadmium resistance determinant of *Staphylococcus aureus* plasmid pI258. *J Bacteriol.* 173(23):7643-9 (1991).

Yurimoto H., Hirai R., Matsuno N., Yasueda H., Kato N. & Sakai Y., HxlR, a member of the DUF24 protein family, is a DNA-binding protein that acts as a positive regulator of the formaldehyde-inducible *hxlAB* operon in *Bacillus subtilis*. *Mol Microbiol*. 57(2):511-9 (2005).

Zaharik M. L., Cullen V. L., Fung A. M., Libby S. J., Choy S. L. K., Coburn B., Kehres D. G., Maguire M. E., Fang F. C. & Finlay B. B., The *Salmonella enterica serovar typhimurium* divalent cation transport systems MntH and SitABCD are essential for virulence in an Nramp1 murine typhoid model. *Infect Immun*. 72(9):5522-25 (2004).

Zhang Y. & Gladyshev V. N., Molybdoproteomes and evolution of molybdenum utilization. *J. Mol Biol.* 379(4):881-99 (2008).