

ADDENDUM

Figure 4-2 : Caption read as “Surface electrostatics of (A) *VcDsbA* and (B) *EcDsbA*”

page 182: Comment: Reference 171 has been cited for “oxidized and reduced *VcDsbA* assignments”

page 185, Figure 4-9: white cells represent non-selected hydrophobic cells.

page 187, Line 5: read as “Compared to ¹H and ¹⁵N, ¹³C CSP are often.....”

Equation 4-2: Square root should be read as $\frac{1}{2}$.

Figure 4-18 : Caption last line read as “Additional or larger CSP were observed in binding site residues of *VcDsbA* spectra in presence of **M2** as compared to shifts induced by **M3**.”

page 207/ Figure 4-25: Read N6 (1), 1 as 1a.

Table 4-5: Read 4 as 1 and 1 as 1a. All LE values listed are in negative.

Figure 5-10: These analogues have been introduced in subsequent sections of this chapter and docking protocol and optimization described before these.

Figure 5-13: NMR-docking pose of A20 was superimposed on the X-ray structure of 2 and this suggested a close agreement with X-ray structure with a RMSD of 1.79 Å.

Figure 5-14: This figure is displayed here (close to analogue binding information table) for the reader's convenient and to consider binding mode of parent compound when discussing the analogues.

Figure 5-16: NMR-docking pose of A28 was superimposed on the X-ray structure of 2 and this suggested a close agreement with X-ray structure with a RMSD of 2.1 Å.

page 261: Extra text removed.

page 269/270: The NMR-docked structures of compounds **A28** and **E8** were significantly similar, additionally X-ray crystal structure of **E8** is already provided in figure 5-22 with NMR binding information. Therefore it's been directed to refer to figure 5-10 to observe docking information of compound **E8**.

Figure 6-6: Yellow residues are active site Cysteine (Cys 31 and Cys 33).

page 308 : ITC data presented in Appendix 8.4.

page 318 , Figure 6-13: This figure displays the comparison of NMR based binding location and crystallographic coordinates of compounds **9** and **9b**. NMR docking supported inner orientation of **9b** as observed in preliminary analysis of experimental binding.

Appendices, Figure 8-4: Chemical name of compound **10** “trifluoromethylthiophenol”

ERRATA

page VII, line 3: “DsbA wide” to “DsbA has “a” wide”

page VII, Second paragraph: “EcDsbA” to “*EcDsbA*”

page 3, line 6: “overcome” to “overcoming”

page 4, line 4: “*E.coli*” to “*E. coli*”

page 6, 1.2.1: “shares” to “share”, “an soluble” to “a soluble”

page 7, Last line: “*E.coli*” to “*E. coli*”

page 14, line 2: structures “has” been to “have” been

page 17, line 7: “loop” to “loops”, line 18: reoxidised “by”

page 24, line 2: “isomerise” to “isomerase”

Section 1.4.1.2, line 6: “quenching” to “quencher”

page 28, line 6: “by absence” to “by the absence”

page 59, line 16: “selectively” to “tagged selectively”

page 59: Deleted repeated lines 1-12

page 61: “labelling” to “labeling”

page 74 : Line 12: “*p*-hydroxybenzoic acid” to “*p*-hydroxybenzoic acid”

page 75: “5kDa Molecular weight cut-off concentrators” to Molecular weight cut-off concentrators of five kDa

page 76: “121 °” to “121°C”

page 82, Line 3: corrected to “The supernatant was filtered and applied to...”

Figure 2-1: B and C corrected to “SDS-PAGE among X1-X5 and C11-D10.....”

Figure 2-2: Circle/Arrow inserted.

Page 89 & 90: “Alkaline phosphatase (AP)”, “Alkaline phosphatase activity: (PhoA)”

Figure 2-7: “Inhibitor (trifluoromethylthiophenol)...”

page 104, last line: “Figure 2-9” to “Figure 2-11”

page 106: “C alpha” to “CA”

Equation 2-4 : “Kd” to “KD”

Figure 2-13: “CBCA (CO)NH” to “CBCA (CO)NH spectra” and “white” to “red”

Page 123: Equation 2-6 (Please refer to Equation 2-5 on page 108)

Figure 2-15: Caption corrected to “needles and twinned shaped crystals and completed to “shown in panel (A) and (C)”

page 137: Last paragraph “rational” to “rationale”

page S144: Equation S2 “Kd” to “KD”

Fig 4-23: “Green” corrected to “pink”.

page 211, Line 4: “(Figure 4-6 and 4-11)” to “(Figure 4-10 and 4-12)”

page 239, Line 7: “fragments 1, 3 and 4” corrected to fragments “1, 2 and 4”

page 121, Last paragraph, line 3: delete “PHENIX”.

page 200, First paragraph, line 1 “...M1-M2...” to “...M1-M7...”.

Page 268, paragraph one, line 2: ...“approximately 27-fold increase in binding affinity...” to “...approximately 17-fold increase in binding affinity...”

page 301, paragraph one, line 7:“showing a 64 fold increase in affinity..” corrected to “...showing a 6.14 fold increase in affinity...”.

page 278, paragraph one, line 1: “about a 10-fold higher affinity” corrected to “about a 2-fold higher affinity”

**Inhibition of Oxidative Protein Folding in
Gram-Negative Bacteria: Fragment-Based Design of DsbA Inhibitors**

Pooja Sharma

Medicinal Chemistry and Drug Action

Monash Institute of Pharmaceutical Sciences

Monash University

PART A: General Declaration

Monash University
Monash Research Graduate School

Declaration for thesis based or partially based on conjointly published or unpublished work

General Declaration

In agreement with Monash University Doctorate Regulation 17/ Doctor of Philosophy and Master of Philosophy (MPhil) regulations the following declarations are made:

I hereby declare that this thesis includes no material that has been acknowledged for the award of any other degree or diploma in any university or academic institution and, to the best of my knowledge and belief, this thesis contains no study or material previously published or written by another individual, except where due reference is made in the text of the thesis.

This thesis includes one unpublished publication. The core theme of the thesis is the identification of small molecule inhibitors of DsbA. The ideas, development and writing up of the paper in the thesis were the principal responsibility of myself, the candidate, working within the Department of Medicinal Chemistry and Drug Action under the supervision of A/Prof. Martin J. Scanlon and Dr Jamie S. Simpson and Dr David T Manallak.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of Chapter 3 my contribution to the work involved the following:

| Thesis chapter | Publication title | Publication status | Nature and extent of candidate's contribution |
|----------------|--|--------------------|---|
| 3 | Inhibition of Oxidative Protein Folding in <i>Escherichia coli</i> | To be Submitted | Experimental design, conduct of all laboratory work and data analysis. Initial preparation of manuscripts and subsequent revisions. |

Signed:

Date:

Acknowledgements

The years of my PhD at Monash University has had mentorship from numerous wonderful people both from within the university and outside of it. It is a pleasure to express appreciation to many people who made this thesis, a rewarding and unforgettable experience of my life.

First and foremost I would like to express my gratitude to my supervisors A/Prof Martin Scanlon, Dr Jamie Simpson and Dr David Manallack. It is difficult to overstate my thank to Martin for giving me an opportunity to work in his research group, his invaluable guidance, perceptiveness, professional support, excellent teaching & funding for the project and to attend several conferences and lab visits. I thank him for allowing me to work on the project of my interest and stay focused to analyse what is important and why I am doing what I am doing. I am greatly indebted to Jamie for his enthusiastic inspiration, constructive criticism and his great efforts to explain things clearly and simply at each stage of this project and his help to avoid getting lost in my exploration. I am thankful to David for his valuable guidance and sharing his experience especially in the computational aspects of this project. Throughout my thesis-writing period, my supervisors provided encouragement, great advice, good teaching and lots of good ideas. I would have been lost without them.

I am grateful to the Faculty of Pharmacy and Pharmaceutical Sciences for financial support and Monash Research Graduate School for MIPRS Scholarship and travel grant during the course of my PhD. Thanks to all the staff and the colleagues at MIPS for providing a stimulating and fun environment in which to learn and grow. I would like to express my gratitude to the members of Scanlon's and Simpson's group individually who have made working on this project through the years so enjoyable.

Dr James Horne, for his guidance in my early days in the lab and for patiently sharing his experience in the biophysical screening of libraries, and teaching a range of wonderful

NMR experiments. You seemed to be an extremely reliable and informative source of NMR and I missed you at various stages of this project since you have shifted to Tasmania. Dr Stephen Headey, for sharing his NMR and protein expression skills and useful discussions. Additionally thank you awesome Steve, for enlighten (or updating) me in thousands of philosophical, cultural, historical and artistic facts (or your own theories) and motivating me with unlimited enthusiasm to do every positive thing in the life.

Dr Tony Velkov, for teaching me protein purification techniques and inspiring me with your constantly optimistic attitude for science. Dr Kieran Rimmer, for helping me to analyse biochemical assays and for making things easier. Dr Jerome Wielens, for the valuable suggestions and explaining many smart methods for X-ray crystallographic data analysis.

I am thankful to Ms Yanni Chin and Dr Aisha Laguerre, my best friends as a PhD student, for helping me to learn many lab techniques. Big thanks to both of you for the years of memorable camaraderie, laughter, sharing each and every thought for 'life and science' and for helping me get through the difficult times with constant encouragement. Thank to Mr Sandeep Chhabra, my great friend and a student colleague from undergraduate to PhD years, for your outstanding support, inspiration and for always believing in me.

I am thankful to Mr Bradley Doak, Mr. Christopher Schreurs and Ms Joan Ho for the chemical synthesis of numerous required compounds for this project and useful discussions. Thanks to Martin Williams, Amelia Vom, Mansha Vazirani, Urmi Dhagat, Dr Eleanor Leung, Dr Maria Hughes, Fiona McRobb, Mark Agostino and Nick Barlow, for their support and sharing good times through the years.

Thanks to Dr David Chalmers who manages the computational chemistry lab at MIPS and has always had the time to help me for solving technical issues and discuss aspects of computational chemistry with me. I could never proceed a step without your

programming help to run all the softwares used in this project. Thanks to Dr James Swarbrick for his useful suggestions and support for NMR experiments.

A special thanks goes to our collaborators; I worked with at the University of Queensland to mainly perform the X-ray crystallographic studies. Prof Jenny Martin, the group leader and my role model too, for her valuable insights and support. Dr Begoña Heras, for her great working efforts for crystallographic studies and for teaching me many useful techniques. Thanks a lot for your encouragement, appreciation and specially, for supporting me in those extraordinary times. Also thanks to Ms Patricia Walden and Dr Stephen Shouldice for helping me in the lab.

Lastly, and most importantly, I wish to thank God and my parents, Late Mr Pawan Sharma and Mrs Arti Sharma, for this life. Thanks to my father, whose remembrance has only increased after so many years of his demise, promised you, when I was a girl wishing to be a doctor and save the world. I am working on it, still. Despite the geographical distances my family was always nearby. Thanks to my elder brother Praveen and his wife Geeta, and my younger brother Shantnu, who through the study career and years had always supported me to follow my heart and inquisitive mind in any direction this took me. Especially Praveen, your encouragement and concern through the years were consistent and a big support to me. If we ever had a family motto that would have been – *If there's a will, there's a way* – a philosophy of life I have been carrying with me every day. To them I dedicate this thesis.

Publication Record

PUBLICATIONS:

Shouldice SR, Heras B, Jarrott R, **Sharma P**, Scanlon MJ, Martin JL.

Characterization of the DsbA oxidative folding catalyst from *Pseudomonas aeruginosa* reveals a highly oxidizing protein that binds small molecules.

Antioxid Redox Signal. **2010** Apr 15;12(8):921-31.

Paxman JJ, Borg NA, Horne J, Thompson PE, Chin Y, **Sharma P**, Simpson JS, Wielens J, Piek S, Kahler CM, Sakellaris H, Pearce M, Bottomley SP, Rossjohn J, Scanlon MJ

The structure of the bacterial oxidoreductase enzyme DsbA in complex with a peptide reveals a basis for substrate specificity in the catalytic cycle of DsbA enzymes.

J Biol Chem. **2009** Jun 26;284(26):17835-45.

COMMUNICATIONS:

Sharma P, Heras B, Doak B, Manallack D, Simpson JS, Martin JL, Scanlon MJ

Inhibition of Oxidative Protein Folding in Gram Negative Bacteria: Fragment Elaboration Coupled with Structure Based Design (Poster)

Lorne Conference on Protein structure and function, **2011** (Lorne, Victoria, Australia)

Sharma P, Heras B, Manallack D, Simpson JS, Martin JL, Scanlon MJ

Inhibition of Oxidative Protein Folding in Gram Negative Bacteria: Fragment-Based Design of EcDsbA Inhibitors (Talk and poster)

Keystone Symposia - New Directions in Small Molecule Drug Discovery (Z8), **2010** (Whistler, British Columbia, Canada)

Sharma P, Heras B, Manallack D, Simpson JS, Martin JL, Scanlon MJ

Structural Insights into Protein-Ligand Interactions of Small Molecule Inhibitors of EcDsbA (Poster)

Lorne Conference on Protein structure and function, **2010** (Lorne, Victoria, Australia)

VIIN Student Symposium, Monash University, **2010** (Clayton, Victoria, Australia)

Sharma P, Heras B, Manallack D, Simpson JS, Martin JL, Scanlon MJ

Discovery of Novel Inhibitors of EcDsbA by NMR-based fragment screening (Poster)

Intersect 09, Fragment Based Design workshop, Latrobe University, **2009** (Victoria, Australia)

Sharma P, Horne J, Manallack D, Simpson JS, Scanlon MJ

Combined NMR spectroscopy & molecular docking studies to investigate the binding of scaffolds to EcDsbA (Poster),

7th Biennial Conference of the Australian and New Zealand Society for Magnetic Resonance, **2008** (Brisbane, QL, Australia)

Sharma P, Horne J, Chin Y, Chalmers D, Manallack D, Simpson JS, Scanlon MJ

Fragment-Based Screening using Molecular Modeling and NMR Spectroscopy : Application to *Vibrio Cholerae* Periplasmic Oxidoreductase Protein VcDsbA (Poster)

VCPPA HDR Research Symposium, Monash University, **2008**, (Parkville, Australia)

Molecular Modeling Conference, **2007**, Bio 21, (Melbourne, Australia)

Abstract

The disulfide-dithiol oxidoreductase enzyme DsbA is an oxidative folding catalyst found in the bacterial periplasm and is a key determinant of virulence in Gram-negative pathogens. DsbA has wide spectrum of substrate specificity and has been demonstrated to catalyse the formation of disulfides in numerous secreted proteins. DsbA operates at a central point in the production of virulence-determinants because most virulence factors are proteins that require disulfide bonds to be active. DsbA knockout mutants in *E. coli*, have been shown to be impaired in a range of processes related to virulence. Significantly, DsbA mutations in many pathogenic bacteria attenuate their virulence, demonstrating the value of targeting DsbA to develop anti-bacterial agents to counteract virulence. DsbA is a structurally as well as functionally well characterized protein which has no direct homologue in eukaryotes. To date there are no known small molecule inhibitors of this protein. Hence DsbA represents an attractive antibacterial target and is the protein of interest for this project.

The current thesis reports the identification of scaffolds that are suitable starting points for designing inhibitors of DsbA. These scaffolds were identified using a fragment-based drug design (FBDD) approach. This study describes a structure assisted-FBDD process performed to validate DsbA as an antibacterial drug target. A broad range of complementary techniques were employed to identify inhibitors of DsbA from two different Gram negative bacterial species; *Escherichia coli* and *Vibrio cholerae*. NMR based screening of a fragment library was carried out in two rounds. The primary screen employed STD (Saturation Transfer Difference) experiments to identify candidate hits

and ^1H - ^{15}N HSQC (Heteronuclear Single Quantum Coherence) experiments were used to confirm binding. A number of strong hits were identified and the binding locations were identified from the HSQC data. The binding efficiency of the fragments was determined by calculating NMR-based dissociation constants. Several of these initial binding fragments also show inhibitory activity in a phenotype assays in *E. coli*. X-ray crystallographic studies revealed that the most potent fragment hits bind in the hydrophobic groove of *EcDsbA* in adjacent or overlapping positions.

Characterisation of the structures of complexes of *EcDsbA* with the original NMR hits was followed by fragment elaboration through structure-guided medicinal chemistry efforts. Preliminary structure-activity relationships (SAR) were identified for several series. This SAR will facilitate the development of novel DsbA inhibitors that specifically target functionally important protein surface sites.

In conclusion, this study presents the discovery of novel small molecule inhibitors of DsbA and provides insights into the development of potential DsbA inhibitors as antibacterial drug candidates.

Table of Contents

| | |
|--|------------|
| PART A: General Declaration..... | II |
| Acknowledgements | III |
| Publication Record | VI |
| Abstract | VII |
| Table of Contents | IX |
| List of Abbreviations | XVI |
| 1. General Introduction | 1 |
| 1. General Introduction..... | 2 |
| 1.1 Gram-negative bacterial infection..... | 2 |
| 1.2 DsbA and virulence in bacteria | 3 |
| 1.2.1 Oxidative catalytic cycle of DsbA..... | 6 |
| 1.3 Structural studies..... | 11 |
| 1.3.1 DsbA structure..... | 11 |
| 1.3.2 Difference among DsbA homologues | 14 |
| 1.3.3 DsbA-DsbB complex..... | 17 |
| 1.3.4 DsbA - substrate | 19 |
| 1.4 Functional and biochemical characterization of DsbA..... | 23 |
| 1.4.1 Biochemical in vitro assay..... | 23 |
| 1.4.1.1 Insulin reduction assay | 23 |
| 1.4.1.2 Fluorescence quenching assay | 24 |
| 1.4.2 Phenotypic assays..... | 26 |
| 1.4.2.1 Bacterial motility assay..... | 26 |
| 1.4.2.2 Heavy metal viability assay | 27 |
| 1.4.2.3 Alkaline phosphatase assay | 27 |
| 1.4.3 DsbA as a suitable target..... | 28 |
| 1.5 Process of drug discovery | 30 |

| | | |
|----------|--|----|
| 1.5.1 | <i>Historical methods</i> | 31 |
| 1.5.1.1. | Nature as a drug source..... | 31 |
| 1.5.1.2 | Combinatorial chemistry..... | 32 |
| 1.5.1.3 | High throughput screen (HTS) | 32 |
| 1.5.2 | <i>Fragment based drug design</i> | 35 |
| 1.5.1.1 | Library design and rule of three..... | 37 |
| 1.5.2.2 | Ligand efficiency | 40 |
| 1.5.2.3 | Advantages of FBDD over conventional methods | 41 |
| 1.5.2.4 | Fragment elaboration | 43 |
| 1.6 | <i>Project aims</i> | 48 |
| 1.7 | <i>Analytical techniques applied for FBDD</i> | 49 |
| 1.7.1 | <i>NMR for fragment screening and structural biology</i> | 50 |
| 1.7.1.1 | Ligand based methods for hit identification | 52 |
| 1.7.1.2 | Hit validation using target based methods..... | 56 |
| 1.7.1.3 | Multidimensional NMR of biological macromolecules | 60 |
| 1.7.1.4 | Structure assignment..... | 61 |
| 1.7.2 | <i>X-ray crystallography for FBDD</i> | 62 |
| 1.7.3 | <i>Other biophysical screening methods in FBDD</i> | 66 |
| 1.8 | <i>Inhibition analysis</i> | 69 |
| 1.9 | <i>Structure of thesis</i> | 70 |
| 2 . | Materials and method | 72 |
| 2.1 | <i>Materials</i> | 73 |
| 2.1.1 | <i>Chemicals</i> | 73 |
| 2.1.1.1 | Fragment Library | 73 |
| 2.1.1.2 | Miscellaneous Chemicals | 74 |
| 2.1.1.4 | Columns | 75 |
| 2.2 | <i>Molecular Methods</i> | 75 |
| 2.2.1 | <i>Transformation of E. coli</i> | 75 |
| 2.2.2 | <i>Glycerol stocks</i> | 76 |
| 2.2.3 | <i>Buffer and media preparation</i> | 76 |
| 2.2.4 | <i>SDS PAGE gel</i> | 77 |

| | | |
|---------|---|-----|
| 2.2.5 | <i>Protein expression</i> | 78 |
| 2.2.5.1 | Expression of unlabeled <i>EcDsbA</i> | 78 |
| 2.2.5.2 | Expression of labeled <i>EcDsbA</i> | 81 |
| 2.2.6 | <i>Protein purification</i> | 82 |
| 2.2.7 | <i>Oxidation, reduction & estimation of protein concentration</i> | 83 |
| 2.2.8 | <i>Protein stability analysis</i> | 88 |
| 2.3 | <i>Functional assays</i> | 89 |
| 2.3.1 | <i>Alkaline phosphatase activity assay</i> | 89 |
| 2.3.2 | <i>Insulin Reduction Assay</i> | 92 |
| 2.3.3 | <i>Fluorescence (FRET) Peptide assay</i> | 94 |
| 2.3.4 | <i>Complementation/Motility Assay</i> | 97 |
| 2.4 | <i>NMR Methods</i> | 102 |
| 2.4.1. | <i>STD NMR experiments</i> | 102 |
| 2.4.2 | <i>HSQC NMR experiments</i> | 104 |
| 2.4.3 | <i>Chemical Shift Mapping</i> | 105 |
| 2.4.4 | <i>K_D and LE measurements</i> | 106 |
| 2.4.5 | <i>3D NMR</i> | 108 |
| 2.4.5.1 | Sequential assignments | 109 |
| 2.5 | <i>X-Ray Crystallography</i> | 116 |
| 2.5.1 | <i>Crystallization</i> | 116 |
| 2.5.2 | <i>Data Collection and Processing</i> | 119 |
| 2.5.2.1 | In-house data collection and processing | 119 |
| 2.5.2.2 | Synchrotron data collection and processing | 120 |
| 2.5.2.3 | Structure determination and Refinement | 121 |
| 2.6 | <i>ITC</i> | 123 |
| 2.7 | <i>Computational Methods</i> | 125 |
| 2.7.1 | <i>Library and Protein Preparation</i> | 125 |
| 2.7.2 | <i>Grid Generation and Docking</i> | 126 |
| 2.7.2.1 | NMR based docking | 127 |
| 2.7.3 | <i>Surface Grid Map Generation</i> | 128 |
| 2.7.4 | <i>Similarity Searching</i> | 129 |

| | | |
|---------|--|-----|
| 2.7.5 | <i>ParaSurf Surface Properties</i> | 130 |
| 3. | Inhibition of Oxidative Protein Folding in <i>Escherichia coli</i> | 132 |
| 4. | Application of NMR and computational approaches in structure assisted fragment screening to identify potential inhibitors of VcDsbA | 163 |
| 4.1 | Introduction | 164 |
| 4.1.1 | <i>Structural and functional features of VcDsbA</i> | 168 |
| 4.1.2 | <i>Previous work</i> | 170 |
| 4.1.2.1 | NMR-based fragment screening | 170 |
| 4.2 | Result and Discussion..... | 173 |
| 4.2.1 | <i>Comparative analysis of VcDsbA binders with EcDsbA FBS results</i> ... | 173 |
| 4.2.2 | <i>Computational Screening</i> | 176 |
| 4.2.3 | <i>NMR-based docking</i> | 181 |
| 4.2.4 | <i>3D NMR experiments as further validation</i> | 187 |
| 4.2.4 | <i>Surface grid map generation</i> | 191 |
| 4.2.7 | <i>Identification of novel VcDsbA hits</i> | 194 |
| 4.2.8 | <i>Inhibition analysis of VcDsbA hits</i> | 198 |
| 4.2.9 | <i>Application of Grid Maps</i> | 202 |
| 4.2.9.1 | Example 1 | 202 |
| 4.2.9.2 | Example 2 | 205 |
| 4.2.9.3 | Example 3 | 210 |
| 4.2.9.2 | Selectivity over EcDsbA inhibition | 215 |
| 4.3 | Conclusion | 217 |
| 5. | Elaboration of phenylthiophene and the phenoxybenzene fragments as potential inhibitors of EcDsbA | 219 |
| 5.1 | Introduction | 220 |
| 5.2 | Results and Discussion | 223 |

| | | |
|-----------|---|------------|
| 5.2.1 | <i>Characterization of phenylthiophene hits 1-3</i> | 223 |
| 5.2.2 | <i>Elaboration of phenylthiophene compounds 1-3</i> | 229 |
| 5.2.2.1 | <i>Analogues of the 2-phenylthiophene series</i> | 231 |
| 5.2.3 | <i>NMR guided molecular docking</i> | 238 |
| 5.2.4 | <i>Analogues of the 3-phenylthiophene series</i> | 246 |
| 5.2.5 | <i>Analogues of the phenoxybenzene series</i> | 259 |
| 5.2.5.1 | <i>Characterization of the phenoxybenzene hit.....</i> | 259 |
| 5.2.5.2 | <i>Elaboration of phenoxybenzene hits 4 and 5</i> | 264 |
| 5.3 | <i>Conclusion</i> | 281 |
| 6. | Structural Characterisation of Diverse Scaffolds Binding to Adjacent Sites of EcDsbA | 283 |
| 6.1 | <i>Introduction</i> | 284 |
| 6.2 | <i>Results and Discussion.....</i> | 286 |
| 6.2.1 | <i>Characterisation and elaboration of 6,5-fused rings</i> | 286 |
| 6.2.2.1 | <i>Characterization of benzofuran hit 6.....</i> | 287 |
| 6.3.2 | <i>Elaboration of benzofuran fragment 6.....</i> | 290 |
| 6.3.3. | <i>Elaboration of benzothiophene hit.....</i> | 297 |
| 6.3.3.1 | <i>Preliminary crystallographic trials for benzothiophenes</i> | 304 |
| 6.3.4. | <i>Elaboration of Chroman hit.....</i> | 308 |
| 6.3.5. | <i>Characterization and elaboration of phenylthiazoles</i> | 316 |
| 6.3.6. | <i>Characterization of phenol hit fragment 10.....</i> | 318 |
| 7. | Summary and Future Direction | 327 |
| 8. | Appendices | 333 |
| 8.1.1 | <i>E.coli DsbA-Peptide Complex.....</i> | 331 |
| 8.1.2 | <i>Pseudomonas aeruginosa DsbA – Structure and Characterization... ..</i> | 331 |
| 8.2 | <i>Mass spectrometry for EcDsbA and VcDsbA</i> | 332 |
| 8.3 | <i>Preliminary Assay Results</i> | 334 |
| 8.3.1. | <i>Alkaline Phosphatase Activity Assay.....</i> | 334 |
| 8.3.2. | <i>Fluorescence (FRET) Peptide assay</i> | 335 |

| | |
|---|------------|
| 8.3.2. <i>Insulin precipitation assay</i> | 335 |
| 8.3.3 <i>Motility Assay</i> | 336 |
| 8.4. Isothermal Calorimetry binding studies | 337 |
| 8.5. STD screening results for phenylthiophene hits..... | 338 |
| 8.6. <i>EcDsbA</i> -HSQC titration results for phenylthiophene hits | 339 |
| 8.7. <i>EcDsbA</i> -fragment and peptide binding shifts | 341 |
| 8.7.1 ^1H - ^{15}N HSQC chemical shifts of <i>EcDsbA</i> -fragments..... | 341 |
| 8.7.2 ^1H - ^{15}N HSQ C chemical shifts of <i>EcDsbA</i> -peptide..... | 346 |
| 8.8 <i>VcDsbA</i> -fragment binding shifts..... | 349 |
| 8.8.1 ^1H ^{15}N HSQC chemical shifts of <i>VcDsbA</i> -fragments | 349 |
| 8.8.2 ^1H , ^{15}N , $^{13}\text{C}_\alpha$ assignments of <i>VcDsbA</i> -fragment..... | 387 |
| References | 393 |

List of Abbreviations

| | |
|------------------|---|
| α | Alpha |
| β | Beta |
| δ | Difference |
| Δ CSP | Average chemical shift changes (perturbation) |
| $^{\circ}$ C | Degrees celsius |
| ClogP | Octanol-water partition coefficient |
| CT | Cholera toxin |
| Da/kDa | Dalton/ kiloDalton |
| DMSO | Dimethyl sulfoxide |
| DSB | Disulfide bond |
| <i>Ec</i> DsbA | <i>Escherichia coli</i> DsbA |
| EC ₅₀ | Half maximal effective concentration |
| FPLC | Fast protein liquid chromatography |
| FBS | Fragment based screen |
| FBDD | Fragment based drug discovery |
| Δ G | Changes in Gibb's free energy |
| g | Gram |
| GSSG | Glutathione (Oxidised) |
| GSH | Glutathione (Reduced) |
| Δ H | Changes in enthalpy |
| HAC | Heavy atom count |

| | |
|-------------------------------------|---|
| HPLC | High performance liquid chromatography |
| ^1H - ^{15}N HSQC | ^1H - ^{15}N Heteronuclear single quantum correlation |
| HTS | High throughput Screening |
| K_D | Dissociation constants |
| K_i | Inhibition constant |
| kJ | kilo Joules |
| k_{off} | Dissociation rate constant |
| k_{on} | Association rate constant |
| μg | Microgram |
| μL | Micro litre |
| μM | Micro molar |
| L | Litre |
| LE | Ligand efficiency |
| M | Molar |
| MAD | Multi-wavelength anomalous diffraction |
| MHz | Megahertz |
| min | Minute |
| mL | Milli litre |
| mM | Milli molar |
| mol | Moles |
| ms | Milli seconds |
| MW | Molecular weight |
| nm | Nanometers |

| | |
|---------------|---|
| NMR | Nuclear magnetic resonance |
| NOE | Nuclear overhauser effect |
| NOESY | Nuclear overhauser enhancement spectroscopy |
| <i>PaDsbA</i> | <i>Pseudomonas aeruginosa</i> DsbA |
| PDB | Protein Data Bank |
| PDI | Protein disulfide isomerase |
| ppm | Parts per million |
| R | Gas constant |
| rms | Root mean square |
| RMSD | Root mean square deviation |
| rpm | Revolutions per minute |
| ΔS | Changes in entropy |
| SBDD | Structure based drug design |
| STD | Saturation transfer difference |
| SPR | Surface Plasmon Resonance |
| 3D | Three dimensional |
| 2D | Two dimensional |
| T | Temperature |
| TS | Thermal shifts |
| <i>VcDsbA</i> | <i>Vibrio Cholerae</i> DsbA |
| V | Volts |
| vdW | van der <i>Waals</i> |
| <i>XfDsbA</i> | <i>X. fastidiosa</i> DsbA |

Chapter 1

General Introduction

1. General Introduction

1.1 Gram-negative bacterial infection

Despite the discovery of antibiotics more than half a century ago microbial infection remains a major cause of death worldwide.¹ The rapid emergence of antimicrobial resistance in bacteria to many current antibiotic agents is a serious threat to global public health.² In fact, some bacterial pathogens have developed resistance to almost all currently used antibiotics. The increase in the worldwide occurrence of resistance in pathogenic Gram-negative bacteria is of particular importance due to lack of adequate therapies against these organisms.³ A recent wide-ranging study found that about 51% of hospital patients in intensive care units on a given day had some sort of bacterial or fungal infection.⁴ Gram-negative bacterial infections constituted 63% of these.⁵ The pharmaceutical industry model favors drug development for chronic needs due to economic pressures, rather than the acute use typical of antibiotics. There is also a disincentive to enter the field of antimicrobial drug development due to the likelihood of rapid development of resistance, which limits the useful lifespan of any potential drug. Therefore the likelihood for the discovery of new antibiotics acting against Gram-negative pathogens in the near future is not good.⁶ The molecular targets and mechanisms of action of most of the currently existing antibiotics are based on the interruption of the growth cycle by restraining the synthesis and assembly of key components of bacterial processes such as DNA replication, cell wall and protein synthesis.^{2b} Thus current antibiotics elicit their effects by either killing (bacteriocidal) or preventing the growth (bacteriostatic) of bacteria. Even though these strategies are

highly successful, they impose selection pressures on the target bacterium, which rapidly develop resistant mutants that are able to grow in the presence of the antibiotic.⁷ The majority of the new Gram-negative antibacterial drugs in preclinical or clinical trials are either derivatives of existing antibiotics which are moderately more stable to degradation by bacterial enzymes such as beta-lactamase, or drugs which are formulated in such a way as to attain higher local therapeutic levels to facilitate overcome resistance but only transiently.⁷ Antibacterial agents that act by completely novel mechanisms are in particularly short supply. Novel strategies are therefore needed to discover and validate new antibiotic targets as well as alternatives to conventional antibiotics to overcome antibiotic resistance in Gram-negative bacteria.

1.2 DsbA and virulence in bacteria

Gram-negative bacteria produce and secrete a variety of proteins in order to be biologically active as well as to cause infection and disease in the human host.⁸ These specific proteins are termed virulence factors, and they play an essential role at one or more stages of the bacterial infection to facilitate host pathogen interactions. The majorities of virulence factors are found on the surface of the bacterial cell or are secreted into their immediate environment. In order to activate a detrimental effect on the host and contribute to the symptoms of disease, pathogenic bacteria use a cascade of virulence determinants such as enterotoxins, haemolysins, herotoxins, adhesions and fimbriae.^{1b} These secreted toxins and enzymes such as proteases can manipulate, damage and even kill host cells. Examples include cholera toxin from the enteric

pathogen *Vibrio cholerae*⁹, pertussis toxin from the respiratory pathogen *Bordetella pertussis*,¹⁰ and heat-labile enterotoxin from enterotoxigenic *Escherichia coli*.¹¹ The targeting of virulence factors is controlled by a relatively small number of secretion systems which direct functional virulence factors to their site of action at the suitable time to cause an infection.¹² For example, many Gram-negative pathogens contain the type II secretion pathway to secrete enzymes that contribute to their ability to cause disease *i.e.* elastase and lipase of *P. aeruginosa* during lung infection in cystic fibrosis patients.¹³

The wide diversity and differing functional mechanisms of virulence factors within the same or between different bacteria, represents a significant challenge in developing anti-virulence agents as anti-microbial drugs. However, most virulence factors are proteins secreted into the harsh extracellular surroundings and thus share a requirement that they remain stably folded in order to maintain their activity. Perhaps as a consequence, the tertiary structure of many of these exported proteins is stabilized by the formation of structural disulfide bonds (DSB) between participating cysteine residues within the polypeptide chain.¹⁴ Without an efficient disulfide bond formation these proteins are subject to rapid degradation and loss of activity.^{1b, 15} For instance the molecular chaperone of fimbriae produced by uropathogenic *E.coli* (UPEC), the pertussis toxin of *Bordetella pertussis*, the secretin component of the type III secretion system in *Yersinia pestis* and a flagellar component in *E. coli* all contain disulfide bonds that require oxidative folding to be functionally active and stable.¹⁶

Both intramolecular, and to a lesser extent intermolecular, disulfide bonds may be formed spontaneously through the air oxidation of thiols in susceptible pairs of cysteine residues without involvement of any enzymatic catalyst. However this is an extremely slow process *in vitro* and does not necessarily occur between the correct combinations of cysteines.¹⁷ *In vivo* this process occurs much more rapidly and reliably through the catalytic intervention of a class of enzymes called oxidoreductases, particularly those of the glutaredoxin and thioredoxin super families.¹⁸ This super family includes protein disulfide isomerase (PDI) in eukaryotes and DSB (disulfide bond formation) protein family in prokaryotes.^{15a}

The periplasmic formation of disulfide bonds in Gram-negative bacteria, requires the removal of two protons and two electrons from two participating cysteine residues, and is catalyzed by the dithiol-disulfide oxidoreductase enzyme DsbA.^{1b, 14a, 15a} DsbA catalyzes this process by a disulfide exchange mechanism with a pair of cysteines within its own structure. Therefore this enzyme is responsible for the correct oxidative folding of many substrates and consequently the virulence of a wide range of Gram-negative pathogenic bacteria. Bacteria deficient in a functional DsbA (*dsbA*⁻ mutants) display pleiotropic avirulent phenotypes since the oxidative folding of a large number of disulfide-bonded proteins is impaired.⁹ For example, *dsbA*⁻ *V. cholerae* are unable to produce cholera toxin and the toxin-coregulated pili; *dsbA*⁻ strains exhibit a non-motile phenotype caused by misfolding of the P-ring protein of the flagella motor in *E. coli*. Furthermore, *dsbA*⁻ *E. coli* exhibit reduced levels of β -lactamase activity and are hypersensitive to benzylpenicillin, dithiothreitol and some divalent metal cations.^{13b, 19}

Previous studies have suggested that DsbA acts as central regulator of virulence in many Gram-negative bacteria and therefore represents a potential drug target for the rational drug design of inhibitors that could counteract the virulence of these organisms and may provide a clinical benefit as effective antibacterial drugs.^{1b, 20} Moreover, bacterial virulence properties have been suggested as attractive targets for the development of new antibacterial agents.²¹ Since inhibitors of virulence may not inhibit general bacterial growth, this approach could decrease the selection pressure that drives the development of resistance to currently available antibiotics.²² To date there are no reported selective and potent inhibitors of the DsbA enzyme, thus there exists the potential to develop novel lead molecules against pathogenic bacteria that have been difficult to treat with existing antibiotics.

1.2.1 Oxidative catalytic cycle of DsbA

All of members the DSB protein family shares a conserved redox active-site sequence Cys-X-X-Cys which have some of the highest redox potential values (~ -120 mV) among all the known thiol–disulfide oxidoreductases and contributes to its oxidative reactivity.²³ In the Gram-negative bacterium *E. coli*, DsbA is an soluble protein of 21 kDa (*EcDsbA*) and is responsible for oxidative folding of substrate polypeptides within the periplasmic space.^{20a} *EcDsbA* represents the prototype DsbA enzyme as it has been the most studied system among all DSB proteins. In the catalytic cycle of *EcDsbA* two cysteine residues (Cys30 and Cys33) in the active site shuffle between their dithiol and

disulfide forms.^{1b, 20} The two residues between the active site cysteines Pro31 and His32 in *EcDsbA* along with a conserved proline residue with a *cis* configuration of its amide bond (*cis*-Pro151) have been revealed to be important in the catalytic activity of *EcDsbA*.²⁴ Previously, site-directed mutagenesis studies by Zapun A. *et al*, identified the specific role of the active site Cys30 in the catalytic cycle of DsbA.²⁵ The replacement of either one or both cysteine residues by serine was carried out and Cys30 was shown to have an accessible thiol group, while Cys33 is shielded from the solvent.²⁵ Therefore through its reactive Cys30, DsbA forms a mixed disulfide bond to a cysteine within a substrate protein. The mixed disulfide is resolved to produce an oxidised substrate protein and release reduced (dithiol form) DsbA. Biochemical studies have shown that the role of DsbA *in vivo* mainly to act as an oxidant²⁶ however DsbA can also act as catalyst *in vitro* to reduce insulin in the presence of DTT similar to other members of the thioredoxin super family.^{18, 27}

The oxidising properties of DsbA result in part from a strained conformation of its disulfide bond as well as the extremely low pKa (3–3.5) of Cys30 at the active site.²⁸ In order to complete the catalytic cycle the active site cysteine pair needs to be reoxidised back to the disulfide form. In *E. coli*, the reduced DsbA selectively reacts with the cytoplasmic membrane-bound protein DsbB and regains its critical intramolecular disulfide bond to return to its oxidized, active form as depicted in Figure 1-1 below.^{17, 29}

Apart from DsbA and DsbB, five other redox proteins – DsbC to DsbG – have been identified in *E.coli* which performs various redox and isomerase roles. They are

located either in the periplasm or embedded in the inner membrane with a functional domain on a periplasmic loop.^{13b, 27, 30} These redox proteins lack any overall homology with DsbA but share the putative active site -Cys-X-X-Cys motif. In the protein folding cycle there are least two functional pairs of enzymes, DsbA/DsbB and DsbC/DsbD that results in two different pathways (Figure 1-2).³¹ In first redox pathway DsbA catalyses oxidative protein folding while DsbB selectively reoxidises DsbA. In the second pathway of isomerisation, the periplasmic DsbC protein has been shown to act as a disulfide bond isomerase to reduce misfolded proteins.³² In this isomerisation pathway DsbC reshuffles incorrectly formed disulfide bonds in exported proteins, whilst DsbD, an inner membrane protein, maintains DsbC in its reduced state.³²

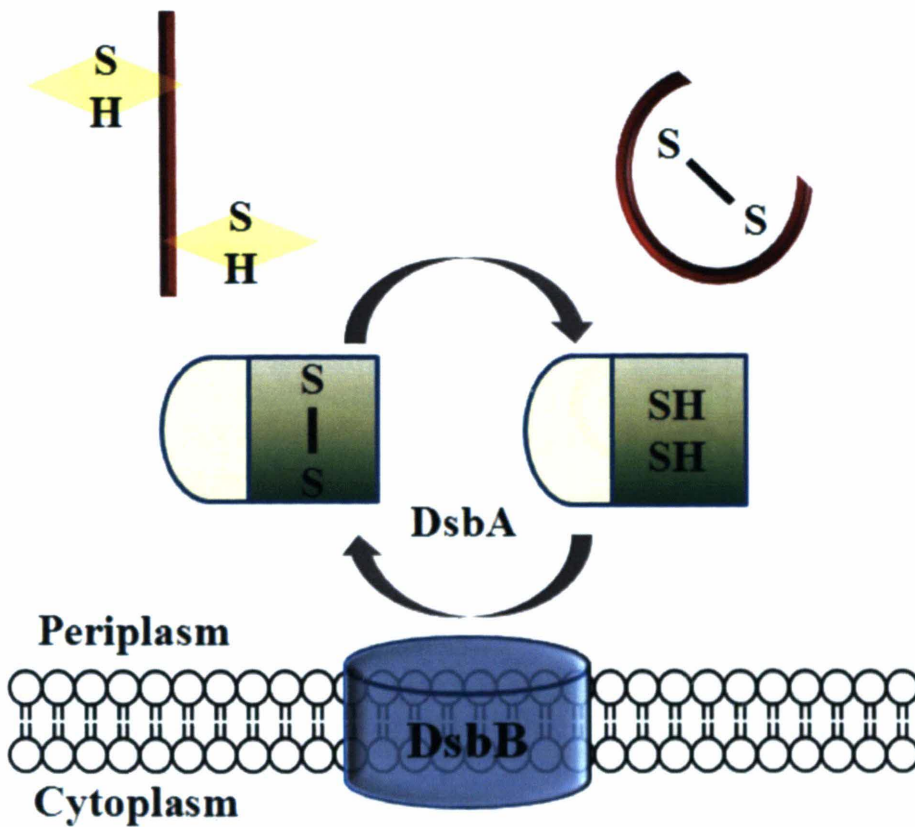


Figure1-1: Schematic representation of the catalytic cycle of DsbA. Oxidised DsbA reacts with a variety of reduced substrate proteins to generate a folded and oxidised substrate and reduced DsbA. Reduced DsbA reacts specifically with membrane-bound DsbB to form a mixed disulfide DsbA-DsbB complex which results in reoxidation of DsbA and reduction of DsbB. (reproduced from reference 1b)

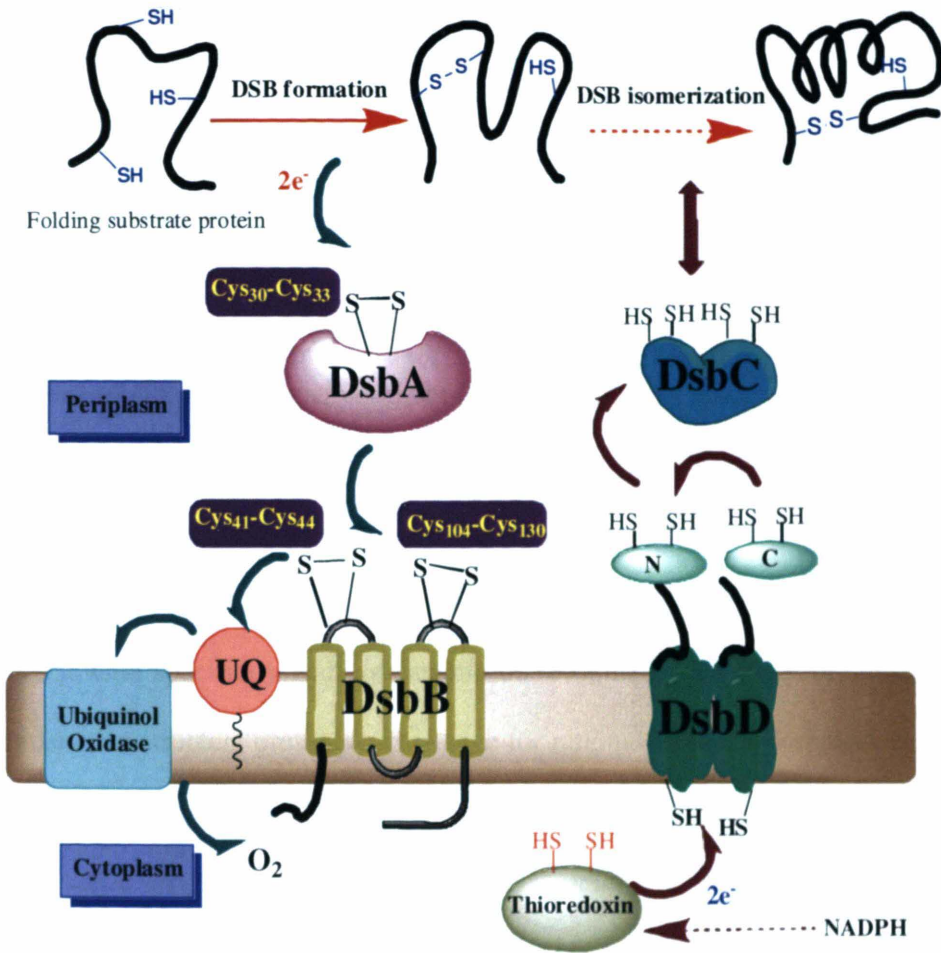


Figure 1-2: Oxidation and isomerisation pathways in the *E. coli* periplasm. Reduced proteins are oxidized by periplasmic DsbA in a rapid and unidirectional process to catalyse their folding. DsbA is reoxidized by the inner membrane protein DsbB. Under aerobic conditions, DsbB is reoxidized by a bound quinone, which ultimately derives oxidative power from molecular oxygen in a cytochrome oxidase- and ubiquinone-dependent reaction. DsbC reshuffles incorrectly formed disulfide bonds in exported proteins and DsbD, an inner membrane protein, maintains the activity of DsbC. Green and brown arrows show the directions of electron flow in DsbA-DsbB and DsbC-DsbD pathways respectively. (Reference 17)

1.3 Structural studies

1.3.1 *DsbA structure*

Protein crystallography has provided a detailed view of the 3D structures of DsbA enzymes available from several different Gram-negative bacteria.^{17, 20a, 33} These enzymes often have low sequence identity, but maintain a conserved structure. These structures share an analogous motif, known as the thioredoxin fold, which is common to many eukaryotic and prokaryotic oxidoreductase enzymes.^{19b, 20a, 31a, 33} Most oxidoreductases have very wide specificity and are capable of recognizing a variety of substrates. This resemblance in structure presents a challenge for rational inhibitor design. A combination of the broad specificity, widespread occurrence and the similarity in the structures of oxidoreductase enzymes makes it apparent that inhibitors targeting surfaces contained entirely within the thioredoxin domain might be expected to have poor selectivity. However, there are a number of areas in which DsbA enzymes are distinct from many other eukaryotic and prokaryotic oxidoreductase enzymes.^{31a} These significant differences indicate that it may be possible to develop DsbA inhibitors selective for the bacterial enzyme.

X-ray crystallographic studies of *EcDsbA* identified the presence of the thioredoxin-like domain (comprising residues 1-62 and 139-189) and notably, revealed an additional globular domain which forms a cap over the active site as depicted in Figure 1-3.^{20a} The structure of DsbA contains a unique α -helical domain (residues 63-138) which is inserted into the thioredoxin fold.^{20a} . The redox active site of *EcDsbA* lies in a cleft at the interface of the two domains and comprises a highly conserved primary sequence motif Cys30-Pro31-His32-Cys33 at the *N*-terminal end of the α 1 helix. In the oxidized form a disulfide bond links Cys30 and Cys33. Through the catalytic cycle, these active site cysteine residues shuttle between their oxidised, substrate-bound, reduced dithiol and DsbB bound forms. Additionally, DsbA contains a number of surface features which have been implicated in functional activity and appear to be at least partially conserved across DsbA enzymes from different bacteria.³⁴ The four striking surface features include a peptide binding groove, a hydrophobic pocket and a hydrophobic patch, which are all on the same face of the protein surrounding the catalytic site.³⁴ On the opposite face to the active site disulfide, there is a region formed by mainly acidic residues called the acidic patch (Figure 1-4).

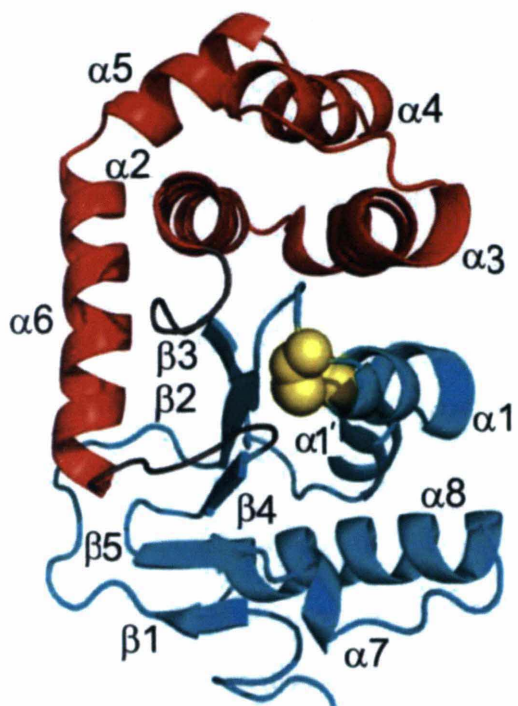


Figure 1-3: Ribbon diagram depicting the structure of *EcDsbA* (PDB code 1fvk, reference 20a). In *EcDsbA* thethioredoxin (cyan) and α -helical (red) domains are linked at two insertion points (gray). Elements of structure are numbered sequentially from the *N*-terminus. The Cys30-Cys33 residues of the active site are shown in (yellow) CPK representation.

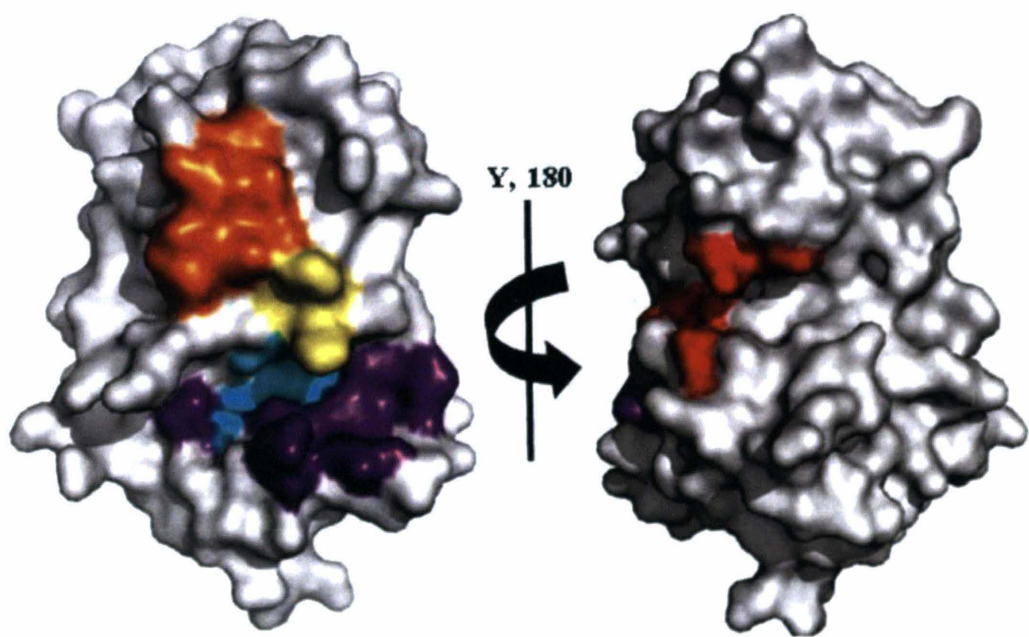


Figure 1-4: The surface features of *EcDsbA* highlighting active site (yellow), hydrophobic patch (orange), hydrophobic pocket (cyan), hydrophobic groove (violet) and acidic patch (red) which is on the opposite side to the active site. (Reference 34)

1.3.2 *Difference among DsbA homologues*

Recent structural and biochemical findings demonstrate significant differences exist amongst DsbA homologues. Several DsbA structures has been reported from different species of pathogenic bacteria such as Gram-negative *E.coli*^{20a}, *Vibrio cholerae*²⁶, *Pseudomonas aeruginosa*³⁵, *Salmonella typhi*³⁶, *Neisseria meningitides*³⁷, *Bordetella parapertussis*³⁸, *Wolbachia pipientis*³⁹ and Gram-positive *Staphylococcus aureus*⁴⁰ and *Bacillus subtilis*.⁴¹ *EcDsbA* is the best characterized member of the DsbA family. All the structurally characterized DsbA molecules contain both the thioredoxin domain and an inserted helical domain along with the catalytic site (CXXC) and *cis*-Pro motifs. These similarities are apparent when the tertiary structure of *EcDsbA* is compared with those of two gram-negative bacteria, *P. aeruginosa* and *V. cholera* as shown in Figure 1-5. The structure of each of these DsbA enzymes has been solved in the oxidised form as well as the reduced form for *VcDsbA* and *EcDsbA*. Each contains a thioredoxin (TRX) domain, an inserted helical domain. *VcDsbA* and *PaDsbA* share relatively low sequence similarity of ~40% and 30% respectively with *EcDsbA*, however retain conserved surface features around their active sites.^{26, 35} However these structures have some significant differences in their surface features and the most obvious region of variability among DsbA proteins is in the loop connecting $\beta 5$ and $\alpha 7$ (Figures 1-3 and 1-5) which has been proposed to be involved in substrate binding and contributes to a region referred to as the "peptide binding groove". For example in *VcDsbA* this groove is considerably shorter due to a six-residue deletion, and the hydrophobic pocket is shallower as compared to *EcDsbA*. In the case of *PaDsbA* this region is flattened and truncated, in

addition to having different *cis*-Pro loop residues, a smaller hydrophobic patch and a more basic surface in contrast to *EcDsbA*. The surface electrostatics of these three DsbA proteins (Figure 1-5) differs tremendously.^{26, 35} Despite these differences in the structures of the enzymes, functional studies have shown that *VcDsbA* and *PaDsbA* are able to complement a *dsbA*⁻ strain of *E. coli*. (Please refer Appendices section 8.1.2 for more details on this study) This suggests that each DsbA is able to bind to and oxidise the wide variety of substrate proteins that are secreted to the *E. coli* periplasm.

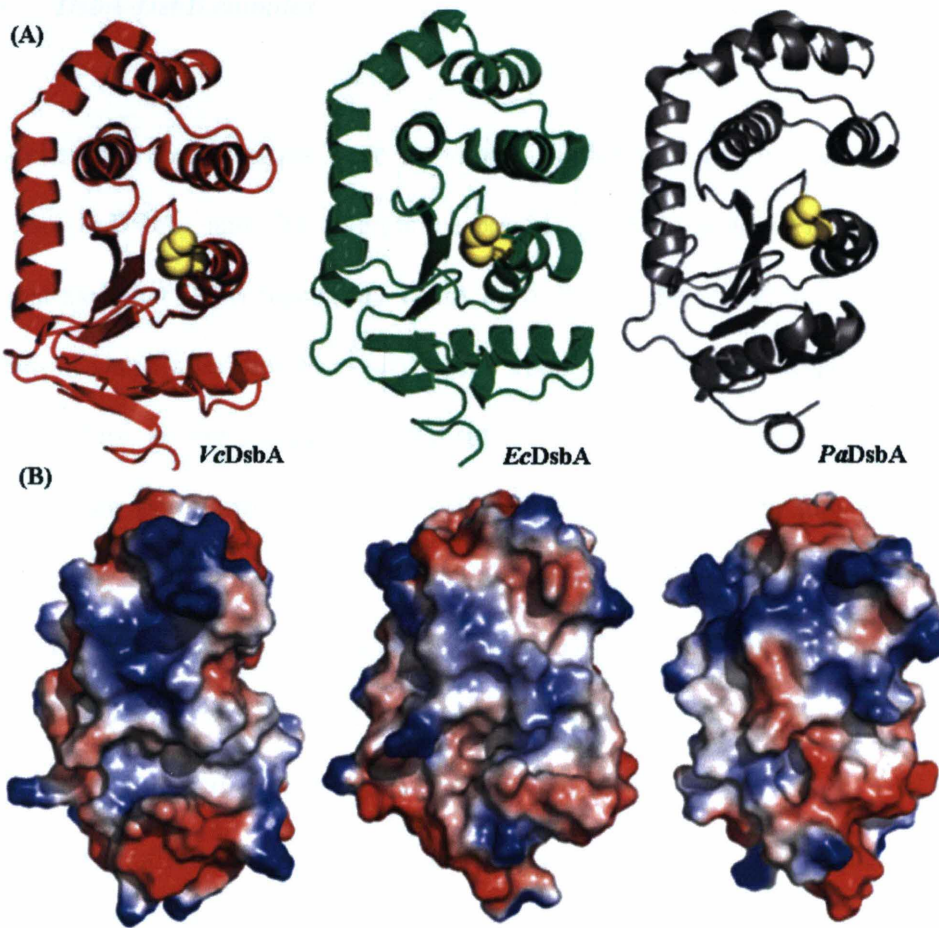


Figure 1-5: A comparative structural view of three different DsbA enzymes. Ribbon diagram of (A) *VcDsbA*-red (PDB code: 1BED), *EcDsbA*-green, (PDB code: 1FVK) and *PaDsbA*-grey (PDB code: 3H93). Sulfur atoms of the catalytic cysteines of all structures are shown as yellow spheres. (B) Surface electrostatics of above three structures in similar order and orientation.

1.3.3 *DsbA-DsbB complex*

A number of recent studies have provided insights into the interaction mechanisms between *EcDsbA* and its membrane-bound partner protein *EcDsbB*.⁴² It was demonstrated that DsbB reactivates DsbA by oxidizing the Cys30 and Cys33 active-site residues of DsbA after they become reduced upon substrate protein oxidation.^{42b, 43} *dsbB* mutants exhibit severe deficiencies in disulfide bond formation and accumulate reduced DsbA in their periplasm.^{15b, 44} DsbB (MW 20 kDa) is integrated into the cytoplasmic membrane by its four transmembrane helices which are linked by two periplasmic loop (P1 and P2). These loops contains a pair of cysteines, Cys41 and Cys44 in the *N*-terminal loop P1 and Cys104 and Cys130 in the *C*-terminal loop P2 that are essential for function.⁴² In the resting state DsbB contains Cys104-Cys130 disulfide. The active site residue Cys30 of reduced DsbA exists as a thiolate anion due to its low pKa value, which attacks Cys104 of DsbB to form a mixed disulfide complex. This is then followed by nucleophilic attack at Cys30 by Cys33 of DsbA to form oxidized DsbA and release reduced DsbB. A conformational change in DsbB following reaction with DsbA results in separation of Cys130 from the mixed disulfide and this cysteine relocation prevents backward resolution of the complex and allows Cys130 to approach and activate the disulfide-generating reaction center composed of Cys41, Cys44, Arg48 (DsbB) and ubiquinone (Figure 1-6). Cys41 and the Cys44 of DsbB is reoxidised ubiquinone or menaquinone dependent on whether the reaction occurs under aerobic and anaerobic conditions. The flexible and broad hydrophobic groove of DsbA allows DsbB Cys104 to get close enough to DsbA Cys30 for the reaction to occur and the Pro100 - Phe106

portion of the P2 loop of DsbB binds with Pro151, Pro163, Gln164, Thr168, Met171 and Phe174 of this groove.^{42b} Additionally the periplasmic loop of DsbB interacts with Arg148-Gly149-Val150 in the *cis*-Pro loop of DsbA and side chain of His32 makes vdW interactions with Ala102-Thr103 in the DsbB loop. The structural data of the DsbA–DsbB complex revealed the mechanism by which protein disulfide bonds are generated *de novo* by the cooperation of DsbB and how these bonds are transferred selectively to DsbA.^{42b}

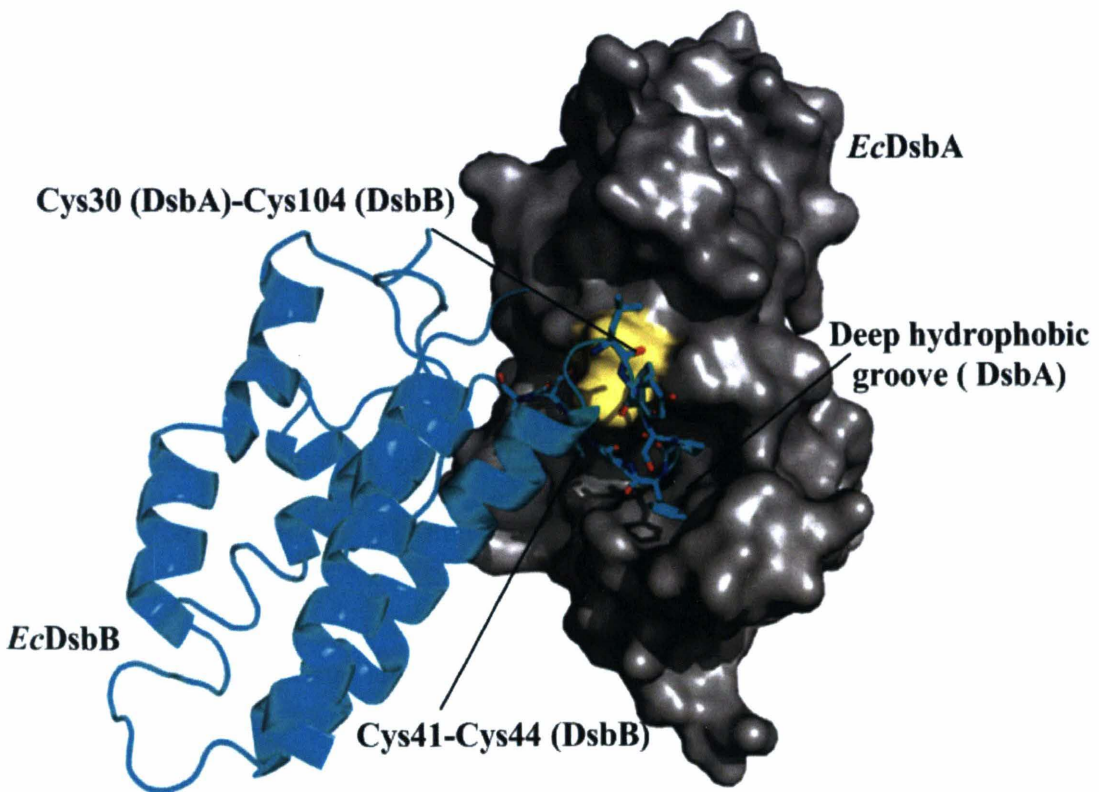


Figure 1-6: X-ray crystallographic complex of membrane protein DsbB (cyan) binding to periplasmic DsbA (grey) inside the hydrophobic groove. The active site Cys30 of DsbA is depicted in yellow. (PDB code 2ZUP) [Reference 29]

1.3.4 DsbA - substrate

Biochemical studies have revealed that DsbA reacts preferentially with unfolded polypeptide substrates and that the non-covalent interactions between DsbA and peptide substrates are reasonably weak.^{29a} Furthermore, *E. coli* have express hundreds of proteins that are predicted to enter the periplasm which contain at least one pair of cysteines and are thus potential substrates of *EcDsbA*.³⁴ Thus *EcDsbA* is expected to have a very broad substrate specificity.⁴⁵ Comparison of the structure of *EcDsbA* with those of thioredoxin bound to two different substrates led to the suggestion substrate proteins bound to DsbA via interaction with the hydrophobic groove, which is also the site of interaction with membrane bound protein DsbB.^{42b} To investigate the interaction of *EcDsbA* with its substrates our group recently reported the structure of a covalently trapped *EcDsbA*-peptide complex.^{19c} These data suggested that instead of binding within the hydrophobic groove, the peptide binds to *EcDsbA* at the interface between the α -helical and thioredoxin domains in the proximity of the hydrophobic patch. The role of the hydrophobic patch and the acidic patch in the catalytic activity of DsbA was previously unclear; however these data suggest that oxidized DsbA may bind to substrates in a different way than reduced DsbA binds DsbB in *E. coli* (Figure 1-6). The crystallographic complex revealed that binding of the peptide and DsbA was stabilised by several H-bonds and van der Waals interactions along with the covalent linkage to Cys30 of *EcDsbA*. The binding location of the peptide was also determined by measuring chemical shift perturbations in ^1H - ^{15}N heteronuclear single quantum correlation NMR experiments on uniformly ^{15}N -labeled *EcDsbA* in the absence and presence of the peptide (Figure 1-7).^{19c}

A peptide containing homoserine in place of cysteine (SigA) was used in the NMR studies to prevent formation of a covalent complex. Analysis of those resonances that were perturbed upon formation of the non covalent complex revealed that many of the most significant perturbations were observed for residues that formed a continuous surface at the interface between the α -helical and TRX domains (Figure 1-7). (Please refer Appendices section 8.1.1 and 8.7 for more details on this study). Additionally, few of the residues in the flexible hydrophobic groove that is the binding site for the periplasmic loop of DsbB were observed in the NMR spectrum but not significantly perturbed. These studies suggested that the DsbA substrate binds in a location that is distinct from the binding site observed in the crystal structure of the DsbA-DsbB complex.^{42b}

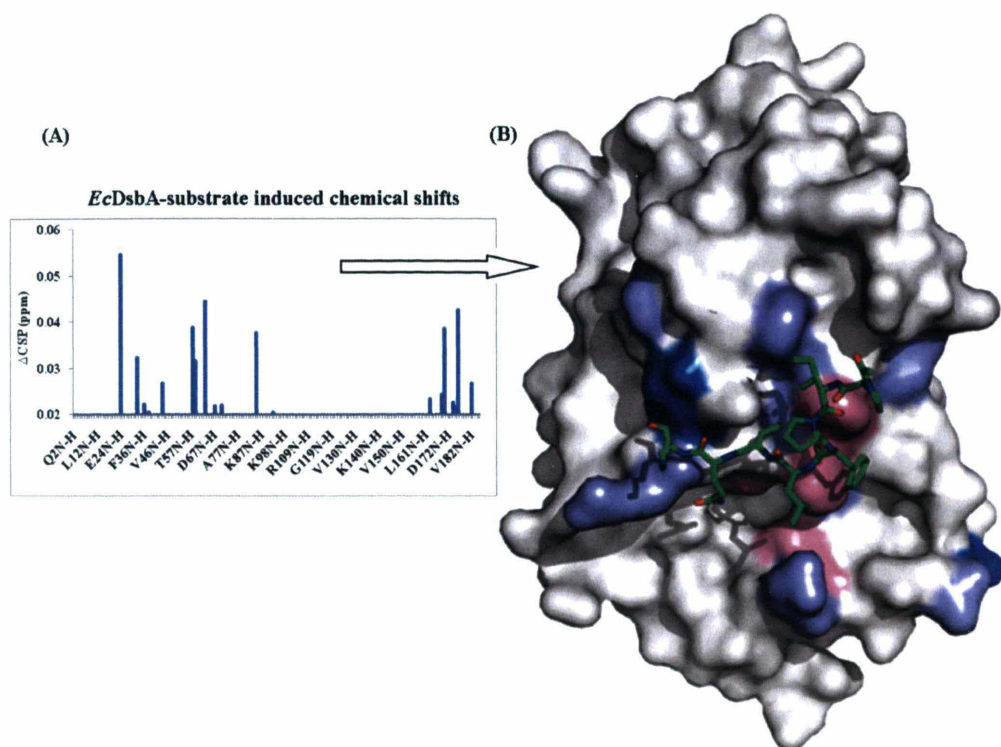


Figure 1-7: (A) Histogram analysis of chemical shift changes induced in the HSQC spectra of *EcDsbA* protein upon the addition of a peptide substrate at 1 mM concentration. (B) Structure of the *EcDsbA* - peptide complex (PDB Code 3DKS; Chain C & E) shown as a white surface. The peptide is shown in green stick format. Residues whose chemical shifts displayed the greatest and modest perturbations are coloured in dark blue and light blue and those are unassigned are coloured in pink. The perturbed residues from the non-covalent complex in solution form an extended surface that coincides with the binding site defined in the crystal structure.

More recently, the crystal structure of a plant pathogen *X. fastidiosa* DsbA (*XfDsbA*) was reported, which suggested that peptide interactions occur in the active site region and the helical domain mainly through hydrophobic contact.⁴⁶ However this complex was formed serendipitously as peptide was not included in the crystallization condition and hence its exact sequence was not known.⁴⁶ As a result it could not be unambiguously fitted into the extra electron density observed in the structure of the *XfDsbA*.⁴⁶ Nonetheless these studies suggested that the peptide bound to *XfDsbA* in a similar site to that described for the P2 loop of *EcDsbB* interacting with *EcDsbA*, specifically with regard to anti-parallel interactions with residues of the *cis*-Pro loop.⁴⁶ Thus, the precise mechanisms of molecular recognition of substrates by *EcDsbA* remain unresolved.

A common finding from the DsbA-DsbB and DsbA-peptide complexes solved to date is that all substrate interactions involve regions surrounding the CXXC catalytic motif.⁴⁷ Thus, the active site motif not only helps control the redox potential of the enzyme, it also contributes to interactions with redox partners.⁴⁸ The *cis*-Pro loop near the active site is also important for substrate binding, including the interaction with *EcDsbB*.^{42b} The hydrogen bonding pattern between bound peptides and the *cis*-Pro loop is similar for both *EcDsbA* and *XfDsbA*. Together the CXXC and *cis*-Pro loop form the central binding site for substrates, with the parts of the nearby groove and the surface of the helical domain occasionally providing additional interaction surfaces.⁴⁹

1.4 Functional and biochemical characterization of DsbA

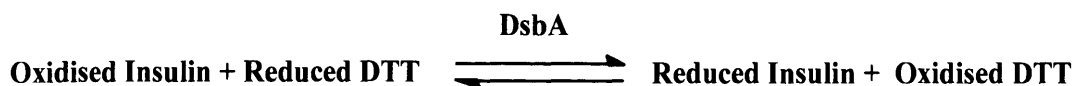
Thioredoxin (TRX) proteins such as DsbA react with a wide range of protein substrates. TRX-containing oxidoreductase enzymes catalyse disulfide bond formation, isomerisation and reduction, depending on the nature of the protein disulfide and the overall thiol/disulfide redox potential. In all cases they can be regarded as catalysing interchange reactions between thiols and protein disulfides. A wide range of assays have been reported that characterize the catalytic properties and oxidoreductase activity of DsbA *in vitro*. Additionally several phenotypic assays have been reported to estimate the activity of DsbA within bacterial cells. A few of these assays will be described here briefly (please refer methods chapter 2 for more details).

1.4.1 Biochemical *in vitro* assay

1.4.1.1 Insulin reduction assay

The insulin assay was developed by Holmgren to study the reaction mechanism of thioredoxin with dithiols in insulin.⁵⁰ This is a facile spectrophotometric assay which can directly measure disulfide oxidoreductase activity.⁵⁰ This assay is based on the activity of oxidoreductase enzymes such as DsbA to catalyse the reduction by DTT of the inter-chain disulfide bonds between the two insulin subunits. Insulin is comprised of two peptide chains namely, the A and B chain, connected by two disulfide bonds. Upon reduction, individual A and B chains are generated and once the free B chain reaches a high enough concentration it precipitates out of the solution. This is the driving force of

this reaction which shifts the equilibrium towards the formation of reduced insulin and oxidised DTT in the presence of DsbA.



Precipitation of the B-chain increases turbidity in the solution, which can be quantified spectrophotometrically. This assay is efficient and simple and can be performed as 96-well plate format.⁵⁰

1.4.1.2 Fluorescence quenching assay

A fluorescence assay has been developed for the study of the oxidative activity of protein disulfide isomerase and DsbA, that uses a specially designed and synthesized fluorescent peptide.⁵¹ This assay is based on the fluorescence change in the peptide upon oxidation. The peptide contains a natural substrate sequence with one pair of cysteine residues. This peptide contains a fluorescent group attached at one terminus and an appropriate fluorescence quenching attached to the other. Upon oxidation by DsbA, the quencher and the fluorophore are brought to close proximity and the fluorescence is quenched leading to a decrease in fluorescence intensity as shown in Figure 1-8.⁵² This assay can be used to measure the oxidoreductase activity of DsbA.⁵³

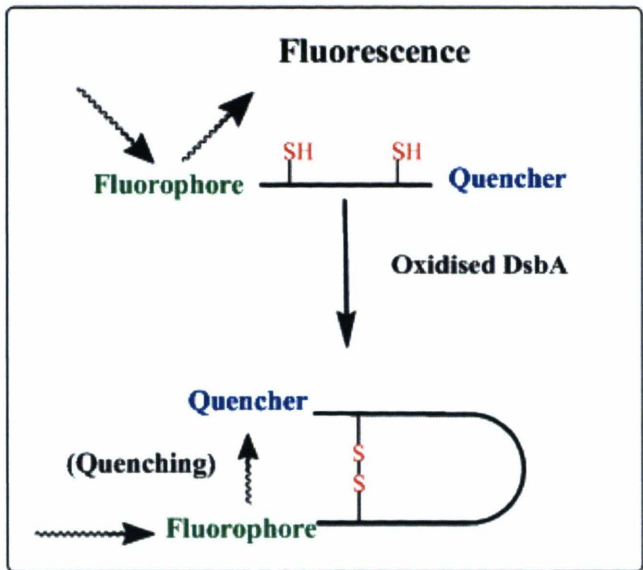


Figure 1-8: Schematic of oxidation of fluoro-peptide by DsbA in fluorophore-quencher system

A limiting factor with each of these assays is that neither is strictly quantitative. In the insulin assay, neither the time taken to generate observable precipitate, nor the rate of insulin precipitation is directly proportional to the concentration (or activity) of DsbA. In the peptide assay, it is under most conditions the reoxidation of DsbA to complete the catalytic cycle that is rate-limiting. Therefore neither assay provides a robust and straightforward measure of DsbA activity.

1.4.2 Phenotypic assays

1.4.2.1 Bacterial motility assay

The bacterial motility assay measures the swarming motility of *E. coli*. In many bacterial pathogens including *E. coli* motility is crucial for virulence as it allows the infecting population to spread within the host. This phenotype is regulated by a cell organelle called the flagellar motor.⁵⁴ The motility phenotype is modulated by the DsbA-dependant folding of a protein subunit FlgI that forms the periplasmic (P-) ring of the flagellar motor^{1b, 55} (Figure 1-9).

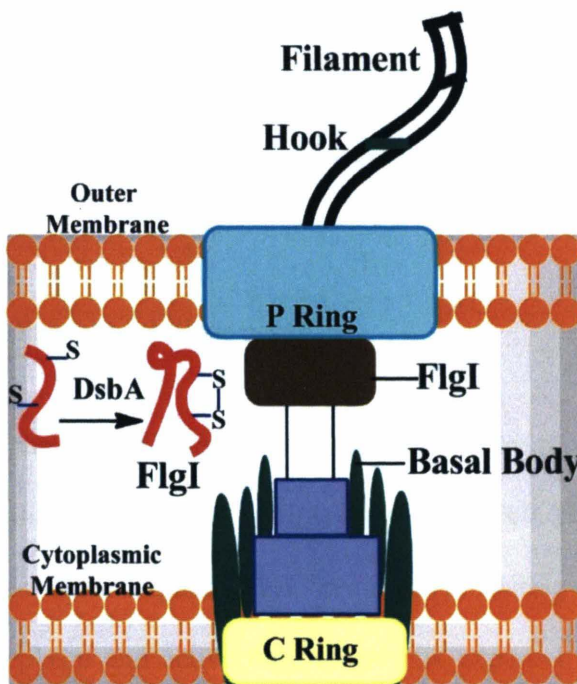


Figure 1-9: Depiction of bacterial flagella system in a simplified format. In many gram negative bacteria cellular spread and survival is related to motility hence virulence. DsbA catalyses the formation of a disulfide bond for the correct folding of the FlgI protein (highlighted in red) which is crucial for virulence. (Adapted from references 1b and 55)

Bacteria deficient in a functional DsbA are non-motile since the oxidative folding of protein FlgI is impaired and without its disulfide bond, FlgI is rapidly degraded.⁴⁷ Therefore motility can be a direct measure to observe DsbA activity.

1.4.2.2 Heavy metal viability assay

The heavy metal viability assay measures the effects of enrichment of Cd^{2+} in the media on the viability of *E. coli*.⁵⁶ In bacterial cells, DsbA has been demonstrated to reduce Cd^{2+} induced toxicity and oxidative stress that causes disruption in the proper folding in many proteins.⁵⁷ DsbA relieves this toxicity, to a certain extent, by forming DSB rapidly in the proteins which reduces the exposure of the free thiol to cadmium.⁵⁸ Consequently, *dsbA*⁻ strains of bacteria are more sensitive to cadmium toxicity.

1.4.2.3 Alkaline phosphatase assay

The alkaline phosphatase (AP) assay measures the activity of DsbA indirectly by observing its capability to oxidise the periplasmic protein AP. Native AP contains a pair of cysteines residues which require DsbA for oxidative folding and activity.^{55b, 59} AP is inactive in the cytoplasm and active only when it is exported to the periplasm, where formation of intramolecular disulfide is catalysed by DsbA. This reaction can be quantified and measured spectrophotometrically in presence of a substrate of AP (e.g. p-nitro phenol).^{55b, 59}

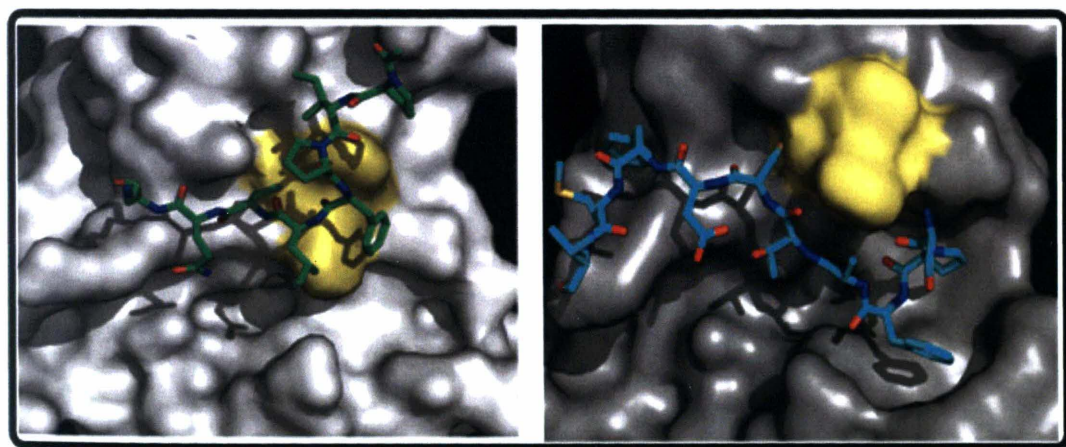
The phenotypic assays described above can be utilised for the quantitative study of these DsbA dependant bacterial cell properties. Comparison of wt vs *dsbA*⁻ strains of *E. coli* enable the sensitivity of the phenotype to DsbA activity to be determined. Furthermore, measurements in the presence of potential inhibitors allow EC50 values for the inhibitors to be determined. However each of these assays can provide insights for only one type of phenotype affected by absence or presence of DsbA.

1.4.3 *DsbA as a suitable target*

Overall, a number of structural and functional characterization studies described in the above sections 1.2 to 1.4 suggests that DsbA is a suitable target to pursue for antibacterial research. DsbA acts as a central regulator of bacterial virulence and the availability of crystal structures as well as its tractability for NMR experiments and the detailed understanding of the mechanism of action and biological activity make it a suitable target for rational drug design. Although DsbA has been suggested to be a suitable antibacterial target, to date there are no reported inhibitors of this enzyme. Regardless of recent breakthroughs in the understanding of the structure and function of DsbA, the rational design of small molecule inhibitors of DsbA remains a challenge. This may be attributed to its wide substrate specificity and the lack of deep binding cavities on its surface which make it a challenging target for inhibition by small molecules.^{1b} Although protein-protein interaction surfaces are more difficult to target, the number of successful studies of small

molecule inhibitors that target such sites is increasing, so that this problem, though technically demanding, may be tractable.⁶⁰

As discussed above, DsbA contains a number of surface features that are implicated in functional activity and appear to be conserved across DsbA enzymes from different bacteria (Figure 1- 4). Our hypothesis is that by targeting these regions, which are unique to DsbA enzymes and necessary for enzyme activity in the catalytic cycle and substrate oxidation, that it will be possible to generate potent and specific inhibitors of DsbA function that could counteract the virulence of pathogenic organisms.



(A) Inhibition Approach -1
No DSB transfer

(B) Inhibition Approach -2
Catalytic Cycle Interruption

Figure 1-10: (A) *EcDsbA*-substrate peptide complex, where the protein structure is shown as a white surface, the active site is coloured yellow and the bound peptide as a green carbon stick model (B) *EcDsbA*-*EcDsbB* complex where *EcDsbA* structure is shown as a grey surface and the interacting loop of *EcDsbB* is shown as a cyan carbon stick model. These complexes suggest two functionally important regions as target sites to be blocked by potential inhibitor interactions with DsbA.

Considering the experimental data described in the previous sections, there exists a possibility of *EcDsbA* inhibition via two pathways (Figure 1-10). One would be to inhibit the interaction of substrate polypeptides with DsbA, which directly stops the transfer of the disulfide bonds. The other potential approach would be through using inhibitors that bind to the DsbB interaction site. In this case DsbA would be active only in oxidized form and until it was all transformed into the reduced form, at which point it would not be re-oxidized by DsbB and hence the DSB redox cycle would not be active. Due to the absence of small molecule inhibitors it is unclear if these represent two entirely separate possible mechanisms. Therefore this study is aimed towards identifying inhibitors of DsbA activity *via* either pathway.

1.5 Process of drug discovery

The small-molecule drug discovery process involves the identification of organic compounds (or “hits”), and usually requires synthesis, characterization, screening, and assays for therapeutic efficacy against a particular biological target.⁶¹ Currently, the research and development cost of each new molecular entity is approximately US\$ 1.8 billion with a time period of 10-15 years.⁶² Despite advances in understanding of biological systems and technology, drug discovery is still a lengthy, expensive and difficult process with a low rate of new therapeutic discovery.⁶³ Historically, most drugs have been discovered either based on an already existing one or by random screening of chemical collections. Some of the most common approaches to drug discovery have been discussed below.

1.5.1 *Historical methods*

1.5.1.1. Nature as a drug source

Natural products generally represent a useful source of starting compounds especially in therapy areas such as antimicrobials, antineoplastics, antihypertensive and anti-inflammatory drugs.⁶⁴ According to a study that reported the contribution of naturally occurring chemicals in drug development in years 1981- 2006, of the 974 new chemical entities, 63% were natural-product derived or semi synthetic derivatives of natural products.⁶⁵ Natural products as starting compounds can represent natural substrates, cofactors or inhibitors of proteins, and in many cases they are complex compounds that challenge current synthetic medicinal chemistry efforts. Lead discovery using natural products involves the modification of compounds of known activity against a drug target in a stepwise manner to improve characteristics such as affinity, potency or bioavailability parameters like solubility or membrane permeability.^{64b} Regardless of their implied potential, only a relatively small portion of Earth's natural products has been tested for bioactivity.^{64b}

1.5.1.2 Combinatorial chemistry

Combinatorial chemistry involves the rapid synthesis of a large number of structurally related compounds with numerous combinations of different functionalities.^{64b} It has been a key technology since the 1990s enabling efficient generation of large screening collections for the requirement of vastly enhanced high-throughput screening⁶⁶ (described below). Technology advances in robotics and computational methods have enabled pharmaceutical industries to produce over 100,000 new and unique compounds per year with a computational enumeration of all possible structures in the form of virtual libraries.⁶⁷ Such a library can consist of thousands to millions of 'virtual' compounds which can be clustered for actual synthesis, based upon various calculations and criteria like ADME and QSAR.⁶⁷ However, after two decades of combinatorial chemistry, it has been pointed out that despite the increased efficiency in chemical synthesis, no increase in drug candidates has been realised.⁶⁹

1.5.1.3 High throughput screen (HTS)

HTS is an established method within large pharmaceutical companies for hit identification and involves the screening of very large libraries comprised of up to a million relatively complex drug-sized compounds (average MW \approx 400).⁷⁰ These large libraries are supported by combinatorial chemistry and are assayed for activity on a target using robotics, data processing and control software, liquid handling devices, and sensitive detectors to quickly conduct millions of biochemical, genetic or

pharmacological tests.⁷¹ HTS based identification of “hits” is followed by activity optimization to develop a “lead” by chemical modification and the corresponding analysis of structure–activity relationships (SAR).⁷² This “hit to lead” process thus generates a series of lead compounds that are then optimized with regard to the properties desirable for the development of candidate compounds for clinical evaluation.

Problems with conventional methods:

Although these conventional approaches have been able to identify many high-value hits, the limitations of screening drug-sized compounds have become apparent. These limitations include low hit rates and frequent failure to progress into hit optimization, which contributes to both time and expense. In addition to the development of an automated and robust assays⁷⁰, these hits require extensive validation in a secondary bioassay.⁷³ Larger HTS compounds can have complicated structures and conformational restrictions often don’t allow optimal binding of the presented functional groups.⁷⁴ Molecules failing in drug discovery trials can often be the consequence of unacceptable physicochemical properties related to lipophilicity and size.⁷⁴ Broadly, a possible reason for this lack of success has been suggested as lack of understanding of what the right types of molecules to make actually are.⁷⁵ It is also associated with the percentage of required “chemical space” to be sampled in these conventional techniques to identify a lead.

The concept of “chemical space” is considered to be central to the process of drug discovery and represents the sum of physical and chemical attributes of all realistic compounds.⁷⁶ The theoretical number of possible “drug-like” compounds has been estimated to be of the order of 10^{40} ,^{76a} while the chemical space of potential compounds that can be synthesized (MW < 500, containing up to 30 non-hydrogen atoms) has been estimated to be over 10^{60} .⁷⁷ Therefore, irrespective of how big the library is, the section of chemical space that can practically be sampled is very small and even the largest HTS libraries ($\sim 10^6$ molecules) are only $10^{-34}\%$ of ‘drug-like’ space.⁷⁸ Notably, an exponential decrease in chemical space is observed with a linear decrease in molecular size. Therefore for molecule less than 160 Da, chemical space is estimated about 4.4×10^7 molecules.⁷⁸ Screening a library of 1000 compounds would cover 0.0023% of the possible space, 30 orders of magnitude more than a traditional HTS screen.⁷⁹ Apart from the simple impracticality of testing every one, or even large mixtures, of drug-like compounds for affinity with a given target, it makes logical and intuitive sense to narrow the search by the preliminary application of various selection criteria.^{76a, 79}

The above mentioned downsides associated with conventional drug discovery methods and the ever increasing costs necessitate alternative methods of discovering efficient small molecule ligands for a given target. Fragment based drug design (FBDD) approaches correspond to a basic departure from conventional methods. It has been adopted by academia and industry in the last 10 to 15 years as an alternate method for lead generation and has yielded many drug candidates that are currently in clinical trials.⁸⁰

1.5.2 *Fragment based drug design*

The basic concept in the FBDD approach is to start the hit and lead optimisation process with small molecular “fragments” with molecular weight in the 120 – 250 Da range. In principle, fragments can be considered to be the components of a larger molecules. The individual fragments that represent the larger molecule each bind with an intrinsic free energy of binding (ΔG^i).⁸¹ However, in order to bind to a protein a small molecule has to overcome considerable translational and rotational entropy penalty and displace solvent before establishing an interaction with target. In practice, this entropic penalty is similar for both fragments and larger drug-like molecules. This means that if two fragments can be linked together such that they still bind in a similar manner as they did in isolation, the joined fragments enjoy an entropic advantage. This results in a third term to describe conformation strain of the ligand and target as well as loss of translational and rotational entropy (ΔG^s).⁸¹ Hence binding of a larger compound (AB) can be described based on their fragments A and B as shown in Figure 1-11.

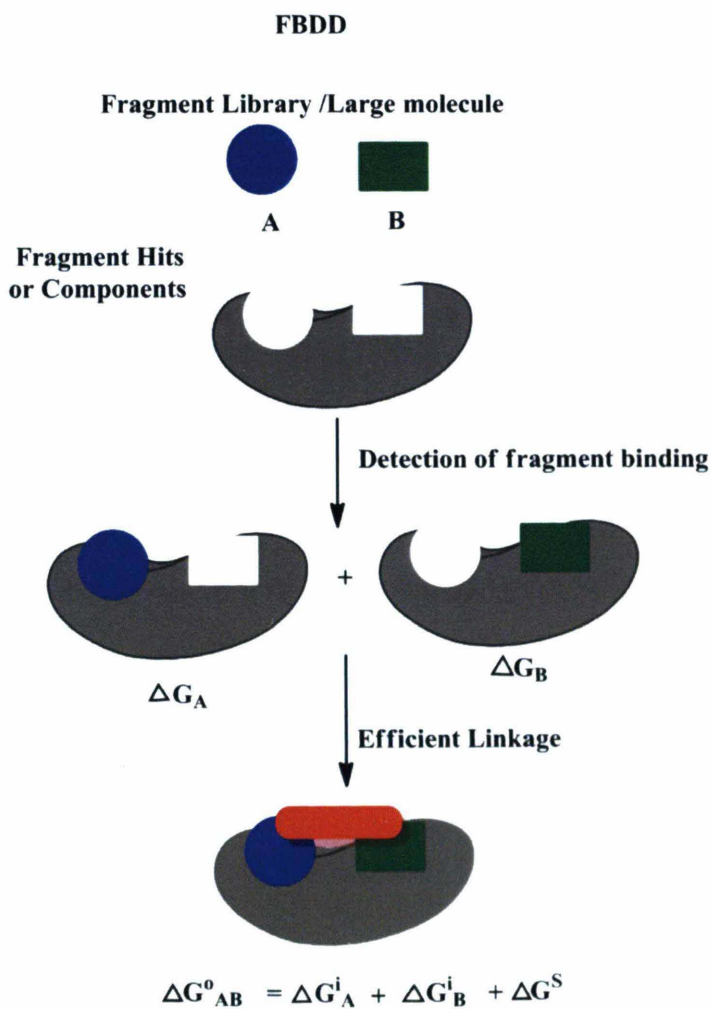


Figure 1-11: Schematic of fragment based drug discovery (adapted from reference 81).

After the molecule (or fragment) has achieved an interaction, additional functionality can be added which may potentially provide higher affinity. This 'fragment assembly' has been proposed as a method for probing a large percentage of the chemical space while synthesizing a minimal number of compounds.⁸²

Over the last decade, FBDD has shown its value by increasing the possibility of successful lead development against selected drug targets.^{77b, 83} FBDD involves the screening of fragments⁸³ to identify low affinity hits often with a high proportion of their atoms expected to make favorable key interactions with surface of target.^{77b} Fragment-based hits are typically weak inhibitors (10 μ M - mM), and therefore need to be screened at higher concentration using very sensitive biophysical detection techniques such as protein crystallography and NMR as the primary screening techniques, rather than bioassays.⁸² These hits can be developed into larger ligands by exploiting adjacent interactions in the protein binding site.⁸⁴

1.5.1.1 Library design and rule of three

The objective of the design of an efficient fragment library is to cover the largest chemical space in as compact a collection as possible.⁷⁹ Three wide-ranging selection criteria, or "filters," can be utilized in the formulation of fragment libraries.⁸⁵ These filters include:

1. Compounds have to meet the suitable drug like physico-chemical properties such as solubility, size, MW, log P or shape.

2. Compounds should not contain or result in evidently toxic or reactive templates.
3. Compounds should have features suitable for optimisation into larger, higher-affinity molecules. This synthetic feasibility is usually based upon medicinal chemistry expertise.

Additionally in some cases fragment library design may include a shape based strategy, using information of known drug-like molecules,⁸⁶ or may be enriched in particular pharmacophores or virtual screening based selections according to desired target classes such as kinases or phosphatases.⁸⁷ However in all cases, the members of a library must be chosen to give as diverse a representation of chemical space as possible.⁸⁵ Few commercial libraries, such as the Maybridge screening collection of ~56,000 lead-like molecules, the “Hitfinder” Collection of ~14,400 molecules and the Fragment Library of ~500 molecules, minimise the groundwork required of the medicinal chemist to synthesise and elaborate suitable candidates for drug development of any kind. These collections however may represent only a relatively small subset of the available chemical space.⁸⁸ Large pharmaceutical companies, who develop proprietary fragment libraries for FBDD – oriented research, are likely to maintain collections on the scale of 10^3 – 10^4 small compounds, conforming to certain selection criteria discussed above. However such collections may still only be scratching the surface of the “fragment universe”, which for fragments of MW < 160 Da is still estimated to be ~14 million compounds.^{77b} Broadly the FBDD library selection process can be categorized in two groups. The first method aims to generate a focused or scaffold based library which contains 10,000–20,000 compounds with an average MW ~ 300 Da.^{87b} Compounds of

this size can be detected in conventional binding or activity assays however the libraries contain less complex molecules than those typically found in HTS screening collections.^{87b} The second approach is the standard fragment approach to use a comparatively small collection of 500–2000 compounds, with an average MW between 110 and 250 Da. Hits from these libraries exhibit a typical binding affinity of micromolar to millimolar range, that represent potential starting points.⁸⁴

These selection criteria discussed for fragment library have evolved from the observation of the properties of drug candidates or drug-like molecules empirically. Therefore the concepts “drug-like”, “lead-like” and “fragment-like” generally refer to molecular size and complexity, and suitability of a compound for its pharmacological and pharmaceutical purpose.⁶⁸ Lipinski’s ‘rule of 5’ corresponds to “drug-like” chemicals, while the more recently proposed ‘rule of 3’ has been suggested for fragments or scaffolds suitable for FBDD.⁸⁹ The ‘rule of 5’ evolved from Lipinski’s empirical observations that poor oral availability of pharmaceutical compounds correlated with a small number of physicochemical parameters, including the molecular weight, the log of the partition coefficient between octanol and water (CLogP) and the number of H-bond donors and H-bond acceptors.⁹⁰ Other properties which have been suggested by others to correlate with oral availability include the polar surface area (PSA), number of rotatable bonds (NROT) and number of fused rings.⁹¹

The ‘rule of 3’ is an adaptation of Lipinski’s guidelines to the field of FBDD, such that possible fragment leads be likely to meet the criteria of MW < 300, number of H-bond donors ≤ 3, number of H-bond acceptors ≤ 3, CLogP ≤ 3, NROT ≤ 3 and PSA ≤ 60.⁹²

1.5.2.2 Ligand efficiency

Fragments are small enough to minimize the chances of unfavourable steric or electrostatic interactions that would prevent them from binding efficiently and are therefore less likely to be hindered by unnecessary or redundant chemical moieties during the course of binding.⁹² This can be quantified in terms of ligand efficiency (LE), which relates the free energy of binding to the number of non-hydrogen atoms in a ligand, calculated as the binding energy contribution per heavy atom.⁹³ As illustrated in Figure 1-12, smaller ligands tend to have a higher binding energy per heavy atom when comparing to the more complex hits that are observed in a high-throughput screen.⁹⁴ This plot suggests that LE drops off very rapidly as molecular size increases until a plateau is reached at large sizes (i.e. above 45 heavy atoms).^{94b} Therefore in the case of fragments although binding affinity is low, these hits usually exhibit high ligand efficiency.^{94b}

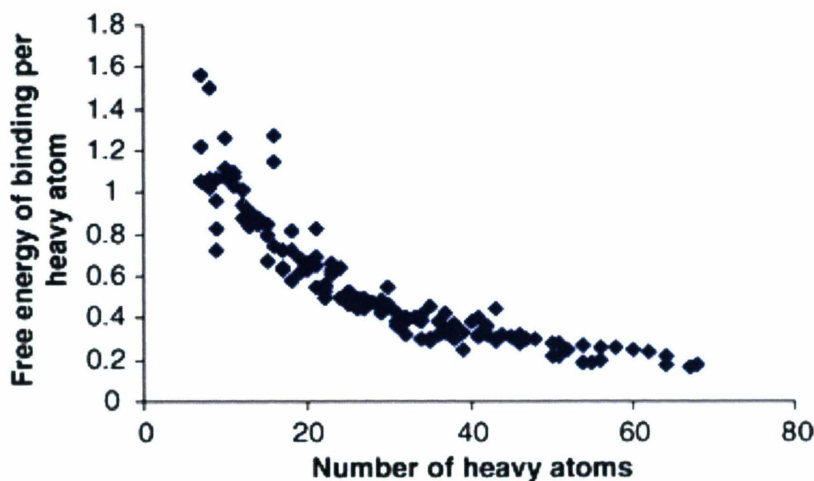


Figure 1-12: Plot of free energy of binding per heavy atom vs. number of heavy atom in 160 identified ligands [reference 94] (excluding molecules with less than six heavy atoms as they are not considered as representatives of drug interactions).

1.5.2.3 Advantages of FBDD over conventional methods

1. FBDD is a powerful strategy for evaluating the druggability of protein surfaces (especially novel targets) and identifying sites that interact with ligands.⁹⁵ The tendency of fragments to interact with a protein is indicative of whether a protein can accommodate a molecule with drug-like properties.⁹⁶
2. Compared with HTS hits, fragments are simpler, less functionalized compounds with correspondingly lower affinities. However, these hits generally bind with higher LE are appropriate for optimization into clinical candidates before exceeding the limits defined by the “Lipinski rules”.^{77b, 83, 97}

3. Targeted libraries based on known scaffolds may provide improved quality of the fragment library to give a greatly enhanced diversity of screening and higher hit rates for a given class of target.
4. FBDD has a better chance of identifying the optimal fragment as they are less complex than larger molecules and thus the functional groups they contain are more freely able to translate and rotate to form their optimal interactions with the target. The use of rational design by NMR or X-ray crystallography then allows for the evolution of fragments into lead compounds based on binding site information.
5. The hit rate from screening fragments is typically much higher than observed with HTS, as there is an inverse relationship between the molecular complexity of compounds screened and the probability of a compound possessing good complementarity with the target protein.^{96a}

FBDD also has some shortcomings, such as the requirement of sensitive and robust detection techniques for low affinity binders. Additionally optimization of these hits also requires use of sensitive biophysical methods and biochemical assay to detect scaffolds in binding in secondary sites. Current technological advances in several useful detection methods like NMR, SPR and X-ray crystallography can overcome this drawback.⁹⁸ Secondly suitable optimization/linking of fragments can also become difficult and almost always requires both structural and binding information to be employed. This limits the range of suitable targets for FBDD.^{80b} Even though FBDD has become a feasible alternative or complimentary method it is often run in parallel with HTS.^{80d}

1.5.2.4 Fragment elaboration

After identification, fragment hit must be evolved into a suitable lead candidate by systematic elaboration using structural insights obtained from biophysical methods like X-ray or NMR.⁹⁹ Fragment affinity for a given target as well as properties such as pharmacokinetics and toxicity can often be modified by substitution at one or more positions with additional functionality or interactions. When structural information is not accessible then the screening of suitable analogues of the original fragment would be carried out in an attempt to establish a conventional SAR to guide the medicinal chemistry efforts.¹⁰⁰ In most cases a prerequisite for success is that the fragment acts as an “anchor”, which binds with high ligand efficiency without changing its binding mode during its progression to a lead candidate.⁹⁸ Methods to elaborate fragments into higher affinity ligands can be grouped into the following, somewhat overlapping, approaches: fragment evolution, linking, merging and optimization and some other reported techniques.^{94b, 96a, 100}

[A] Fragment Evolution: Fragment evolution is generally based on conventional hit to lead medicinal chemistry efforts.^{96c, 100} It requires the addition of functionality that binds to adjacent regions of the target site or improving the interactions of an existing core to increase potency.¹⁰¹ This process is directed by information about the binding mode of the original fragment, as well as knowledge of the topology of the binding site itself (Figures 1-13 [B] and 1-14).

[B] Fragment Merging: Fragment merging or fusion involves the development of larger molecules using optimal features of two small fragment binding in adjacent and overlapping positions. An example is shown in Figure 1-13 [C] for larger hit developed from fragment merging approach.⁸¹

[C] Fragment Linking: Fragment linking involves the joining of two weakly binding fragments that have been identified to bind at adjacent or proximal sites in a given target protein surface. The linking process may be performed either in a stepwise fashion, by addition of linkers of increasing length, or by reference to structural knowledge of the active site in an analogous fashion to fragment evolution. Successful linking can be very challenging, due to the fact that the linker has to be just the right length and the right conformation to be able to link fragments so that they can reach their respective binding sites in an optimal way^{81, 102} (Figure 1-14 and 1-13 [A]). Structural information is usually required for this strategy to succeed since it can avoid exhaustive and very large combinatorial and indiscriminate searching to find a suitable linking scheme for the combination of two fragments. Theoretically, the resulting lead can exhibit improved binding equal to the sum of the free energies of binding of the individual fragments, assuming the energetic contribution of the linker is negligible.¹⁰³

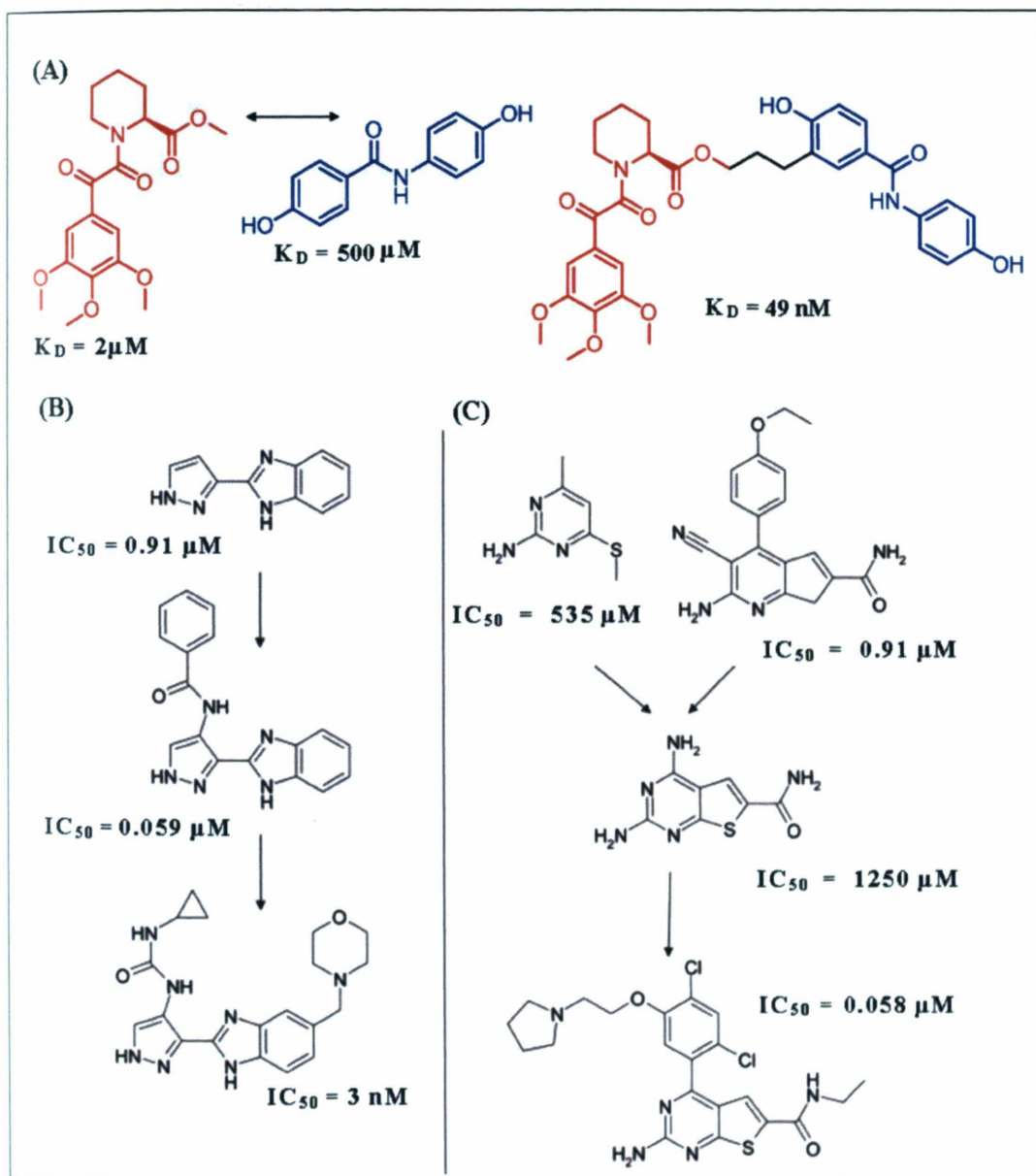


Figure 1-13: Successful examples of potent binders developed from fragment elaboration strategies [A] Linking [B] Evolution and [C] Merging.

[D] Fragment Optimization: Fragment optimization is somewhat similar to the conventional hit to lead process and fragment evolution as described above. It involves the optimisation of only a particular part of the molecule, in order to enhance the properties other than the inherent affinity of the original fragment or to deal with some specific issue, such as an ADME issue, oral bioavailability, cell penetration, and stability *in vivo* (Figure 1-14). Overall this approach is very similar to that applied in the final phase drug discovery process, where slight changes to a given lead series are investigated to adjust ADME parameters and affinity.^{94b, 100, 104}

Additionally there are some reports of other methods of fragment elaboration such that self-assembly and tethering. The fragment assembly method, which is usually termed dynamic combinatorial chemistry, involves reactions between fragments having complementary functional groups, in the presence of the template protein target. [Figure 1-14] Effectively, the protein self-selects or even catalyses the synthesis of its own inhibitor through the process of non-covalent binding. These can then be detected using a biochemical assay.¹⁰⁰

Tethering is another fragment elaboration method which allows the identification of a small molecule or fragment that bind to specific regions of a protein target site.¹⁰⁵ This is followed by fragment enhancement by combining with another molecule, moiety or functionality to provide high-affinity drug leads (Figure 1-14). This method is unique in using a covalent and reversible bond to stabilize the protein-ligand interaction and usually exploits a cysteine residue at or in the proximity of the binding site.¹⁰⁰ In absence of a

suitable cysteine, the residue can be included via single point mutation. The cysteine is able to capture fragments that bind close to it, provided the fragment has the chemical functionality to bind to a cysteine. There are a number of successful examples employing this method in FBDD.^{83, 105-106}

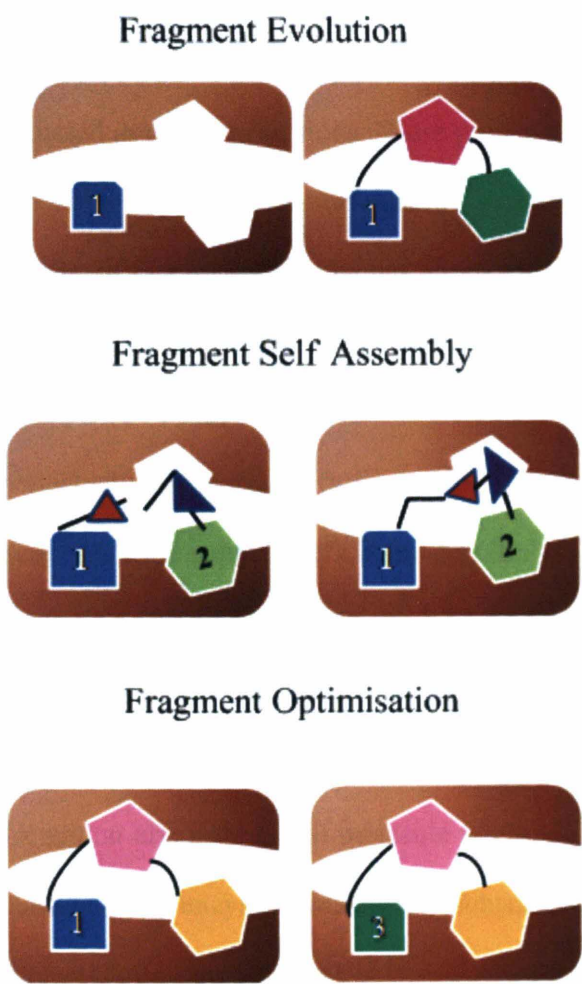


Figure 1-14: Representation of several FBDD elaboration methodologies: fragment evolution, self-assembly and optimization respectively.

1.6 Project aims

The principal hypothesis of this project was that FBDD could be exploited to develop inhibitors of DsbA activity. The accessibility of NMR spectroscopy and X-ray crystallographic information along with an efficient protein expression system suggest that DsbA is amenable to FBDD. This approach can identify weakly binding fragments that bind to pockets on the DsbA surface around active site and are likely to inhibit protein activity as discussed in previous sections. These potential fragment hits may then be developed into the larger ligands with substantially greater affinity and increased selectivity by using fragment elaboration techniques. Application of the fragment-based approach to identifying potential inhibitors of the suitable targets, *Ec*DsbA and *Vc*DsbA, is expected to have multiple advantages. First is the application of a rather new scheme to a novel enzyme target, which will represent further example of the FBDD approach for antibacterial drug discovery. The second benefit is the potential discovery of inhibitors of an enzyme that is known to be of functional importance to several pathogenic bacterial species; moreover targeting bacterial virulence is one approach that has yet to be fully exploited for antibacterial drug discovery. Third, the project will contribute to the growing body of information about the small molecule binding and selectivity of thiol-disulfide oxidoreductase class of enzymes. In order to substantiate the validity of these hypotheses the major project aims are listed below.

1. Identification of DsbA binding fragments using NMR based FBDD approach.
2. Validation and biophysical characterization of identified fragment hits.

3. DsbA inhibition activity assessment for potential hits by optimizing and modifying existing assays for small molecule testing.
4. Structural studies to characterize binding modes of DsbA ligands by X-ray crystallography, NMR and/or molecular docking.
5. Applying fragment elaboration strategies to develop better ligands and characterization of newer molecules to generate SAR for most promising lead compound series.

1.7 Analytical techniques applied for FBDD

Sensitive detection of target- fragment binding (affinity range of 10 μ M – mM) is the key to successful FBDD.⁸³ Binding can be assessed in a variety of different ways, using assays that provide different levels of information. These range from (i) demonstrating a functional effect (IC₅₀/EC₅₀) (ii) biophysical characterisation of direct binding (most commonly via NMR or SPR) (iii) kinetically and/or thermodynamically characterising the fragment-target interaction (using ITC and/or SPR) to (iv) generating structural models of binding at atomic resolution.¹⁰⁷ Various techniques are possible to provide different data to support FBDD and these include NMR spectroscopy, X-ray crystallography, surface plasmon resonance (SPR), fluorescence spectroscopy and mass spectrometry (MS).⁸⁵ These techniques provide different levels of information and vary from in terms of their throughput. Some approaches are low-throughput and limited to a few test compounds per day (ITC and DSC) whilst others are medium-throughput allowing analysis of a few dozens to hundreds of compounds per day (SPR and NMR) whereas higher throughput approaches may accomodate a few thousands (mass

spectrometry, TS) or more (waveguide-grating based techniques).¹⁰⁷ Different approaches provide somewhat complementary roles and allow precise pictures to be built up of both conformational and dynamic aspects of active site interactions with fragments¹⁰⁸. Several analytical techniques commonly used in FBBD efforts are described briefly in the sections below.

1.7.1 NMR for fragment screening and structural biology

NMR is a physical phenomenon based upon the magnetic properties of an atom's nucleus. NMR scrutinizes a magnetic nucleus, like that of a hydrogen atom, by aligning it with a magnetic field and perturbing this alignment using an electromagnetic field. The response to the field is what is exploited in NMR spectroscopy. Magnetized nuclei of a molecule such as ^1H or ^{15}N are affected by their surrounding nuclei, those that are chemically bonded (spin-spin coupling) and those that are nearby in space (dipole-dipole coupling). Therefore in presence of a magnetic field, the specific features of a molecule can be explored to assign resonance signals to the nuclei within the molecule and to determine the distances between them.¹⁰⁹ The majority of NMR based techniques have been developed to monitor and characterise intermolecular interactions based on these basic principles.

Advantages: Biological NMR is a suitable technique for use in FBDD based on following reasons: direct physical interactions between ligand and target are detected, the protein does not need to be modified for analysis, the samples can mimic the native or

biological environment in terms of temperature, salt concentration and pH as closely as needed, either ligand-based or target-based NMR experiments can be used to acquire data, and the throughput can be relatively high (up to hundreds of compounds a day).¹¹⁰ An important feature of NMR is its capability to use different parameters to detect the association, including chemical shifts (^1H , ^{15}N or ^{13}C), chemical-shift anisotropy (^{19}F), transverse and longitudinal relaxation, cross-relaxation in the protein–fragment complex (transferred-NOE or STD-NMR) or cross-relaxation between the fragment and protein-bound water molecules (water-LOGSY).¹¹¹ NMR is highly sensitive to changes in the slight effects of the local electronic environment on the resonance frequencies of nuclear spin, as well as the ability to evaluate the structures of protein-fragment complexes that do not give diffraction-quality crystal and are thus not amenable to X-ray crystallography.^{84, 112}

NMR facilitates the estimation of the potency and kinetics of binding processes, and can provide insights on competitive binding and allosteric effects. Moreover, recent developments in immobilised-protein NMR in the past decade have made it possible to study the interactions of ligands with membrane-bound proteins and intact viruses.¹¹³ Hence NMR is suitable for the study of a vast range of protein-ligand interactions.

Limitations: Although NMR is well suited to medium-throughput analysis, structural analysis of protein-ligand complexes effectively remains limited to macromolecules having MW less than ca. 35-40 kDa.¹¹³ Additionally, NMR is inherently insensitive, which necessitates the use of relatively high sample concentrations. This can be improved by using cryogenically cooled probes, miniaturised receiver coils and samples¹¹⁴ or by

enhancing NMR relaxation effects using paramagnetic ions.¹¹⁵ Digital recording, auto samplers, and higher magnetic fields cut down the time for data acquisition and improve the spectral quality. As a consequence, new screening technologies involving NMR have become indispensable for modern structure-based drug discovery.

There are a number of experimental NMR methods that can be used to detect the binding of fragments to protein targets. These can be divided into two general approaches:

1. Ligand - detected methods which measure changes in NMR signals of the ligand upon binding to a larger target and 2. Protein - detected methods which measure the changes in the spectra of the protein target due to the influences of ligands.^{84, 111b, 116} Both techniques have their intrinsic advantages for diverse applications, making them largely complementary. These aspects of biomolecular NMR that are applied in the current study are described briefly in the sections below.

1.7.1.1 Ligand based methods for hit identification

Ligand-based methods mainly use one-dimensional (1D) NMR spectra where fast detection of the ligand signals enables the screening of mixtures of compounds without the need for deconvolution, so long as the signals of the ligands do not overlap. These methods require relatively low concentrations of unlabelled protein, and are useful for rapid cost-effective acquisition of data while there are fewer limitations on the properties of the target. Usually, the target is not isotopically labeled and its molecular mass can be practically unlimited in some techniques (*e.g.* saturation transfer difference (STD),

discussed below). The target protein can even be immobilized or membrane-bound. Additionally some simple information about the binding epitope and the interaction mode can be obtained in some cases. As analysis occurs by 1D proton methods, interpretation of spectra is a relatively straightforward process which is amenable to automation. Most of the ligand-detected methods are limited to low and medium affinities because they rely on relatively rapid exchange of the ligand between its bound and free state. Ligands that bind too tightly are indistinguishable from the target and thus are broadened or suppressed, which can result in false negatives. In contrast, non-specific binding can result in the appearance of false positives.¹¹⁷

Saturation Transfer Difference Spectroscopy (STD) is a widely applied ligand-based method for FBDD screening. The technique is based on “on-resonance” frequency selective irradiation of the target protein in solution which causes selective ¹H saturation of the target protein via spin diffusion but no perturbation of small molecule resonances (Figure 1-15).¹¹⁸ Magnetization is subsequently propagated with great efficiency throughout the protein via intramolecular proton-proton cross-relaxation pathways, ultimately being transferred through intermolecular cross-relaxation to any ligand molecules binding weakly to the protein.¹¹⁸ When this ligand dissociates from the target into solution, the magnetization change transferred in the bound state is retained in the free ligand. Non-binding molecules are not saturated and so will not register on the 1D spectrum (Figure 1-15).

The difference from a reference spectrum taken without on-resonant irradiation thus yields a spectrum containing only those ligands that have been perturbed by binding to the target. The reference spectrum is obtained by off-resonance irradiation in a spectral window where no target signals appear ¹¹⁹ (Figure 1-15). In this way binders are distinguished from non-binders resulting in primary filtration of fragment hits for a given target from the library sampled as mixtures of diverse compounds.

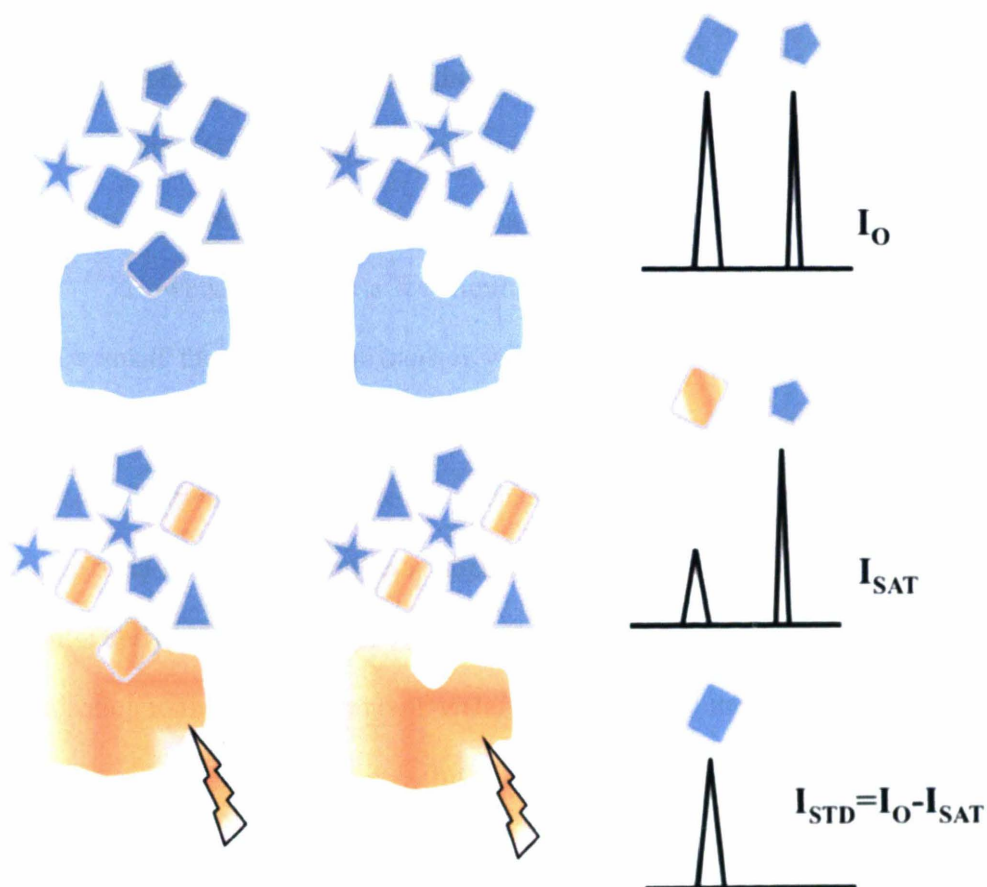


Figure 1-15. Illustration of the STD-NMR based experiment used for FBDD. Rectangles and other smaller shapes represent binding and non binding compounds, respectively. During the off-resonance experiment (top panel), no magnetization is transferred to the binding ligand from the target (shown as biggest blue shape) and signal intensity remains unchanged (I_O). If the irradiation changed to “on-resonance frequencies” (middle panel), magnetization transfer occurs which lowers the signal intensity of the ligand (I_{SAT}). The difference between the two spectra only contains the signals of binding components (I_{STD}). (Reproduced from reference 118)

The use of STD experiments to detect protein ligand binding offers a range of advantages:

(1) the quantity of protein and ligand needed to give significant signals is relatively small
(2) there is no upper limit for the size of the target (3) they can be easily extended to 2D or 3D spectra if required (4) ligand conformational information can also be obtained for a given complex (5) assessment of the binding constants can also be drawn from STD

spectra.^{111b, 120} However there are a few limitations associated with this approach: (1) if

the target is small then saturation transfer will be poor and it is possible that no STD

signals will be obtained (2) it is not applicable to high-affinity binders (3) the STD build-up has to compete with other relaxation mechanisms, meaning that some targets are less

amenable to STD-NMR.^{82, 121} In these situations, an effective alternative may be found in

the form of a derivative of STD, the WaterLOGSY experiment which involves the transfer of magnetisation to the ligand from bulk water associated with the receptor active site.^{108b,}

¹²¹⁻¹²² This approach has been successfully employed to elucidate binding interactions between enzyme and cofactors and illustrating the sensitivity of the approach and its application in combination with other binding assays.^{108b}

1.7.1.2 Hit validation using target based methods

Protein-detected NMR screening is a more precise and sensitive method in comparison to the ligand-based approach for the study of fragment binding to a given protein target.

Although these experiments have a lower throughput than STD NMR, they generate fewer false positive hits and hence are more reliable. This approach was the starting point for the original SAR by NMR studies conducted by the Fesik group in the Abbott

Laboratories in the early 1990s.⁸² These methods can easily differentiate between non-specific and specific binding events and spectral changes caused by aggregation and pH-alterations can be excluded. Protein -detected experiments require large amounts of protein in the range of 0.1–1 mM to accomplish an effective screening process.^{96b} For fragment screening, the implementation of these experiments involves recording two-dimensional (2D) spectra, which often require longer acquisition times (10-30 min per sample). The greatest advantage of this approach is the possibility of extracting detailed structural information about the ligand–target complex from the spectra. Protein-based screening commonly takes the form of a comparison of *e.g.* [^1H - ^{15}N]-HSQC spectra of an isotope-labeled protein in the absence and then the presence of a potential ligand.¹¹⁰ For a protein smaller than ~30kDa, the ^1H - ^{15}N HSQC spectrum is usually well resolved, where each cross peak is diagnostic of an individual amino acid residue (Figure 1-16).

Here, in contrast to ligand-detected screening methods, it is not possible to screen mixtures of compounds without deconvolution.¹¹⁰ Perturbations in particular HSQC resonances, induced by the influence of the ligand on the proteins' spectrum, allow identification of the binding site in combination with resonance assignment data (Figure 1-16). The ligand can be approximately located on the protein surface via ligand induced CSP based mapping by acquiring a ^1H - ^{15}N HSQC or ^1H - ^{13}C spectra for the protein – ligand complex.^{117a} A binding event induces changes in the chemical shifts of both protein and ligand while the nuclei of atoms participating directly generally show the largest changes in resonances. These differences in chemical shift are calculated as weighted average changes in chemical shift¹²³:

Equation 1-1

$$\Delta_{avg} = \sqrt{((\Delta\delta H)^2 + (\delta\Delta N \times 0.154)^2)/2}$$

Equation 1-2

$$\Delta_{avg} = \sqrt{((\delta\Delta C_{\alpha})^2 + (\delta\Delta C_{\beta})^2)}$$

Where $\Delta\delta H$, $\Delta\delta N$, $\Delta\delta C_{\alpha}$ and $\Delta\delta C_{\beta}$ corresponds the changes in chemical shift between apo and ligand bound proteins for amide proton and nitrogen, C_{α} and C_{β} resonances respectively.

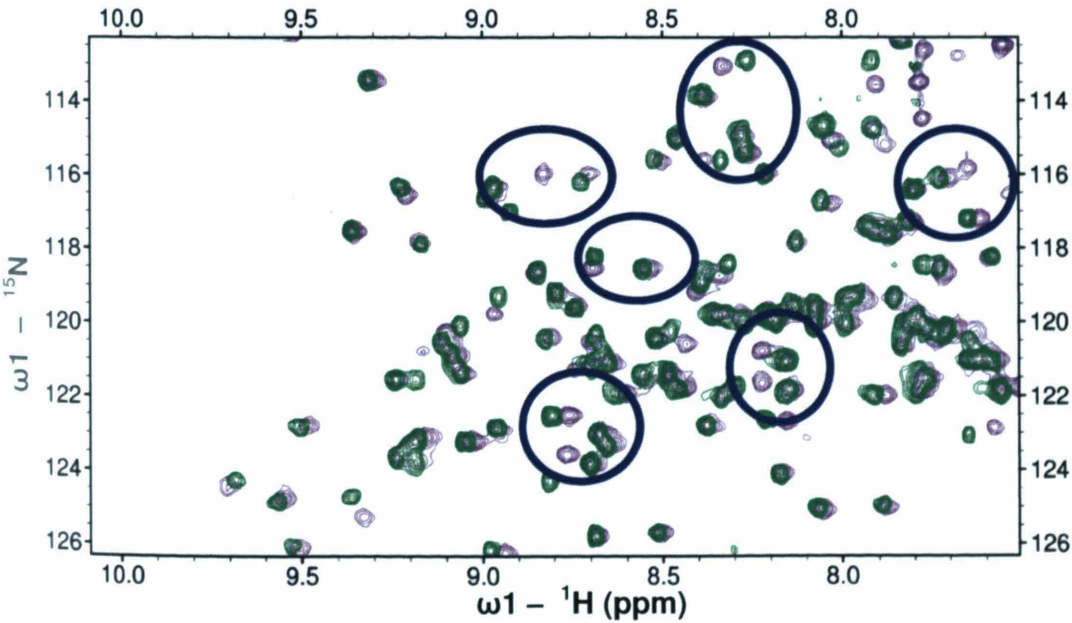


Figure 1-16: Overlay view of the ^1H - ^{15}N HSQC spectra of 150 μm Apo *EcDsbA* (green resonances) in the presence of a STD hit fragment at 1 mM (purple resonances). Blue circles identify a set of ^1H - ^{15}N resonances that undergo significant shifts changes in the presence of ligand, hence this fragment is a confirmed hit.

An additional advantage of protein-based methods is that they enable binding constants to be calculated by undertaking HSQC titrations justifies the added cost and effort involved in the generation of labeled protein and the increased complexity of receptor-based methods.⁸² HSQC titrations can be used to measure binding affinity and stoichiometry of the protein-ligand interaction.¹²⁴ Protein-ligand binding interactions are also explored by evaluating relaxation times (T1 and T2) and differentiating inter and intra molecular NOEs between a residue and a ligand. These experiments can quantify conformational changes upon protein-ligand complex formation as occasionally the dynamics of a protein may be changed upon binding. This analysis requires a full backbone assignment, and so necessitates the expression and purification of ¹⁵N- and usually ¹³C-enriched protein. As the targets usually have high molecular masses (>10 kDa), the signals are subject to fast relaxation.¹¹³ This restraint limits the size of targets that can be observed to MW <100 kDa, even if techniques like transverse relaxation-optimized spectroscopy (TROSY) or cross-relaxation-induced polarization transfer (CRIPT) are included.¹¹³ Although this limit may be overcome with specific methyl group labeling or if the binding site can be selectively labeled with amino acids, in the future, interesting targets might still remain intractable with these techniques.

The further advantage of allowing binding constants to be calculated by ¹H-¹⁵N HSQC titrations justifies the added cost and effort involved in the generation of labeled protein and the increased complexity of receptor-based methods.⁸² HSQC titrations can be used to measure binding affinity and stoichiometry of the protein-ligand interaction.¹²⁴

1.7.1.3 Multidimensional NMR of biological macromolecules

To facilitate structural analysis of protein-ligand complexes it is a requirement that proteins must be labeled with NMR sensitive isotopes in order to enable heteronuclear NMR spectra to be acquired.^{117b, 125} For the study of biological complexes, this restricts the analysis to molecules consisting of active isotopes of the species common in organic systems (notably ^1H , ^{13}C , ^{15}N , ^{19}F and ^{31}P). The high abundance of ^1H , as well as its sensitivity, makes it highly amenable to NMR spectroscopy, and hence ^1H NMR is by far the most widely applied technique. As ^{15}N is present in a natural abundance of only

0.37%, and ^{13}C in a natural abundance of 1.11%, this imposes a requirement for efficient methods of enriching proteins with these NMR-active nuclei before multidimensional heteronuclear NMR is feasible.^{117b, 125}

Enrichment can provide uniform labeling of a given nuclide throughout a protein or selective incorporation of isotope labels for particular amino acids. With an efficient over-expression system, ^{15}N -labeling is comparatively cheap while in contrast; uniformly ^{13}C -labelling is relatively more expensive. Additionally several newer labeling methods allow the selective assimilation of ^{13}C -labels for example a common scheme selectively enriches ^{13}C in the methyl groups of isoleucine, leucine and valine.^{77b, 95b} The apparent signal intensity of methyl groups is higher than that of amide protons therefore isotopic enrichment with ^{13}C at specific locations forms a convenient alternative to ^{15}N -labelling.^{84, 125a} Accessible methods in expression and purification of a wide range of proteins in bacterial hosts now make production of isotopically enriched proteins a routine exercise and so a range of multidimensional NMR techniques has been incorporated into the suite of methods used in FBDD.^{83, 117b, 125}

1.7.1.4 Structure assignment

NMR based solution structure determination is the only way to obtain high resolution information on proteins that are not amenable to crystallography and is most useful for smaller proteins. Due to their small size, DsbAs are amenable to an analysis of their protein structure via NMR. The details of all of the multidimensional NMR experiments utilized in this study have been described in next chapter.

The first step towards full structure determination by NMR spectroscopy involves the assignment of resonance peaks using 2-D (HSQC) and a range of 3-D triple-resonance spectra. Protein backbone amide assignment is usually a prerequisite for FBDD programs utilising the “SAR by NMR” approach.¹⁰²⁻¹⁰³ Using a combination of assignment data of side chain protons and heavy atoms, tertiary structure determination can be carried out, although the structures are often of lower resolution than those obtained using X-ray crystallography (described below) and the process of structure determination by NMR is more time-consuming, which makes X-ray crystallography the preferred technique for generating information on protein-fragment complexes. Notwithstanding, where crystallography is not feasible, NMR is capable of providing high-resolution details of protein-fragment structures that can guide fragment evolution.

1.7.2 X-ray crystallography for FBDD

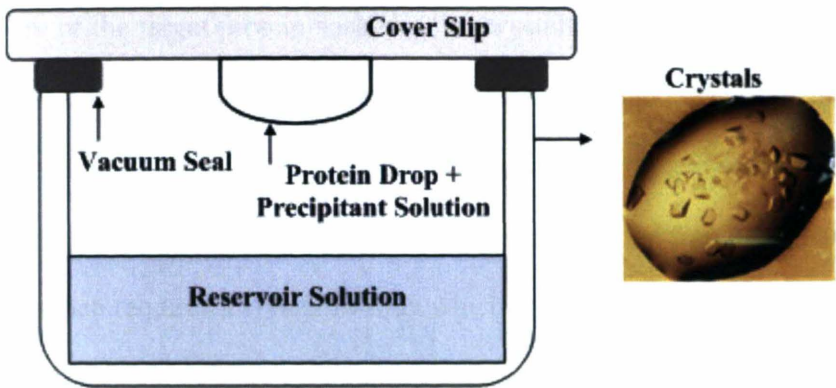
X-ray crystallography is the preferred method for revealing details of the binding of fragments to their target proteins. Even in the case of very weakly binding fragments, there are some successful studies which demonstrate unique and specific binding of small fragments at high concentration.¹²⁶

A range of technological advances such as new methods in molecular biology and novel computer software for analysis of X-ray data resulted in time and resource-efficient methods to obtain crystal structures of therapeutic target proteins. The increase in the rate of obtaining crystal structures of protein–ligand complexes has now allowed X-ray crystallography to be used for lead discovery and, in particular, as the method of choice for the structural characterisation of complexes that is required in FBDD approaches. Moreover, recent developments in automation and informatics have led to the development of this technique to the study of dynamic aspects of structure-based drug discovery.¹²⁷ At the start of 2011, there were over 71000 three-dimensional protein structures deposited in the Protein Data Bank – of which around 62,000 have been solved by X-ray crystallography. Whilst not all represent therapeutic targets for drug design, the PDB is rich in structures of protein-ligand complexes. These structures can provide useful insights on binding mechanisms that can contribute to the design of novel ligands. Where structures are available of small fragments bound to their targets, specific binding interactions can be elucidated; this can lead to highly-targeted elaboration strategies, which then can be tested and refined in an iterative process.⁸⁵

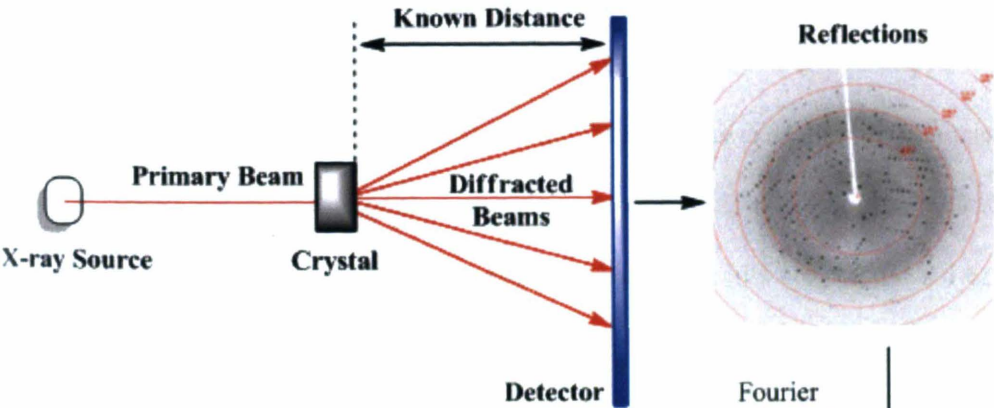
A condition for this approach is producing crystals of high quality to solve the structure of a complicated arrangement of atoms within the target protein or complex.¹²⁸ Typically

crystals are grown using one of the three conventional systems: sitting drop, hanging drop or microbatch.¹²⁸ The hanging drop method is the most widely used technique to generate crystals.¹²⁹ In this crystal forming method, a drop of highly concentrated purified protein mixed with a precipitant solution is incubated above a reservoir of the precipitant solution (Figure 1-17). The final composition of precipitant solution is generally the outcome of extremely difficult condition optimization cycles performed for the given target.¹²⁸ These conditions vary between proteins and must be identified through extensive screening trials depending upon the nature of the study.¹²⁹ Diffusion of water from the concentrated drop and the reservoir causes crystal formation through super saturation of the protein within the drop (Figure 1-17). Crystals are separated and are subject to X-ray diffraction.¹²⁸ In an X-ray diffraction measurement, a crystal is mounted on a goniometer and gradually rotated while being bombarded with X-rays, producing a diffraction pattern of regularly spaced spots known as reflections (Figure 1-17). These 2D images taken at different rotations are converted into a 3D model of the density of electrons within the crystal using the mathematical method of Fourier transforms, combined with chemical data known for the sample. Poor resolution (fuzziness) or even errors may result if the crystals are too small, or not uniform enough in their internal makeup. Amplitude and phases are calculated using methods like molecular replacement and single and multi-wavelength anomalous diffraction (SAD and MAD)¹²⁹ and are used to build an electron density map. Protein structure can be determined from the electron density map.

1. Crystal Generation- Hanging Drop



2. Measure Crystal Diffraction



3. Refinement

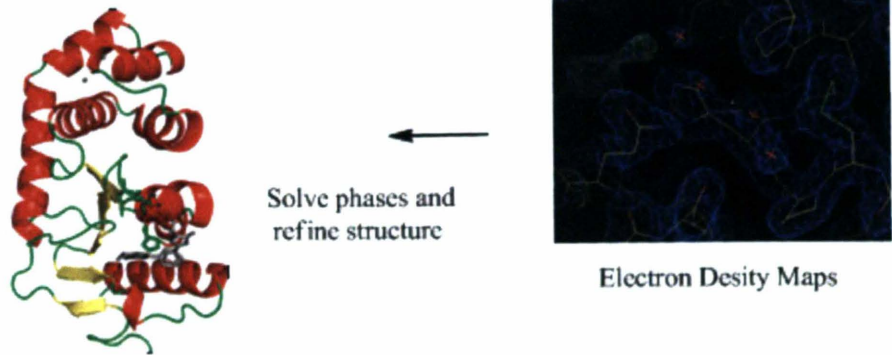


Figure 1-17: Schematic of the main steps involved in X-ray crystallography studies for protein or complex structure determination.

The crystallographic screening of fragment libraries involves determining the crystal structure of the target protein soaked or co-crystallized with cocktails of fragments. The crystal structure is solved for each cocktail separately and any compound that has bound will be detected directly and can be used to drive structure-based drug design.^{120a} In contrast to NMR based screening, the binding mode of the fragment is available however this approach requires a crystal form in which the protein target has an accessible binding site and which can endure high concentrations of fragment mixtures.⁸⁵ Sometimes crystal packing may result in non favourable binding conditions that require crystal engineering.¹²⁹ Additionally, to screen a reasonable size fragment library more resources are required for data collection and interpretation. Therefore most commonly, a range of cost- and time-effective biophysical methods like NMR, SPR, high concentration screening (HCS) assay and virtual screening are preferred as the primary screening methods for fragment libraries.^{86a, 98} Subsequently promising hits are progressed to X-ray crystallography and an iterative structure-based drug design can be used to develop lead molecules.

1.7.3 Other biophysical screening methods in FBDD

In addition to NMR, there is a range of screening techniques that have been applied in this initial phase of the FBDD process.^{111a, 130} However, it is also evident that NMR alone cannot provide the high-content information required in an FBDD process.¹³¹ Several other useful biophysical techniques may be appropriate. These biophysical, label-free detection techniques are generic, direct and informative, and there is a growing interest in

applying these during hit qualification.¹³² Few of these techniques which are relevant to this study have been discussed here briefly.

Surface plasmon resonance (SPR) has gained increased popularity in recent years for the initial screening of fragment binding.¹³³ Here fragment-protein binding is observed as a change of the refractive index of a sensor surface as a function of sample concentration. If the protein itself is attached to the surface the measurement is direct and can provide kinetics of binding and a K_D value. Alternatively, in a competition assay small-molecule ligands attached to the surface enable the determination of the concentration at which a fragment prevents protein binding to the attached ligand and can be used to determine inhibition constants K_i .¹³⁴ The major challenge in SPR concerns the reliable and robust immobilization (target or fragment) to the sensor chip surface without influencing binding properties. If appropriate immobilization can be achieved, SPR is a powerful technique for FBDD, which has the advantage that it uses relatively small amounts of protein.¹³⁴

Fluorescence-based thermal shift assays (also termed as thermofluor and differential scanning fluorimetry) are becoming useful as demonstrated by several groups for focused small molecule screening¹³⁵ as well as hit validation after HTS.¹³⁶ These studies suggest that the thermal shift (TS) method can be used for the fast, reagent friendly and cost-effective screening and detection of small molecules with low micromolar to submicromolar target affinities.¹³² In principle this assay measures the binding through associated fluorescence quantum yield enhancement of a fluorescent dye sensitive to the

chemical environment (e.g. Sypro Orange) in the presence of a melting protein and tested compound. Melting is induced by the exposure of the target protein to heat and ligand binding typically adds to the stabilization energy of the protein, thereby shifting the cooperative melting transition to higher temperatures.¹³⁵ Screening experiments can be conducted in 96 or 384 well plates and several 384-well plates can be screened per day therefore its throughput can be in the low thousands of samples per day.

The TS method is widely applicable to all sorts of soluble target proteins but it is not universally applicable. There have been cases where it has been unsuccessful for some proteins.¹³² Moreover, there are some unresolved issues associated with this screening method, such as; no clear evidence of direct correlation between melting temperature transition and concentration and affinity of the ligand in case of different enthalpic and entropic components of target interaction.¹³⁷ Furthermore, to quantify thermal shifts as a thermodynamic read-out technique, the folding transition monitored in TS has to be reversible and an interaction of the ligand with the unfolded state of the target protein has to be excluded.^{132, 137} Therefore this method is still in early stages of development for extensive use in FBDD.

Isothermal titration calorimetry (ITC) is another biophysical technique which can provide K_D values and other thermodynamic parameters of ligand binding.^{108b} ITC is used to determine the directly measure of the thermodynamic parameters such as binding affinity (K_a), enthalpy changes (ΔH), and binding stoichiometry (n) of the interaction

between two or more molecules in solution.¹³⁸ From these initial measurements Gibbs energy changes (ΔG), and entropy changes (ΔS), can be determined using the relationship:

$$\Delta G = -RT\ln K_a = \Delta H - T\Delta S$$

(Where R is the gas constant and T is the absolute temperature).

Isothermal titration calorimetry provides the very accurate and direct measurement of the enthalpy of any reaction under isothermal and isobaric conditions. However the quantities of protein and ligand required for such analyses are generally excessive for use in the screening stage, but there continue to be advances in efficiency and the method can be useful in validation and for the selection of fragments for development.¹³⁹

1.8 Inhibition analysis

In general, the aim of this study is to identify small molecule inhibitors of DsbA to inhibit oxidative folding processes in bacterial periplasm. NMR-based fragment screening in combination with structural techniques can identify the ligands that bind to the site of interest in a given target. However all of the ligands that bind are not essentially inhibitors of an enzyme or pathway. Inhibitory or functional activity of a ligand depends on the binding mechanism and site of action as well as affinity. In FBDD, after identification and characterization of promising fragment hits it is necessary to determine the inhibitory activity before carrying out the further elaboration process. Therefore

functional assays that describe the effects of a fragment on the DsbA activity are a prerequisite of current project. The inhibition measurement assays need to be reproducible and ideally robust.

When considering relatively low affinity binders such as fragments, the assay should be pharmacologically sensitive enough to detect the effects of fragment binding on the activity of DsbA. Several of the *in vitro* and cell based phenotypic functional assays as described in section 1.4 could be optimized and explored further to measure DsbA inhibition to identify potential fragment hits.

1.9 Structure of thesis

The overall focus of this thesis is to establish inhibition of oxidative folding in Gram-negative bacteria by application of FBDD techniques. This has been conducted with a range of screening approaches, structural studies and inhibition assays to identify inhibitors of bacterial periplasmic enzyme DsbA.

Chapter 2 describes the complete experimental protocols for the following four results chapters. Chapter 3 has been directly reproduced from a manuscript to be submitted and includes results from the studies conducted. These results were obtained by a systematic strategy for fragment library screening and characterization using biophysical techniques and assays in order to target *EcDsbA*. This chapter reports the investigation of *EcDsbA* as an antivirulence target by discovering its small molecule inhibitors binding to functionally important region of this enzyme. NMR based screening, X-ray

crystallographic binding and phenotypic assays were described to characterized fragment hit for several classes.

Chapter 4 presents a combination of computational and experimental approaches applied in FBDD was used to identify and characterize fragment-based inhibitors of *VcDsbA*. This chapter investigates the specifics of binding of a range of structurally diverse fragments to *VcDsbA* using NMR CSP in conjunction with molecular docking and inhibition profiles.

Chapters 5 and 6 are complete experimental chapters and reports application of NMR based fragment screening coupled with X-ray crystallography and structure-based design to discover novel μm leads for the development of *EcDsbA* inhibitors. High resolution and preliminary crystallographic complexes of *EcDsbA*-ligand were used to guide the fragment progression process in chapters 5 and 6 respectively.

The data and findings of this thesis are reviewed in Chapter 7 which addresses the conclusions for the hypothesis and aims detailed in Chapter 1. Future directions for the project are also presented in this chapter. Formatting has been standardized across the entire thesis (except chapter 3) with a common reference list at the end of the thesis. Additional experimental results not shown in Chapters 3-6 are presented in the Appendices.

Chapter 2

Materials and methods

2.1 Materials

2.1.1 Chemicals

2.1.1.1 Fragment Library

For the fragment based screening a library was purchased from Maybridge (Trevillet, Cornwall, UK) and comprised 1132 compounds of average molecular mass 208 Da. The library comprised of structurally diverse and low complexity pharmacophoric compounds. Most of these fragments consist of a heterocyclic, rigid single core with 1-3 small side groups; therefore with limited torsional degrees of freedom. Chemical cores include benzenes, indoles, pyrimidines, indazoles, pyridines, benzimidazoles, benzofurans, pyrazoles, pteridines, cyclohexanes, cyclohexanes, quinolines, naphthalenes, isoxazoles, oxazoles, thiazoles, piperidines, quinazolines and isoquinolines. The fragment library was constructed using the following properties as filters: $150 \leq \text{molecular mass} \leq 250\text{Da}$, $-2 \leq \text{CLogP} \leq 3$, number of rotatable bonds ≤ 3 , number of H-bond donors ≤ 2 , and acceptors ≤ 4 along with high aqueous solubility.^{96b, 140} All compounds were provided as pure solids or oils in either 10 mg or 25 mg quantities. These compounds were dissolved in $^2\text{H}_6\text{-DMSO}$ to generate a concentrated stock solution that was stored at 4°C . Compounds were assessed for purity and solubility by recording 1D ^1H NMR spectra of single fragments (1 mM) in 100 mM phosphate buffer pH 7.0 in 99% $^2\text{H}_2\text{O}$, 1% $^2\text{H}_6\text{-DMSO}$. Those that were $< 90\%$ purity, questionable identity or not soluble under these conditions based on the chemical shifts, coupling and intensity of peaks in the NMR spectra were discarded before screening. For follow-up experiments, the desired compounds were purchased either from Maybridge, or where available from other

suppliers including Apollo Scientific Ltd (UK) and Sigma-Aldrich (Castle Hill, NSW, Australia).

2.1.1.2 Miscellaneous Chemicals

Escherichia coli strain BL21 Codon Plus (DE3)-RIL was purchased from Stratagene (Sydney, NSW, Australia). The following chemicals were purchased from Sigma-Aldrich (Castle Hill, NSW, Australia): glucose, glycine, imidazole, Isopropyl β -D-thiogalactopyranoside (IPTG), 1-butanol (99.8% HPLC grade), dithiothreitol (DTT), sodium dodecyl sulfate (SDS), tetramethylethylenediamine (TEMED), kanamycin B sulfate salt, 2-(N-morpholino)ethanesulfonic acid hydrate (MES), *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES), polymyxin B sulfate, Tris(hydroxymethyl)aminomethane hydrochloride (Tris HCl), potassium phosphate monobasic (KH_2PO_4), potassium phosphate dibasic (K_2HPO_4), polyethylene glycol (PEG) 8000, magnesium sulfate heptahydrate ($\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$), *p*-nitrophenyl phosphate, alanine, arginine, glutamine, histidine, isoleucine, leucine, lysine, proline, serine, threonine, valine, phenylalanine, tyrosine, *p*-aminobenzoic acid, *p*-hydroxybenzoic acid, tryptophan, asparagine, aspartic acid, glutamic acid, hypoxanthine, uracil, biotin, nicotinamide, riboflavin and thiamine and Bovine Serum Albumin (BSA) protein. The following chemicals were purchased from Proscience (Blackburn, VIC, Australia): acetic acid, acetic acid (glacial), ammonium persulfate, ammonium sulfate, 2-mercaptoethanol, sodium chloride, and bacteriological agar. Acrylamide-Bis solution and Bradford reagents were purchased from Bio-Rad (Regents Park, NSW, Australia).

Ammonium chloride ^{15}N ($^{15}\text{NH}_4\text{Cl}$) and ^{13}C glucose were purchased from Cambridge Isotopes Laboratories (Novachem Pty Ltd, Collingwood, VIC, Australia). Tryptone, yeast extract and Luria-Bertani (LB) medium were purchased from Oxoid (Adelaide, SA, Australia).

2.1.1.4 Columns

Columns for protein purification included a Phenyl Sepharose High Performance 16/10 (PHP) hydrophobic interaction column, ion exchange columns MonoQ 10/10 HR and MonoQ 5/5 HR, a gel filtration column (Superdex 75 HR 10/30) and a 5 mL HisTrap (nickel affinity column) which were purchased from GE Healthcare (Rydalmere, NSW Australia). Protein purification was undertaken on ÄKTA basic and ÄKTA purifier systems (GE Healthcare). 5 kDa molecular weight cut-off concentrators were purchased from Millipore (North Ryde, NSW, Australia).

2.2 Molecular Methods

2.2.1 Transformation of *E. coli*

Transformations of *E. coli* with expression plasmids for *EcDsbA* and *VcDsbA* were carried out chemically according to protocol adapted from Chung *et al.*¹⁴¹ The plasmid B0013-(5644bb) encoding DsbA from *E. coli* was kindly provided by our collaborators at University of Queensland, Brisbane, Australia. The *E. coli* expression system for *VcDsbA* was a kind gift from Ronald Taylor (Dartmouth University, New Hampshire). Competent

E. coli cells (50 μ L) were combined with the sample of 1 μ L plasmid DNA in a 1 mL centrifuge tube and incubated for 30 min on ice. Cells were heat shocked by placing them in a 42°C water bath for 45 seconds then returned to ice. 950 μ L of room temperature expression/recovery media (LB) was added to the competent cells in the culture tube and cells were grown for 1 hr with shaking at 250 rpm at 37 °C. The resulting bacterial culture was pelleted by centrifugation (at 14000 g and 4 °C for 1 min.) The supernatant was discarded and the pellet resuspended in 0.5 mL of fresh LB. The *E. coli* cells were plated onto LB agar (100 μ g/mL Kanamycin) and incubated for 16 hr at 37 °C.

2.2.2 *Glycerol stocks*

Glycerol stocks were prepared by resuspending *E. coli* overnight cultures in sterilised LB and glycerol media in 1:1 ratio. 1 ml cultures were transferred to microtubes and stored at -80 °C. This glycerol stock preparation protocol was adapted from Qiagen Pty Ltd (Doncaster, VIC, Australia) instructions.

2.2.3 *Buffer and media preparation*

All nutrient media used in protein expression and assays were autoclaved at 121 ° for 20 min and stored at 4 °C or room temperature (RT) prior to use. Heat liable additives were sterile-filtered prior to inoculation. Minimal media were prepared by combining a '1 x salt' solution and additive solutions after autoclaving. The compositions of different nutrient media are shown in tables 2-1, 2-2 2-3, and 2-8. All buffers were 0.22 μ M

filtered and degassed prior to use. Buffer compositions are described in the relevant sections below.

2.2.4 SDS PAGE gel

Protein samples were routinely analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) on a Protean3 Mini Gel apparatus (Biorad Reagents Park, NSW, Australia). SDS-polyacrylamide gel electrophoresis (SDS-PAGE) separates proteins according to MW and by their different running speed through an acrylamide polymer.¹⁴² In this study, a 15% polyacrylamide resolving gel was used with a sample loading volume usually 5-10 μ L. Samples of soluble fractions were mixed in 1:1 ratio with Laemmli buffer (50mM TrisHCl pH 6.8 10% glycerol 2.0% SDS and 0.01% w/v bromophenol blue). To prepare cell pallet samples for gel, 100 μ L sample of cell lysate or culture was spun down by centrifugation at 14000 g for 10 min at 4 °C. Samples containing cell pallets were heated to 100 °C for 10 min prior to loading in double volume of Laemmli buffer with reduced samples also containing β -mercaptoethanol (0.05%). Loaded samples were stacked by running at constant volts (100V or 20mA) until the sample reached the resolving gel and strength was then increase to 280V or 90 mA. Gels were stained by a solution containing 1 g/L Coomassie blue, 100 mL/L acetic acid and 300 mL/L ethanol for 15 min. Gels were destained in solution containing 30% methanol and 10% acetic acid to visualize the proteins. Additionally gels were silver stained using a protocol Adapted from Shevchenko A *et al.*¹⁴³

For silver staining, gels were soaked in a fixing solution (40% methanol, 13.5% formalin) twice for 10 min prior to washing them in water for 5 min. In next step gels were soaked for 1 min in 0.02% $\text{Na}_2\text{S}_2\text{O}_3$ (Sodium Thiosulphate) twice and washed in water for 20 sec. The gels were soaked for 10 min in 0.1% AgNO_3 and rinsed with water and again with small volume of developing solutions (3% sodium carbonate, 0.05% formalin, 0.000016% $\text{Na}_2\text{S}_2\text{O}_3$). After that gels were soaked in fresh developing solution (fresh formalin) until band intensities were visible (1-3 min) and 12.5 mL of 2.3 M citric acid (or 6 g of solid citric acid) was added with continue shaking for 10-15 min. Finally gels were washed (for 10 min) and soaked in water for 30 min or longer before drying. Molecular weights were estimated by comparison to a set of known standards (Precision Plus Protein, Biorad) as markers from 10 kDa to 250 kDa.

2.2.5 Protein expression

2.2.5.1 Expression of unlabeled *EcDsbA*

Unlabeled *EcDsbA* was expressed using the autoinduction method adapted from Studier's protocol.¹⁴⁴ The method involves the expression of recombinant proteins using the T7 polymerase system without requiring inducers such as IPTG during mid-log phase of the cultures. A buffered medium containing a mixture of carbon sources is utilized. The composition of components and media used for autoinduction is described in table 2-1 and table 2-2. Autoinduction media was prepared using LB (Lucia broth) media, 1M MgSO_4 , 20x NPS, 50x5052 and 1000x trace metals mixture with antibiotic kanamycin.

Starter cultures were prepared by inoculation of autoclaved media with from glycerol stocks of *E.coli* B0013 and grown over night at 37°C, 260 rpm. These cultures were transferred to the large autoinduction media and shaken and incubated at 37°C, 260 rpm until an *A*₆₀₀ (Absorbance at 600 nm) of saturation point was reached. Saturation point of cell growth was found after ~ 24 hrs at OD (1:10) 1.49 for *E. coli* B0013 cells protein expression culture. This culture was spun down by centrifugation at 14000 g for 30 min at 4 °C. The pellet was resuspended in 50 mL of polymyxin sulfate B (4 mg/mL in 20mM Tris pH 8.0), with shaking at 200 rpm for 24 hrs at 20 °C to release the periplasmic fraction containing DsbA.¹⁴⁵ The periplasmic fraction was clarified by centrifugation at 25000 g for 20 mins at 4°C to obtain the supernatant.

Table 2-1: Composition of the media used for the auto-induction method.

| Component | 1 liter | Concentration |
|-----------------------|---------|---------------|
| ZY | ~928 ml | - |
| 1 M MgSO ₄ | 1 ml | 1 mM |
| 50x 5052 | 20 ml | 1x |
| 20x NPS | 50 ml | 1x |
| kanamycin (25 mg/ml) | 4 ml | 100 µg/ml |

Table 2-2: Composition of media components used in auto induction method

| 20x NPS pH ~6.75 | | 1000x trace metals mixture (100 ml in ~50 mM HCl) | |
|---|----------------|---|-------|
| Component | Quantity (1 L) | H ₂ O | 36 ml |
| Sterile Water | 900 ml | 0.1 M FeCl ₃ •6H ₂ O (dissolved in ~0.1 M HCl) | 50 ml |
| (NH ₄) ₂ SO ₄ | 66 g | 1M CaCl ₂ | 2 ml |
| KH ₂ PO ₄ | 136 g | 1M MnCl ₂ •4H ₂ O | 1 ml |
| Na ₂ HPO ₄ | 142 g | 1 M ZnSO ₄ •7H ₂ O | 1 ml |
| 50x5052 | | 0.2 M CoCl ₂ •6H ₂ O | 1 ml |
| Glycerol | 250 g | 0.1 M CuCl ₂ •2H ₂ O | 2 ml |
| H ₂ O | 730 ml | 0.2 M NiCl ₂ •6H ₂ O | 1 ml |
| Glucose (| 25 g | 0.1 M Na ₂ MoO ₄ •2H ₂ O | 2 ml |
| α-lactose | 100 g | 0.1 M Na ₂ SeO ₃ •5H ₂ O | 2 ml |
| 1 M MgSO ₄ (100 ml) | | 0.1 M H ₃ BO ₃ | 2 ml |
| MgSO ₄ •7H ₂ O | 24.65 g | Kanamycin (25 mg/ml in water, filter sterilize) | |
| Sterile Water | to make 100 ml | LB (in MilliQ H ₂ O) to make 1 L) | |
| | | Tryptone | 10 g |
| | | NaCl | 5 g |
| | | Yeast Extract | 5 g |

2.2.5.2 Expression of labeled *EcDsbA*

Expression of uniformly ^{15}N and ^{15}N ^{13}C isotope-labeled *EcDsbA* and *VcDsbA* proteins was done according to the protocol described by Marley *et al.*¹⁴⁶ An overnight streak culture was grown from glycerol stock on LB/agar *kan*⁺ plates. An isolated colony was used to inoculate a 100 ml starter culture of LB/*kan*⁺ and grown overnight at 37 °C. These starter cultures were used for the inoculation of large cultures consist of 3L LB and then shaken and incubated at 37°C, 260 rpm until an A_{600} OD of 0.6 was reached. These cells were centrifuged at 14000 g for 30 min at 4 °C and resuspended in expression media. Minimal media (M9 salts) was used for the labeled protein expression and composition of this media listed in table 2-3. To achieve the ^{15}N and ^{13}C labeling ^{15}N -labeled ammonium chloride and ^{13}C -labeled glucose were present as the sole nitrogen and carbon source, respectively. Following transfer to the labelling media the cultures were grown at 37°C for 1 hour and induced with 1mM IPTG. The culture was grown post-induction 5-8 hrs to allow protein expression, followed by pelleting by centrifugation at 14000 g for 30 min at 4 °C. Expressed labeled DsbA was released from the harvested cells as described for unlabeled *EcDsbA*.

Table 2-3: Composition of minimal media used for IPTG induction method

| 1 x M9 salt | | Additives | | Heat liable additives (0.22 µM filtered) | |
|---------------------------------|----------------------|---------------------------------|--------------|--|----------|
| Component | Quantity (100 ml) | Component | 1L | Component | Quantity |
| KH ₂ PO ₄ | 1 g | FeCl ₃ (10 mg/mL) | 0.2 mL | D-glucose (20 g/100 mL) | 20 mL |
| K ₂ HPO ₄ | 6.7 g | 1M MgSO ₄ | 2 mL | Thiamine (2 mg/mL) | 5 mL |
| NaCl | 0.5 g | 1M CaCl ₂ | 2 mL | Kanamycin (100 mg/mL) | 1 mL |
| NH ₄ Cl | 1 g | ZnCl ₂ (34 mg/mL) | 0.2 mL | | |
| MilliQ H ₂ O | to 100 mL | | to 874 mL | | |

2.2.6 Protein purification

Purification of labeled and unlabeled *EcDsbA* and *VcDsbA* proteins was carried out using 2 chromatographic steps on FPLC.¹⁴⁷ Firstly, ammonium sulfate was added to the supernatant to a final concentration of 1M. The supernatant was filtered (I hope! - put in details here) and applied to a Phenyl Sepharose High Performance 16/10 hydrophobic interaction¹⁴⁸ column which had been equilibrated in 20 mM Tris (pH 8.0), 50 mM NaCl, 1 M (NH₄)₂SO₄.

A linear gradient from 100% 1 M ammonium sulfate buffer B to 100% low salt buffer A was run over three column volumes to elute the protein (Figure 2-1).

Buffer A: 20 mM Tris, 50 mM NaCl (pH 8.0)

Buffer B: 20 mM Tris, 50 mM NaCl, 1 M $(\text{NH}_4)_2\text{SO}_4$ (pH 8.0)

Fractions were collected automatically and DsbA was located by the UV profile of the elution and the appropriate fractions analysed by SDS-PAGE followed by silver staining.¹⁴⁹ Fractions containing DsbA were pooled and buffer exchanged into 20 mM Tris (pH 8.0) using HiPrep 26/10 desalting column. This PHP purified protein was applied to MonoQ HR5/5 cation-exchange column.¹⁵⁰ The protein was eluted with a gradient from 100% low salt buffer C to 100% high salt buffer D (Figure 2-2).

Buffer C: 20 mM Tris (pH 8.0).

Buffer D: 20 mM Tris + 1 M NaCl (pH 8.0)

This two-step purification protocol was sufficient to pure protein, which was estimated to be > 95% pure from analysis by SDS-PAGE followed and silver staining (Figure 2-2).

The final yield was about 50-70 mg of protein/L of medium.

2.2.7 Oxidation, reduction & estimation of protein concentration

Purified DsbA from pooled fractions was oxidized by treatment with a solution of copper phenanthroline (30: 9: 61 of 100 mM CuCl_2 : 1 M Phenanthroline: Tris 20 mM) at 4°C for 4-5 hrs. Excess copper phenanthroline was removed by buffer exchanged using a PD-10 (supplier) or HiPrep 26/10 desalting columns (Figure 2-3). DsbA samples were reduced by the addition of DDT to the final concentration of 1 mM into the protein samples which were incubated for 1hr at 4°C for complete reduction. Buffer exchange was performed at various stages to transfer DsbA protein sample into the appropriate

buffer systems for different experiments including NMR, crystallography and assays. For instance, in STD NMR experiments protein was buffer exchanged using desalting column equilibrated with 10mM HEPES (pH 7.0), 150 mM NaCl.

The concentration of DsbA protein samples was estimated using UV-visible spectrophotometry where the absorbance of each sample was recorded at a wavelength of 280 nm. The protein concentration was calculated using an extinction coefficient of $28545 \text{ M}^{-1}\text{cm}^{-1}$ and $10555 \text{ M}^{-1}\text{cm}^{-1}$ for *EcDsbA* and *VcDsbA* respectively, calculated from amino acid composition using the ExPASy site (<http://au.expasy.org/tools/protparam.html>). In addition, Bradford assays were performed using BioRad kit according to manufactures' instructions. UV measurements were made either on a Cary50 spectrophotometer (Varian, Mulgrave, Vic) or a Nanoview Spectrophotometer (GE Healthcare). Mass spectroscopy analysis results for protein characterization have been presented in Appendices section 8.2. For 2D, 3D NMR and crystallographic analysis, labeled and unlabeled *EcDsbA* was concentrated to 150 μM , 500 μM and 1200 μM respectively in the corresponding buffer systems required for each experiment described in relevant section.

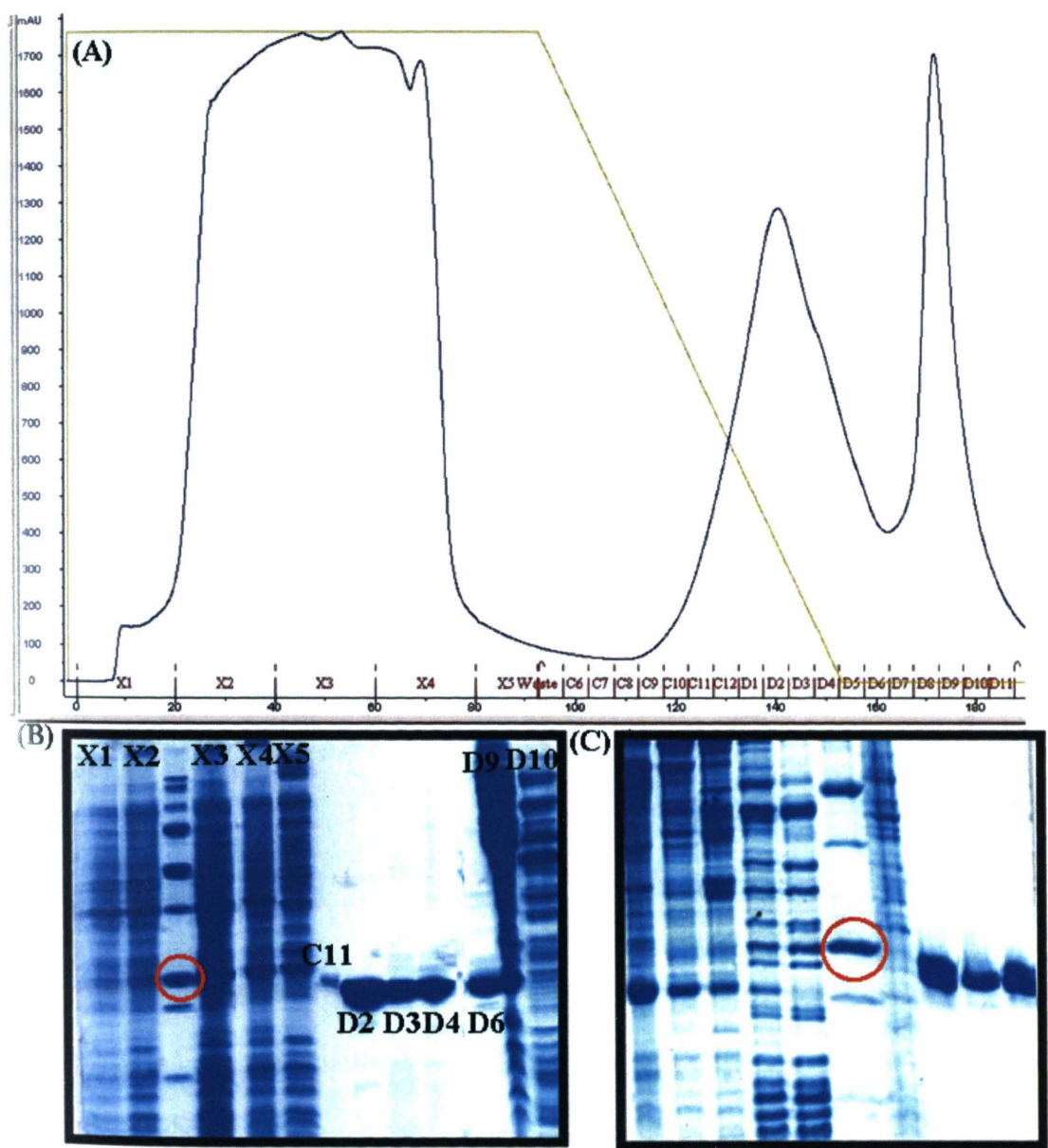


Figure 2-1. (A) A representative chromatogram depicting the result of injecting unlabeled *EcDsbA* over the PHP 16/10 hydrophobic interaction column. The brown and green lines represent the UV₂₈₀ reading of the eluted fractions and linear buffer gradient, respectively. (B) and (C) SDS-PAGE of selected fractions among X1-X4 and C9-C11, from two different PHP FPLC runs for ¹⁵N and unlabeled *EcDsbA* respectively. Red circle represents 21 kDa MW standard and arrowed lanes suggestive of purified DsbA.

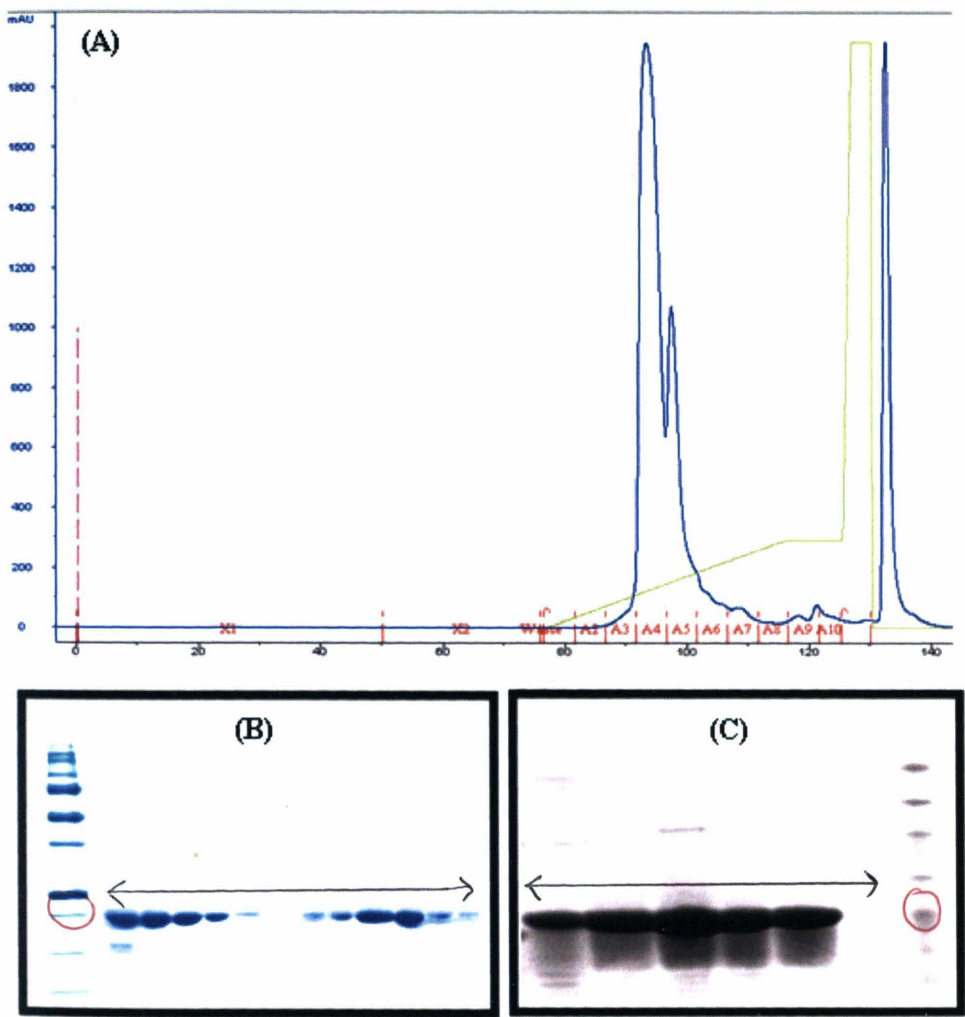


Figure 2-2. (A) A representative chromatogram depicting the result of injecting PHP purified ^{15}N labeled protein over the MonoQ 5/5 cation-exchange column. The blue and green line represents the UV₂₈₀ reading of the eluted fractions and linear gradient respectively (B) and (C) SDS-PAGE and silver stained gel for few selected fractions among A3-A10 from two different MonoQ 5/5 FPLC runs for $^{13}\text{C}^{15}\text{N}$ and ^{15}N *EcDsbA* respectively. Red circle represents 21 kDa MW standard and arrowed lanes suggestive of purified DsbA.

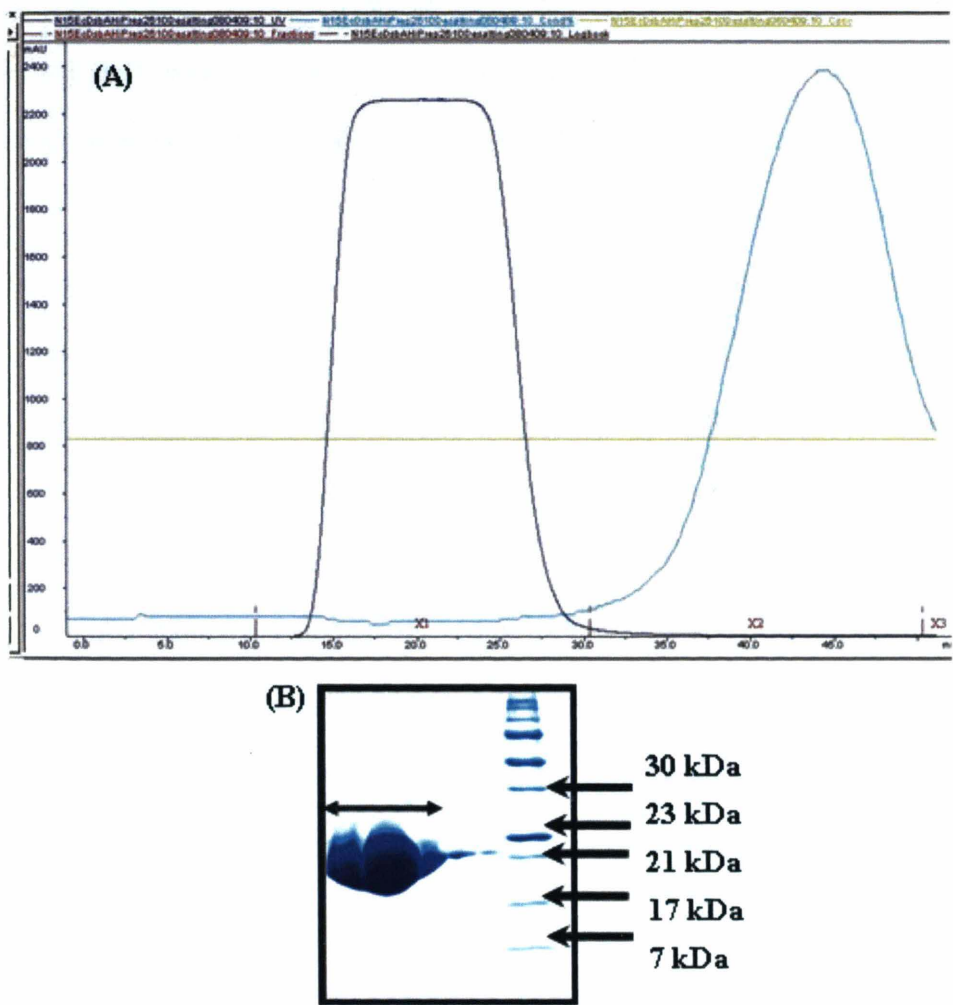


Figure 2-3. (A) A standard chromatogram of DsbA depicting the result of injecting MonoQ purified and oxidized ^{15}N labeled DsbA protein over the HiPrep 26/10 desalting column. The brown, green and cyan line represents the UV₂₈₀ reading of the eluted fractions, linear gradient and conductivity respectively. (B) SDS-PAGE fractions X1 for ^{15}N *EcDsbA* loaded in different volumes (7 μL , 10 μL and 4 μL). Arrowed lanes suggestive of purified DsbA whereas MW standard have been shown in the right side lane.

2.2.8 Protein stability analysis

A thermal shift assay was performed to find out buffer condition for which *EcDsbA* exhibited the greatest thermal stability.¹⁵¹ This fluorescence-based assay was performed using an Applied Biosystems Step One Plus Thermal Cycler (Melbourne, VIC, Australia). Solutions of 5 μ L *EcDsbA* (30 μ M stock, final concentration 5 μ M), 3 μ L 100X Sypro orange dye (final concentration 10X) and 22 μ L tested buffer were added to the wells of the 96-well plate. The total well volume used was 30 μ L. The plate was heated from 20°C to 99°C with a heating rate of 1.0 °C/minute. The fluorescence intensity was measured with excitation/emission = 583/610 nm. The different buffer solutions tested were Acetate (pH=4-4.5), Citrate (pH=5.0), MES (pH=6-6.5), HEPES (pH=7.5-8.0), Tris (pH 8.0-8.5) containing different additives (50 mM – 150 mM NaCl, 0.5 mM arginine, 0.5 mM glutamine amino acids and combinations of additives) at varying ratio. The transition region of the raw thermal melting curve data were fitted by non-linear least squares regression to determine the mid-point temperature of the protein un-folding transition (T_m). After testing several buffer conditions, *EcDsbA* was found to be most stable in 50mM MES (pH 6.5), 150 mM NaCl, 0.5 mM RE (Arginine, Glutamine Acid) as shown in melting curves (Figure 2-4).

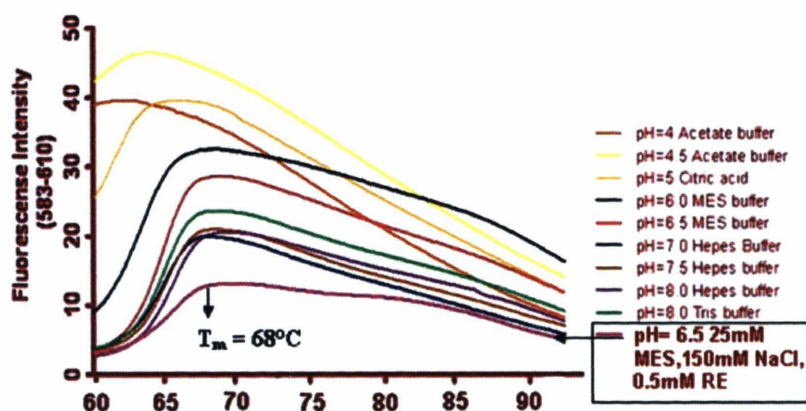


Figure 2-4: Protein stability buffer screening for *EcDsbA* using thermal shift assays.

2.3 Functional assays

2.3.1 Alkaline phosphatase activity assay

Alkaline phosphatase (AP) assay can measure the activity of DsbA *in vivo* by observing its capability to oxidise bacterial periplasmic protein alkaline phosphatase. Native AP contains a pair of cysteines residues which require DsbA for oxidative folding.^{55b, 59} AP is inactive in the cytoplasm and active only when it is exported to the periplasm, where formation of intramolecular disulfide occurs by DsbA. This reaction can be quantified and measured spectrophotometrically in presence of a substrate of AP (e.g. p-nitro phenol).^{55b, 59} Alkaline phosphatase activity of bacterial strains with or without DsbA can be compared in these conditions along with compounds to be tested for inhibition of DsbA activity.

Alkaline Phosphatase (PhoA) activities were measured essentially according to the protocol described by Brickman & Beckwith.¹⁵² Briefly, bacteria were grown overnight in 20 ml minimal medium A which is a low phosphate media. (Table 2-4) The low phosphate concentration induces the synthesis of alkaline phosphatase (PhoA) from the chromosomal copy of the alkaline phosphatase gene (*phoA*).⁵⁹ Cells were incubated on ice for 20 min, pelleted by centrifugation for 10 min at 2350 g and washed twice with 20 ml 1 M Tris, pH 8.0. After the last centrifugation, cells were resuspended in 2 ml 1 M Tris pH 8.0, and the OD600 was measured in a spectrophotometer (Cecil CE2021). The cells were then permeabilized by adding 50 μ l 0.1% (w/v) SDS and 100 μ l chloroform and mixed by vortexing. Samples were equilibrated for 5 min at 28 °C after which 500 μ l nitrophenol phosphate [0.4% (w/v) in 1 M Tris, pH 8.0] was added and the time recorded. The reaction was allowed to proceed at 28 °C until a yellow colour developed. As soon as the colour change was noted, the reaction was stopped by adding 1 ml 1 M KH_2PO_4 . Samples were vortexed before the A420 and A550 were measured and units of PhoA activity were calculated from the absorbances. Preliminary results obtained from the analysis of fragment inhibitory activity studies using AP assay have been shown in Appendices section 8.3.1.

Table 2-4: Composition of minimal media used for alkaline phosphatase activity assay

| Components (for 1L) | Quantity |
|---|-----------------------|
| K ₂ HPO ₄ | 10.5 g |
| KH ₂ PO ₄ | 4.5 g |
| (NH ₄) ₂ SO ₄ | 1g |
| sodium citrate | 1g |
| glucose | 0.4% w/v |
| vitamin B1(thiamin) | 1 ug ml ⁻¹ |
| MgSO ₄ | 1 mM |
| MilliQ | To make 1L |

Units of PhoA activity were calculated from the following formula:

1000* (A420 –1.75*A550) / (Time* OD600 of culture) * Dilution factor

Where A420 and A550 are measured absorbances at the end of the assay, OD600 is the OD reading of the culture. Time is the time from addition of AP substrate to making the readings. Calculations were correlated with "dilution factor" in cases where cells were grown to a high OD in the overnight culture and were diluted in same buffer (1 M Tris pH 8.0).⁵⁹

2.3.2 Insulin Reduction Assay

The insulin reduction assay was employed to measure disulfide oxidoreductase activity of DsbA in presence of selected potential inhibitors.⁵⁰ DsbA catalyses the reduction of the disulfide bonds in insulin, which leads to the precipitation of the insulin B chain. Hence DsbA activity can be measured turbidometrically by recording the increase in absorbance at 650 nm. Reagents used in the present assay are listed in table 2-5. Stock solutions of 10 mM insulin were prepared by dissolving insulin in 0.1 M HCl. These stocks were then diluted to the desired concentration in buffer containing 200 mM sodium phosphate and 1.2 mM EDTA, pH 8.

Table 2-5. Reagents used in insulin precipitation assay

| Reagents | Final concentration |
|-----------------------------------|--|
| <i>EcDsbA</i> 50 μ M | 5 μ M |
| DTT 100 mM | 1 mM |
| Fragments (varying Concentration) | Depended on the K_D for DsbA (100 μ M – 3 mM) |
| Insulin 10 mM | 0.5 mg/mL |

This assay was carried out initially in a 1 ml cuvette to optimize concentration of each component including insulin, DTT and DsbA. After the parameters had been optimised in a cuvette-based assay, it was transferred to a 96-well plate format. The assay buffer

contained 200 mM sodium phosphate and 1.2 mM EDTA, pH 8.0. *EcDsbA* was added to a final concentration of 5 μ M. Fragments were added from concentrated stocks in DMSO, such that the DMSO concentration in each well was constant. A DMSO-only control containing no fragment was run on each plate. DTT was added to a final concentration of 1 mM from a stock solution at a concentration of 100 mM.

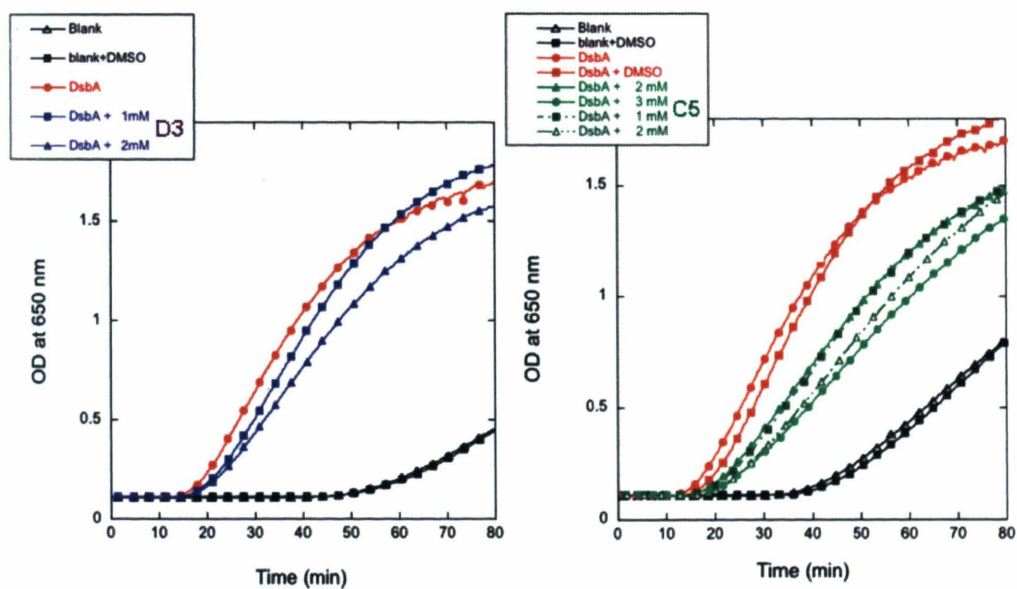


Figure 2-5: UV spectrum obtained from insulin reduction assay in presence two of the *EcDsbA* binding fragments **D3** and **C5** respectively at different concentrations 1-3 mM.

The reactions were initiated by addition of insulin from a 10 mM stock to a final concentration of 0.5 mg/mL. The reactions were performed at 298 K and were monitored at 650 nm on a UV plate reader (Molecular Devices Versa Max plate reader) for 90 mins. The resultant data¹⁵³ were analysed using the initial slopes during precipitation, to give an

apparent rate of precipitate formation during precipitation phase (Figure 2-5). The steeper the slope, the faster the precipitation occurred. These initial rates were plotted against the tested fragment concentrations to measure any inhibitory effects on the rate of precipitation. Preliminary results obtained from the analysis of fragment inhibitory activity studies using insulin assay have been shown in Appendices section 8.3.2.

2.3.3 Fluorescence (FRET) Peptide assay

A fluorescence resonance energy transfer (FRET) peptide based *in vitro* assay was optimized by Dr Kieran Rimmer (MIPS) in order to measure DsbA inhibitory activity of selected compounds. This assay is based on an approach described by Ruddock et al and measures the ability of DsbA to oxidise a peptide substrate containing appropriate fluorogenic moieties⁵¹ (Figure 2-6) that was synthesized by Mr Bradley Doak (MIPS). All the initial fluorescence measurements were performed at 298 K on a Varian Cary Eclipse Spectrophotometer with excitation wavelength of 340 nm and emission wavelength of 620 nm. After optimization the assay was performed in 384 well Perkin Elmer white OptiPlate-384 (Part #: 6007290) on a Perkin-Elmer Envision Plate Reader. The composition of assay buffers has been listed in table 2-6.

Each well contained 25 μ L of buffer A which contained DsbA plus DsbB, a zero DsbB control or a zero DsbA control, respectively. Experiments were undertaken in the presence of increasing concentrations of test fragments, added from appropriate stock solutions such that the concentration of DMSO was kept constant. Control experiments

were performed containing either DMSO only (i.e. no fragment) or no-DMSO no-fragment on each plate.

Finally the FRET-Peptide (22.5 μ L at 5 μ M) was added to begin the reaction. Fluorescent readings were begun immediately following addition of the peptide. Figure 8-5 (and Appendices section 8.3.2) shows the kinetic inhibition plot of DsbA oxidation activity (DMSO blank) in various concentrations (0.4 mM to 20 mM) of an inhibitor in the FRET-peptide assay.

Table 2-6. Reagents used in fluorescence peptide assay

| Components | Final concentration |
|--|---|
| Buffer A (50 mM MES, 50 mM NaCl, 0.2 % glycerol, 2 mM EDTA, pH 5.5) | 1.7 μ M <i>EcDsbB</i> (in membrane prep) 40 nM <i>EcDsbA</i> |
| Buffer B (zero DsbB control) (50 mM MES, 50 mM NaCl, 0.2 % glycerol, 2 mM EDTA, pH 5.5) | 40 nM <i>EcDsbA</i> |
| Buffer C (zero DsbA control): (50 mM MES, 50 mM NaCl, 0.2 % glycerol, 2 mM EDTA, pH 5.5) | 1.7 μ M <i>EcDsbB</i> (in membrane) |
| Fragment solutions (in 100% DMSO) | 300 μ M- 20 mM or varied |
| FRET-Peptide | 5 μ M |

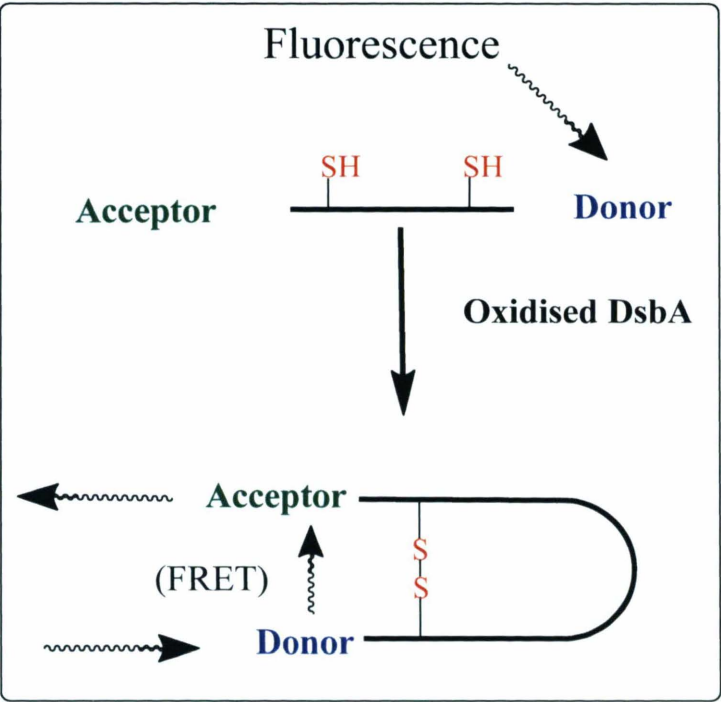


Figure 2-6: Schematic of oxidation of fluoro-peptide by DsbA in FRET system

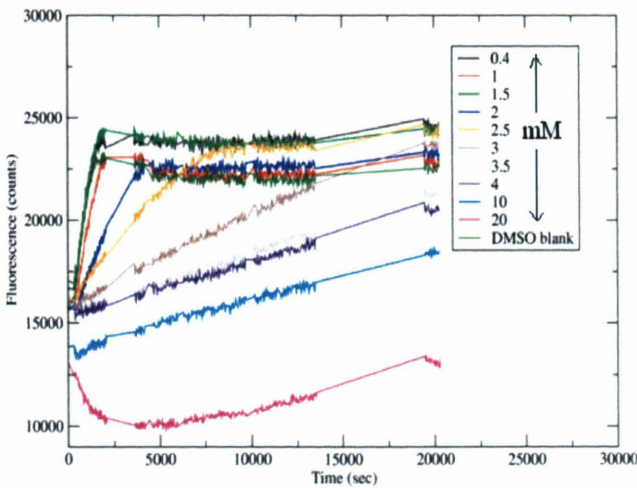


Figure 2-7: Inhibitor concentration-dependant plots of DsbA oxidising activity in FRET-peptide assay.

2.3.4 Complementation/Motility Assay

Bacterial strains and plasmids for this cell based assay were kindly provided by Prof J. C. Bardwell (Dept of Microbiology and Molecular Genetics, Harvard Medical School, Boston, USA). The *E. coli* strains used for this study^{15a} are described in Table 2-7. The protocol of this bacterial motility based DsbA complementation assay was adapted from previous studies^{55b, 154} and was optimized to measure the motility inhibition activity of potential inhibitors of DsbA. The assay was performed in M63 minimal media containing 40 µg ml⁻¹ of each essential amino acid (except cysteine and cystine) and 0.3% agar. The pH of the media was adjusted to 7.0-7.5 and additives as listed in Table 2-8 were added.

Table 2-7: *E. coli* strains used for Motility assay

| <i>E. coli</i> JCB Strains | Genotype |
|----------------------------|---|
| JCB570 | Wild-type <i>E. coli</i> with DsbA normal function (wt) |
| JCB571 | <i>kan</i> ⁺ Mutated <i>E.coli</i> cells with DsbA gene deactivated (<i>dsbA</i> ⁻) |
| JCB571- VcDsbA | <i>kan</i> ⁺ Mutated <i>E. coli</i> cells with VcDsbA function (<i>pVcDsbA</i>) <i>E. coli dsbA</i> ⁻ mutant transformed with VcDsbA plasmid |
| JCB816 | Wild-type <i>E. coli</i> with DsbA normal function |
| JCB817 | Mutated <i>E. coli</i> with DsbA gene deactivated (<i>dsbA</i> ⁻) |

After being autoclaved the agar was allowed to cool to ~65 °C at which point the appropriate additives were added with thorough mixing. This media was poured in 12 well culture plates (Greiner Multi well Plates, Sigma) to final volume 1 ml in each well.

Compounds to be tested were added from concentrated stocks in DMSO and mixed thoroughly. These plates were allowed to cool at room temperature before inoculation of *E.coli* strains. Negative controls (JCB817), DMSO controls and blank experiments were performed within each experimental trial as triplicates.

For end point experiments, the compounds (or fragments) to be tested were added from to final concentrations of 500 μM or 1 mM. The amount of DMSO was kept constant as 4 μL in all experiments including end point, positive control and inhibition calculation experiments. Serial dilutions of the original compound stock were prepared to give final concentrations of the compounds in media ranging from 0.1 μM to 1 mM for the titration and effective inhibition (EC_{50}) calculation. All assays were performed in triplicate.

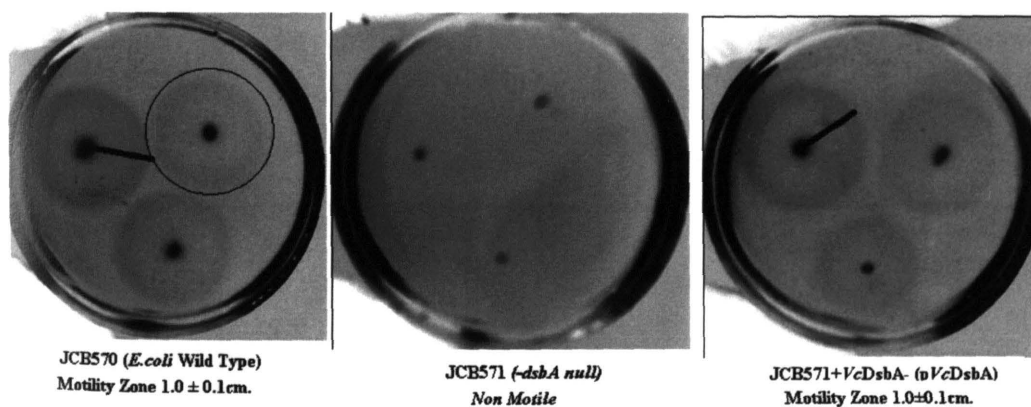


Figure 2-8: Motility of *E. coli* wild type and mutant strains on soft agar with cysteine free minimal media.

Table 2-8: Composition of minimal media used for complementation/ motility assay.

| Components | Composition | Final concentration (v/v) | Sterilisation and storage conditions |
|--|--|---------------------------|--------------------------------------|
| M63 Stock Salts MilliQ water to 1L | KH ₂ PO ₄ (15 g) K ₂ HPO ₄ (45.6 g) (NH ₄) ₂ SO ₄ (10g) FeSO ₄ (6.25 mg) | 20% | Autoclaved Stored at 4 °C |
| MgSO ₄ 1M MilliQ to 50 ml | MgSO ₄ (1.32g) | 0.1% | Autoclaved Stored at 4 °C |
| AAA 100X MilliQ water to 50 ml pH to 7-7.5 | Alanine, Arginine, Glutamine, Glycine, Histidine, Isoleucine, Leucine, Lysine, Proline, Serine, Threonine, Valine (200 mg each) | 1% | Autoclaved Stored at 4 °C |
| BAA 100X 0.1MKOH to 50 ml | Phenylalanine, Tyrosine, p-aminobenzoic acid, p hydroxybenzoic acid (200mg each) Tryptophan (100mg) | 1% | Filter sterilise Stored at 4 °C |
| FAA 100X MilliQ water to 50 ml pH to 7-7.5 | Asparagine, Aspartic acid, Glutamic acid (200mg each) | 1% | Filter sterilise Store at 4 °C |
| Hypo /Uracil DMSO to 50 ml | Hypoxanthine, Uracil (200mg each) | 1% | Filter sterilise Store at -20 °C |
| Vitamins MilliQ water to 50 ml | Biotin, Nicotinamide, Riboflavin, Thiamine (10mg each) | 1.67% | Filter Sterilise Store at 4 °C |
| Agar Solution MilliQ water to 460 ml | Agar (1.5 gm) | 72.8% | Autoclave |

| | | | |
|------------------------|----------------|----|---------------------------------------|
| Glycerol 50ml (20%) | 10 ml Glycerol | 2% | MilliQ water to 50 ml Autoclave |
|------------------------|----------------|----|---------------------------------------|

E. coli wild type (JCB570 and JCB816), *dsbA*⁻ (JCB571 and JCB817) and *VcDsbA* complemented *dsbA*⁻ (JCB571 - p*VcDsbA*) mutant strains were grown individually as streak cultures in LB plates at 37 °C overnight to obtain single colonies of each type of strain. Triplicates of fresh single colonies were inoculated by stabbing at the centre of each well containing desired amount of tested compounds in MM media. Plates were incubated for 7-10 hours at 37°C and a lawn of cells was allowed to grow before analysis of the swarming zone of motile *E. coli* cells. Figure 2-8 shows an example of control experiment in motility assay performed in 5 ml culture plates instead of 12 well plates for picture clarity. The diameter of bacterial swarming around the center (inoculation point) was measured manually. An example of the analysis of fragment inhibitory activity studies using motility assay has been shown in Appendices section 8.3.3.

Data analysis was performed using GraphPad Prism version 5.0.¹⁵⁵ Swarming zone diameters were normalized to the DMSO-only control and the effect of the compounds was calculated as the percentage inhibition of motility, which were fitted against inhibitor concentration in logarithmic scale to a sigmoidal dose (concentration) response using Equation 2-1 given below (Figure 2-9).

Equation 2-1

Y= Bottom+ (Top-Bottom)/(1+10^((LogEC₅₀-X)))

Where X is the logarithm of concentration, Y is response that starts at Bottom (base line response) and goes to Top (maximum response) with a sigmoid shape (Hill slope).

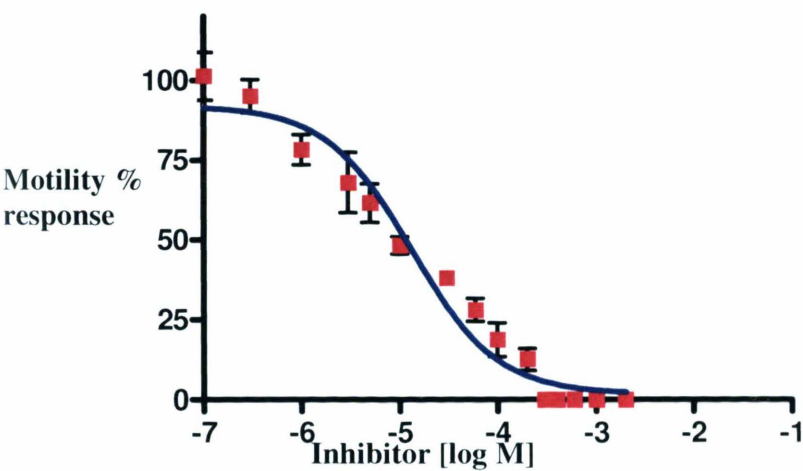


Figure 2-9: Motility inhibition of pVcDsbA *E.coli* mutants shown as sigmoidal dose response plot obtained from titration of a fragment hit **M5** identified by NMR guided computational screening.

2.4 NMR Methods

NMR experiments performed for this project can be divided in two parts:

1. NMR experiments for the biophysical screening of fragment library and elaborated compounds: STD and ^1H -1D followed by validating experiments ^1H - ^{15}N HSQC/HMQC end point and ^1H - ^{15}N HSQC titrations.¹⁵⁶
2. Protein triple resonance NMR experiments¹⁵⁷ for backbone and side chain assignment of DsbA: ^{13}C HSQC, ^{15}N CBCACONH¹⁵⁶, HNCA¹⁵⁸, HNCACB¹⁵⁹, HNCO¹⁵⁶, HNCACO¹⁶⁰, HCCONH-TOCSY, CCONH-TOCSY and HCCH-TOCSY¹⁶¹

2.4.1. STD NMR experiments

STD (Saturation transfer difference) NMR experiments¹⁶² for fragment library screening were performed in standard NMR tubes with 10 μM unlabelled protein and mixtures of up to 5 fragments with a concentration of ~ 300 μM for each fragment. The sample volume was 550 μL with 10% D_2O and 90% buffer containing 10 mM HEPES pH 7.0 and 150 mM NaCl. 1D NMR experiments were performed in 99.9% D_2O for each fragment before screening. Spectra were recorded on a Bruker BioSpin 800 MHz spectrometer (Billerica, MA, USA) using a triple-resonance 5 mm TXI cryoprobe equipped with single-axis gradients and an autosampler. Standard pulse sequences were used for 1D and STD data acquisition and water suppression was achieved using the WATERGATE scheme. STD-NMR experiments were carried out at 10 $^\circ\text{C}$.¹⁶³. ^1H chemical shifts were referenced to the $^1\text{H}_2\text{O}$ signal at 4.70 ppm. The 3s irradiation period

consisted of a train of 50 ms Gaussian pulses separated by a 1 ms delay. A 35 ms T_2 filter was used to suppress residual protein signals.¹⁶⁴ STD data sets were the average of 128 scans. NMR data were processed in TOPSPIN version 3.1 (Bruker BioSpin).¹⁶⁵ Automated baseline smoothing was applied using a polynomial function. Figure 2-10 illustrates an example of hit identification for a mixture of compounds using STD spectra.

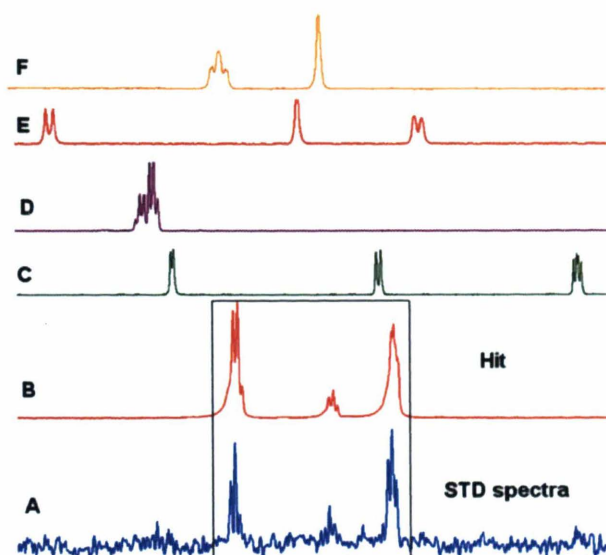


Figure 2-10: Identification of an *EcDsbA* binder from STD NMR based screening by comparing the STD spectrum (A) with the reference 1D ^1H spectra for each fragment (spectra B–F). Rectangular Box connects resonances observed in the STD spectrum with their corresponding signals in the 1D ^1H reference spectrum of the ligand.

2.4.2 HSQC NMR experiments

Second round of fragment screening was performed by recording 2D ^1H - ^{15}N HSQC (Two-dimensional ^1H - ^{15}N heteronuclear single-quantum correlation) experiments for DsbA in the presence of individual fragments hits identified from STD screen. 2D ^1H - ^{15}N HMQC (Heteronuclear Multiple Quantum Coherence) experiments were also performed for the ligand binding studies at the subsequent stages of the project. All HSQC / HMQC spectra were recorded at 298K. Initially, end point ^1H - ^{15}N HSQC experiments were performed in Shigemi (NMR) tubes with 1mM compound (1 mM) to be tested in 300 μL protein sample. Protein samples were prepared with 100 μM - 150 μM ^{15}N -labeled DsbA proteins (*Ec*DsbA or *Vc*DsbA) in buffer containing 10 mM HEPES or 50mM HEPES (pH 7.0) and 150mM NaCl and 1 mM EDTA with 90% H_2O and 10% D_2O . Spectra were recorded on a Varian Unity Inova 600 MHz spectrometer (Varian Inc., USA) equipped with a single axis gradient triple resonance cryoprobe and on a Bruker BioSpin 800MHz spectrometer as mentioned above. Typical acquisition time was 30 min and 15 min per experiment for HSQC and HMQC respectively. Standard pulse sequences were used for data acquisition. NMR data were processed using NMRPipe^{165b} and processed spectra were analyzed initially with NMR Draw and then with Sparky 3.113¹⁶⁶ software. In total, about 200 2D spectra were recorded for the secondary screen, about 125 spectra for fragment elaboration process and another 450 spectra were recorded K_D determinations. A simple overlay of each screening spectrum was measured without any automation at this stage as this was found as most robust and sufficiently quick procedure for data analysis (Figure 2-9).¹⁶⁷

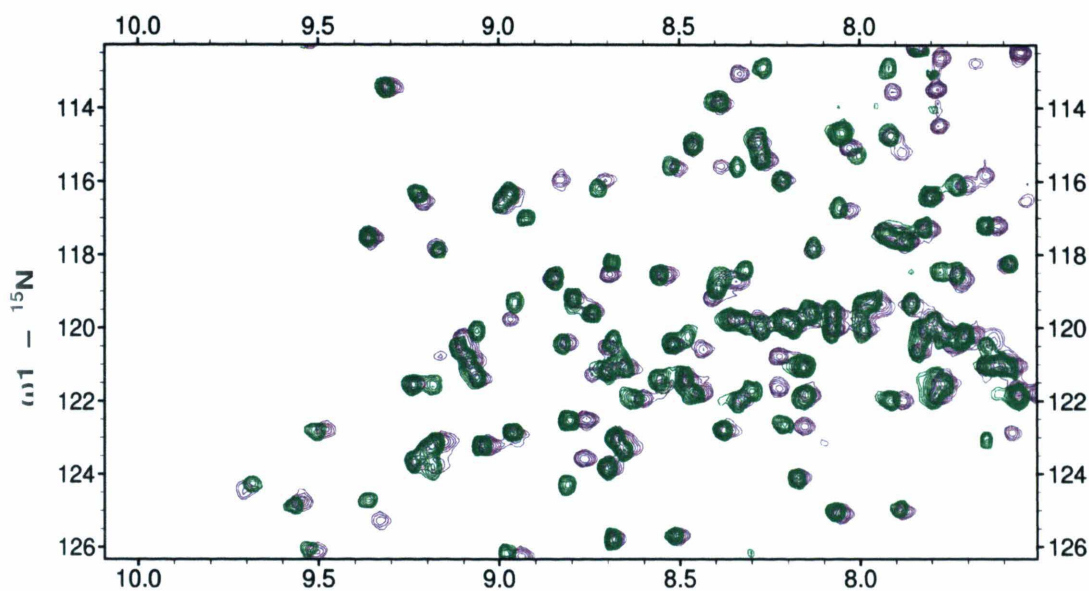


Figure 2-11: Overlay view of the ^1H - ^{15}N HSQC spectra of 150 μm Apo *EcDsbA* (green resonances) in the presence of a STD hit fragment at 1mM (purple resonances). Blue circles identify some of the ^1H - ^{15}N resonances that undergo significant shifts changes in the presence of this ligand, hence this fragment is a confirmed hit.

2.4.3 Chemical Shift Mapping

Overall weighted average chemical shift changes ($\Delta\delta$) in ^1H - ^{15}N HSQC spectra were calculated for assigned residues using equation 2-2 and for 3D spectra using equation 2-3.¹²³

Equation 2-2

$$\Delta\delta = \sqrt{(\delta H_{\text{complexed DsbA}} - \delta H_{\text{Apo DsbA}})^2 + 0.154(\delta N_{\text{complexed DsbA}} - \delta N_{\text{Apo DsbA}})^2}$$

Equation 2-3

$$\Delta\delta = \sqrt{(\delta H_{\text{complexed DsbA}} - \delta H_{\text{Apo DsbA}})^2 + 0.154(\delta N_{\text{complexed DsbA}} - \delta N_{\text{Apo DsbA}})^2 + 0.276(\delta CA_{\text{complexed DsbA}} - \delta CA_{\text{Apo DsbA}})^2}$$

Where δH , δN and δCA denote the chemical shift changes for the amide proton, nitrogen and C^α resonances respectively, between the apo and holo proteins upon addition of the ligands. The resonance assignment process is discussed below. Residues which underwent the largest changes in chemical shift were mapped onto the available crystal structure of the corresponding DsbA. Significant changes were defined as shift changes ranging larger than 0.03 ppm for a weighted shift. PyMOL molecular viewer (DeLano Scientific LLC, USA) was used for visualization of the CSP mapped onto the DsbA structures.¹⁶⁸

2.4.4 K_D and LE measurements

For the calculation of binding constant by 1H - ^{15}N HSQC titration, a range of concentrations of corresponding ligand (fragment/compound) were added sequentially to a protein sample to record spectra at those varying concentrations. Upon completion of a HSQC titration the protein sample was deemed to be saturated if at this point the amide resonances of the 1H - ^{15}N HSQC spectra no longer shifted on addition of more ligand. Stock concentrations of each test ligand were prepared by diluting compounds in a

solution with appropriate amounts of DMSO and protein buffer. Protein-ligand titrations were achieved by adding microliter amounts of each test ligand stocks to the ^{15}N labeled DsbA protein (*Ec*DsbA and *Vc*DsbA) sample. The sample was mixed and allowed to equilibrate prior to ^1H - ^{15}N HSQC data collection. DMSO only control experiments were performed prior to decide the maximum amount of DMSO to be tolerated in each sample without causing any changes in apo DsbA ^1H - ^{15}N HSQC spectra.

The binding affinity of ligands to the DsbA was determined by monitoring the chemical shifts of ^1H - ^{15}N HSQC (from titration series) cross peaks as a function of ligand-to-protein concentration ratio. The curves representing the concentration-dependent perturbation of several different crosspeaks were fitted using the following equation 2-4 which assumes 1:1 binding stoichiometry.

Equation 2-4

$$\Delta\delta = \Delta\delta_{\text{max}} / 2(1 + [\text{L}_0]/[\text{P}_0] + K_d/[\text{P}_0] - \sqrt{1 + [\text{L}_0]/[\text{P}_0] + K_d/[\text{P}_0]^2}) - 4 [\text{L}_0]/[\text{P}_0]$$

Where K_d is the dissociation constant, $\Delta\delta$ and $\Delta\delta_{\text{max}}$ are the observed and maximal chemical shift change, and $[\text{L}_0]$ and $[\text{P}_0]$ are the total ligand and protein concentrations, respectively. Fitting was performed using the non-linear least squares fitting option in Graphpad Prism version 5.0.¹⁵⁵ The change in free energy on binding (ΔG) (kJ mol^{-1}) and ligand efficiency per heavy atom (LE) $\text{kJmol}^{-1} \text{HAC}^{-1}$ were calculated using following equations:¹⁶⁹

Equation 2-5

$$\Delta G = -R \times T \times \ln (K_D)$$

Where R is the gas constant (8.314 J.mol⁻¹.K⁻¹) and T is the temperature (298 K)

Equation 2-6

$$LE = \Delta G / (\text{number of heavy atoms})$$

2.4.5 3D NMR

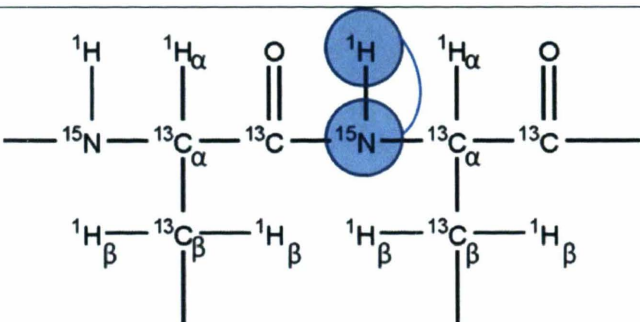
Standard triple resonance experiments¹⁵⁷ CBCACONH¹⁵⁶, HNCA¹⁵⁸, HNCACB¹⁵⁹, HNCO¹⁵⁶ and HNCACO¹⁶⁰ were performed and recorded at 298K on a Varian Unity 600 MHz spectrometer. ¹H-¹⁵N HSQC spectra as well as ¹H-¹³C HSQC were also recorded using double labeled protein samples. HCCONH-TOCSY¹⁶¹ and HCCH-TOCSY experiments were also conducted to identify side-chain carbon and proton shifts¹⁷⁰ for EcDsbA. (Table 2-10) Sequential assignments for apo VcDsbA were previously reported by Dr James Horne (MIPS), although these were recorded under slightly different buffer conditions to those employed for fragment screening.¹⁷¹ The published assignments were used to guide assignments in the fragment binding buffer conditions to determine the protein ligand interactions by observing changes in 3D spectra upon ligand binding. These experiments were performed in Shigemi (NMR) tube with ¹³C¹⁵N- (double) labeled and highly purified 500μM protein with 300μL sample volume, containing 10% D₂O and 90% buffer. Buffer 25mM MES, 0.5mM RE and 150mM NaCl (pH 5.0) and 10 mM

HEPES with 150mM NaCl (pH 7.0) were used for *EcDsbA* and *VcDsbA* samples respectively to retain best stability conditions for proteins. These assignments were later transferred onto all of the multidimensional spectra recorded for protein sample in the relevant buffers used for binding studies. Standard pulse sequences (Varian VnmrJ) were used for data acquisition. NMR data were processed using NMRPipe^{165b} and processed spectra were analyzed with Sparky 3.113¹⁶⁶

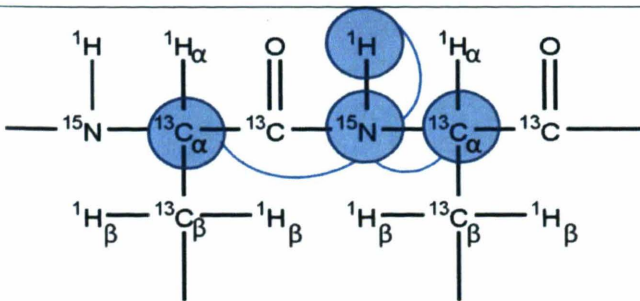
2.4.5.1 Sequential assignments

Sequential assignments for *EcDsbA* and *VcDsbA* were obtained using standard triple resonance approaches.¹⁵⁷ In general these approaches rely on acquiring a pair of experiments that utilise scalar correlations between nuclei to link atoms in adjacent residues. For example, one complementary pair of experiments are the HNCA and the HN(CO)CA. In the HNCA experiment, correlations are observed between the amide proton and amide nitrogen frequencies of the same amino acid residue (denoted i), (H_i and N_i) and the $C\alpha$ carbon frequencies both the same residue ($C\alpha_i$) and the preceding residue ($C\alpha_{i-1}$). The complementary experiment, the HN(CO)CA shows correlations between H_i , N_i and $C\alpha_{i-1}$. In this way, a pair of experiments allows identification of the frequencies of nuclei in adjacent residues and enables the determination of which is the intra-residue and which is the inter-residue signal. Various pairs of such experiments exist (Table 2.9), which enable correlations to be established between different nuclei in adjacent pairs of residues.

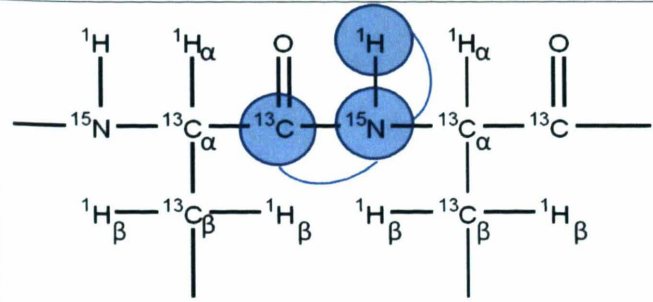
Table 2-9: The magnetization path with names of multidimensional and triple resonance experiments used for sequential assignments. Backbone ^{13}C is represented by C_α and R groups by C_β . The spectra contain correlation between blue circles nuclei. Nuclei involved in the pathway, but not detected in the spectra are circled white. Main residue number represented by i and previous residue as $i-1$.



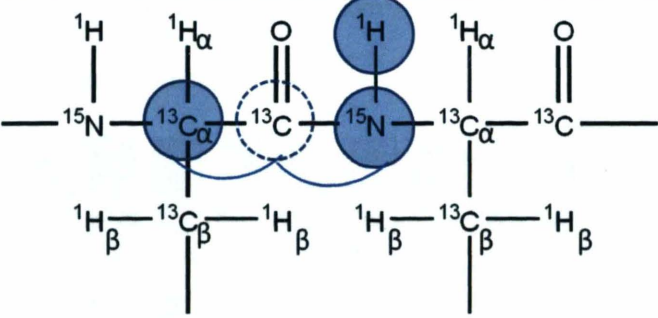
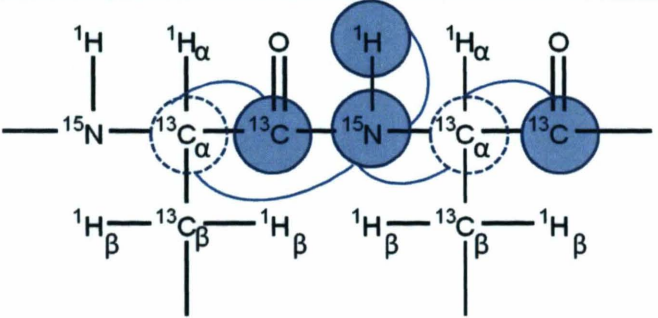
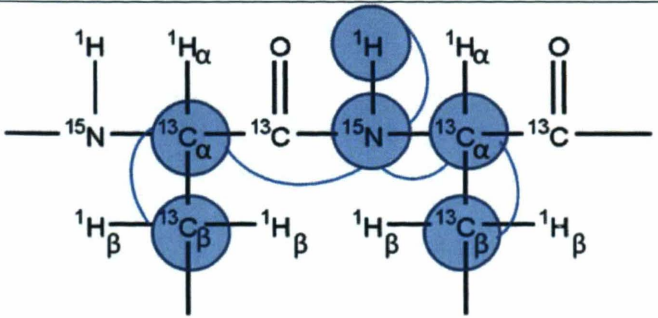
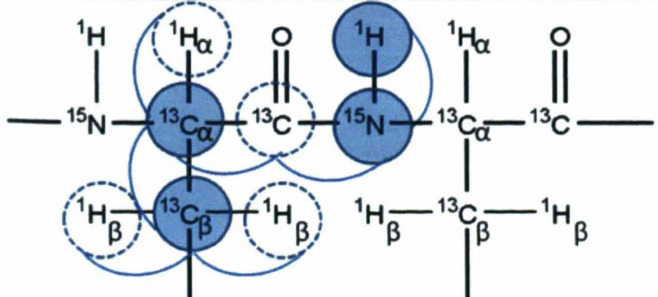
^1H - ^{15}N HSQC
Correlates the chemical shifts of ^1H and ^{15}N



HNCA
Correlates the chemical shifts of ^1H , ^{15}N , $^{13}\text{C}_{\alpha i}$ and $^{13}\text{C}_{\alpha i-1}$



HNCO
Correlates the chemical shifts of ^1H , ^{15}N and $^{13}\text{C}_i$

| | |
|---|--|
|  | <p>HN(CO)CA</p> <p>Correlates the chemical shifts of ^1H, ^{15}N and $^{13}\text{C}_{\alpha i-1}$</p> |
|  | <p>HN(CA)CO</p> <p>Correlates the chemical shifts of ^1H, ^{15}N, $^{13}\text{C}_i$ and $^{13}\text{C}_{i-1}$</p> |
|  | <p>HNCACB</p> <p>Correlates the chemical shifts of ^1H, ^{15}N, $^{13}\text{C}_{\alpha i}$, $^{13}\text{C}_{\alpha i-1}$, $^{13}\text{C}_{\beta i}$ and $^{13}\text{C}_{\beta i-1}$</p> |
|  | <p>CACB(CO)NH</p> <p>Correlates the chemical shifts of ^1H, ^{15}N, $^{13}\text{C}_{\alpha i}$ and $^{13}\text{C}_{\beta i-1}$</p> |

These can be connected as shown in Figure 2-13 to generate sequential assignments of the backbone atoms.

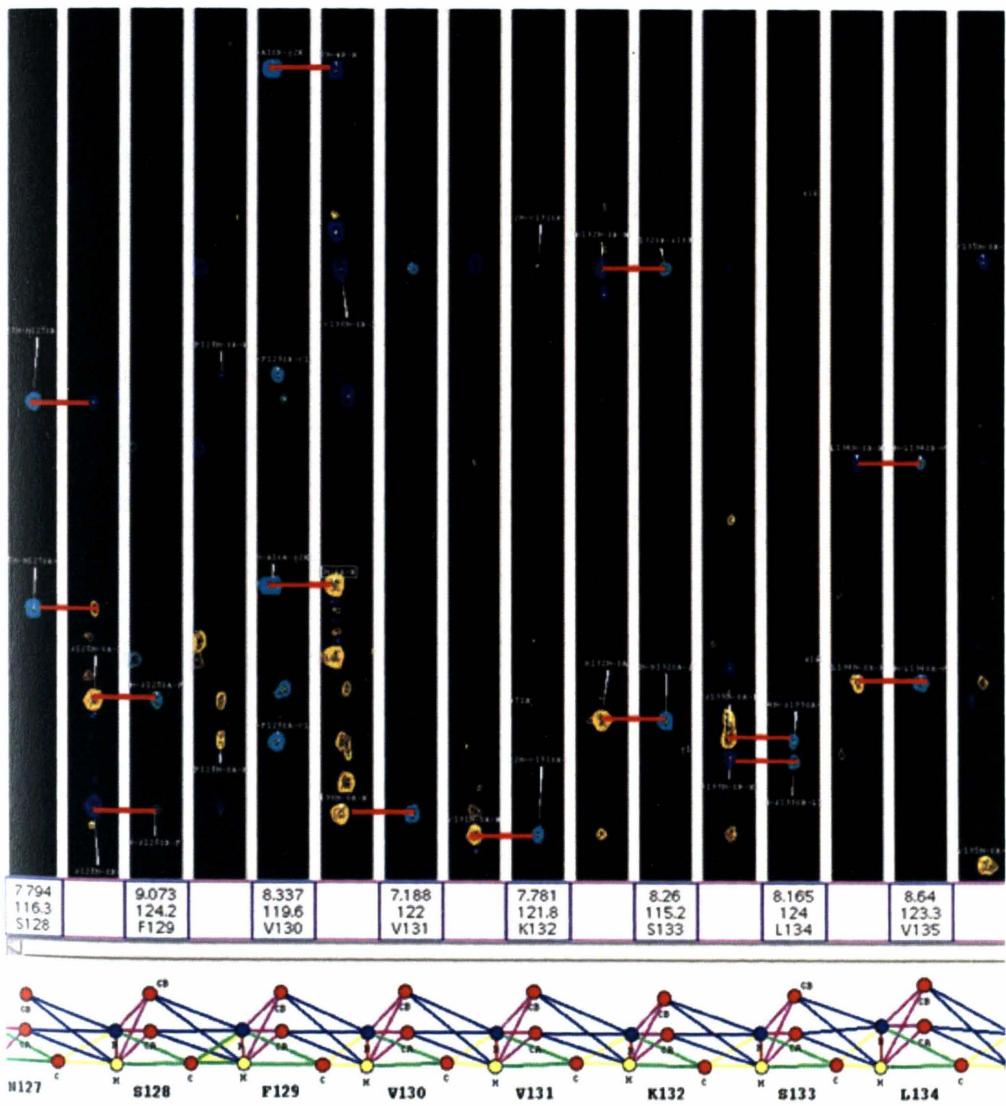


Figure 2-13: Sequential through-bond connections of oxidised *EcDsbA* for residues Ser128-Val135 in CBCA(CO)NH and HNCACB spectra. $C\alpha$ and $C\beta$ peaks are show in CBCA(CO)NH in cyan. Preceding and intra residue peaks of $C\alpha$ and $C\beta$ are shown in yellow and blue respectively. White lines denote sequential connectivity between the individual residues.

The following experiments as listed in table 2-10 were acquired for *EcDsbA*.

Table 2-10: NMR spectra obtained for *EcDsbA*. Data were acquired for resonance assignment and as a foundation for fragment-binding studies. Spectra were acquired at 600MHz, 318K.

| Experiment | Sweep width | | | Number of points | | |
|---|----------------|-----------------|-----------------|------------------|-----------------|-----------------|
| | ¹ H | ¹⁵ N | ¹³ C | ¹ H | ¹⁵ N | ¹³ C |
| [¹⁵ N, ¹ H]-HSQC | 14 | 32 | | 1024 | 64 | |
| [¹⁵ N, ¹ H]-HSQC | 16 | 17 | | 1024 | 64 | |
| [¹³ C, ¹ H]-HSQC | 14 | | 64 | 1024 | | 128 |
| [¹³ C, ¹ H]-HSQC | 14 | | 140 | 1024 | | 128 |
| CBCA(CO)NH_HEPES | 14 | 20 | 64 | 1024 | 24 | 44 |
| CBCA(CO)NH_MES | 16 | 17 | 64 | 1024 | 24 | 48 |
| HNCACB_HEPES | 14 | 20 | 64 | 1024 | 24 | 64 |
| HNCACB_MES | 16 | 17 | 64 | 1024 | 24 | 48 |
| HNCO | 16 | 17 | 16 | 1024 | 32 | 32 |
| HN(CA)CO | 16 | 17 | 36 | 1024 | 24 | 32 |
| HNCA_HEPES | 16 | 17 | 30 | 1024 | 24 | 32 |
| HCCH-TOCSY | 14 | 20 | 32 | 1024 | 32 | 64 |
| CCONH TOCSY | 14 | 20 | 64 | 1024 | 32 | 48 |

The HCCH-TOCSY¹⁷⁰ allows assignment of the ^1H and ^{13}C resonances of the side chains by utilizing the isotropic mixing of ^{13}C magnetization to obtain cross peaks at the frequency of each carbon in an amino acid side chain.¹⁷² In H(CCO)NH –TOCSY magnetisation is transferred from the side-chain ^1H nuclei to their attached ^{13}C nuclei and isotropic ^{13}C mixing. From here, magnetisation is transferred to the carbonyl carbon, on to the amide nitrogen and finally the amide hydrogen for detection. The chemical shift is evolved simultaneously on all side-chain hydrogen nuclei, as well as on the amide nitrogen and hydrogen nuclei, resulting in a three-dimensional spectrum with one nitrogen and two hydrogen dimensions.¹⁷²

In general, these experiments allowed the assignments of backbone (H, N, C', C $_{\alpha}$) and C $_{\beta}$ resonances and these were transferred to the corresponding spectra of protein in buffer used for fragment binding studies. Almost completely assigned ^1H - ^{15}N HSQC spectra of *EcDsbA* have been shown in Figure 2-14. The sequential backbone resonances for *EcDsbA* and *VcDsbA* as in complex with a choice of ligands have been listed in Appendix sections 8.7 and 8.8 respectively.

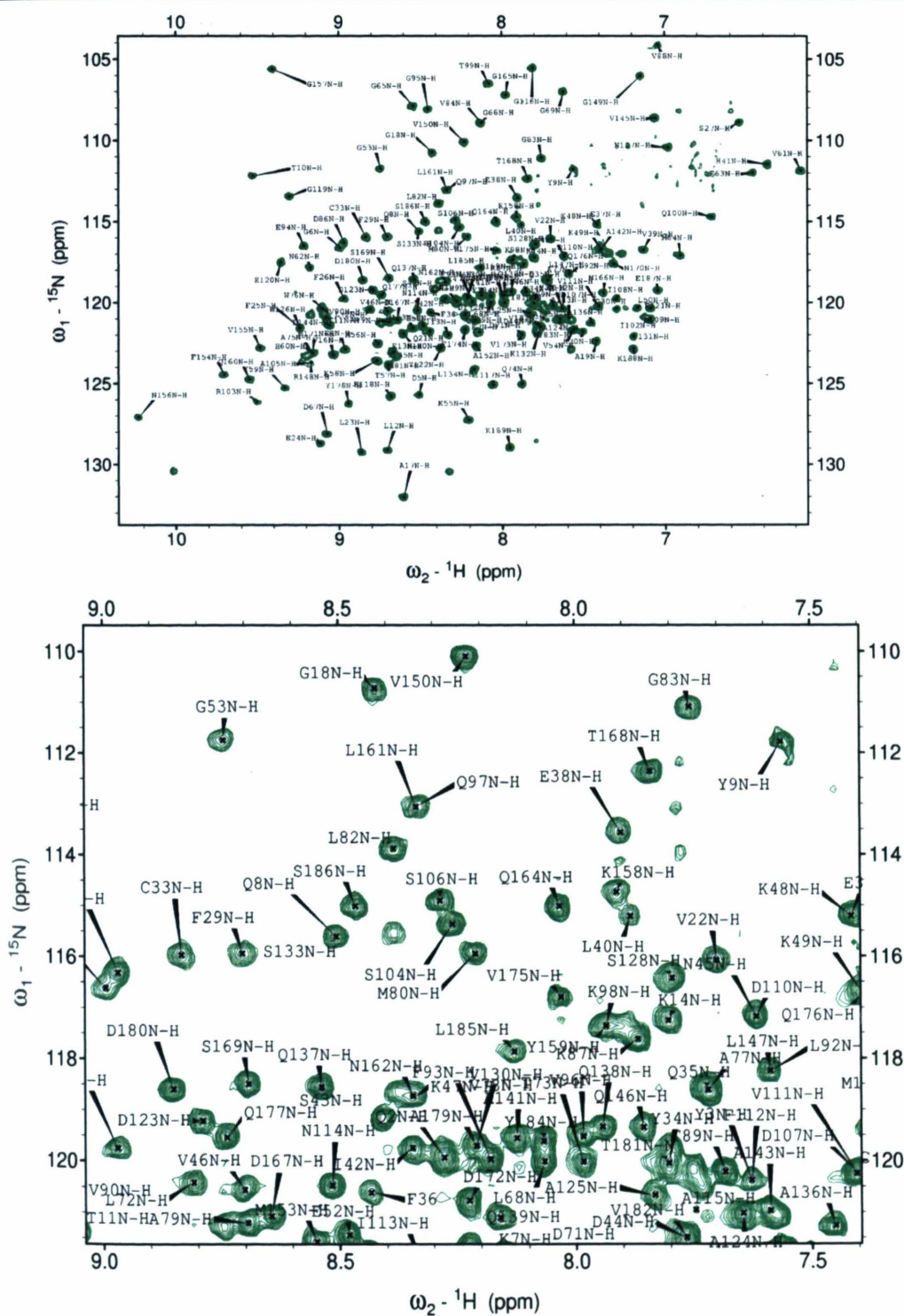


Figure 2-14: (A) ^{15}N -HSQC spectra of oxidised *EcDsbA* with resonance assignments of almost complete backbone amides. (B) Inset of same spectrum illustrating labelling of N-H resonances.

2.5 X-Ray Crystallography

2.5.1 Crystallization

Protein crystallization experiments were performed at the ROCX crystallography facility under the supervision of our collaborators Dr Begoña Heras, Dr Stephen Shouldice and Prof Jenny Martin (Institute for Molecular Bioscience, Queensland Bioscience Precinct, University of Queensland, St Lucia, QLD 4072, Australia).

Crystals of *EcDsbA* were obtained using the hanging-drop vapour-diffusion method.¹⁷³ Crystallization trials were set up in VDXm 24-well hanging-drop plates with 18 mm siliconized cover slips (Hampton Research, San Diego CA, USA). Each cover slip held one 2 μ L drop containing 1 μ L protein solution and 1 μ L well solution while the well solution volume was kept 500 μ L. The crystallization plates were incubated at 293 K and imaged using a temperature controlled RockImager (Formulatrix, Waltham, MA USA). Single and large rectangular crystals along with a few twinned crystals of *EcDsbA* appeared within two days in the conditions contained PEG of average molecular weight 8000 Da (PEG8000), Cacodylate 100 mM pH 6.1, CuCl_2 1 mM, Glycerol 20%. Focused screens were performed to find optimized concentrations for glycerol and PEG8000 within the range of 4% - 20% (v/v) glycerol and 12% - 17% (w/v) PEG8000. (Table 2-10) The best diffracting crystals were obtained in Cacodylate 100 mM pH 6.1, CuCl_2 1 mM, PEG8000 14%-15% (w/v), glycerol 4%-8% (v/v) and *EcDsbA* concentration of 16 mg/mL (Figure 2-15).

Table 2-10: Focused screening crystallographic conditions to optimize PEG8000 and Glycerol percentage for *EcDsbA* crystals.

| | 1 | 2 | 3 | 4 | 5 | 6 |
|--|------|-----|-----|-----|-----|-----|
| A PEG 8000 Glycerol | 12 % | 13% | 14% | 15% | 16% | 17% |
| | 4% | 4% | 4% | 4% | 4% | 4% |
| B PEG 8000 Glycerol | 12 % | 13% | 14% | 15% | 16% | 17% |
| | 6% | 6% | 6% | 6% | 6% | 6% |
| C PEG 8000 Glycerol | 12 % | 13% | 14% | 15% | 16% | 17% |
| | 8% | 8% | 8% | 8% | 8% | 8% |
| D PEG 8000 Glycerol | 12 % | 13% | 14% | 15% | 16% | 17% |
| | 10% | 10% | 10% | 10% | 10% | 10% |
| Each wall contains Cacodylate pH 6.1 (100 µl), CuCl ₂ 1mM, H ₂ O upto 550 µl | | | | | | |

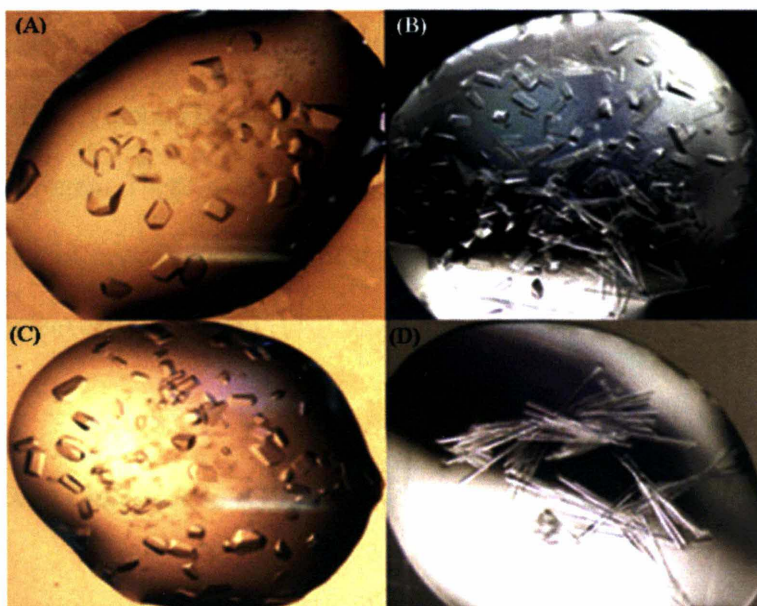


Figure 2-15: Crystals of *EcDsbA* at varying concentrations of PEG 8000 and Glycerol grown as (A) single and large (B) rectangular or needle shaped (C) small and prismatic (D) or twinned shaped crystals. A typical diffraction quality crystal shaped as prismatic or rectangular with average single crystals of dimensions approx. $0.3 \times 0.15 \times 0.1 \text{ mm}^3$ at PEG 8000 14%-15% (w/v) and glycerol 4%-8% (v/v) shown in

In order to obtain *EcDsbA* crystals with a desired compound, apo protein crystals were soaked into the crystallographic stabilizing solution (24% (w/v) PEG8000, Cacodylate 100 mM pH 6.1, CuCl_2 1 mM and glycerol 20% (v/v)) mixed with concentrated stocks of individual fragments in DMSO. The final concentration of fragments in *EcDsbA* stabilizing solution ranged between 3 mM -10 mM depending upon the solubility of fragment.¹⁷³ A typical soaking time was 2-4 hrs for most of the fragments, while longer

or shorter soaks were performed in a few cases to improve the quality of data. *EcDsbA* crystals were harvested from hanging drops using a 0.2-0.3 mm CrystalCap Copper Magnetic HT Cryoloop (Hampton Research, San Diego CA, USA) prior to flash cooling in liquid nitrogen. These crystals were maintained at 100 K throughout data collection and the trials of crystal cooling indicated that both glycerol and PEG were necessary to attain sufficient cryoprotection.¹⁷³

2.5.2 Data Collection and Processing

Crystals were visually inspected for quality, and their dimensions were measured using an in-lens graticule on a stereomicroscope (NIKON, SMZ-U, USA). Images were acquired on a Leica Z16 APO (KLI500 LCD) stereomicroscope, using the IM500 program (Leica Microsystems, Wetzlar, Germany). Diffraction data from the crystal produced using the soaking experiments were collected both on the in house radiation source at UQ-ROCX and the Australian Synchrotron.

2.5.2.1 In-house data collection and processing

Preliminary diffraction data from the initial soaking experiments were collected at the X-ray diffraction and radiation source UQ-ROCX. This is an Rigaku FR-E copper rotating anode generator operating at 45 kV, 45 mA with Osmic Confocal Vari-Max HF optics at a wavelength of 1.5418 Å (Rigaku Americas, Houston USA). The crystal-to-detector distance was 150 mm and images were collected over 90° (between 0° and 180°), with

0.5° image widths (oscillation angle) per frame, to make a total of 360 data frames with the exposure time for each frame as 1 min. Reflections were measured with a Rigaku Saturn 944 CCD area detector (Rigaku Americas, Houston USA). A Cryo Industries CryoCool LN₂ (Cryo Industries, Manchester, NH USA) was used for keeping crystals at 100 K during data collection.¹⁷⁴ Diffraction data were processed using and CrystalClear 1.4 (Rigaku Corporation). The diffraction pattern of *EcDsbA* crystals is shown in Figure 2-16.

The image frames collected were indexed and integrated with Crystal Clear 1.4.1 and scaled and merged with SCALEPACK (Otwinowski & Minor, 1997). Molecular replacement phasing using PHENIX the MolREPL module of CCP4 and the structure of native *EcDsbA* (Martin et al 1997) as starting model, was carried out to obtain an initial model with phases.¹⁷⁵ Data in the resolution range 15 Å – 4 Å were used in the evolutionary search procedure. Model refinement was carried out using input scripts (modified as required) of the CCP4 and PHENIX.

2.5.2.2 Synchrotron data collection and processing

Fragment soaked *EcDsbA* crystals, produced using the optimal soaking conditions were transported in liquid nitrogen filled trays with cryo loops to the Australian Synchrotron Radiation Facility at Clayton, VIC, Australia. Each crystal was placed in a special labelled vial containing liquid nitrogen. Ten of these vials were slotted into a cartridge (or “puck”) for storing the crystals within the microdiffractometer.¹⁷⁶ Subsequently X-ray

diffraction data were collected at the Australian Synchrotron on PX1 beamline 3BM1 (wavelength 0.953645 Å) with maximum intensity of $\sim 7.5 \times 10^{11}$ photons. The image frames collected were indexed, integrated, scaled and merged with Crystal Clear (Rigaku) or HKL2000.¹⁷⁷

2.5.2.3 Structure determination and Refinement

The structures of *Ec*DsbA-fragment complexes were solved by molecular replacement phasing coupled with difference Fourier methods¹⁷⁵ using the structure of oxidized *Ec*DsbA (PDB accession code 1FVK) with waters removed as the preliminary model. The crystal structure contains two molecules of *Ec*DsbA in the asymmetric unit nominated DsbA_A and DsbA_B (corresponding to molecule A and molecule B in the asymmetric unit of *Ec*DsbA crystals). Initially data were refined via a rigid body refinement in Refmac and subsequent refinement cycles were performed PHENIX refinement.^{175, 178}

Modeling and viewing of the *Ec*DsbA crystal structures was performed with WinCoot version 0.6.¹⁷⁹ Co-ordinates for the ligands were built using the Molsketch portion of PHENIX and restraints files were generated using PHENIX (conformer).^{178, 180} The process used implicit modeling and refining the structure of the protein first, followed by the addition of water molecules, buffer or glycerol molecules and finally the fragment or ligand into ($F_o - F_c$) difference maps. The criteria used for including a water molecule were the presence of at least one possible hydrogen bond within 3.2 Å and $2F_o - F_c$ density at 1 σ and $F_o - F_c$ density at 3 σ . The model was improved with cycles of iterative

model refitting and building with Wincoot^{179b} density modification and refinement using the phenix.refine wizard.^{175, 178} *R*-free analysis using 5–10% of reflections was used for cross-validation. Some surface residues with weak electron density were modeled with reduced occupancies for the side chain atoms and some residues were modeled with alternate conformations.

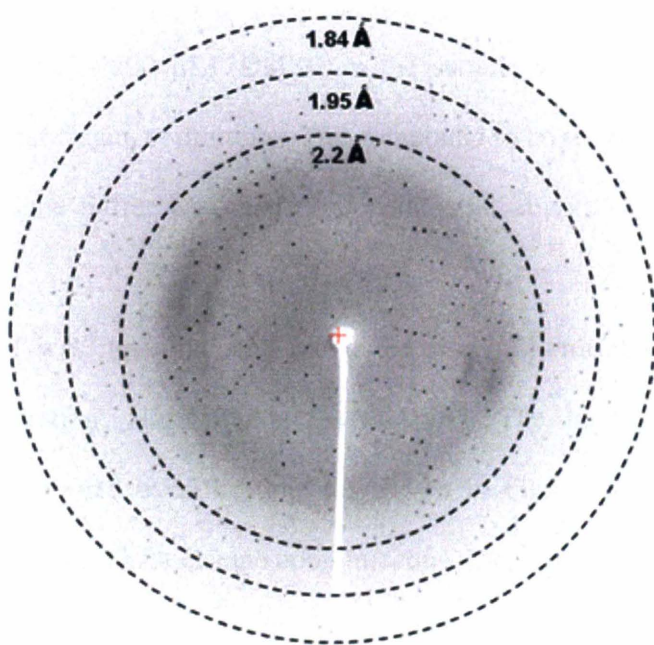


Figure 2-16: Diffraction pattern image of a typical *EcDsbA* crystal (space group C2) used for fragment binding studies.

2.6 ITC

Calorimetric experiments were conducted to measure fragment binding affinities to *EcDsbA*, using an isothermal titration microcalorimeter instrument iTC200 (MicroCal LLC, Northampton, MA, USA). Experiments were conducted at 25 °C and all the solutions and samples were degassed under vacuum prior to titrations. The reference cell was filled with Milli-Q water. Purified *EcDsbA* protein at a concentration of 200 µM in the buffer 10mM HEPES pH 7.5 with 0.6% (v/v) DMSO were loaded into the sample cell (reaction cell volume ~ 200 µL). DMSO in the sample was calibrated to match the residual DMSO after fragment dilutions. The compound to be tested at a concentration of 1–3 mM in the same buffer was then titrated into the sample cell with 20 × 2.5 µL injections.

The heat released was measured and integrated using MicroCal Origin 7.0 software (OriginLab Corporation, Northampton, MA, USA). The heat released upon their interaction was monitored over the time (as shown in Figure 2-17), where each peak represented the associated heat change upon injecting a few microliters volume of sample into the ITC reaction cell. Three independent trials ± the standard error were used to derive the average values of reported equilibrium thermodynamic parameters. The association constant (K_a , $1/K_d$), enthalpy (ΔH) and stoichiometry (N) were calculated using a single-site binding model. The energies of interaction were calculated from K_a values using the Gibbs free energy equation 2-6

(Equation 2-6)

$$\Delta G = -RT (\ln K_a)$$

where R is the gas constant 1.9872 kcal/mol, T is the temperature of the experiment in degrees, where ΔG is the free energy of binding (kcal/mol) Kelvin (298 K), and K_a is the association constant for the compound.

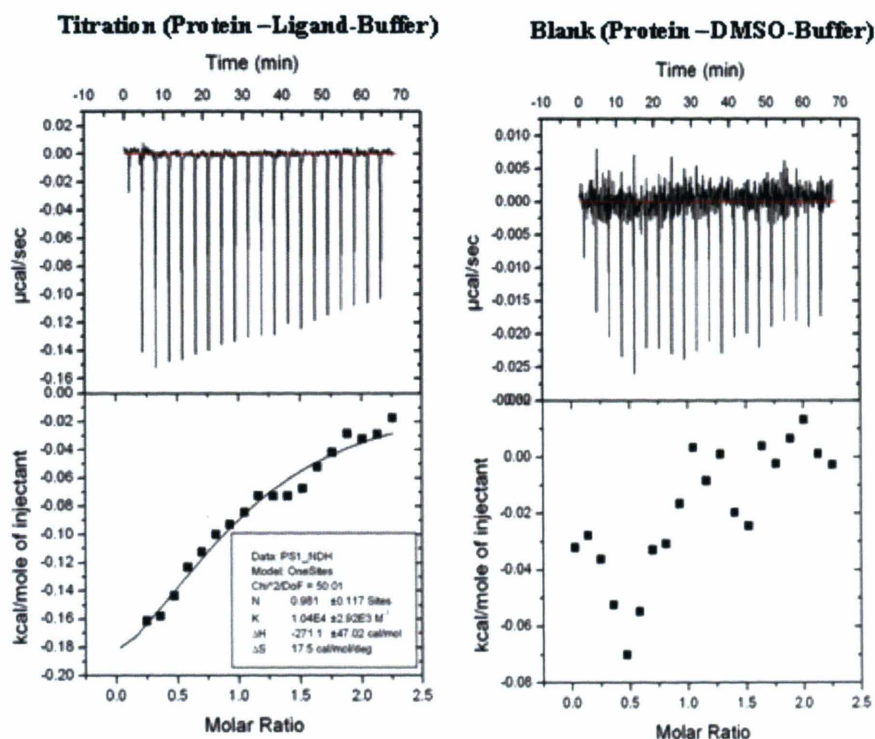


Figure 2-17: ITC based affinity measurement of one weak binding and enthalpically driven fragment hit of *EcDsbA*, identified from NMR based screening.

2.7 Computational Methods

2.7.1 Library and Protein Preparation

A commercially available fragment library from Maybridge was selected for the virtual screening purpose consisting of 500 heterocyclic compounds. A smiles to 3D coordinates conversion was performed using CORINA.¹⁸¹ 3D structures of fragments were optimized using the OPLASA force field to obtain lowest energy conformations and appropriate protonation states were selected for computational screening.¹⁸¹ Fragments were optimized for docking using the LigPrep tool with their protonation states fixed at pH 7.0 ± 2.0 units and no tautomeric states were generated (Maestro software).¹⁸² The coordinates for the VcDsbA (PDB code 1BED)²⁶ and EcDsbA ((PDB code 1FVK)^{20a} structures were downloaded from the Protein Data Bank and were processed with the Protein Preparation Wizard in Maestro for subsequent grid generation and docking.^{182a} Using this tool, hydrogen atoms were added to the protein and the protonation states for histidine residues were optimized. Crystallographic waters and ions not deemed to be important for ligand binding were deleted and the entire protein was minimized. Default options within the Protein Preparation Wizard were used to assign bond order and minimization using the OPLS_2005 force field with 50 steps of steepest-descent minimization to correct bond lengths and angles.^{182a} Additionally some residues containing structural errors were fixed manually.

After protein preparation, the SiteMap tool was used to explore possible binding sites on the entire protein structure.¹⁸³ This feature identifies top-ranked potential receptor binding

sites those are most likely to contribute to tight protein–ligand or protein–protein binding.¹⁸³ Default parameters were used to report a maximum of five of the best ranked sites. Subsequently site analysis and scoring was undertaken for the NMR identified binding region as described under the site mapping method section.

2.7.2 *Grid Generation and Docking*

Grid based molecular docking studies were performed using Maestro, version 8.0, and XP Glide, version 4.5¹⁸⁴ For the preliminary docking of the fragment library into DsbA proteins four grids were prepared using the Receptor Grid panel. Amino acid residues located at four top ranked largest sites were used to define a centroid of selected residues when generating the enclosed box for the docking of ligands. Default parameters were used for van der Waals radii scaling factor (0.8), per-atom van der Waals radius and charge scaling values (cutoff of 0.25).¹⁸⁴ Distance-dependent dielectric constant was defined as 2.0 and maximum number of steps was set to 500 during the conjugate gradient minimization cycles for the energy minimization stages for selected poses. Extra-precision (XP) mode was used for flexible docking with OPLS_2001 force field top ranked docked pose was saved for each input ligand.^{184b, 185} Enrichment factors¹⁸¹ were calculated at 2, 5, and 10% of the total database screened, using the following equation:¹⁸⁶

$$EF \% = (\text{Hits}_{\text{sampled}} / N_{\text{sampled}}) / (\text{Hits}_{\text{total}} / N_{\text{total}})$$

2.7.2.1 NMR based docking

NMR data was obtained from the distinct pairs of protein-fragments complexes and used for the constraint docking of the fragments hits and the fragment library. The binding site residues those were highly perturbed (CSP > 0.03 ppm) upon fragment binding were used to generate parameters for NMR-based constraint docking. Constraints were added in these perturbed regions of the protein to identify most likely binding location for fragment hits.¹⁸⁷ These constraints were added by locating regions or 'cells' those were derived based on the properties of selected residues.^{184b, 185} Thus a set of hydrophobic cells were located in the proximity of most perturbed residues. A rectangular docking grid (or an enclosing box) was generated using the centre of the perturbed residues with the inclusion of selected hydrophobic constraints. The grid gives a measure of the effective size of the search space in defined binding region. The enclosing box length was set at value of 10 Å in X, Y and Z directions with 8 Å ligand docking boundaries to observe more specific binding positions of fragments according to NMR results. Other parameters were kept similar as described for preliminary docking. Final scoring and ranking was performed with Glide XP score. A number of generated docked solutions were optimized to obtain the best scoring pose by the Glide selection program.^{185a, 188} Information obtained from finally selected complexes was compared with ¹H-¹⁵N HSQC data where distances of the docked ligand were measured to surrounding residues within 5 Å. In other words, CSP (distances to -NH of highly perturbed residues) of each ligand pose were estimated and were found to be in good agreement with their experimental values. A simple linear relationship was considered between the magnitude of the CSP

and the distance from the amino acids NH atoms and the nearest ligand atom according to the protocol described by Stark *et al.*¹⁸⁷

2.7.3 Surface Grid Map Generation

In the next step, the NMR-based site was mapped and analyzed using the Hydrophobic/philic panel and Sitemap tools in Maestro. The binding site mapping procedure operates in a method equivalent to Goodford's GRID algorithm.¹⁸⁹ The first step for grid map computation was to generate a rectilinear box that contained NMR-based identified set of amino acid residues and was defined as the binding region. This box was composed of grid points with a typical grid spacing of 1 Å within the box. Components of the electrostatic field in the x, y, z dimensions and van der Waals energies were calculated at each of the grid points using the OPLS-AA force field for partial charges and vdW parameters. A probe can be defined by a vdW sphere of radius 1.5 Å and well depth 0.2 kcal/mol with a point dipole moment of 2.3 Debye. A putative van der Waals energy and the magnitude and direction of the electric field were computed for this probe centered at each grid point by considering interactions with all defined atoms of the receptor site within a cutoff distance. Site maps depend on the character of the nearest receptor atom and explicitly show the shape and extent of hydrophilic and hydrophobic regions.¹⁹⁰ These regions are defined in a way that considers both spatial proximity to the receptor and suitability for occupancy by solvent water. Finally, the hydrophilic and hydrophobic grids were read by Maestro, which contoured the grids at the empirically selected default isosurface values (-6 and -0.5

kcal/mol).¹⁹⁰ These values have been adjusted in such a way that the hydrophilic and hydrophobic maps should not overlay but rather should be separated by some "neither" space.¹⁹¹ Constraints based XP docked conformations of all of the 500 compounds in the library were manually screened using the hydrophobic and hydrophilic surface maps (grids) for the binding site. In this screening process, fragments were separated on the basis of presence of all favourable and at least one unfavourable interactions with the surface grids.

2.7.4 Similarity Searching

Similarity searches were performed using the computer software Instant Jchem (Chemaxon).¹⁹² These searches were performed on several sets of the compounds listed here:

- i. Maybridge fragment library 1 (1132 compounds)
- ii. Chembridge Screening collection (4100 compounds)
- iii. *In house* collection of compounds (> 3000) at MIPS, Monash University

Additionally, CAS (Chemical Abstracts Service) was explored using the Scifinder software 2007 [Columbus, Ohio: American Chemical Society] to obtain scientific information for numerous compounds. Babel software was used for calculating the Tanimoto similarity score for each tested analogue against parent compounds.¹⁹³

2.7.5 *ParaSurf Surface Properties*

The ParaSurf software was used to study the surface properties of fragments, with the aim of studying the molecular similarities of the binding fragments by their surface characteristics.¹⁹⁴ ParaSurf™ (Cepos InSilico Ltd., Erlangen, Germany) is a program that generates isodensity or solvent-excluded surfaces from the results of semiempirical molecular orbital calculations. The surface may be generated by “shrink wrap” or “marching-cube algorithms” and the former may be fitted to a spherical harmonic series.¹⁹⁵ Parasurf calculates four local properties, at the points on the surface (1) the molecular electrostatic potential (2) the local ionization energy (3) the local electron affinity (4) the local polarizability, together with a standard set of 40 molecular descriptors. Finally ParaFit was used to superimpose sets of compounds using each of the properties generated by ParaSurf in turn.¹⁹⁵ The calculation of these descriptors was followed by principle component analysis and cluster analysis for the selection of most important surface descriptors and classifying the fragment library respectively.

Chapter 3

Inhibition of Oxidative Protein Folding in *Escherichia coli*

PART B: Suggested Declaration for Thesis Chapter

Monash University

Declaration for Thesis Chapter Three

Declaration by candidate

In the case of Chapter three, the nature and extent of my contribution to the work was the following:

| Nature of contribution | Extent of contribution (%) |
|--|----------------------------|
| Planning and undertaking the experimental work. Preparing the manuscript | 70% |

The following co-authors contributed to the work. Co-authors who are students at Monash University must also indicate the extent of their contribution in percentage terms:

| Name | Nature of contribution | Extent of contribution (%) for student co-authors only |
|-------------------|--|--|
| Begoña Heras | Assistance in acquisition and analysis of X-ray crystallographic data. Assistance with refinement of structures. | |
| Stephen Shouldice | Assistance in acquisition and analysis of X-ray crystallographic data. | |
| James Horne | Assistance in acquisition and analysis of NMR data | |
| David Manallack | Supervision of the project. Preparation of the manuscript. | |
| Jennifer L Martin | Experimental design. Supervision of the project. Preparation of the manuscript. | |
| Jamie S Simpson | Experimental design. Supervision of the project. Preparation of the manuscript. | |
| Martin J Scanlon | Experimental design. Supervision of the project. Preparation of the manuscript. | |

Candidate's
Signature

Declaration by co-authors

The undersigned hereby certify that:

- (1) The above declaration correctly reflects the nature and extent of the candidate's contribution to this work, and the nature of the contribution of each of the co-authors.

Inhibition of Oxidative Protein Folding in *Escherichia coli*

Pooja Sharma, Begoña Heras, James Horne, Stephen Shouldice, David Manallack, Jennifer L. Martin, Jamie S. Simpson and Martin J. Scanlon*

There is now overwhelming evidence that the disulfide-dithiol oxidoreductase enzymes that catalyse the formation of disulfide bonds in the periplasm of many Gram negative bacteria^[1] are key determinants of virulence.^[2] In the prototypical disulfide bond (dsb) system of *E. coli* K12 there are two distinct pathways. The first is the oxidative pathway, which consists of a soluble periplasmic protein DsbA and a membrane bound protein DsbB. DsbA introduces disulfide bonds into nascent polypeptide chains containing pairs of cysteine residues as they are translocated to the periplasm.^[3] In this process a disulfide bond in the active site of DsbA is reduced. DsbA is reoxidised by DsbB, which derives its oxidising power from bound quinones that couple disulfide bond formation to the electron transport chain.^[4] Secreted proteins of *E. coli* show a strong bias for even numbers of cysteines and most contain a single pair of cysteines.^[5] However, in proteins containing more than two cysteine residues there exists the possibility for incorrect disulfide formation. This is compounded by the fact that DsbA appears in most cases to introduce disulfides between consecutive cysteine residues, which in proteins with multiple non-consecutive disulfide bonds can lead to misfolding.^[6] In such cases the isomerase pathway is required. This second pathway consists of two soluble proteins, DsbC and DsbG, which are maintained in the periplasm in their active reduced state by a membrane bound partner DsbD. Together these proteins correct non-native disulfides.^[6]

DsbA has broad substrate specificity and has been

shown to catalyse the formation of disulfides in numerous secreted proteins.^[7] In *E. coli* for example there are more than 250 proteins that are predicted to enter the periplasm which contain at least two cysteines and are therefore potential substrates for DsbA.^[5] For this reason DsbA operates at a pivotal point in virulence-factor production because most virulence factors are proteins that require disulfide bonds to fold and function. Mutants of DsbA in *E. coli* have been shown to be impaired in a range of processes related to virulence including toxin secretion,^[8] motility,^[9] pilus biogenesis^[2a] and the formation of functional type III secretion systems.^[2c] Furthermore, without a functional dsb system, it has been shown that *E. coli* become hypersensitive to reducing agents (such as dithiothreitol (DTT)), divalent metal ions and benzylpenicillin.^[10] Significantly, DsbA mutations in many pathogenic bacteria attenuate their virulence, demonstrating the value of targeting DsbA to develop anti-virulence agents.^[1]

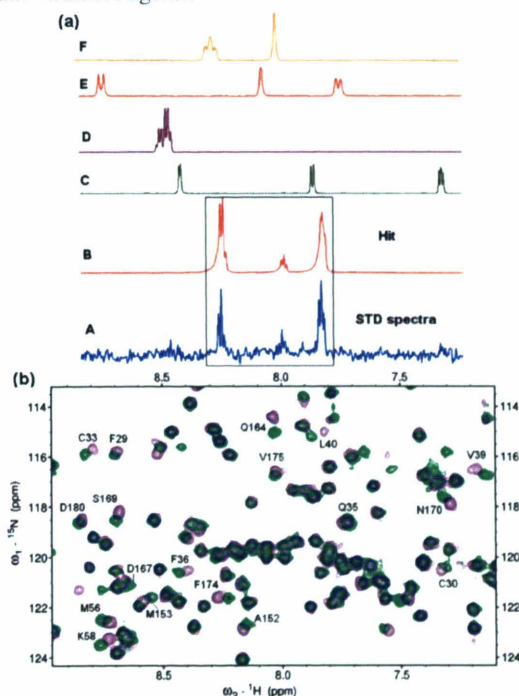


Figure 1. a) Identification of phenoxybenzene **2** as an EcDsbA binder from STD NMR based screening by comparing the STD spectrum (A) with the reference 1D ¹H spectra for each fragment (spectra B–F). b) HSQC analysis of fragment binding to EcDsbA. ¹H-¹⁵N HSQC spectrum of EcDsbA in the absence (green) and presence (purple) of fragment **2**.

A/Prof M.J. Scanlon, Dr J.S. Simpson, Dr D. Manallack, P. Sharma, Dr J. Horne
Monash Institute of Pharmaceutical Sciences
Monash University, 381 Royal Parade, Parkville, Victoria,
Australia
Fax: (+61 3 9903 9582)
E-mail: martin.scanlon@monash.edu.

Prof. J. L. Martin, Dr B. Heras, Dr S. Shouldice
Institute for Molecular Bioscience,
University of Queensland, St Lucia, Brisbane, Queensland,
Australia

[*] This work was supported by NHMRC funding to MJS, BH and JSS. PS gratefully acknowledges to MIPS for financial support and thank to Dr Jerome Wielens for discussion.

Supporting information for this article is available on the WWW under <http://www.angewandte.org>

Fragment-based screening (FBS) is a relatively new approach for identifying ligands of protein targets. The strategy involves identifying small, low-affinity ligands ('fragments') and combining these to produce larger, higher-affinity compounds.^[11] The major advantage of FBS over more traditional high-throughput screening is that FBS provides a more rapid and effective means of identifying ligand 'hits' at a protein target. It samples chemical space more efficiently than traditional approaches^[12] and therefore requires far fewer compounds to be tested to identify suitable hits as starting points for development. In the current study we have employed a library of 1132 compounds with average $M_r = 208$. A major challenge associated with FBS-based approaches to drug design is that due to the small size of the fragments, they generally bind to their target with low affinity and therefore structural details of their complexes are generally required to support compound optimisation.^[13] In this respect, DsbA is well suited to an FBS approach. Structural details of *E. coli* DsbA (referred to herein as *EcDsbA*) in different stages of its catalytic cycle^[14] have provided key insights into the features that regulate the activity of this enzyme. The structure of *EcDsbA* consists of a thioredoxin (TRX) fold, which is common to many disulfide-dithiol oxidoreductases,^[15] with an inserted helical domain that is unique to DsbA enzymes. (Figure S2 in supporting information) The TRX domain contains the redox-active pair of cysteines (C30-P31-H32-C33) and a *cis*-proline residue (P151) that is adjacent to the active site. The surface surrounding the active site of *EcDsbA* is largely hydrophobic and has been shown to contribute to the binding of unfolded polypeptides.^[16] This is consistent with kinetic studies that have demonstrated that reduced polypeptides are better substrates of *EcDsbA* than organic dithiols such as DTT and glutathione.^[17] In addition, the crystal structure of a complex between *EcDsbA* and *EcDsbB* revealed that a groove adjacent to the active site of *EcDsbA* is the binding site for a periplasmic loop of *EcDsbB*^[14b] and provided key details of the mechanism of disulfide formation in *E. coli*.

| Fragment Screening | # Results |
|---|-----------|
| STD screening hits | 120/1132 |
| HSQC screening (CSP > 0.03 ppm) hits | 20/1132 |
| Ligand Efficiency >1.4 kJ mol ⁻¹ HAC ⁻¹ | 11/1132 |
| EC ₅₀ < 100 μM | 6/1132 |
| Promising hits in phenotype microarray | 4/1132 |
| Hits with High Quality Crystal Complex | 10/1132 |

Table 1. Fragment library screening to identify *EcDsbA* ligands

In this study we have performed FBS on our in-house compound library using saturation transfer difference (STD) NMR spectroscopy (Figure 1a).^[18] Approximately 10% of the compounds in the library gave a positive STD result (Table 1).

The binding of these fragments was validated by measuring chemical shift changes of *EcDsbA* resonances in ¹H-¹⁵N heteronuclear single quantum coherence (HSQC) spectra (Figure 1b). Twenty fragments produced chemical shift perturbations (CSP) of > 0.03 ppm in ¹H-¹⁵N HSQC spectra of uniformly ¹⁵N labelled *EcDsbA* and these were retained for further characterisation (Table 1).

Dissociation constants (K_D) and ligand efficiencies (LE) for these fragments were calculated by titrating *EcDsbA* with increasing concentrations of each fragment (Figure 2a and Table 2). This analysis revealed that eleven fragments bound with a LE > 1.4 kJ mol⁻¹ HAC⁻¹ (Table 1), which suggests that they are amenable to being developed into high-affinity inhibitors of *EcDsbA*.^[19] These fragments could be classified into seven distinct chemical classes (Table 2). Sequence-specific backbone resonance assignments for *EcDsbA* were generated using standard triple-resonance NMR methods and these were used to identify the binding site for each of the fragments. Each of the fragments caused significant perturbations of residues in the hydrophobic groove of *EcDsbA*, however the patterns of perturbations varied for the each fragment listed in Table 2 (see Table S1 in the Supporting Information).

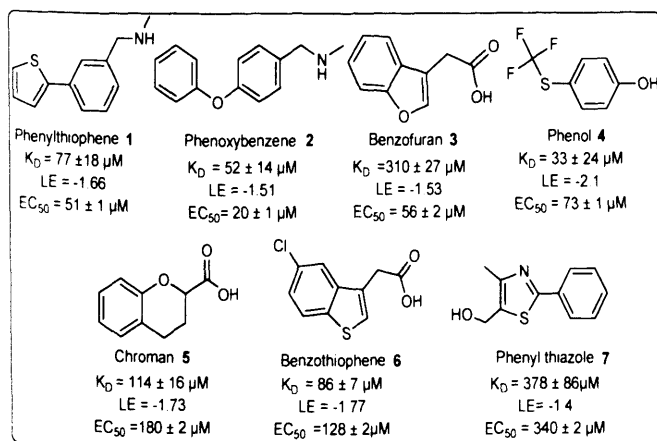


Table 2: Seven chemical classes of NMR based fragment screening hits of *EcDsbA* with representative structure of each class shown with NMR based K_D , ligand efficiency (LE) units kJ mol⁻¹ HAC⁻¹ and EC_{50} determined in bacterial motility assay.

To characterise their binding, fragments were soaked into preformed crystals of *EcDsbA*. Complexes were observed for four of the fragments, which produced diffraction data in the range of 1.7 - 2.0 Å resolution

(Figure 3). *EcDsbA* crystallises with two monomers in the asymmetric unit.

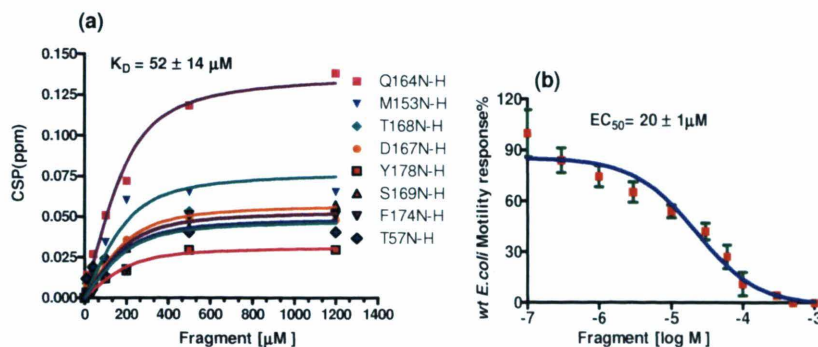


Figure 2. a) ^1H - ^{15}N HSQC titration curve obtained for fragment **2** for those residues that undergo dose dependant chemical shift changes upon addition to *EcDsbA*. The measured CSP was analysed as a function of fragment concentration to calculate the dissociation constant K_D b) Dose-dependent inhibition of bacterial motility was observed and analysed to calculate EC_{50} values for the fragments. A representative plot for phenoxybenzene fragment **2** is shown.

In each case, the fragment was observed to bind in the hydrophobic groove region comprising residues P163, Q164, T168, S169, M171, F 174, F36, V39 and L40 near to H32, Q35 and *cis*-P151 of the active site region of only one of the two *EcDsbA* molecules. Data collection and refinement statistics for the complexes are listed in Table S2 and electron density maps of binding region have been shown Figure S3 of the supporting information. Phenylthiophene **1** and phenoxybenzene **2** were observed to bind in a similar region of the hydrophobic groove adjacent to the CPHC motif of *EcDsbA*. Binding was dominated by hydrophobic contacts with numerous vdW interactions identified (through analysis of the complexes using CCP4).^[20] In both cases the fragment was surrounded by H32, F36, V39 and Q164 with partial stacking of F174 over the phenyl ring of the fragment **1** and the benzene ring of fragment **2**, respectively. The methyl aminomethyl substituent was in close enough proximity (3.8 Å for **2** and 3.7 Å for **1**) to P151, Q164 and H32 to make vdW interactions in both cases. The thiophene ring of fragment **1** and the phenoxy substituent of fragment **2** mainly interact with residues L40, T168 and M171 hydrophobically. The phenyl ring of the fragment **1** and the central benzene ring of fragment **2** make hydrophobic contacts with residues F36 and P163 in addition to partial stacking with F174.

The benzofuran fragment **3** was found to bind in a slightly different location, with the aromatic core of this fragment occupying a similar location to the methylaminomethyl portion of phenoxybenzene **2** and phenylthiophene **1** (Figure S4 of supporting information), surrounded by F36 and P151 (Figure 3).

A water-mediated H-bond was present between the carboxylate of benzofuran **3** and the backbone carbonyl oxygen of V150 of *EcDsbA* (**3** - 2.7 Å - H_2O - 2.5 Å - V150). vdW interactions were observed between benzofuran **3** and the residues (R148, G149, V150) preceding P151 in the *cis*-proline loop of *EcDsbA*. 4-((Trifluoromethyl)thio)phenol (**4**) was observed to bind at a third site, also within the hydrophobic groove but further from the P151 loop than the other three fragments. Binding was stabilised by an H-bond between the phenol OH and the backbone amide of Q35 (1.9 Å, N-H - O). The aromatic ring of phenol **4** lies on a largely hydrophobic surface formed by the side chains of V39, L40 and M171, whilst the trifluoromethyl group makes hydrophobic interactions with F174.

The binding locations observed in the crystal structures are consistent with the patterns of CSP observed in the ^1H - ^{15}N HSQC spectra of the fragments (Figure 4). Thus the patterns of CSP induced by phenylthiophene **1** and phenoxybenzene **2** are very similar (Figure 4 and Table S1 in the Supporting Information), whereas benzofuran **3** causes relatively larger perturbations in residues in the loop connecting the TRX and helical domains of *EcDsbA*, which is adjacent to the *cis*-proline loop, and phenol **4** shows the largest perturbations in the loop containing F174 that interacts with the trifluoromethyl group. The large CSP observed in this loop upon titration with phenol **4** are also consistent with the change in the conformation of the residues in this region (P163-F174) that is observed in the crystal structure of the complex (see Figure S5 in the Supporting Information).

The ability of these fragments to inhibit the activity of *EcDsbA* was tested by measuring their effects on the motility of *E. coli* in soft agar. Mutants in *dsbA* fail to

assemble the FlgI subunit that forms the P-ring of the flagellar motor and are therefore non-motile in soft agar.^[21]

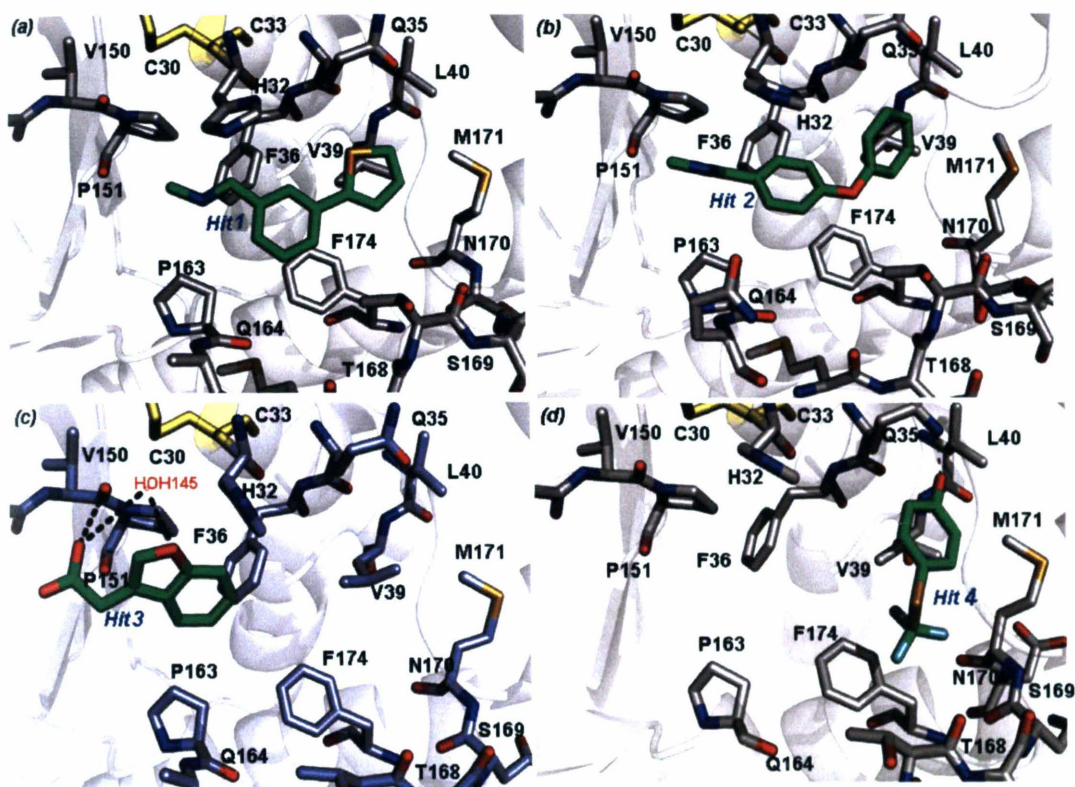


Figure 3: Binding modes of identified hits 1, 2, 3 and 4 from crystal complexes with *EcDsbA*, (grey stick and white ribbon representation). Fragments which bind in hydrophobic groove near C30-C33 (yellow) have been shown as green stick representation. Oxygen atoms, nitrogen atoms and sulfur atom are colored in red, blue and yellow, respectively

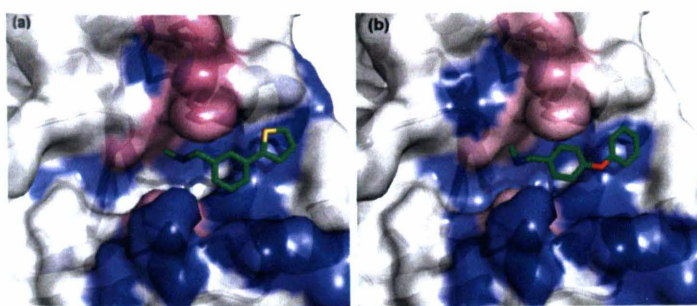


Figure 4. *EcDsbA* surface mapped with CSP (> 0.03 ppm) induced and calculated from ¹H-¹⁵N HSQC spectra, upon protein-binding of phenyl thiophene hit 1 (a) and phenoxybenzene 2 (b) at 1 mM (green stick representation). The CSP and unassigned residues (P151, P31 and H32) have been colored as blue and pink respectively.

Each of the fragments (1, 2, 3 and 4) was found to inhibit *E. coli* motility in a dose-dependent fashion (Figure 2, Table 2). In order to determine that this effect was due to inhibition of *EcDsbA*, a *dsbA* strain of *E. coli* was complemented with DsbA from *Vibrio cholerae* (*VcDsbA*).

Although *EcDsbA* and *VcDsbA* share low sequence identity the surface features around the active site of the two enzymes are largely conserved^[16, 22] and

VcDsbA can functionally complement the motility defect in *dsbA* *E. coli* (Figure 5). Nonetheless, the residues that constitute the binding site for the fragments in *EcDsbA* are not well conserved between the two enzymes (see Figure S6 and S7 in the supporting information). The fragments (1, 2, 3 and 4) either do not bind or bind weakly to *VcDsbA* as determined by their ability to induce CSP in the ¹H ¹⁵N HSQC spectrum of uniformly ¹⁵N-labelled *VcDsbA*. The *VcDsbA* non-binding fragments were found to have no effect on the motility of *dsbA* *E. coli*

complemented with *VcDsbA*, suggesting that the loss of motility is due to specific inhibition of *EcDsbA* (Figure 5).

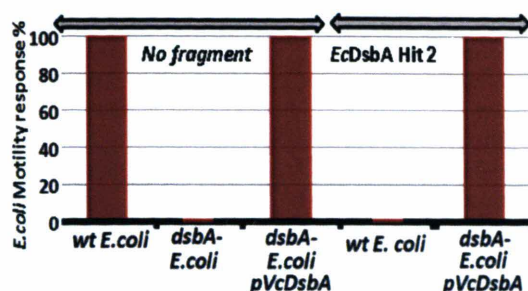


Figure 5. Motility restoration by *VcDsbA* and selective inhibition of wt *E. coli* motility by *EcDsbA* binding fragment 2.

Keywords: Bacterial virulence · Disulfide bonds · Fragment based drug design · NMR · X-Ray Crystallography. Phenotype assays

In conclusion, through a combination of FBS and X-ray crystallographic structure determination we have identified a chemically diverse group of fragments that bind in the hydrophobic groove of *EcDsbA*. The fragments show dose-dependent effects on the DsbA-dependent phenotype of motility. The structural data indicates that the fragments occupy different regions of the hydrophobic groove, thereby providing a structural rationale for their optimisation. Taken together the current data indicate that *EcDsbA* is a suitable target for the development of narrow spectrum anti-virulence agents. It has been suggested that compounds with this spectrum of activity will exert less selective pressure for the development of resistance than traditional antibiotics that either kill bacteria or inhibit their growth and may therefore represent a future treatment of multiply-resistant bacterial infections.^[23]

Experimental Section

The expression and purification of *EcDsbA* and *VcDsbA* was performed essentially as described previously.^[24] Full details are provided in the supporting information. Structural models of the complexes of *EcDsbA* have been deposited with the Protein Data Bank under the accession codes (1AAA, 2BBB, 3CCC and 4DDD, which will be updated later). Further details of the NMR binding and affinity studies, details of protein crystallisation and structure refinement and phenotypic analysis of the fragments are provided in the supporting information.

Received: ((will be filled in by the editorial staff))

Published online on ((will be filled in by the editorial staff))

References

- [1] B. Heras, S. R. Shouldice, M. Totsika, M. J. Scanlon, M. A. Schembri, J. L. Martin, *Nat Rev Microbiol* **2009**, 7, 215-225.
- [2] aF. Jacob-Dubuisson, J. Pinkner, Z. Xu, R. Striker, A. Padmanabhan, S. J. Hultgren, *Proc Natl Acad Sci U S A* **1994**, 91, 11552-11556; bT. Miki, N. Okada, H. Danbara, *J Biol Chem* **2004**, 279, 34631-34642; cT. Miki, N. Okada, Y. Kim, A. Abe, H. Danbara, *Microb Pathog* **2008**, 44, 151-158; dJ. A. Peek, R. K. Taylor, *Proc Natl Acad Sci U S A* **1992**, 89, 6210-6214; eC. V. Rosadini, S. M. Wong, B. J. Akerley, *Infect Immun* **2008**, 76, 1498-1508; fV. E. Shevchik, G. Condemine, J. Robert-Baudouy, *Embo J* **1994**, 13, 2007-2012; gT. H. Stenson, A. A. Weiss, *Infect Immun* **2002**, 70, 2297-2303; hC. R. Tinsley, R. Voulhoux, J. L. Beretti, J. Tommassen, X. Nassif, *J Biol Chem* **2004**, 279, 27078-27087; iJ. F. Tomb, *Proc Natl Acad Sci U S A* **1992**, 89, 10252-10256; jM. Watarai, T. Tobe, M. Yoshikawa, C. Sasakawa, *Proc Natl Acad Sci U S A* **1995**, 92, 4927-4931; kV. P. Zav'yalov, T. V. Chernovskaya, D. A. Chapman, A. V. Karlyshev, S. MacIntyre, A. V. Zavialov, A. M. Vasiliev, A. I. Denesyuk, G. A. Zav'yalova, I. V. Dudich, T. Korpela, V. M. Abramov, *Biochem J* **1997**, 324 (Pt 2), 571-578; lH. Z. Zhang, M. S. Donnenberg, *Mol Microbiol* **1996**, 21, 787-797.

- [3] aY. Akiyama, S. Kamitani, N. Kusakawa, K. Ito, *J Biol Chem* **1992**, 267, 22440-22445; bJ. C. Bardwell, K. McGovern, J. Beckwith, *Cell* **1991**, 67, 581-589.
- [4] M. Bader, W. Muse, D. P. Ballou, C. Gassner, J. C. Bardwell, *Cell* **1999**, 98, 217-227.
- [5] R. J. Dutton, D. Boyd, M. Berkmen, J. Beckwith, *Proc Natl Acad Sci U S A* **2008**, 105, 11933-11938.
- [6] M. Berkmen, D. Boyd, J. Beckwith, *J Biol Chem* **2005**.
- [7] H. Kadokura, H. Tian, T. Zander, J. C. Bardwell, J. Beckwith, *Science* **2004**, 303, 534-537.
- [8] aD. T. Foreman, Y. Martinez, G. Coombs, A. Torres, Y. M. Kupersztoch, *Mol Microbiol* **1995**, 18, 237-245; bC. Wulfig, R. Rappuoli, *Arch Microbiol* **1997**, 167, 280-283.
- [9] A. Hiniker, J. C. Bardwell, *J Biol Chem* **2004**, 279, 12967-12973.
- [10] D. Missiakas, C. Georgopoulos, S. Raina, *Proc Natl Acad Sci U S A* **1993**, 90, 7084-7088.
- [11] aD. Lesuisse, G. Lange, P. Deprez, D. Benard, B. Schoot, G. Delettre, J. P. Marquette, P. Broto, V. Jean-Baptiste, P. Bichet, E. Sarubbi, E. Mandine, *J Med Chem* **2002**, 45, 2379-2387; bD. C. Rees, M. Congreve, C. W. Murray, R. Carr, *Nat Rev Drug Discov* **2004**, 3, 660-672.
- [12] aR. A. Carr, M. Congreve, C. W. Murray, D. C. Rees, *Drug discovery today* **2005**, 10, 987-992; bM. J. Hartshorn, C. W. Murray, A. Cleasby, M. Frederickson, I. J. Tickle, H. Jhoti, *J Med Chem* **2005**, 48, 403-413.
- [13] C. W. Murray, T. L. Blundell, *Curr Opin Struct Biol* **2010**, 20, 497-507.
- [14] aL. W. Guddat, J. C. Bardwell, J. L. Martin, *Structure* **1998**, 6, 757-767; bK. Inaba, S. Murakami, M. Suzuki, A. Nakagawa, E. Yamashita, K. Okada, K. Ito, *Cell* **2006**, 127, 789-801; cJ. L. Martin, J. C. Bardwell, J. Kuriyan, *Nature* **1993**, 365, 464-468; dJ. J. Paxman, N. A. Borg, J. Horne, P. E. Thompson, Y. Chin, P. Sharma, J. S. Simpson, J. Wielens, S. Piek, C. M. Kahler, H. Sakellaris, M. Pearce, S. P. Bottomley, J. Rossjohn, M. J. Scanlon, *J Biol Chem* **2009**, 284, 17835-17845; eH. J. Schirra, C. Renner, M. Czisch, M. Huber-Wunderlich, T. A. Holak, R. Glockshuber, *Biochemistry* **1998**, 37, 6263-6276.
- [15] J. L. Martin, *Structure* **1995**, 3, 245-250.
- [16] L. W. Guddat, J. C. Bardwell, T. Zander, J. L. Martin, *Protein Sci* **1997**, 6, 1148-1156.
- [17] aN. J. Darby, T. E. Creighton, *Biochemistry* **1995**, 34, 3576-3587; bA. Zapun, J. C. Bardwell, T. E. Creighton, *Biochemistry* **1993**, 32, 5083-5092.
- [18] B. Meyer, J. Klein, M. Mayer, R. Meinecke, H. Moller, A. Neffe, O. Schuster, J. Wulfsen, Y. Ding, O. Knaie, J. Labbe, M. M. Palcic, O. Hindsgaul, B. Wagner, B. Ernst, *Ernst Schering Res Found Workshop* **2004**, 149-167.
- [19] P. J. Hajduk, *J Med Chem* **2006**, 49, 6972-6976.
- [20] aS. Bailey, *Acta Crystallographica Section D-Biological Crystallography* **1994**, 50, 760-763; bE. B. Krissinel, M. D. Winn, C. C. Ballard, A. W. Ashton, P. Patel, E. A. Potterton, S. J. McNicholas, K. D. Cowtan, P. Emsley, *Acta Crystallogr D Biol Crystallogr* **2004**, 60, 2250-2255.
- [21] F. E. Dailey, H. C. Berg, *Proc Natl Acad Sci U S A* **1993**, 90, 1043-1047.
- [22] S. H. Hu, J. A. Peek, E. Rattigan, R. K. Taylor, J. L. Martin, *J Mol Biol* **1997**, 268, 137-146.
- [23] D. A. Rasko, V. Sperandio, *Nat Rev Drug Discov* **2010**, 9, 117-128.
- [24] J. Marley, M. Lu, C. Bracken, *J Biomol NMR* **2001**, 20, 71-75.

Supporting Information

Inhibition of Oxidative Protein Folding in *Escherichia coli*

Pooja Sharma¹, Begoña Heras², James Horne¹, Stephen Shouldice², David Manallack¹, Jennifer L. Martin², Jamie S. Simpson¹ and Martin J Scanlon^{1*}

¹*Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville 3052, (Australia)*

²*Institute of Molecular Bioscience, University of Queensland, Chancellors Place, St Lucia, Brisbane, Queensland, 4067 (Australia)*

Expression and purification of unlabeled and isotope-labeled EcDsbA

Expression and purification of uniformly ^{15}N isotope-labeled and unlabeled protein was performed according to the protocol described by Marley *et al* ^[1] and Studier ^[2] respectively. The protocol for isotope-labeled protein expression is described here briefly.

An overnight streak culture was grown from glycerol stock on LB/agar *kan*⁺ plates. An isolated colony was used to inoculate a 100 ml starter culture of LB/*kan*⁺ (50 $\mu\text{g ml}^{-1}$) and grown overnight at 37 °C. These starter cultures were used for the inoculation of large cultures consisting of LB (3 L) and then shaken and incubated at 37°C, 260 rpm until an A_{600} OD of 0.6 was reached. These cells were centrifuged at 14000 g for 30 min at 4 °C and resuspended in minimal media (M9 salts) used for protein expression. To achieve the ^{15}N and ^{13}C labelling, ^{15}N -labeled ammonium chloride and ^{13}C -labeled glucose were present as the sole nitrogen and carbon source. Bacteria were grown at 37 °C for 1 hour and induced with 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG). The culture was then grown post-induction for 5-6 hrs to allow protein expression, followed by pelleting by centrifugation at 14000 g for 30 min at 4 °C. Expressed EcDsbA was released from the harvested cells by periplasmic fractionation: the cells were resuspended in 50 mL of 4 mg/mL polymyxin sulfate B in 20 mM Tris pH 8.0 with shaking at 200 rpm for 24 hrs at 20 °C to release the periplasmic fraction containing EcDsbA.^[3] The lysate was centrifuged at 25000 g for 20 min at 4 °C and the supernatant (cell lysate) was retained.

Purification of proteins was carried out using 2 chromatographic steps on fast protein liquid chromatography (FPLC) (Amersham Pharmacia Biotech). ^[4] Firstly, ammonium sulfate was added to the supernatant to a final concentration of 1 M. The supernatant was filtered and applied to a Phenyl Sepharose High Performance 16/10 hydrophobic interaction^[5] column (Amersham Biosciences, Sydney, Australia) which had been equilibrated in 20 mM Tris (pH 8.0), 50mM NaCl, 1 M $(\text{NH}_4)_2\text{SO}_4$. A linear gradient from 100% buffer B (20 mM Tris (pH 8.0), 50 mM NaCl, 1 M $(\text{NH}_4)_2\text{SO}_4$) to 100% buffer A (20 mM Tris (pH 8.0), 50mM NaCl) was used to elute the protein. Fractions were collected automatically and DsbA was located by the UV profile of the

elution and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) followed by silver staining.^[6] Fractions of DsbA were pooled and buffer exchanged into 20 mM Tris (pH 8.0). This PHP purified protein was applied to MonoQ HR5/5 cation-exchange column^[7] (Amersham Biosciences, Sydney, Australia). The protein was eluted with a gradient from 100% buffer C (20mM Tris (pH 8.0)) to 100% buffer D (20mM Tris (pH 8.0) and 1 M NaCl). This two-step purification protocol was sufficient to obtain pure protein, which was estimated to be > 95% pure from analysis by SDS-PAGE followed and silver staining. The final yield was about 80 mg of protein/L of medium. DsbA was oxidized by 1 mL of copper phenanthroline solution (30:9:61 of 100 mM CuCl₂: 1 M Phenanthroline: Tris 20mM), which was dialysed away using a stirred cell apparatus (Omega 3K 150 mL). Finally, protein was buffer exchanged using a desalting column equilibrated with 50 mM HEPES (pH 7.0), 150 mM NaCl to get ~60 mg DsbA from 1L culture. Unlabeled and isotope-labeled protein samples were stored at 4 °C and used for crystallography and NMR studies. The concentration of DsbA protein samples was estimated from the absorbance at 280 nm, using calculated extinction coefficients of 28545 M⁻¹cm⁻¹ and 10555 M⁻¹cm⁻¹ for *Ec*DsbA and *Vc*DsbA.^[8] For 2D, 3D NMR and crystallographic analysis, labeled and unlabeled *Ec*DsbA was concentrated to 150 μM, 500 μM and 1200 μM respectively.

Fragment library

A library comprised of 1132 fragments was purchased from Maybridge (Trevillet, Cornwall, UK) (<http://www.maybridge.com>). For storage, stocks of each fragment were prepared by diluting the individual fragments (10 mg) in ²H₆-dimethyl sulfoxide (DMSO, 80 μl) to give a concentration of approximately 600 mM, based on the average molecular weight of compounds in the library (208 Da). 1D ¹H NMR spectra for each fragment (1 mM in 100 mM phosphate buffer pH 7.0, 99% D₂O, 1% ²H₆-DMSO) were collected at 283 K on a Bruker-Biospin Avance 800 MHz spectrometer fitted with a cryoprobe and sample changer. Fragments were then mixed

in groups of up to five compounds with well resolved resonances in their ^1H 1D NMR spectra for screening purposes.

NMR Spectroscopy

Saturation Transfer Difference (STD) NMR

STD (Saturation transfer difference) NMR experiments^[9] for fragment library screening were performed in standard NMR tubes with 10 μM unlabelled protein and mixture of up to 5 fragments with an approximate concentration of 300 μM for each fragment. The sample volume was 550 μL with 10% D_2O and 90% buffer containing 10 mM HEPES pH 7.0 and 150 mM NaCl. Spectra were recorded on a Bruker BioSpin 800 MHz spectrometer (Billerica, MA, USA). Standard pulse sequences were used for 1D and STD data acquisition and water suppression was achieved using the WATERGATE selection scheme. Fragment screening experiments were carried out at 10 °C. ^1H chemical shifts were referenced to the $^1\text{H}_2\text{O}$ signal at 4.70 ppm.^[10] STD data sets were the average of 128 scans with a 5s saturation pulse. NMR data were processed in TOPSPIN version 3.1 (Bruker BioSpin)^[11]

^{15}N -HSQC Spectra

Second round of fragment screening was performed by recording ^1H - ^{15}N HSQC (Two-dimensional ^1H - ^{15}N heteronuclear single-quantum correlation) experiments for individual fragments hits identified from the STD screen. All spectra for the screening experiments were recorded at 298 K and end point titration experiments were performed in Shigemi (NMR) tubes with 1 mM test compound in 300 μL protein sample. Protein samples were prepared with 100 μM - 150 μM ^{15}N -labeled DsbA protein (*EcDsbA* or *VcDsbA*) in buffer containing 50 mM HEPES (pH 7.0), 150 mM NaCl and 1 mM EDTA with 90% H_2O and 10% D_2O . Spectra were recorded on a Varian Inova 600 or Bruker Avance 800, respectively, both of which were equipped with cryogenically cooled probes. NMR data were processed using NMRPipe.^[11b] Processed spectra were analysed with NMRDraw and Sparky 3.113^[12] software.

Compounds were regarded as positive hits if significant chemical shift perturbations were observed in the HSQC spectrum upon addition of the fragment. Overall weighted average chemical shift changes ($\Delta\delta$) in 2D ^1H - ^{15}N HSQC spectra were calculated for all assigned residues using Equation S1.^[13]

Equation S1

$$\Delta\delta = \sqrt{(\delta H_{\text{complexed DsbA}} - \delta H_{\text{Apo DsbA}})^2 + 0.154(\delta N_{\text{complexed DsbA}} - \delta N_{\text{Apo DsbA}})^2}$$

K_D and LE measurements

For the calculation of binding constant by HSQC titration, spectra were recorded at a range of concentrations of the ligand (fragment/compound), by sequentially adding higher concentrations of the ligand to the protein solution. A titration was considered complete (saturated) when the amide resonances of the ^1H - ^{15}N HSQC spectra no longer shifted on addition of ligand, or the solubility limit of the fragment was reached. Stock concentrations of each test ligand were prepared by diluting compounds in a solution with consistent amounts of DMSO and protein buffer. DMSO only control experiments were performed prior to decide the maximum amount of DMSO tolerated in each sample without causing any changes in apo DsbA HSQC spectra. Protein-ligand titrations were achieved by successively adding microliter amounts of test ligand stock to the ^{15}N labeled *EcDsbA* protein. The sample was mixed and allowed to equilibrate prior to ^1H - ^{15}N HSQC data collection. The binding affinity of ligands to the DsbA was determined by monitoring the chemical shifts of ^1H - ^{15}N HSQC (from titration series) cross peaks as a function of ligand-to-protein concentration ratio. The curves of different cross peaks were fitted using the following Equation S2 which assumes 1:1 binding stoichiometry.

Equation S2

$$\Delta\delta = \Delta\delta_{\text{max}} / 2(1 + [L_0]/[P_0] + K_d/[P_0] - \sqrt{1 + [L_0]/[P_0] + K_d/[P_0]^2}) - 4 [L_0]/[P_0]$$

Where K_D is the dissociation constant, $\Delta\delta$ and $\Delta\delta_{\text{max}}$ are the observed and maximal chemical shift changes, and $[L_0]$ and $[P_0]$ are the total ligand and protein

concentrations, respectively. Fitting was performed using the non-linear least squares fitting option in Graphpad Prism 5.0.^[14] Binding free energy changes (ΔG) (kJ mol^{-1}) and ligand efficiency per heavy atom (LE) unit $\text{kJ mol}^{-1} \text{HAC}^{-1}$ was calculated using Equations S3 and S4:^[15]

Equation S3

$$\Delta G = - RT \ln (K_D)$$

Where R is the gas constant ($8.314 \text{ J.mol}^{-1}\text{K}^{-1}$) and T is the temperature (298 K)

Equation S4

$$\text{LE} = \Delta G / (\text{number of heavy atoms})$$

NMR Assignments

Standard triple resonance experiments CBCACONH^[16], HNCA^[17], HNCACB^[18], HINCO^[16] and HNCACO^[19] using optimized parameter sets were recorded at 298 K on a Varian Unity 600 MHz spectrometer. ^1H - ^{15}N HSQC spectra and ^1H - ^{13}C HSQC were also recorded. These experiments were performed in Shigemi (NMR) tubes with $^{13}\text{C}/^{15}\text{N}$ - labeled protein (500 μM) with 300 μL sample volume, containing 10% D_2O and 90% buffer (25 mM MES, 0.5 mM arginine, 0.5 mM glutamate and 150 mM NaCl, pH 5.0). These assignments were later transferred onto all of the multidimensional protein spectra recorded for binding studies in different buffers. NMR data were processed using NMRPipe. Processed spectra were analysed with NMRDraw and Sparky 3.113 software using AMD Opteron workstations running SUSE Linux.

X-ray Crystallography

Crystallization

Crystals of *EcDsbA* were obtained using the hanging-drop vapour-diffusion method.^[20] Final crystallization trials were set up in VDXm 24-well hanging-drop plates and 18 mm siliconized cover slips (both from Hampton Research, San Diego CA, USA). Each cover slip held one 2 μ L drop containing 1 μ L protein solution and 1 μ L well solution while the well solution volume was kept at 500 μ L. The crystallization plates were incubated at 293 K and imaged using a temperature controlled RockImager (Formulatrix, Waltham, MA USA). The best diffraction quality crystals were obtained in Cacodylate 100 mM pH 6.1, CuCl_2 1 mM, PEG8000 14%-15% (w/v), glycerol 4%-8% (v/v) and *EcDsbA* at a concentration of 16 mg/mL. A typical diffraction quality crystal was prismatic or rectangular with average crystals of *ca.* $0.3 \times 0.15 \times 0.1 \text{ mm}^3$.

Soaking

In order to obtain *EcDsbA* crystals with a desired compound, apo protein crystals were soaked in the crystallographic stabilizing solution mixed with concentrated stocks of individual fragments in DMSO. The final concentration of fragments in *EcDsbA* stabilizing solution (cryoprotectant) ranged between 3 mM -10 mM depending upon the solubility of fragment.^[20] This solution comprised 24% (w/v) PEG8000, cacodylate 100 mM pH 6.1, CuCl_2 1 mM and glycerol 20% (v/v). A typical soaking time was 2-4 hrs for most of the fragments, while longer or shorter soaks were performed in a few cases to improve the quality of data. *EcDsbA* crystals were harvested from hanging drops using a 0.2-0.3 mm CrystalCap Copper Magnetic HT Cryoloop (Hampton Research, San Diego CA, USA) prior to flash cooling in liquid nitrogen. These crystals were maintained at 100 K throughout data collection and the

trials of crystal cooling indicated that both glycerol and PEG were necessary to attain sufficient cryoprotection.^[20]

Data collection

Diffraction data were collected on Rigaku FR-E copper which is a X-ray diffraction and radiation source with rotating anode generator (45 kV, 45 mA) and Osmic Confocal Vari-Max HF optics at a wavelength of 1.5418 Å (Rigaku Americas, Houston USA). The crystal-to-detector distance was 150 mm and images were collected over 90° (between 0° and 180°), with 0.5° image widths (oscillation angle) per frame, to make a total of 360 data frames with the exposure time for each frame as 1 min. Reflections were measured with a Rigaku Saturn 944 CCD area detector (Rigaku Americas, Houston USA). CryoCool LN₂ (Cryo Industries, Manchester, NH USA) was used for keeping crystals at 100 K during data collection.^[21] Diffraction data image frames were collected, indexed and integrated with CrystalClear1.4 (Rigaku Corporation) and scaled and merged with SCALEPACK.^[22] The structure of native *EcDsbA* (PDB code 1FVK)^[23] was used as a starting model during molecular replacement phasing method in CCP4 to obtain an initial model with phases.^[22] Model refinement was carried out using input scripts (modified as required) of the CCP4 and Phenix.^[22, 24] The crystal structure contains two molecules of *EcDsbA* in the asymmetric unit nominated as chain A and chain B (monomer A and monomer B correspondingly). Modelling and viewing of the *EcDsbA* crystal structures was performed with WinCoot.^[25] Co-ordinates for the ligands were built with generated restraints files using the sketcher module CCP4. The process used involved modeling and refining the structure of the protein first, followed by the addition of water buffer/metal molecules, and finally the fragment or ligand into (F_o – F_c) difference maps. The criteria used for including a water molecule were the presence of at least one possible hydrogen bond within 3.2 Å and 2F_o–F_c density at 1 σ and F_o–F_c density at 3 σ. The model was improved with cycles of iterative model refitting and building with density modification and refinement using the phenix.refine wizard. R-free analysis using 5–10% of reflections was used for cross-validation. Data validation parameters were calculated in Molprobit and summarized in **Table S2**.

Bacterial Motility Assay

Bacterial strains and plasmids for this cell based assay were kindly provided by Prof J. C. Bardwell (Dept of Microbiology and Molecular Genetics, Harvard Medical School, Boston, USA). *E. coli* wild type, *dsbA*⁻ and *VcDsbA* complemented mutant strains were grown individually as streak cultures in LB plates at 37 °C overnight to obtain single colonies of each type of strain.^[26] ^[27] Triplicates of fresh single colonies were inoculated by stabbing at the centre of each well. Wells contained the fragments to be tested at an appropriate concentration in MM media in 12 well plates. This media consisted of 0.3% agar, M63 salts with 40 µg ml⁻¹ of each amino acid (except cysteine and cystine).^[27] Plates were incubated for 7-10 hours at 37°C and a lawn of cells was allowed to grow before analysis of the swarming zone of motile *E. coli* cells. The diameter of bacterial swarming around the centre (inoculation point) was then measured. Data processing was performed with Graph Pad Prism 5.0.^[14] Swarming zone diameters were normalized to the DMSO-only control and the effect of the compounds was calculated as the percentage inhibition of motility, which were fitted against inhibitor concentration in logarithmic scale to a sigmoidal dose (concentration) response using given below.

$$Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{((\text{LogEC}_{50} - X))})$$

Where X is the logarithm of concentration, Y is response that starts at Bottom (base line response) and goes to Top (maximum response) with a sigmoid shape (Hill slope)

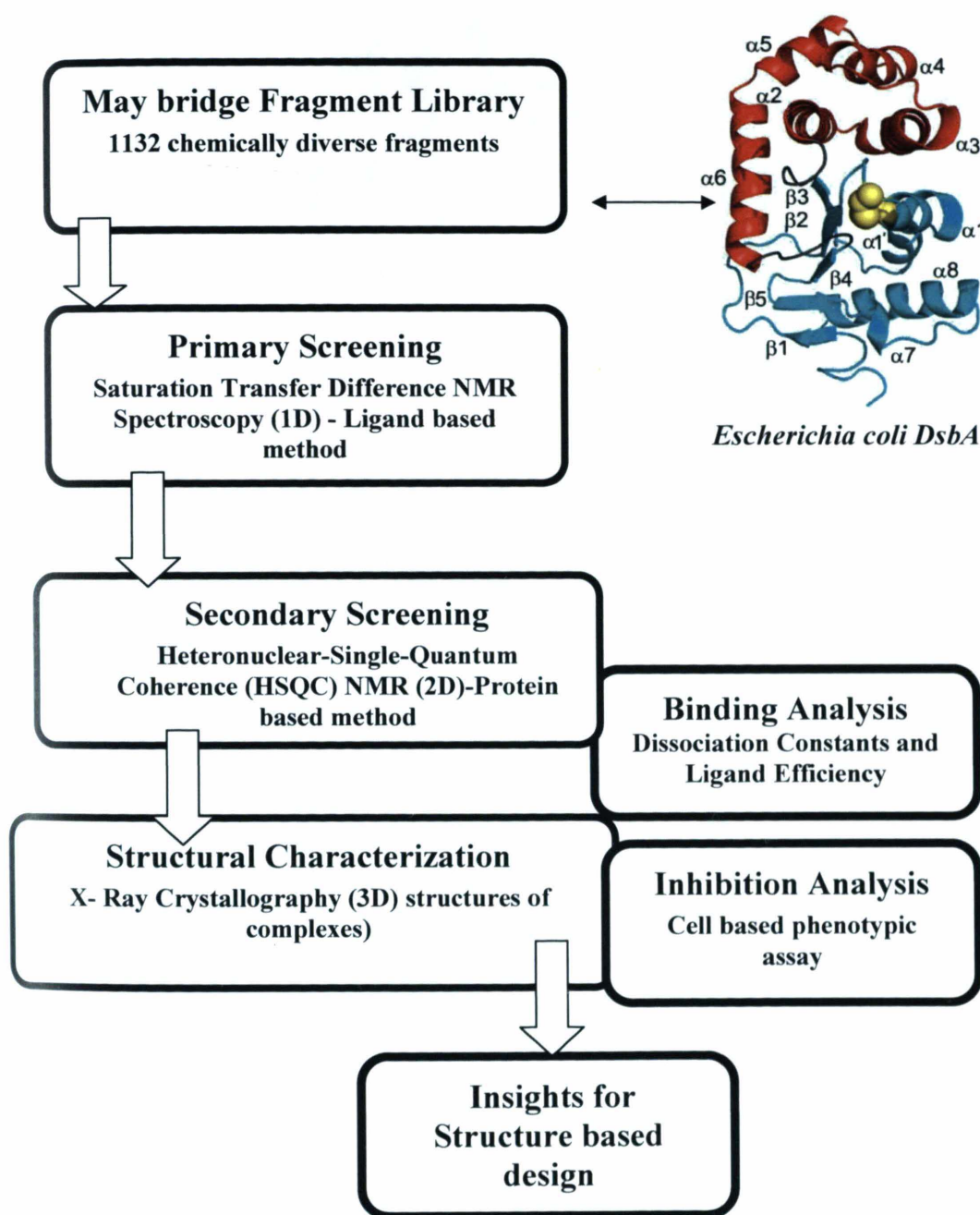


Figure S1: An outline of fragment based screening studies performed to identify *EcDsbA* inhibitors in current manuscript.

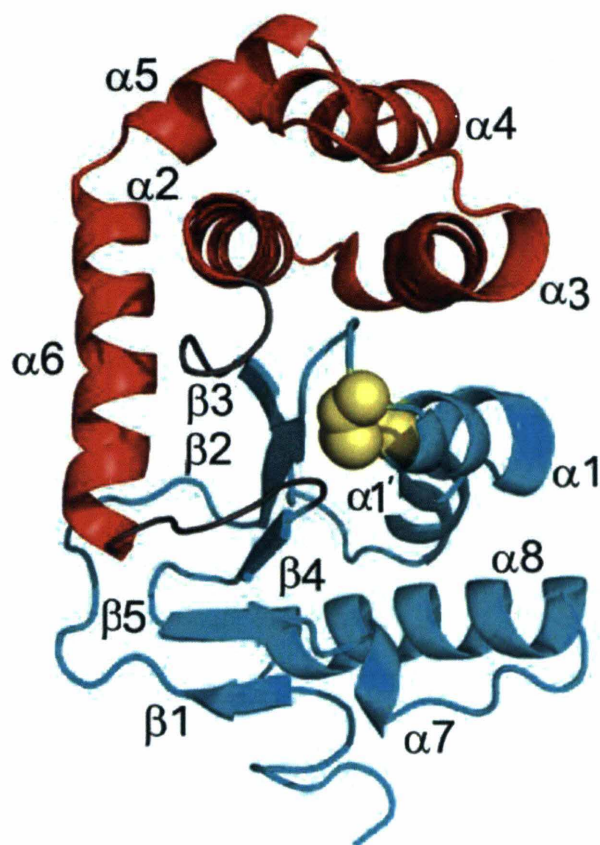
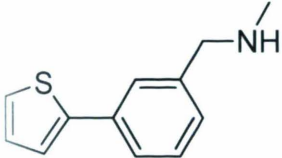
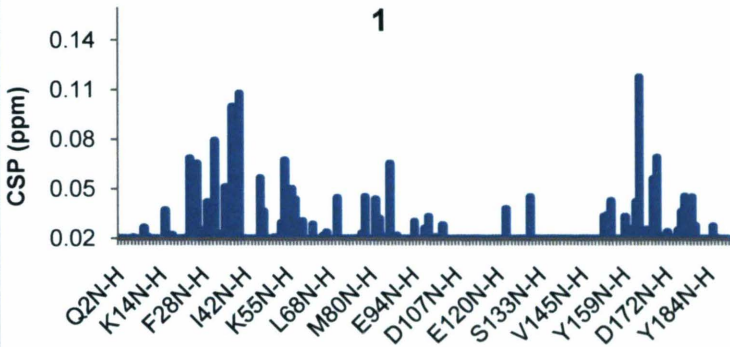
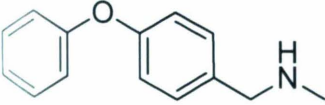
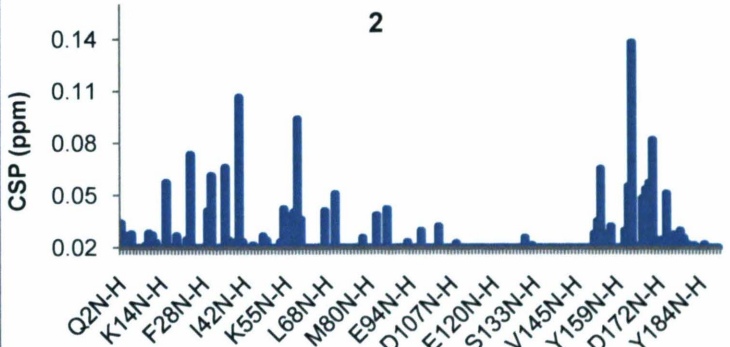
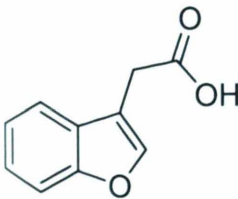
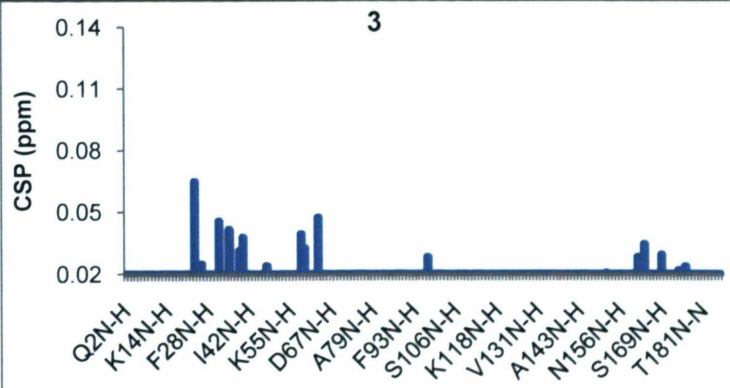
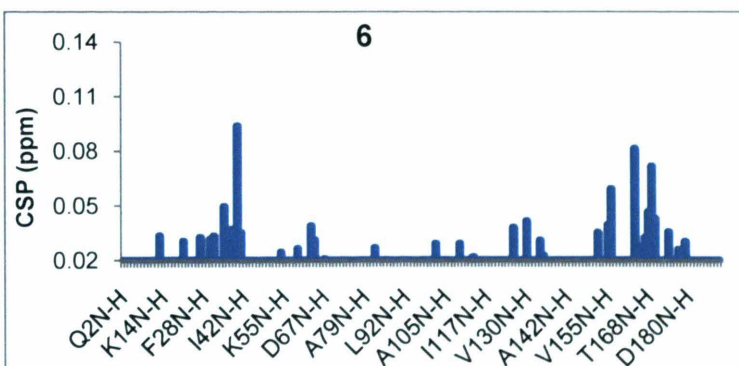
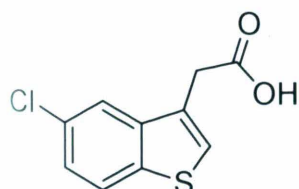
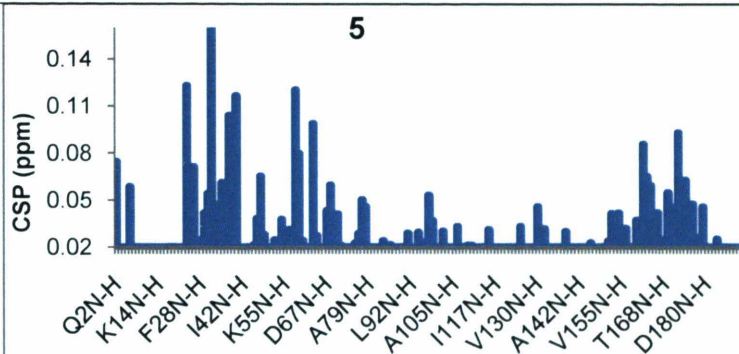
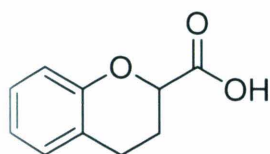
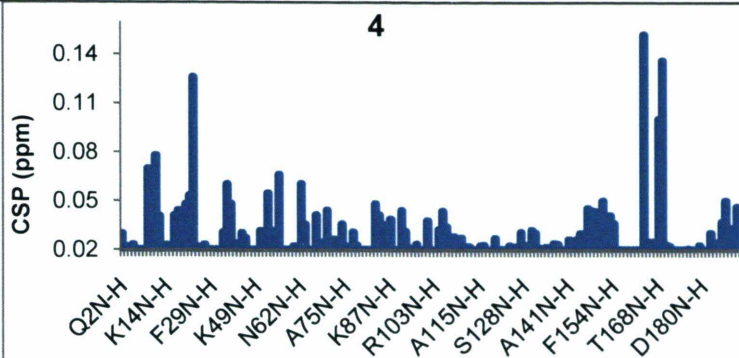
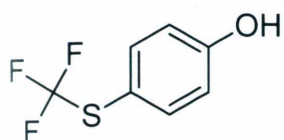
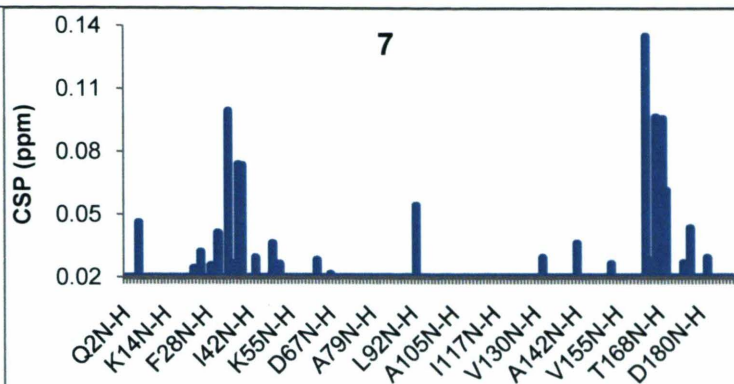
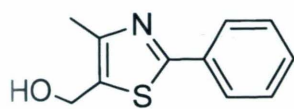


Figure S2: Ribbon diagram of native *EcDsbA* Monomer A in the asymmetric unit (PDB code 1FVK). In the crystal complex each of the monomer A and B adopted the characteristic secondary structure of the DsbA family^[23] comprised of thioredoxin (cyan) and α -helical (red) domains and the insertion points (gray). Elements of secondary structure are numbered sequentially from the *N*-terminus. The cysteine residues of the active site are shown in (yellow) CPK representation.

Table S1 – CSP ($\Delta\delta(\text{NH}) > 0.02$ ppm) induced in ^1H - ^{15}N -HSQC spectra of *EcDsbA* upon the addition of fragment for each backbone amide NH.

| Fragment Structures | HSQC Histogram analysis |
|---|--|
|  | <p style="text-align: center;">1</p>  |
|  | <p style="text-align: center;">2</p>  |
|  | <p style="text-align: center;">3</p>  |





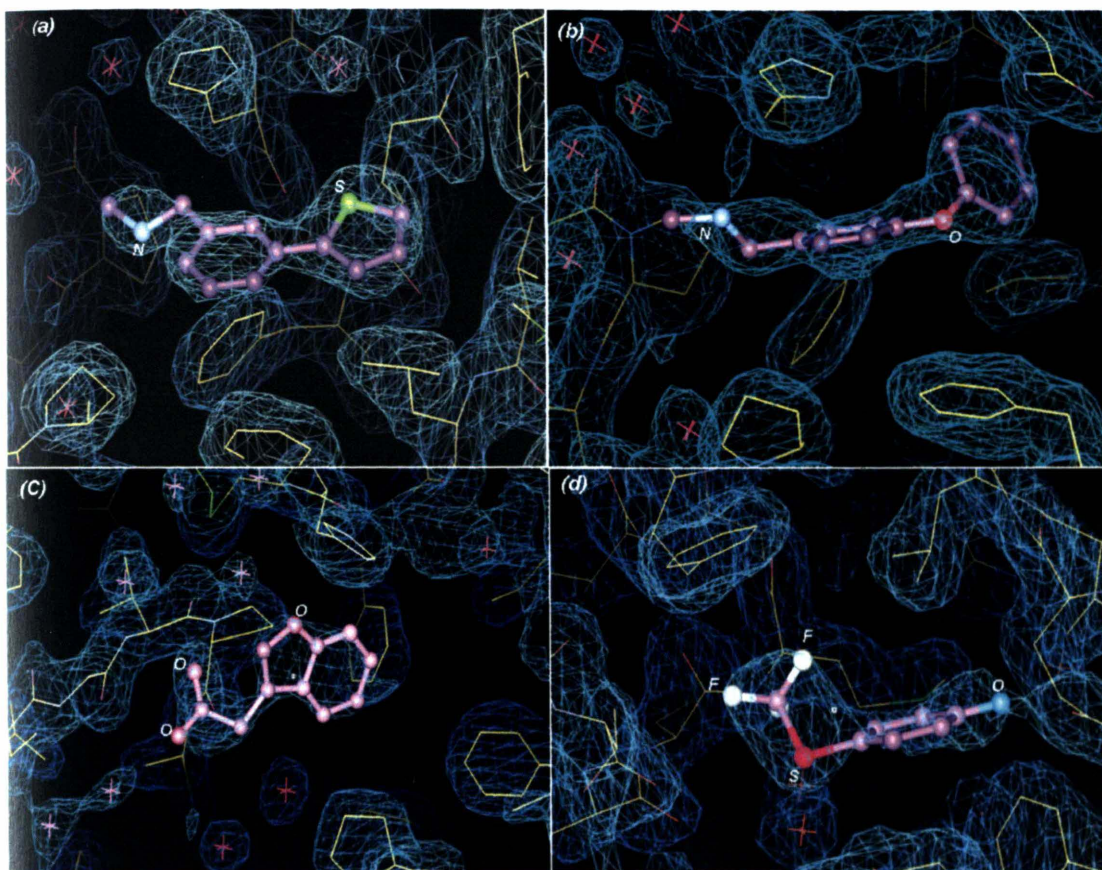


Figure S3: (2Fo - Fc) electron density maps for *EcDsbA*-fragment complexes with *EcDsbA* (contoured at σ level 1.0). Protein is shown as yellow line representation, water molecules shown as red star shaped line model, and fragments are shown as purple sticks representation with labeled heteroatoms.

Table S2: Data Collection and Refinement Statics

| Fragment | | 1 | 2 | 3 | 4 |
|--|---|------------------|------------------|------------------|------------------|
| Space group | | C2 | | | |
| Cell dimensions (Å) (α=γ=90°) | a | 116.22 | 116.23 | 117.4 | 117.03 |
| | b | 63.94 | 64.00 | 63.5 | 62.84 |
| | c | 74.51 | 74.20 | 74.68 | 74.04 |
| | β | 126.15 | 125.78 | 126.46 | 126.57 |
| Radiation Source | | Rotating Anode | | | |
| Wavelength (Å) | | 1.54 | | | |
| Diffraction data (for the highest resolution shell are shown within parentheses) | | | | | |
| Resolution (Å) | | (52.84-1.99) | (34.53-1.8) | (31.77-1.96) | (34.05-2.07) |
| Number of observed reflections | | 109444 | 110612 | 109534 | 93945 |
| Number of unique reflections | | 29801 | 40860 | 31354 | 26300 |
| Redundancy | | 3.67 (3.39) | 3.68 (3.43) | 3.66 (3.41) | 3.57 (3.07) |
| Completeness (%) | | 98.0 (94.1) | 98.2 (99.8) | 99.76 (96.78) | 99.4 (95.0) |
| I/ σ(I) | | 16.8 (3.5) | 14.2 (1.33) | 13.3 (1.36) | 18.4 (8.1) |
| Rmerge (%) | | 0.047 (0.261) | 0.053 (0.331) | 0.037 (0.320) | 0.043 (0.104) |
| Refinement Statistics | | | | | |
| resolution (Å) | | (52.84-1.99) | (34.53-1.8) | (31.77-1.96) | (34.05 - 2.07) |
| protein residues | | 376 | 376 | 376 | 376 |
| solvent molecules | | 308 | 326 | 241 | 317 |
| inhibitor | | 1 | 1 | 1 | 1 |
| R _{free} (%) | | 24.7 | 25.6 | 24.8 | 22.5 |
| R _{cryst} (%) | | 19.8 | 20.6 | 19.3 | 17.0 |
| root mean square deviations | | | | | |
| bonds (Å) | | 0.012 | 0.007 | 0.018 | 0.025 |
| angles (deg) | | 1.277 | 1.104 | 1.613 | 1.925 |
| Ramachandran plot | | | | | |
| residues in most favored regions (%) | | 97.85 | 98.39 | 97.58 | 97.85 |
| residues in allowed regions (%) | | 2.15 | 1.59 | 2.3 | 1.71 |
| estimated coordinated error | | | | | |
| Luzzati mean coordinate error (Å) | | 0.34 | 0.29 | 0.31 | 0.28 |
| Wilson B factor, Å ² | | 34.42 | 24.74 | 26.00 | 20.64 |

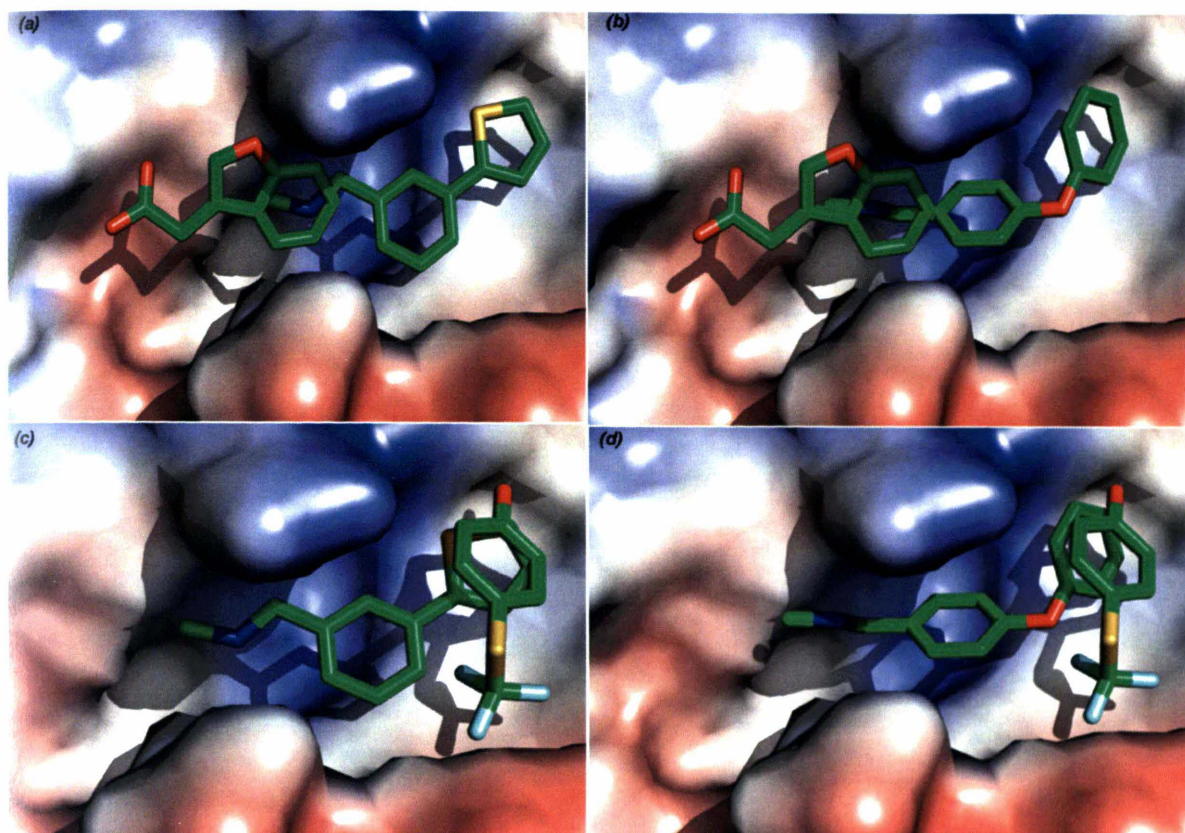


Figure S4: Crystallographic binding positions of fragments (a) **1** and **3** (b) **2** and **3** (c) **1** and **4** and (d) **2** and **4**. All fragments are shown as green carbon stick models in overlapping and adjacent binding sites in the hydrophobic groove on the *EcDsbA* surface, which is shown as an electrostatic surface model.

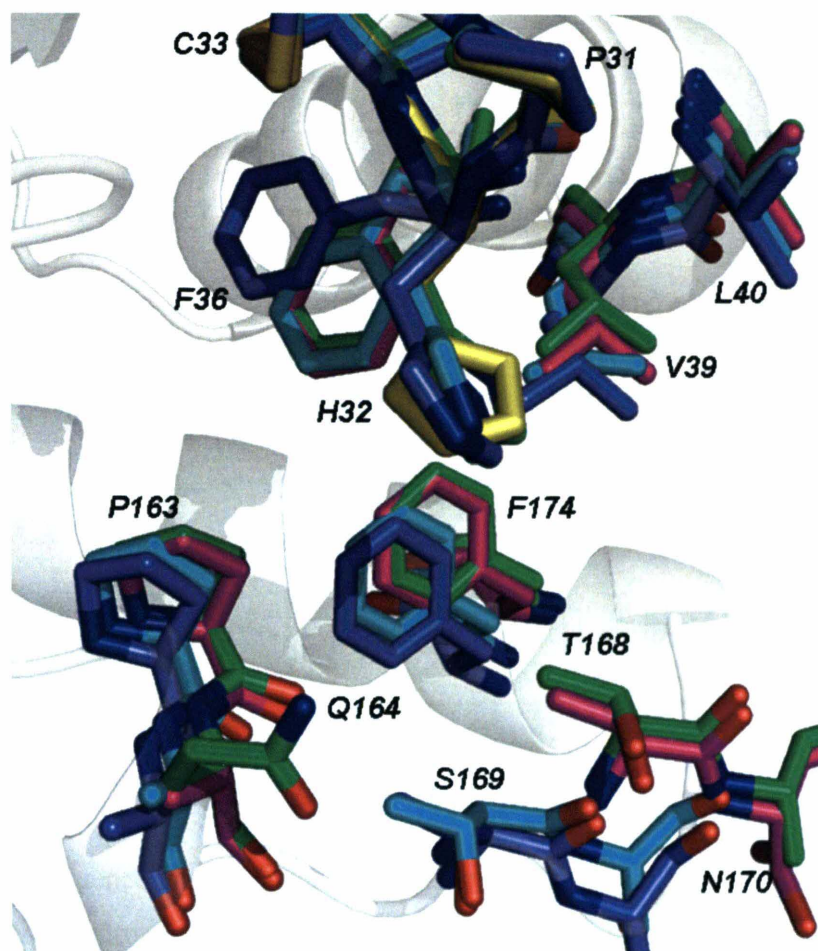


Figure S5: Overlay view showing major conformational transitions in hydrophobic groove, pocket and active site residues upon fragment 1, 2, 4 and 3 binding to EcDsbA (ligand not shown). Blue, cyan, yellow and green stick models represent EcDsbA residue side chains in complex with fragments 4, 3, 2 and 1 respectively whereas *apo* structure is shown as white cartoon representation.

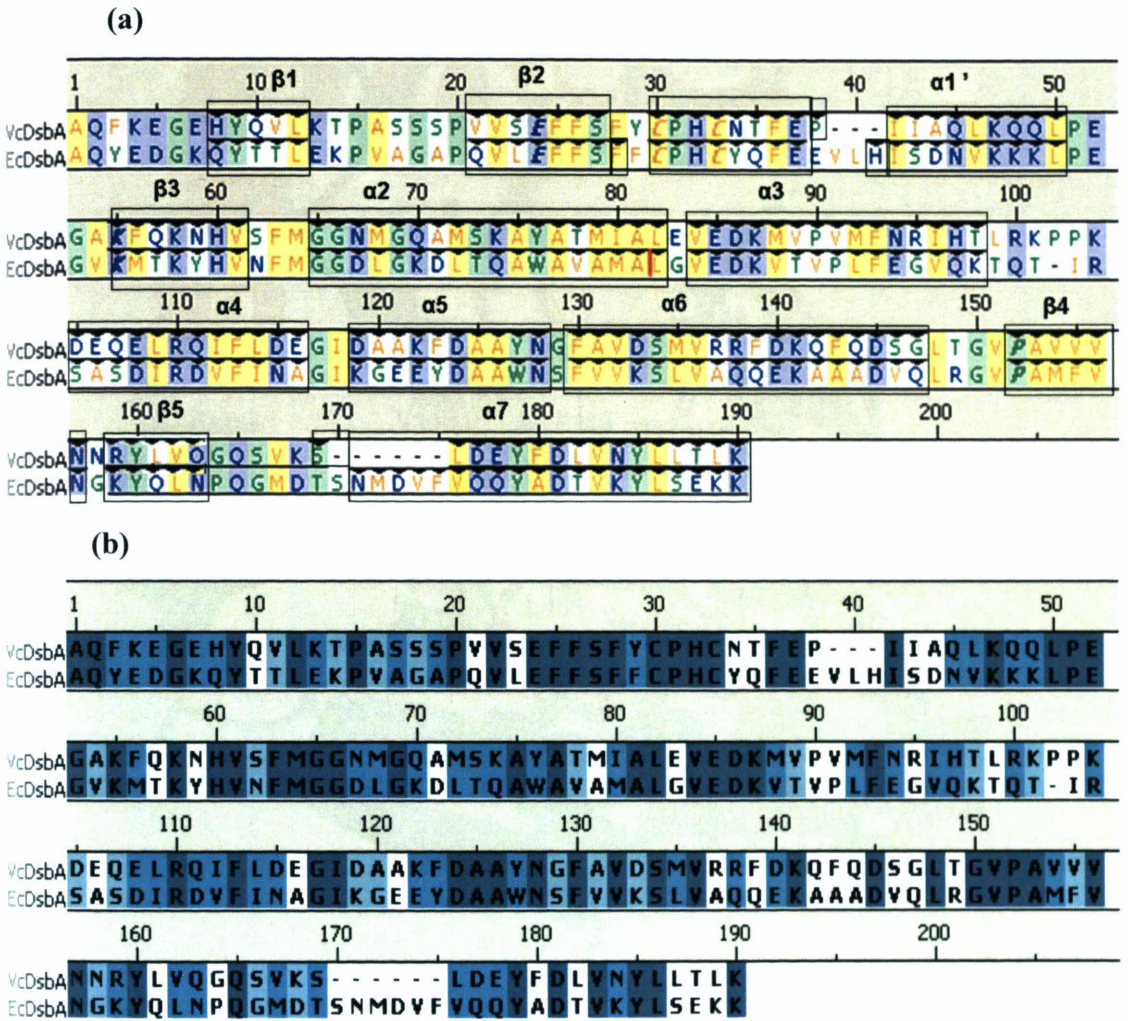


Figure S6: Sequence alignment of *V. cholerae* DsbA^[28] and *E. coli* DsbA^[23] colored based on the hydrophobicity (top panel) and structural (bottom) comparison. (a) Secondary structural units are boxed and labeled, and residues that are generally conserved throughout the thioredoxin superfamily are shown in bold. Residues are colored and highlighted as: hydrophobic- yellow, hydrophilic- blue and neutral- green. (b) Residues are colored based on sequence similarity where gradient of blue refers to similarity levels as identical, strong and weak.

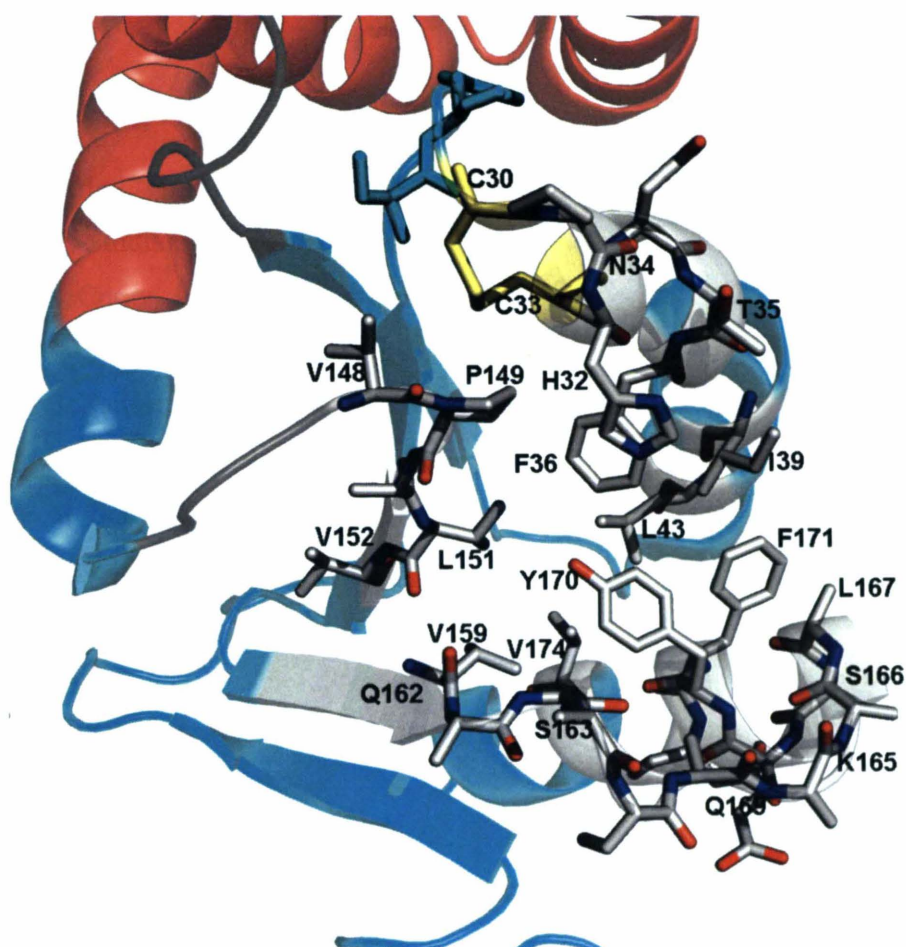


Figure S7: Hydrophobic groove and active site region of *VcDsbA*.^[28] This region also represents the fragment-binding site located by NMR experiments for the majority of *VcDsbA* small molecule binders identified in present study. Protein is shown as ribbon diagram where thioredoxin and α -helical domains are colored as cyan and red respectively. C30, C33 and other residues are shown as yellow and element colored stick models respectively.

References

- [1] J. Marley, M. Lu, C. Bracken, *J Biomol NMR* **2001**, 20, 71.
- [2] F. W. Studier, *Protein Expr Purif* **2005**, 41, 207.
- [3] a)G. Cerny, M. Teuber, *Arch Mikrobiol* **1971**, 78, 166; b)R. A. Dixon, I. Chopra, *Antimicrob Agents Chemother* **1986**, 29, 781.
- [4] M. J. Santos, J. A. Teixeira, L. R. Rodrigues, *J Chromatogr B Analyt Technol Biomed Life Sci* **2011**.
- [5] A. M. Senczuk, R. Klinke, T. Arakawa, G. Vedantham, Y. Yigzaw, *Biotechnol Bioeng* **2009**, 103, 930.
- [6] D. Pan, A. P. Hill, A. Kashou, K. A. Wilson, A. Tan-Wilson, *Anal Biochem* **2011**.
- [7] Y. Yigzaw, P. Hinckley, A. Hewig, G. Vedantham, *Curr Pharm Biotechnol* **2009**, 10, 421.
- [8]
- [9] K. E. Kover, P. Groves, J. Jimenez-Barbero, G. Batta, *J Am Chem Soc* **2007**, 129, 11579.
- [10] M. Piotto, V. Saudek, V. Sklenar, *J Biomol NMR* **1992**, 2, 661.
- [11] a)S. Grzesiek, A. Bax, *J Biomol NMR* **1995**, 6, 335; b)F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J Biomol NMR* **1995**, 6, 277.
- [12] K. D. Goddard TD, **2006**.
- [13] A. Ayed, F. A. Mulder, G. S. Yi, Y. Lu, L. E. Kay, C. H. Arrowsmith, *Nat Struct Biol* **2001**, 8, 756.
- [14] H. J. a. A. C. Motulsky, 5.0 ed., GraphPad Software Inc., , **2007**, p. San Diego CA.
- [15] a)D. Huang, A. Caflisch, *J Mol Recognit* **2010**, 23, 183; b)C. Abad-Zapatero, J. T. Metz, *Drug Discov Today* **2005**, 10, 464; c)S. Barelier, J. Pons, K. Gehring, J.-M. Lancelin, I. Krimm, *J. Med Chem* **2010**, 53, 5256.
- [16] S. Grzesiek, A. Bax, *J Biomol NMR* **1993**, 3, 185.
- [17] S. Grzesiek, Bax, A., *J Magn Reson* **1992**, 96, 432.

- [18] M. Wittekind, Mueller, L. , *J Magn Reson* **1993**, *101*, 201.
- [19] R. T. Clubb, V. Thanabal, G. Wagner, *J Magn. Reson.* **1992**, *97*, 213.
- [20] A. L. Carvalho, J. Trincao, M. J. Romao, *Methods Mol Biol* **2009**, *572*, 31.
- [21] D. Jain, V. Lamour, *Methods Mol Biol* **2010**, *673*, 129.
- [22] a)L. Potterton, S. McNicholas, E. Krissinel, J. Gruber, K. Cowtan, P. Emsley, G. N. Murshudov, S. Cohen, A. Perrakis, M. Noble, *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2288; b)E. B. Krissinel, M. D. Winn, C. C. Ballard, A. W. Ashton, P. Patel, E. A. Potterton, S. J. McNicholas, K. D. Cowtan, P. Emsley, *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2250.
- [23] Martin, *Nature* **1993**, *365*, 464.
- [24] P. D. Adams, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger, *Acta Crystallogr D Biol Crystallogr* **2002**, *58*, 1948.
- [25] a)P. Emsley, K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2126; b)P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallographica Section D* **2010**, *66*, 486.
- [26] J. C. Bardwell, K. McGovern, J. Beckwith, *Cell* **1991**, *67*, 581.
- [27] a)S. Jonda, M. Huber-Wunderlich, R. Glockshuber, E. Mossner, *EMBO J* **1999**, *18*, 3271; b)S. Sinha, P. R. Langford, J. S. Kroll, *Microbiology* **2004**, *150*, 2993.
- [28] S. H. Hu, J. A. Peek, E. Rattigan, R. K. Taylor, J. L. Martin, *J Mol Biol* **1997**, *268*, 137.

Chapter 4

Application of NMR and computational approaches in structure-assisted fragment screening to identify potential inhibitors of *VcDsbA*

4.1 Introduction

An *EcDsbA* homolog from the Gram-negative pathogen *Vibrio cholerae* is a periplasmic protein known as *VcDsbA*.⁹ This regulates the correct oxidative folding, maturation and secretion of virulence determinants such as cholera toxin (CT) and toxin-coregulated pili (TCP) in the cholera pathogen.^{26, 196} Research suggests that *VcDsbA* (MW 21.5 kDa) expression is closely related to *V. cholerae* colonization and toxin production, and in the absence of functional *VcDsbA* this pathogen is avirulent.¹⁹⁷ The importance of *VcDsbA* function for *V. cholerae* colonization and toxin formation makes this enzyme a potential target for the design of novel therapeutic agents that counteract bacterial virulence that might thereby prevent or reduce the symptoms of cholera.²⁶

The current chapter concerns the application of FBDD to investigate *VcDsbA* as an antibacterial target. In chapter 3, the fragment-based discovery of *EcDsbA* inhibitors was presented using experimental data obtained by NMR spectroscopy and X-ray crystallography. Crystallographic complexes were used as a guide to design second-generation ligands, some of which bound with improved potency, which will be described in subsequent chapters 5 and 6. Hence the combination of NMR and crystallography provided useful information for *EcDsbA* inhibitor design.

Acquiring structural information for protein-fragment complexes, usually *via* X-ray crystallography, is an important aspect and also a major limitation to the number and types of target that are amenable to FBDD.^{96a} Unlike *EcDsbA*, structural data on the

complexes of fragments bound to VcDsbA could not be obtained *via* X-ray crystallography as neither soaking nor co-crystallization yielded diffraction-quality crystals. In the absence of any experimental or crystallographic structural data it is necessary to use alternatives such as computational methods in order to generate structural models to describe protein-ligand interactions at the binding interface. In addition, any available information from X-ray crystallography (such as a structure of the *apo*-protein) or NMR (such as the binding site of the ligand) can be directly exploited in computational methods to generate models of complexes that cannot be crystallized. These computational models can be utilized for the design of improved ligands.¹⁹⁸ Secondly, their higher throughput makes computational methods potentially attractive as a way to prioritize fragments from the much larger commercially available sets both for initial screening and subsequent structural characterization.¹⁹⁸

SBDD computational methods can be divided in two categories based on the type of problem and the level of information that is desired from the calculations.¹⁹⁹ The first category includes *de novo* ligand design methods. These involve incremental and interactive combination of small fragments that are predicted to make favourable interactions with residues in the binding site to design new molecules.²⁰⁰ The second category includes molecular docking and scoring protocols.^{193, 201} These methods are based on structural information that has been observed in experimental structures of protein–ligand complexes and can be used to search for new ligands.²⁰²

An approach that has been developed for determining optimal positions of small chemical fragments and the design of elaborated compounds is based on "map-based

computational methods".²⁰³ A drug design program GRID developed by Goodford, can be used to determine diverse functional group binding sites in a target macromolecule.¹⁸⁹ In a GRID based algorithm, functional groups are represented as spherical probes and the interaction energy of the probe with the target can be evaluated at discrete points on a grid which can be defined as an enclosed box containing the site of interest. The interaction energy is evaluated using a potential energy function that is the sum of electrostatic, vdW, and hydrogen bond contributions.¹⁸⁹ This approach involves generating functionality maps that reflect particular properties of the binding site *e.g.* hydrophobicity, hydrophilicity, and charge. Ligands containing functional groups that interact with these adjacent functionality maps or 'site points' are most likely to contribute to tight protein–ligand binding.¹⁸⁹ A somewhat similar approach is the multi-copy simultaneous search (MCSS) method, which optimizes the free energy of small individual functional group replicas simultaneously in such a way as to identify the optimal positions and orientations of different functional groups (such as methanol, methyl ammonium, acetate, propane, benzene, *N*-methylacetamide and water) in the binding site of interest.²⁰⁴ Chemical fragments used for MCSS functionality maps are typically small functionalities that span a range of different possible moieties (polar, charged or hydrophobic) that could be used to build larger and more realistic molecules.

Many protein-ligand complexes contain certain regions of the binding surface that contribute a major amount to the binding free energy and are often termed as "hot spots" or druggable regions.⁹⁵ Biophysical screening of fragment libraries can expose these regions and a reasonably high "hit rate" (~0.5 - 10% of the fragments with $K_D < \text{mM}$) is

taken as an indication that targets have sites that are likely to bind drug-like ligands with high affinity.^{126a, 198, 108b} Using the FBDD approach, screening is generally followed by analogue design or fragment enlargement on the basis of available information such as the structure of the fragment or details of its interaction with the target protein. Fragment libraries typically contain hundreds of compounds, accordingly, generating analogs of the initial hits can result in there being a large number of compounds to be tested for affinity. To make this process time and cost effective, development of a computational, rather than experimental, protocol might be very useful in order to filter out unwanted ligands.²⁰⁵ The generation of molecular interaction maps of a target site can provide valuable insights into the nature of the binding site and key interactions associated with molecular recognition. Functional groups in these energetically favourable positions can then be linked together to construct novel ligands that are complementary to the binding site of the target. Alternatively, these positions can be used to modify known ligands to improve their binding affinity.²⁰⁶

The current chapter describes our approach to the identification and optimization of fragments that bind to and inhibit VcDsbA. The initial binding information was derived from ligand-based NMR experiments and computational docking data, which were analysed with the use of protein surface grid maps. Information on binding affinity and binding site location was obtained from protein-based NMR experiments and this was used to guide docking and generation of structural models of the protein-ligand complexes. Subsequently these models were used to design a series of compounds that

were synthesized (by Mr Bradley Doak, MIPS) and found to be inhibitors of VcDsbA which bound showing excellent ligand efficiency.

4.1.1 A comparison of the structural and functional features of VcDsbA and EcDsbA

As described previously (chapter 1, section 1.3.1.1), the structure of VcDsbA consists of a typical thioredoxin domain (residues 1-59 and 137-181 of the primary sequence) into which an α -helical domain is inserted (residues 60-136) (Figure 4-1). The catalytic site comprises the canonical CXXC-motif (Cys₃₀ProHisCys₃₃), which is adjacent to a *cis*-Pro residue (P₁₄₉).²⁶ The VcDsbA structure also contains a number of functionally important surface features similar to EcDsbA, including the peptide binding hydrophobic groove, hydrophobic pocket and hydrophobic patch (Figure 4-2) and an acidic patch on the opposite face to the active site disulfide.^{20a, 26} However, there are several differences in surface electrostatics compared to EcDsbA. These are more obvious in the flexible and uncharged region of the β 5/ α 7 loop and α 7 helix that contains a high proportion of hydrophobic residues and has been postulated to be the substrate binding site (Figure 4-2).^{49, 207} This groove is shortened as compared to EcDsbA and the hydrophobic pocket of VcDsbA is shallower. Moreover VcDsbA has a much less extensive acidic patch than that of EcDsbA.³⁴ In addition to the structure of oxidised VcDsbA, which was determined by X-ray crystallography, the structure of reduced VcDsbA has been solved using NMR spectroscopy by Dr James Horne (MIPS).²⁰⁸ Although the structures of oxidized and reduced VcDsbA were very similar, several clear differences were observed in the

dynamics and interdomain motions of the different redox states of the protein.²⁰⁸ These differences in the dynamics may be important factors in both the ability of oxidized DsbA to accommodate a diverse range of substrates and the more specific interaction of reduced DsbA with DsbB in the catalytic cycle of the enzyme.²⁰⁸⁻²⁰⁹

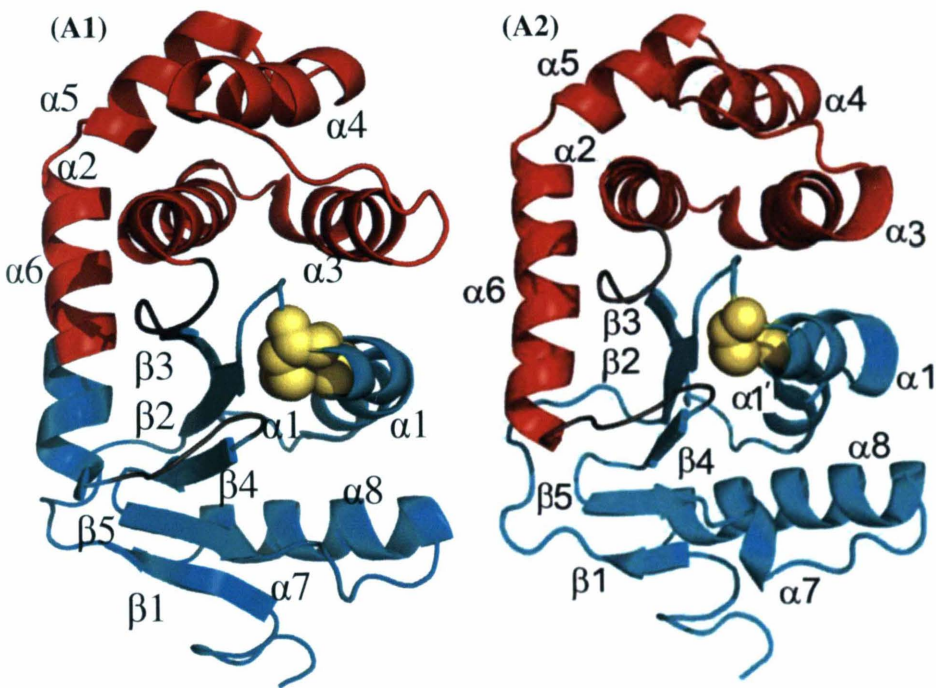


Figure 4-1: Ribbon diagrams of (A1) VcDsbA and (A2) EcDsbA. Both show the characteristic secondary structure of the DsbA family, comprised of thioredoxin (cyan) and α -helical (red) domains and the insertion points (gray). Elements of secondary structure are numbered sequentially from the *N*-terminus. The cysteine residues of the active site are shown in (yellow) CPK representation [references 26, 20a].

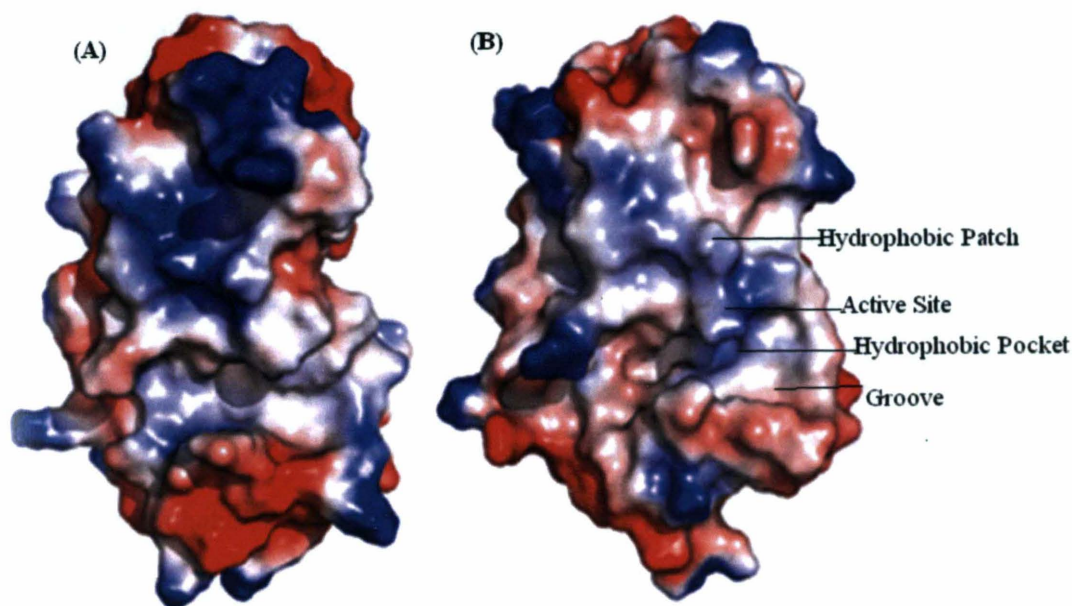


Figure 4-2: Surface electrostatics of (B1) VcDsbA and (B2) EcDsbA. Surface features are labelled for *EcDsbA* [references 26, 34].

4.1.2 Previous work

4.1.2.1 NMR-based fragment screening

In order to identify fragment hits for VcDsbA NMR-based fragment screening was conducted previously by Dr James Horne (MIPS) and Ms Yanni Chin (MIPS). The details of this experimental screening are been presented here briefly.

Saturation transfer difference NMR experiments were conducted as a primary screen on a library of 500 structurally diverse fragments. This approach identified 40 fragments from the library which were ranked initially on the basis of the strength of STD signal. These 40 primary hits were validated by recording ^{15}N - ^1H HSQC experiments and only those fragments that induced significant chemical shift perturbations in the HSQC spectra were considered as real hits. The chemical structures of the hits are shown in Figure 4-3 and were found to be structurally distinct from strongest hits obtained for *Ec*DsbA. Near complete chemical shift assignments of the $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ resonances of the backbone, aliphatic side chains and some aromatic side chains of oxidized and reduced *Vc*DsbA had been previously determined, along with other studies focused on structural and dynamic aspects of the target.¹⁷¹ These resonance assignments were used to locate the binding site of the identified *Vc*DsbA fragment hits. Dissociation constants were also determined from HSQC titrations, yielding an approximate ranking of the fragments, with affinities ranging from 80 μM to 20 μM . Most of these fragments showed major perturbations of the residues at the hydrophobic groove and active site region. At saturating concentration, the greatest perturbations were induced by benzimidazole fragment **N2** and benzophenone fragment **N6** with measured approximate K_D values of 32 μM and 20 μM , respectively.

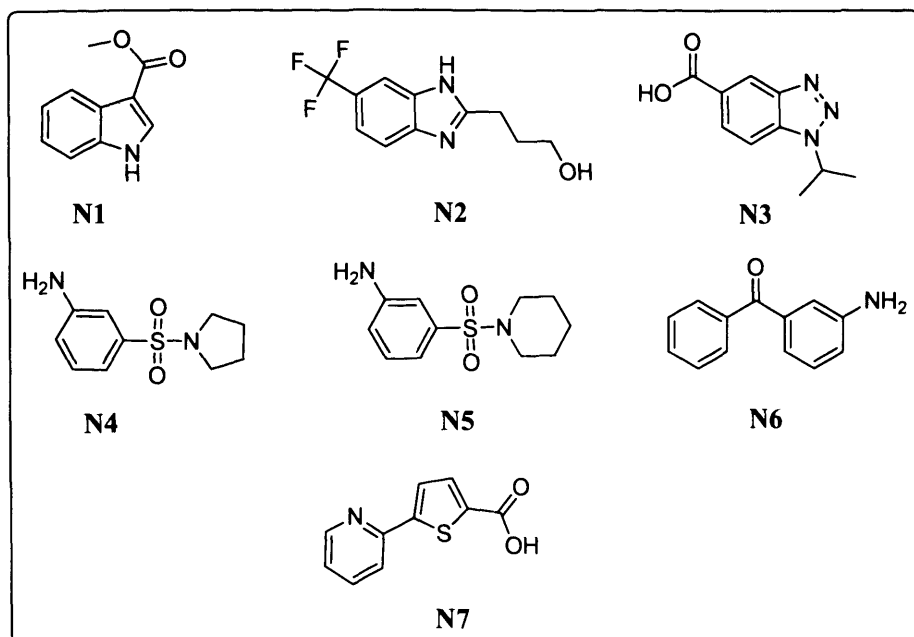


Figure 4-3: NMR-based fragment screening hits for VcDsbA

4.2 Result and Discussion

4.2.1 Comparative analysis of VcDsbA binders with EcDsbA FBS results

Despite the observation that VcDsbA and EcDsbA bound a different subset of fragments, it was found that the binding location for the VcDsbA hits was broadly similar to that of EcDsbA fragment hits (obtained from Chapter 3, 5, and 6 results). To confirm the fragment-binding specificities for these two enzymes, HSQC experiments were performed to test the ability of the VcDsbA hits to bind to EcDsbA and *vice versa*. Figure 4-3 (VcDsbA binders **N1 - N7**) and Figure 4-4 (EcDsbA binding compounds **1- 4, B1, B6, E2, E3, C4, 10** and **SE12**) show the structures of these fragments tested by HSQC experiments at 1 mM against EcDsbA and VcDsbA. End-point HSQC spectra were recorded and the fragments were ranked as strong (+++), medium (++), weak (+) or non binder (-) on the basis of a threshold applied for CSP of ≥ 0.03 ppm. This threshold number was set for this study according to the extent of strongest and weakest HSQC CSP binding data for ligands identified for EcDsbA or VcDsbA. Somewhat surprisingly most of the original VcDsbA hits were found not to bind to EcDsbA; however a subset of the EcDsbA hits (**E3, B1, B6, C4, and 10**) bound to VcDsbA causing weak or moderate shifts (Table 4-1).

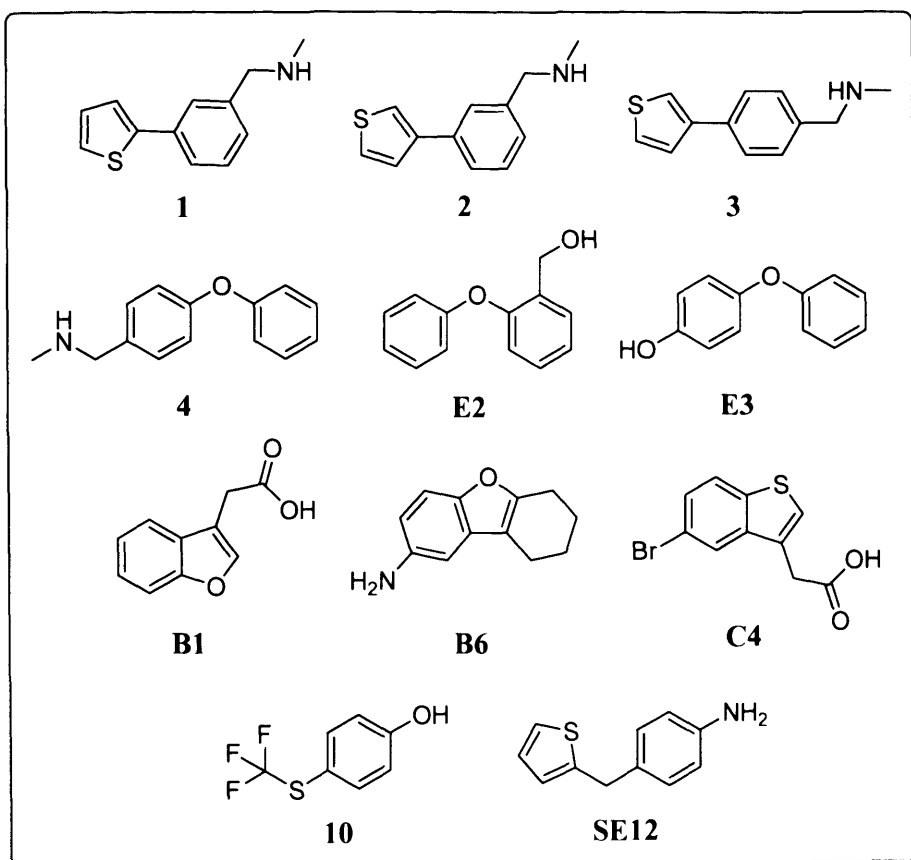


Figure 4-4: NMR based fragment screening hits of *Ec*DsbA tested for binding to *Vc*DsbA.

Table 4-1: HSQC binding for selected fragments for *EcDsbA* and *VcDsbA*

| Fragments | CSP Strength <i>EcDsbA</i> | CSP Strength <i>VcDsbA</i> | Fragments | CSP Strength <i>EcDsbA</i> | CSP Strength <i>VcDsbA</i> |
|-------------|----------------------------------|----------------------------------|-----------|----------------------------------|----------------------------------|
| 1 | +++ | - | N1 | - | ++ |
| 2 | +++ | - | N2 | - | +++ |
| 3 | +++ | - | N3 | - | ++ |
| B1 | ++ | + | N4 | - | ++ |
| B6 | +++ | ++ | N5 | - | ++ |
| C4 | +++ | ++ | N6 | - | ++ |
| E1 | +++ | - | N7 | - | ++ |
| E2 | +++ | - | | | |
| E3 | ++ | ++ | | | |
| 10 | +++ | ++ | | | |
| SE12 | +++ | - | | | |

These observations indicate that *VcDsbA* and *EcDsbA* have different specificities for fragment binding. Although structures of *EcDsbA* in complex with fragments have been obtained (Chapter 3), attempts to soak fragments into crystals of *VcDsbA* proved unsuccessful. Therefore, an alternate approach employing *in silico* computational methods, in conjunction with experimental data largely derived from NMR data, was used in an effort to probe the structures of *VcDsbA*-fragment complexes. The details of this approach are described in this chapter.

4.2.2 Computational Screening

Computational approaches have the advantage that they are rapid, however in order to generate useful structural insights they must provide an accurate description of the protein-ligand complex. To assess the reliability of different computational approaches, docking runs were undertaken where a fragment library of 500 compounds from Maybridge was docked into VcDsbA without using any experimental binding information using Glide (version 5.0; Schrodinger).²¹⁰ Before performing docking runs, the sitemap tool (Maestro, version 8.0; Schrodinger) was used to explore apparent binding sites on the protein structure using the "site finding" process.¹⁹¹ This process identifies the most likely receptor binding sites by linking together 'site points' that are most highly ranked in terms of their ability to contribute to tight protein–ligand or protein–protein binding.^{184b} These sites on VcDsbA correspond to (1) the hydrophobic patch, (2) the acidic patch (3) the hydrophobic groove adjacent to active site and (4) a region near the $\alpha 6$ helix of VcDsbA as shown in Figure 4-5. A threshold site score of 0.80 was set as the average score for the recognition of sites of interest as a 'drug-binding site'.^{185b, 190} The hydrophobic groove site scored 0.981 which was the highest score of the possible binding sites in VcDsbA, and is consistent with the NMR data.

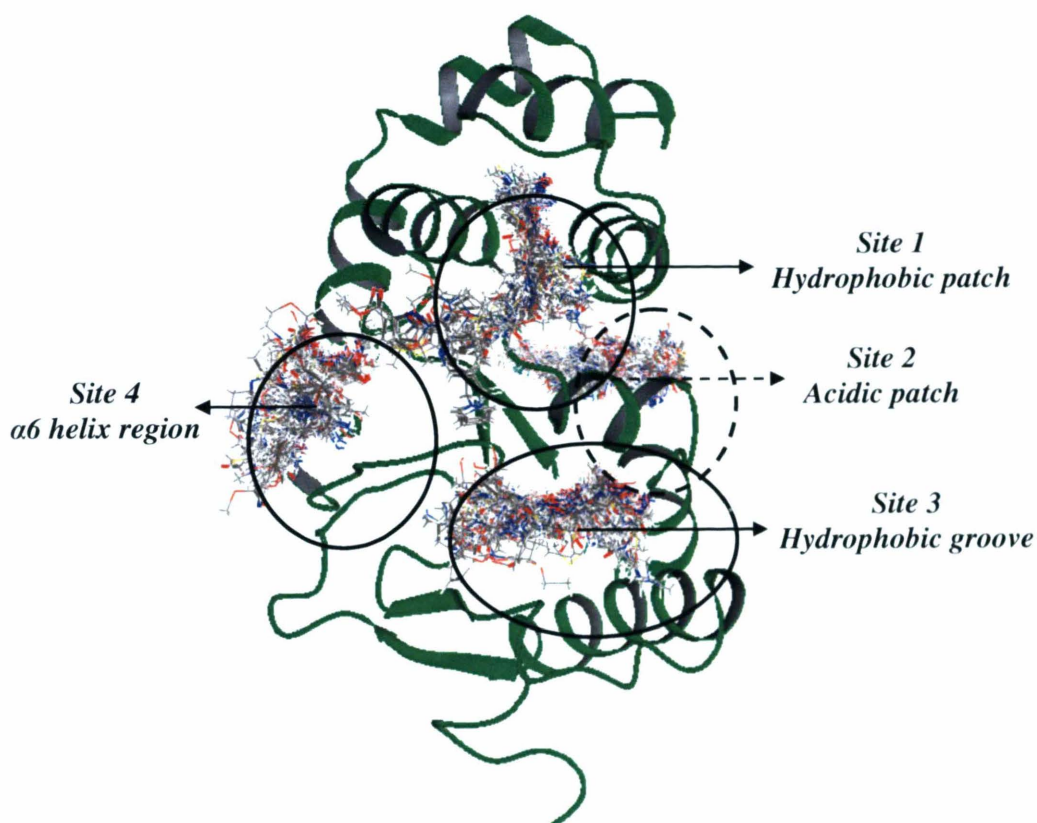


Figure 4-5: Clusters of docked fragments in the four largest different sites on the VcDsbA surface identified in the site finding process used for primary docking runs. Black arrows and circles as solid and dashed lines indicate the location of these sites on the front and rear face of the protein respectively.

Subsequently, all 500 compounds from the fragment library containing the seven binders identified in the NMR screening were docked into each of these four sites separately, using both flexible docking and the XP (extra precision) scoring options. (Notably the EcDsbA binders those has been shown in Figure 4-4 were not part of the docked

fragment library that used for VcDsbA screening). The docking solutions for the fragments were ranked using Glide XP score^{184b, 185a} to identify the top ranked fragments. Enrichment factors for docking were calculated using equation 4-1.^{181, 186}

Equation 1

$$EF \% = (\text{Hits}_{\text{sampled}} / N_{\text{sampled}}) / (\text{Hits}_{\text{total}} / N_{\text{total}})$$

If the docking performed perfectly, the compounds that bound to VcDsbA would be the highest ranked, whereas if the docking performed no better than random selection of compounds from the library, the binders would be distributed evenly - as indicated by the red line in Figure 4-6.

To be useful in the prediction of active compounds, a docking program must be able to rank the known binders higher than the non-binders (or *decoy*) compounds. Accordingly, we docked the fragment library containing seven VcDsbA binders (Figure 4-3) into the crystal structure (PDB code 1BED). The experimental data provided no evidence of binding at any site other than the hydrophobic groove, however the docking predicted that binding was possible for more than 80% fragments of the library in each pocket of VcDsbA, indicating that the docking was neither accurate nor predictive. Although the hydrophobic groove had the highest 'site score' and XP docking scores were consistently higher for the fragments docked in this pocket as compared to other sites (table 4-2), the ranking of binders within the hydrophobic groove was inconsistent with the experimental data. Figure 4-6 shows an enrichment plot of the active compounds recovered against the library ranked by XP score for the hydrophobic groove as the binding site. It can be

observed from the plot that the performance of docking was very poor in terms of identifying binders. In addition, if ranking scores were predictive of binding to VcDsbA, it might be expected that the seven compounds which are known to bind VcDsbA would have higher docking scores for the binding pocket rather than the other sites for which there is no experimental evidence of binding. We found quite similar ranges for the binders and non binders using XP docking scores (-0.2 to -6.0 kcal/mol) for each of the four binding sites. The XP docking scores for all four binding sites for the seven experiment binders (Figure 4-3) have been listed in table 4-2 below.

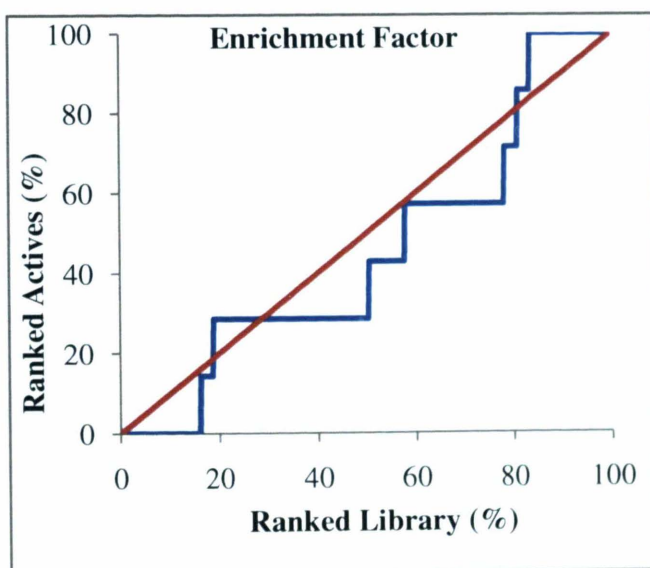


Figure 4-6: Enrichment plot of primary Glide XP docking of fragment library into hydrophobic groove of VcDsbA shown as blue line. Red line indicates the line of zero enrichment.

Table 4-2: Glide XP scores from primary docking into the: (1) Hydrophobic patch (2) Acidic Patch (3) Groove (4) $\alpha 6$ helix region, listed for the NMR based binders of VcDsbA.

| Compounds XP Score (kcal/mol) | N1 | N2 | N3 | N4 | N5 | N6 | N7 |
|-------------------------------------|-------|-------|-------|-------|-------|-------|-------|
| Site 1 | -5.07 | -5.12 | -4.89 | -5.59 | -4.01 | -4.25 | -4.56 |
| Site 2 | -4.32 | -0.18 | -2.54 | -4.72 | -4.78 | -1.42 | -2.53 |
| Site 3 | -6.57 | -6.66 | -6.13 | -5.61 | -5.53 | -5.66 | -6.09 |
| Site 4 | -3.75 | -2.98 | -3.14 | -3.12 | -3.56 | -3.84 | -2.31 |

One possible reason for the apparently poor results when docking fragments may be that the ranking system used in each of the docking protocols was optimized for larger drug-like molecules and that the parameters may therefore be less accurate when dealing with smaller fragments.^{185b, 211} Furthermore, docking programs tend to predict polar interactions more accurately and perform more effectively when the binding pose contains more than a minimum number (1-2) of hydrogen bonds (“anchor points”) since hydrophobic interactions are more difficult to model.^{184c, 185b, 211a, 212} Thus, the compounds in the fragment library which are designed to be low molecular weight and low complexity present a significant challenge to many current docking approaches.²¹³ For these reasons, it was decided to attempt to improve the performance of the docking

program in ranking the fragments by using binding information from the NMR experiments described above in the previous section 4.1.2.

4.2.3 *NMR-based docking*

NMR-guided molecular docking has been demonstrated to be a useful approach to provide structural information on protein–ligand complexes. Using data such as intermolecular NOEs²¹⁴ or chemical shift perturbations¹⁸⁷ detected in selectively labeled proteins it is possible to generate a limited number of constraints that can be used to guide the docking process.²¹⁵ These approaches require less data than is necessary for a full NMR-based structure calculation and therefore can generate details of protein ligand complexes relatively rapidly.^{125a, 187, 214} Recently, a number of approaches have been reported that successfully employ structural restraints to reject or refine the initial binding models through further computational work.^{187, 210, 215a, 216}

Validation of NMR-based docking was considered as a prerequisite, prior to carrying out these studies to investigate VcDsbA-fragment complexes. In the present case we decided to validate this method on EcDsbA fragment hits. For EcDsbA we have both high resolution crystallographic complexes and NMR data. As a consequence NMR-based docking was used to generate binding models of the fragments for which experimental data (from NMR and/or X-ray crystallography) were also available. The HSQC spectrum of EcDsbA with and without the ligand was used to map CSPs and thereby to select the likely binding site on EcDsbA, and docking calculations were performed to obtain a set

of structural models (Method section 2.9.2). Initially, only the confirmed binders were docked into *EcDsbA* using NMR-based docking. Interestingly, the top ranked fragment binding poses were found to be very similar to those observed in the crystallographic complexes of the corresponding fragments with *EcDsbA*. Figure 4-7 illustrates the examples of comparison between the NMR-based docking solutions and the crystallographic coordinates of the *EcDsbA*-fragment complexes. In the case of *EcDsbA* this approach was able to generate a reasonable agreement between most of the crystallographic poses and docked solutions with RMSD values of 1.65 Å - 2.3 Å observed for the heavy atoms of the ligands. Additionally XP scoring performed better in this case (compare to the preliminary library docking against *VcDsbA*) with higher enrichment factors upon docking of a library containing 16 *EcDsbA* experiment binders (Figure 4-8).

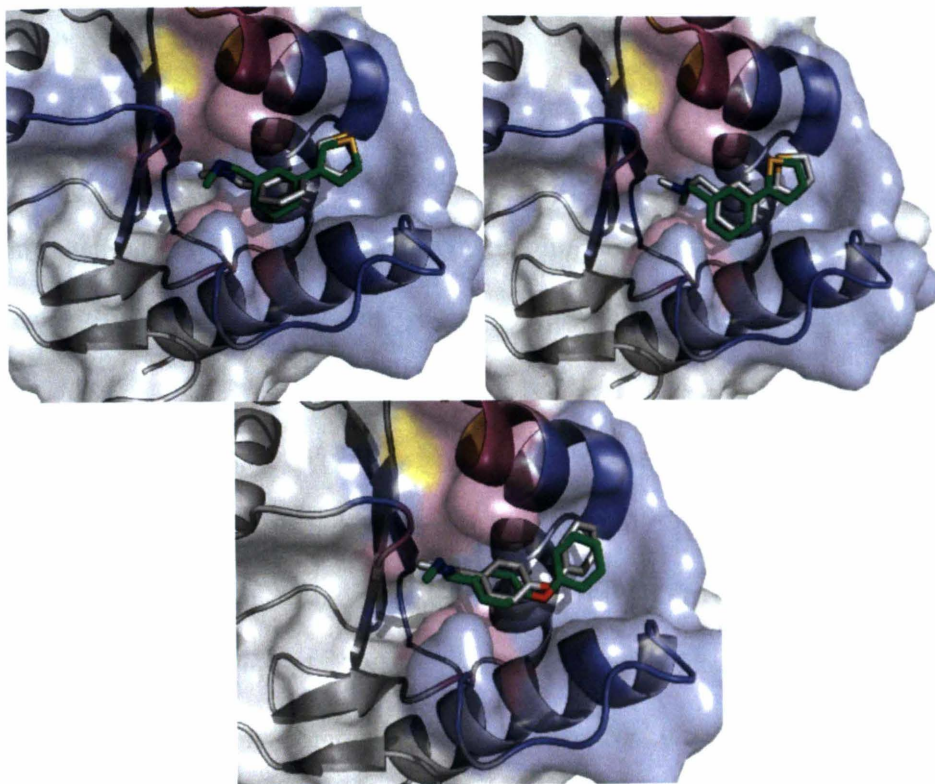


Figure 4-7: Superposition of the crystallographic coordinates of fragments (grey carbon atoms) bound to *EcDsbA*, with the binding mode of the same fragments predicted by the NMR-based docking calculations (green carbon atoms).

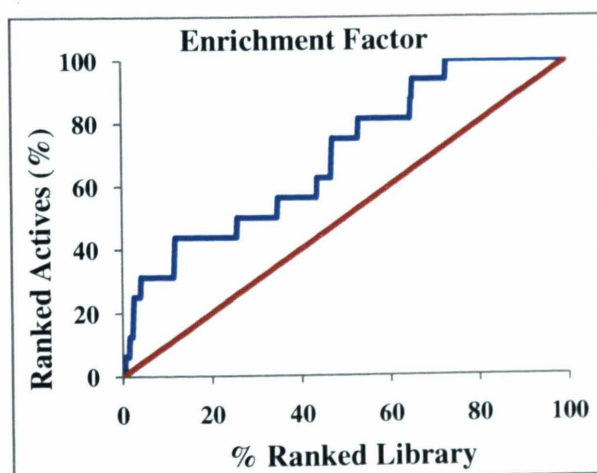


Figure 4-8: Enrichment plot (blue line) of NMR- guided Glide XP docking of the fragment library into the hydrophobic groove of *EcDsbA*. Red line indicates the line of zero enrichment.

This is a significant result as the NMR-based Glide XP docking identified higher numbers of actives compared to that of primary docking and the docked poses conformed to the experimental data. Therefore, these results with *EcDsbA* suggested that this docking method may be able to predict the binding mode of fragments more accurately. Consequently, NMR-based docking was used to predict *VcDsbA*-fragment complexes.

The hydrophobic pocket of *VcDsbA* is formed from residues of the thioredoxin domain and includes Phe36, Ile39 and Ile40, *cis*-Pro149, Val151, Val159 and Tyr170.²⁶ The groove is composed of residues Phe36, *cis*-Pro149, Gln162, Ile39, Leu158, Gly161, Val164, Leu167 and Phe171.²⁶ NMR data was obtained from a total of seven distinct pairs of protein-fragments complex and used for the docking simulation. The binding site residues which were perturbed upon fragment binding were used to generate parameters for NMR-constrained docking. Constraints were added in those regions that showed CSP in the HSQC data in order to identify most likely binding location for fragment hits. These constraints were added by locating regions or 'cells' that were derived based on the properties of selected residues. Thus a set of hydrophobic cells were located in the proximity of the most perturbed residues including Phe36, Ile39, Ile40, *cis*-Pro149, Val151, Val159, Leu158, Gly161, Gln162, and Tyr170 along with the active site residues Cys33 and Cys30 (Figure 4-9). A rectangular docking grid (Figure 4-9) was generated using the centre of the perturbed residues with the inclusion of selected hydrophobic constraints.

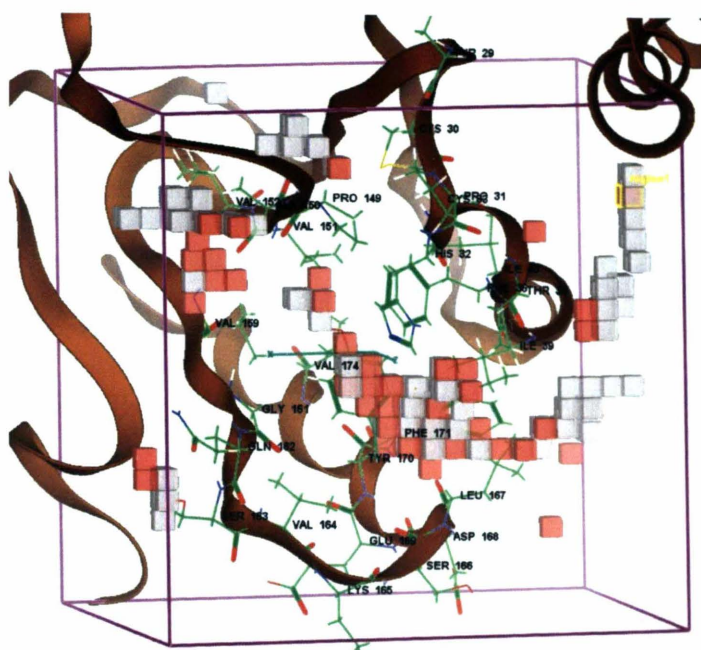


Figure 4-9: NMR binding data based grid generated for docking using CSPs ($\Delta(\text{NH}) > 0.03$ ppm) from the majority of VcDsbA hits. Pink cells in the enclosing box (length 10 Å in X, Y and Z direction) shows the selected hydrophobic cells in close proximity of the perturbed hydrophobic residues in HSQC.

The details of generating parameters for constraint docking using residues that are highly perturbed upon ligand binding, have been described in chapter 2 section 2.9.2.1. Briefly, for each complex, docking grids were generated on the VcDsbA NMR binding region of the corresponding fragment. Within this grid area, hydrophobic constraints were added in those regions that showed CSP in the HSQC spectra of VcDsbA. A number of generated docked solutions were optimized to obtain the best scoring pose by the Glide selection program.^{185a, 188} Information obtained from the selected complexes was compared with HSQC data by measuring the distances from the docked ligand to surrounding residues. Assuming that the size of the CSP observed is related to the proximity of the fragment to

a given residue, the distances measured in the docking solutions were found to be in reasonable agreement with the experimental CSP values. A simple linear relationship was considered between the magnitude of the CSP and the distance from the amino acids NH atoms and the nearest ligand atom according to the protocol described by Stark *et al.*¹⁸⁷ Figure 4-10 illustrates an example of an NMR-constrained docked structure of benzimidazole fragment **N2** which showed strong CSP upon binding to VcDsbA in HSQC-based screening.

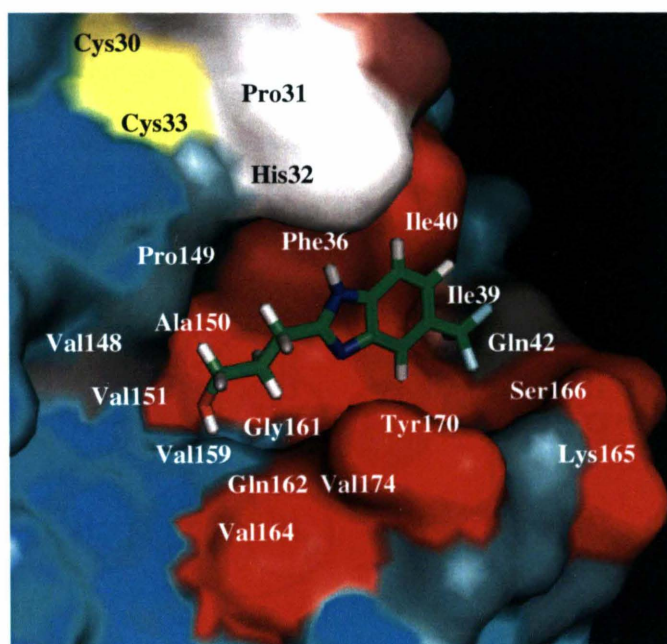


Figure 4-10: NMR-constrained docked structure of a benzimidazole fragment **N2** (shown as green carbon stick model), which showed strong CSPs upon binding to VcDsbA. VcDsbA surface (cyan) is mapped with CSPs calculated from HSQC and 3D NMR spectra, upon protein-binding of this fragment. Perturbed, active site and unassigned residues have been colored as red, yellow and white respectively (PDB code 1BED).

4.2.4 3D NMR experiments as further validation

Analysis of carbon atom chemical shifts (C_α and C_β) of a protein structure can provide a more comprehensive overview of conformational changes in the backbone due to ligand binding. As described in method section 2.4.5, the 3D NMR experiment CBCA(CO)NH correlates the NH group to the preceding C_α and C_β chemical shifts (Figure 4-11). Compared to 1H , ^{15}N shifts, ^{13}C CSP are often more diagnostic for locating the binding site as they are less affected by small changes in hydrogen bonding that can result from small conformational changes caused by ligand binding. It was envisaged that the fragment binding in conjunction with chemical shift perturbations of C_α and C_β would identify important residues in VcDsbA binding.

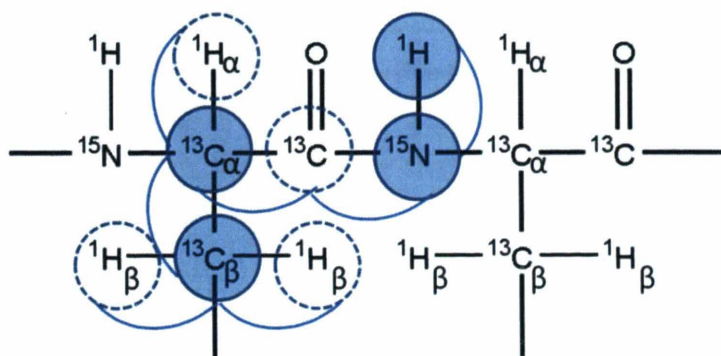


Figure 4-11: Schematic of triple resonance (3D) NMR CACB(CO)NH experiment which correlates the chemical shifts of 1H , ^{15}N , $^{13}C_{\alpha i-1}$ and $^{13}C_{\beta i-1}$.

3D NMR CACB(CO)NH experiments were performed on a $^{13}C^{15}N$ -labeled sample of VcDsbA in complex with two strong binding fragments, benzimidazole **N2** and

benzophenone **N6**, to facilitate a more detailed analysis of key interacting amino acid residues and to validate the location of binding by carbon shifts. Weighted $^{13}\text{C}\alpha^{15}\text{N}^1\text{HN}$ CSP induced by fragment binding for each residue were calculated using Equation 4-2 and are shown as a histogram analysis in Figure 4-12. (Appendices section 8.8 lists the $^{13}\text{C}\alpha^{15}\text{N}^1\text{HN}$ chemical shifts of fragments **N2** upon binning with VcDsbA).

Equation 4-2

$$\Delta\delta = \sqrt{(\delta H_{\text{complexed DsbA}} - \delta H_{\text{Apo DsbA}})^2 + 0.154(\delta N_{\text{complexed DsbA}} - \delta N_{\text{Apo DsbA}})^2 + 0.276(\delta C_{\alpha_{\text{complexed DsbA}}} - \delta C_{\alpha_{\text{Apo DsbA}}})^2}$$

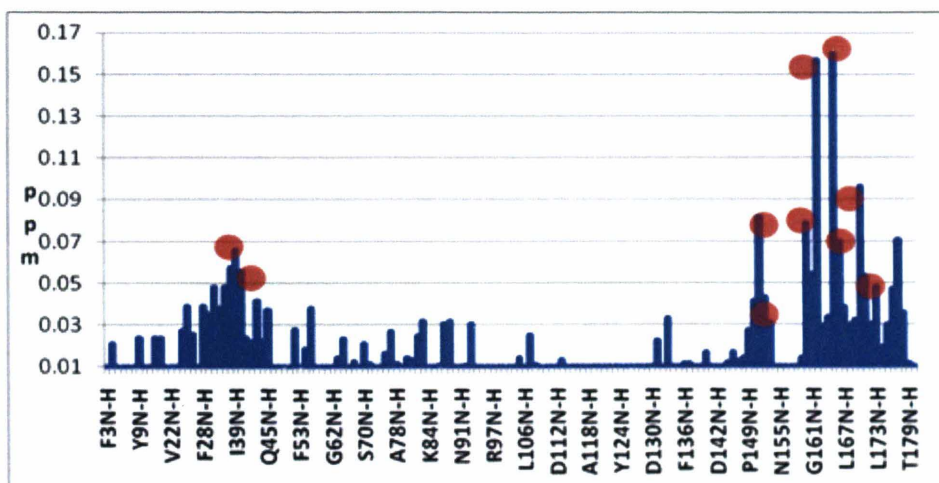


Figure 4-12: Histogram analysis from 3D NMR (CBCA(CO)NH) experiments on double labeled sample of VcDsbA in complex with the benzimidazole fragment **N2**. Red circles represent resonances closest to the docked solution of **N2** as shown in figure 4-10 of its NMR constrained docked pose.

Mapping of carbon perturbations on the VcDsbA structure showed binding of fragments in the hydrophobic groove region, near the active site, consistent with the results described above. These chemical shifts localized on the VcDsbA structure further

validated the NMR-based docked solutions for the corresponding fragments as highlighted in Figure 4-12 and 4-10. However, some additional shifts were found in regions other than the hydrophobic pocket and active site. These shifts may be the result of conformation changes in VcDsbA due to the flexibility of the groove.

After generating NMR optimized parameters to dock these fragment hits to VcDsbA, the whole fragment library was docked using average perturbations of the binders to define the binding site with advanced docking settings, in an attempt to obtain more reliable binding locations and energy optimized poses (Figure 4-13). The Glide XP score for NMR-based docking was matched again with the experimental list (Table 4-3). However, as shown in the enrichment plot in Figure 4-14 NMR guided docking performed not much better than the standard protocol.

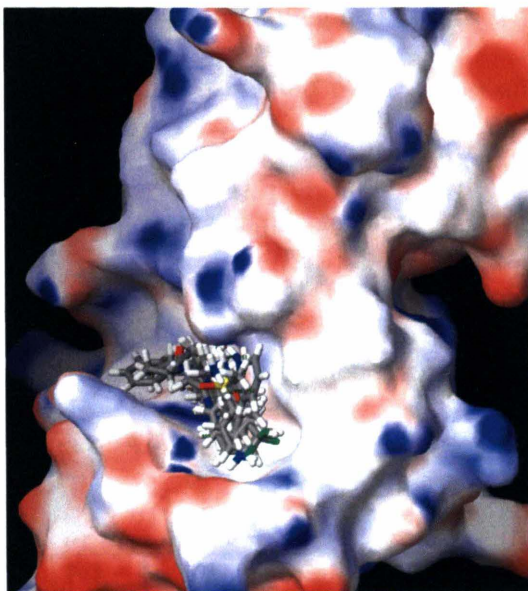


Figure 4-13: NMR guided docked solution of top 5% ranked fragments filtered by XP docking mode from the fragment library as potential binders in the VcDsbA hydrophobic groove. Protein and fragments are shown as electrostatic surface and element coloured stick representation respectively.

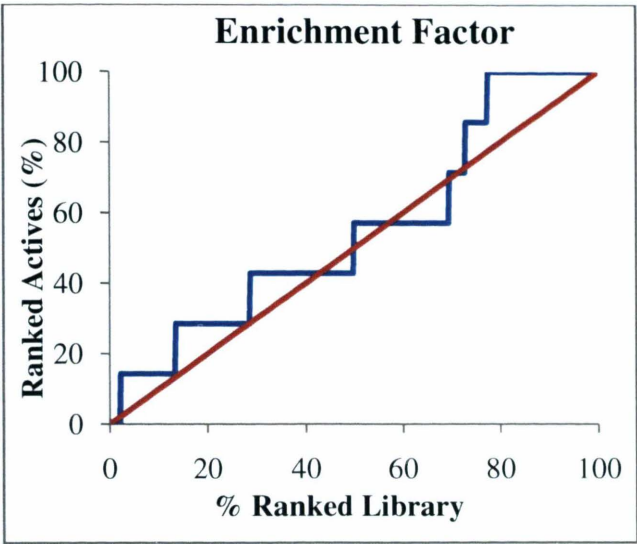


Figure 4- 14: Enrichment plot (blue line) of NMR guided Glide XP docking of fragment library into hydrophobic groove of *VcDsbA*. Red line indicates the line of zero enrichment.

A possible reason for the limitations in ranking real ligands with respect to non ligands is that scoring functions are usually estimated only with positive data, i.e, information relating to known binders. Negative data or information relating to non ligands or inactive molecules is rarely integrated while developing these scoring functions.²¹¹

Table 4-3: Glide XP score from NMR guided docking for fragments hits bound to hydrophobic groove of *VcDsbA*.

| Compounds | N1 | N2 | N3 | N4 | N5 | N6 | N7 |
|--|-------|-------|-------|-------|-------|-------|-------|
| Hydrophobic groove binding XP score (Kcal/mol) | -5.07 | -5.12 | -4.89 | -5.59 | -4.01 | -4.25 | -4.56 |

4.2.4 *Surface grid map generation*

Although NMR-based constraint docking generated the most probable binding solutions for the identified binders, they did not appear to be capable of differentiating binders and non-binders of VcDsbA from the library. Therefore a further screening approach was applied in an effort to select the binders from the fragment library using the molecular interaction features of the binding site. We carried out screening of the docked fragment library using receptor surface grid maps to envisage preferred locations of ligand atoms in the VcDsbA site.

The proposed fragment binding site, as determined by NMR-constrained docking, of VcDsbA was analysed by generating surface grid maps in this region, as described in methods (Chapter 2). These grid maps partitioned the accessible space in this site into three hydrophobic regions with an isosurface value (the value at which the maps are contoured) of -0.5 kcal/mol and four hydrophilic regions with an isosurface value of -6.0 kcal/mol while the remaining regions had mixed properties or neutral electrostatics (Figure 4-15). The hydrophilic map can be further partitioned into separate hydrogen-bond donor and acceptor maps.¹⁹⁰ Hydrophobic and hydrophilic regions were marked by surface contours that enclosed the region in question. The "neither" or mixed regions, in contrast, were implicit; these are simply regions that are accessible to the ligand but are not marked as being either hydrophobic or hydrophilic. The extent and shape of these maps depends upon the properties of the closest amino acid residues.²¹⁷ If the polar region or hydrophilic groups of a ligand interacts with the hydrophilic grid map, and the

non-polar region or hydrophobic groups of a ligand interacts with the hydrophobic grid map (i.e. similar interactions), these interactions are deemed as favourable interactions for ligand occupancy and binding (Figure 4-16). Dissimilar type of interactions with the surface grid maps are considered to be unfavourable interactions (Figure 4-16).

The NMR-derive binding location was used as the site for docking all of the 500 fragments, which subsequently were manually, scored using the hydrophobic and hydrophilic surface grid maps. In this scoring process fragments were separated, depending on whether they showed all favourable or at least one unfavourable interaction with surface grids.

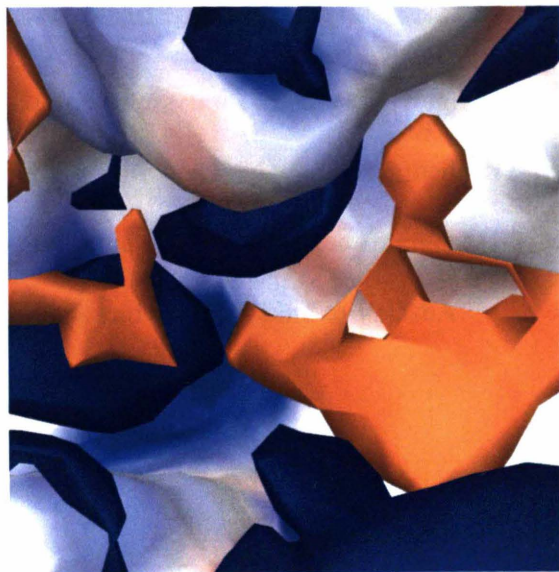


Figure 4-15: Surface grid maps of the NMR identified fragment binding region on the VcDsbA structure shown as an electrostatic solid surface. Hydrophobic and hydrophilic grid maps have been colored as orange and blue respectively.

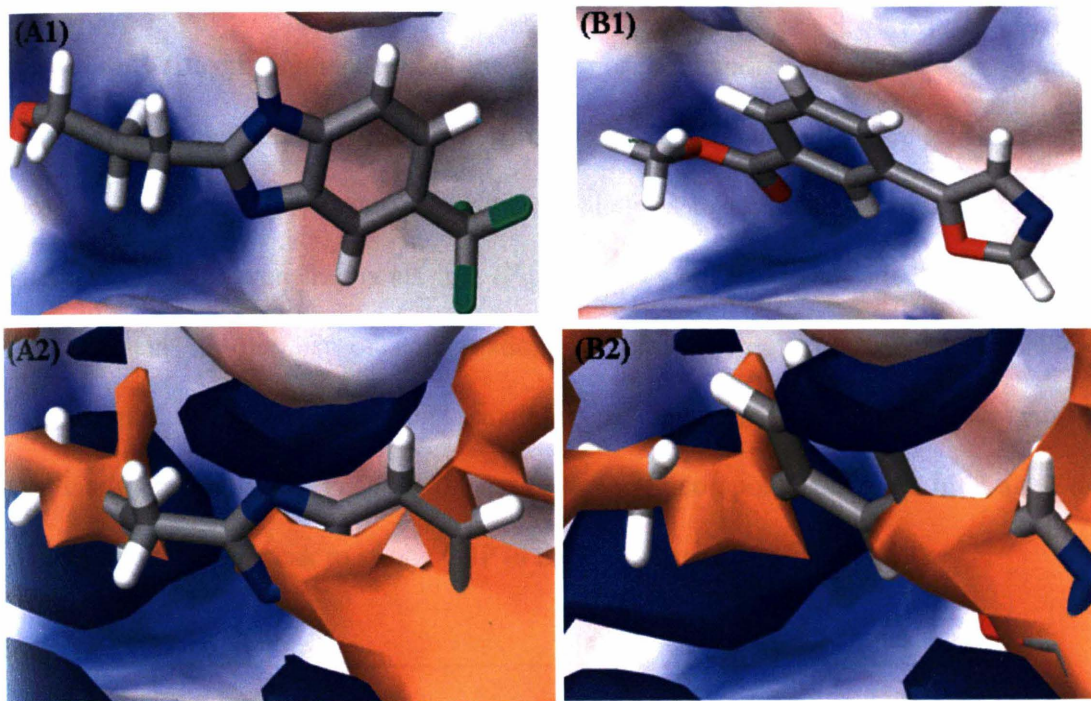


Figure 4-16: Examples of interactions of docked fragments with surface grid maps of VcDsbA fragment binding site. **(A1)** docked fragment **N2** (real binder) and **(A2)** its favourable interaction with maps. **(B1)** a docked fragment **(B2)** which shows unfavourable interactions with grid maps and does not bind to VcDsbA.

4.2.7 Identification of novel VcDsbA hits

In total only fifteen fragments out of the library of five hundred fragments were found to fit in a favourable manner (i.e. no unfavourable interactions) using surface grid maps, and these were selected for further experimental analysis to test their binding to VcDsbA. Interestingly, all seven fragments that (Figure 4-3) were previously identified by NMR were present in this list of computationally selected fragments. End point HSQC experiments were performed on remaining computationally selected fragments to evaluate their binding to VcDsbA. Seven of the eight compounds showed binding to the protein by inducing weak to strong CSP in the HSQC spectra (Figure 4-17).

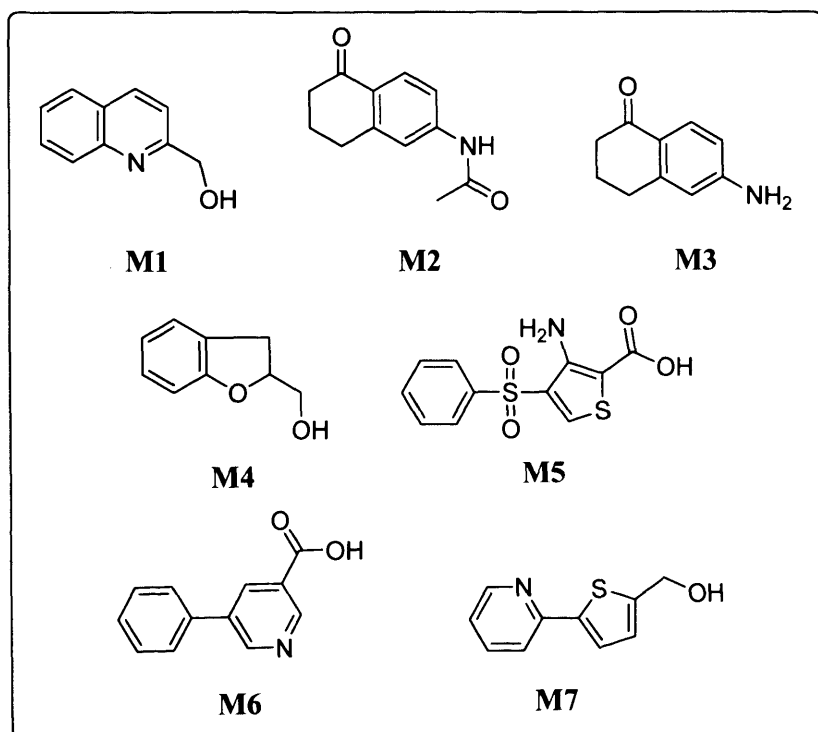


Figure 4-17: Additional fragments selected using molecular modelling and identified as VcDsbA binders by NMR HSQC experiments.

In total, fourteen fragments were identified as VcDsbA binders in the experimental validation of the grid map based screening. It was observed that each binder shows at least one hydrophobic and one hydrophilic type of interaction in the binding pocket. Fragments were rejected for the presence of at least one unfavourable interaction. Experimental validation was also performed for non-binders as a negative control. Five fragments showing predicted unfavourable binding were randomly selected as negative controls (false binders) for the NMR–HSQC experiment. All of them were found to be as non-binders of VcDsbA and one example is shown in Figure 4-16 (B1). In this way, manual screening of the fragment library using surface grid maps identified binders of VcDsbA with an accuracy of > 90 % for virtual selection. Additionally, some VcDsbA hits that were discarded in STD experiments as false negatives (due to anti phase spectra or competitive binding *etc.*) were selected using the site map based method.

These hits consist of 6,5-fused bicyclic, 6,6 or 6,5 ring structures and a number of them were quite similar to the strongest hits obtained from NMR screening of VcDsbA. Additionally one new class of scaffolds with a 6,6-fused bicyclic ring structure was identified as strong binders of VcDsbA. End point ^1H - ^{15}N HSQC experiment on these fragments confirmed their binding location in the hydrophobic groove region with a distinct pattern of CSP maps. In accord with the NMR based original hits, chemical shift mapping suggested that these fragments bind to the hydrophobic groove and interact with the active site region (Figure 4-18). K_D values were also calculated by HSQC titration that yielded an approximate ranking of the fragments, with affinities ranging from 15 μM to 80 μM .

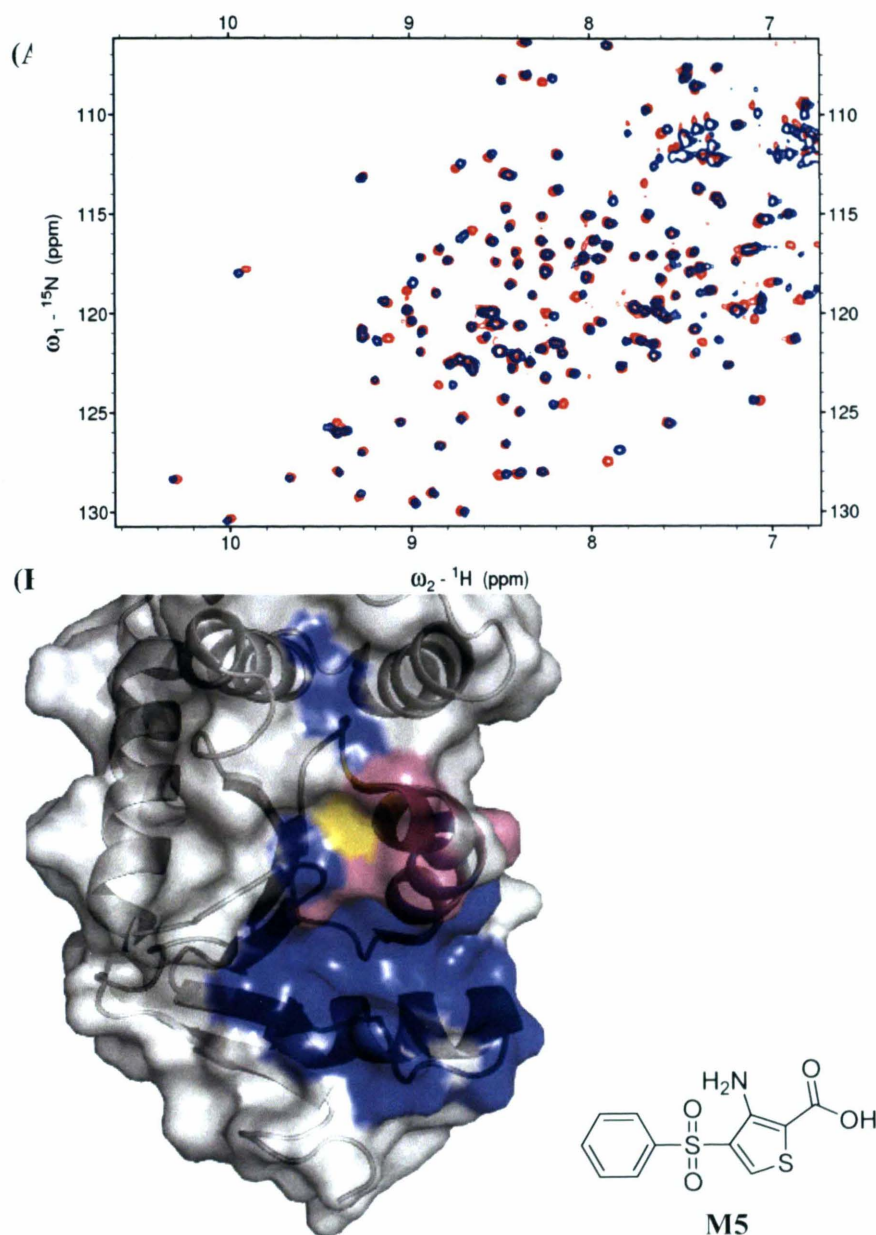


Figure 4-18: (A) ^1H - ^{15}N HSQC spectrum of VcDsbA (100 μM) (red) overlaid with the spectrum of VcDsbA in complex with phenylsulfonyl thiophene **M5** (blue) at 1mM. (B) VcDsbA surface mapped with CSPs induced, upon protein-binding of **M5**. Higher chemical shift perturbations and unassigned residues have been coloured as blue and green respectively on the white surface representation of the VcDsbA structure (PDB code 1BED).

In most cases, the peaks showing the greatest magnitude of CSP in the groove included Gln162, Ser163, Lys165 Ser166 and Leu167. Residues perturbed from the hydrophobic pocket were Phe36, Ile39, Ile40, Val151, Val159 and Tyr170. A few of the identified fragments induced perturbations in those residues of the hydrophobic patch which connect the thioredoxin and helical domains (Figure 4-18). These shifts might arise from conformational changes rather than binding in close proximity. Please refer Appendices section 8.8 for the HSQC chemical shifts details of strong binding fragments **M2 M3** and **M5** upon binding with VcDsbA).

The active site of VcDsbA lies in a cleft at the interface of the two domains and comprises a highly conserved primary sequence motif Cys30-Pro31-His32-Cys33 at the N-terminal end of helix α 1. In its oxidized form, a disulfide bond links Cys30 and Cys33. Interestingly, no peak was observed in the ^{15}N -HSQC spectrum for His32 either in the presence or absence of fragments, which is consistent with this residue undergoing conformational exchange. The flexibility of the binding site has previously been suggested to be important in substrate recognition and catalysis in DsbA enzymes.^{34, 208} Most of the fragments also induced weak to strong perturbations in the exposed active site residues, Cys30 and Cys33, which is encouraging as this suggests that the fragments may be close enough to the active site to perturb the DsbA catalytic cycle. The CSP mapped on the VcDsbA structure clearly show that the fragments bind in close proximity to the catalytic site in the hydrophobic groove (Figure 4-18).

4.2.8 *Inhibition analysis of VcDsbA hits*

The NMR-based fragment screening and computational approaches described above have successfully identified fragments that bind to the hydrophobic groove of VcDsbA. Since the fragments identified appear to bind in functionally important regions of VcDsbA, this suggested that these fragments might be inhibitors of VcDsbA activity. In an attempt to assess inhibition, we sought an *in vivo* assay. Several *in vivo* bacterial assays have been described previously to monitor the DsbA dependant phenotypes and study the functions of this enzyme (Chapter 1, section 1.4). In the current study, some of these have been modified and employed, and one such assay is described here that was used to measure the inhibitory activity of identified VcDsbA fragment hits.

For many bacterial pathogens including *E. coli* and *V. cholera*, motility is a crucial phenotype for virulence as it allows the infecting population to spread within the host as well as enhance colonization of bacteria.²¹⁸ This phenotype is regulated by a cell organelle called the flagellar motor and the periplasmic ring of the motor is composed of FlgI protein subunits. Importantly DsbA is essential for the correct folding and activity of FlgI.¹⁵⁴ Bacteria deficient in a functional DsbA display a non-motile phenotype since the oxidative folding of the protein FlgI is impaired.^{9, 219} We explored the motility of wild type *E. coli* as a direct measure of DsbA dependent folding to determine the motility inhibition by using HSQC identified ligands of EcDsbA as described in Chapter 3.

VcDsbA and *EcDsbA* share relatively low sequence similarity (<40%), yet show conserved surface features around their active sites which have been proposed to form the substrate-binding site. Several experiments have demonstrated that the functions of *EcDsbA* can be complemented by *VcDsbA*.^{19c, 35, 49} Therefore, a complementation experiment was performed using three strains of *E. coli* bacteria as listed below:

-JCB570: Wild-type *E. coli* with DsbA normal function

-JCB571: Mutated *E. coli* with DsbA gene deactivated (*dsbA*⁻)

-p*VcDsbA*: JCB571 transformed with a plasmid which expresses *VcDsbA*.

The results were evaluated by measuring the diameter of swarming or the motility zone produced by the bacteria within the incubation period. The JCB571 (*dsbA*⁻) strain of *E. coli* was unable to swarm on the soft agar plate while the JCB570 (wild-type) *E. coli* strain was motile under the same conditions. The -p*VcDsbA* mutants successfully restored the motility phenotype to JCB571 and it was fully comparable to the wild-type *E. coli* cells (Figure 4-19).

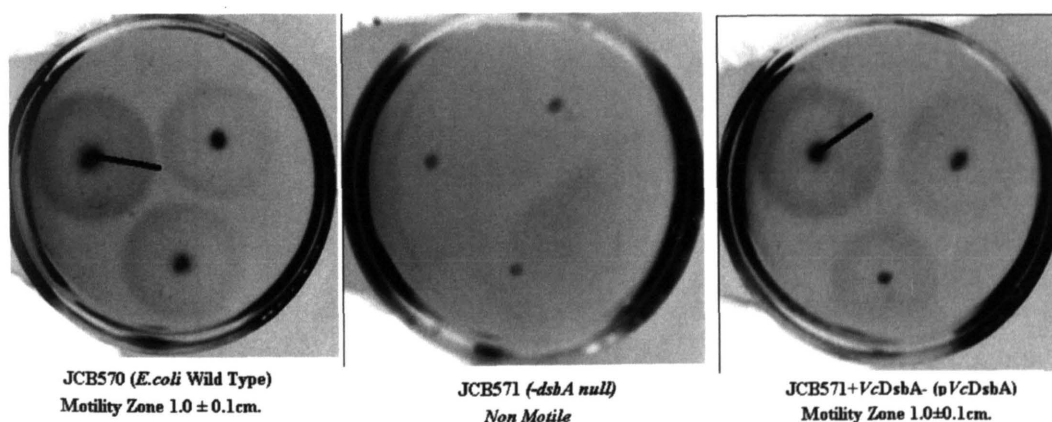


Figure 4-19: Motility of *E. coli* wild type and mutant strains on soft agar with cysteine free minimal media.

The inhibition effect of the VcDsbA hits **N1-N7** and **M1-M2** (Figure 4-3 and 4-12) was determined by undertaking the motility assay in the presence of a single concentration of the compound (500 μM) and measuring the % inhibition of motility for the -pVcDsbA *E.coli* mutants. For those compounds that displayed 100% inhibition of motility at 500 μM , an EC_{50} was measured as shown in Figure 4-20 for the dose response plot for VcDsbA binder **M5**. Two fragments **N6** and **M1** showed poor solubility in minimal media used for the motility assay and were not tested further. Several of the tested fragments showed inhibition of swarming capability of -pVcDsbA *E.coli* mutants. Fragments **M2**, **M3**, **M5** and **N2** showed maximum inhibition in the low μM range and notably benzimidazole **N2** and thiophene **M5** were found to be the most active inhibitors with EC_{50} values of 8 μM and 38 μM respectively. The summary of inhibition assay results along with HSQC binding strength for the tested compounds is listed in Table 4-4.

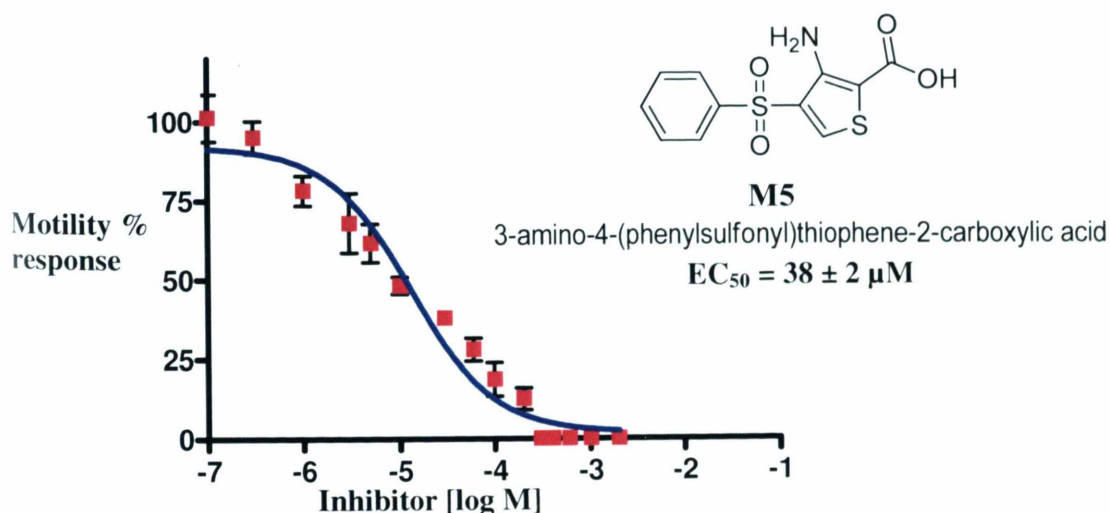
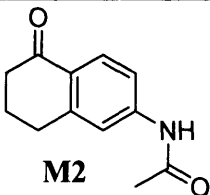
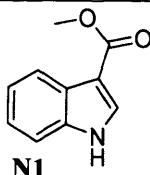
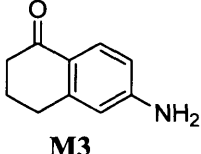
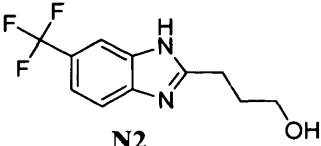
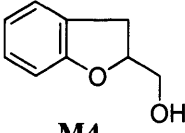
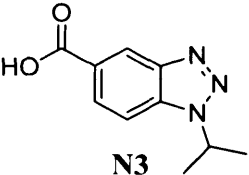
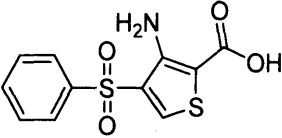
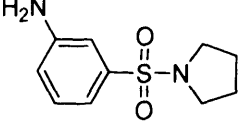
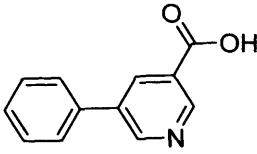
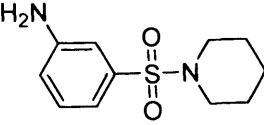
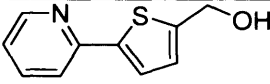
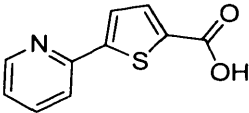


Figure 4-20: Motility inhibition of pVcDsbA *E.coli* mutants shown as sigmoidal dose response plot obtained from the titration of a fragment hit **M5** identified by NMR guided computational screening.

Table 4-4: Chemical structures, intensity of CSP observed and % inhibition of motility of *pVcDsbA* mutants at 500 μ M of *VcDsbA* binding fragments.

| COMPOUND | PROFILE | COMPOUND | PROFILE |
|---|-------------|---|-------------|
|  M2 | +++ 100% |  N1 | ++ 50% |
|  M3 | ++ 100% |  N2 | +++ 100% |
|  M4 | + 0% |  N3 | ++ 80% |
|  M5 | +++ 100% |  N4 | ++ 0% |
|  M6 | ++ 40% |  N5 | ++ 0% |
|  M7 | - 0% |  N7 | ++ 0% |

4.2.9 *Application of Grid Maps*

4.2.9.1 **Example 1**

The concordance between the computational results and the experimental data indicates that the site model is quite reasonable; hence we sought to apply the model to the extension of fragments into larger molecules with higher affinity. Apparently this empirical screening strategy follows one of the basic principles of FBDD by evaluating the druggability of protein surfaces and identifying sites that interacts with ligands.^{185b, 190}

An example is presented here to show the further use of the site model as described above in the case of two fragment hits identified by NMR-guided computational screening. Dihydronaphthalenone fragments **M2** and **M3** are two closely related fragments, where **M2** has an additional acetyl group attached to the amine. When docked poses of these two fragments were fitted to the mapped binding site they showed similar hydrophilic interactions. Acetamide **M2** showed an extra hydrophobic interaction due to the presence of the acetyl group as shown in Figure 4-21. Additionally the interactions within a 5 Å radius of these fragments were classified and compared with HSQC CSP (Figure 4-22). These results were found to correlate well with their NMR data, where acetamide **M2** showed additional interactions, inducing larger shifts and a number of extra perturbed residues as compared to **M3** (Figure 4-23).

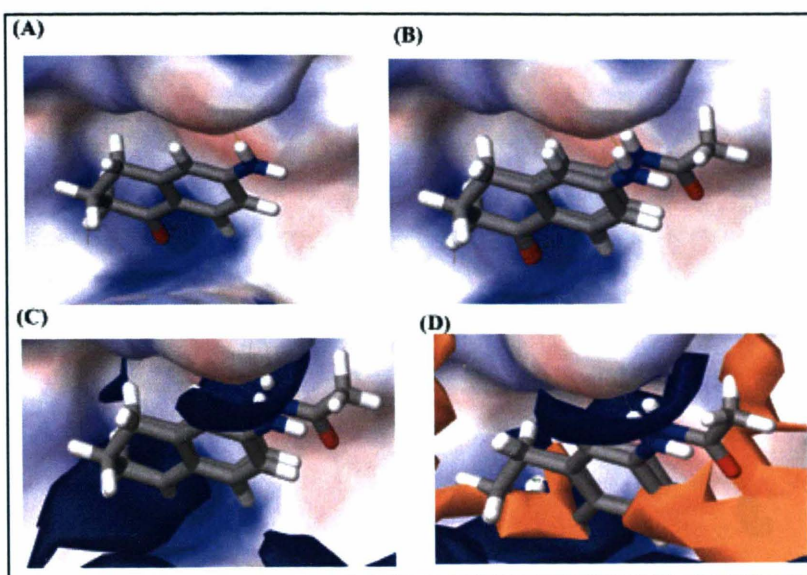


Figure 4-21. NMR-constrained docked pose of dihydronaphthalenone fragments (A) **M3** superimposed (B) on **M2** along with interaction with (C) hydrophilic and (D) hydrophobic surface grid maps of VcDsbA binding pocket.

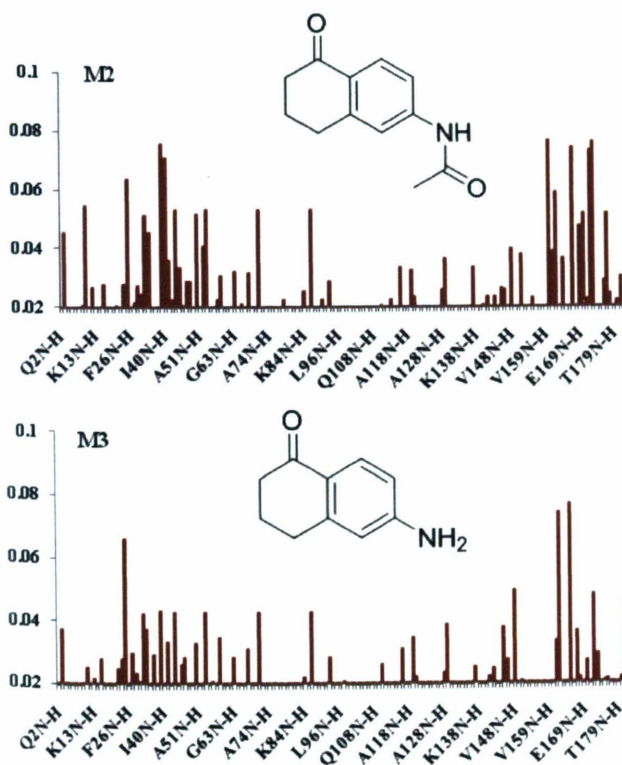


Figure 4-22: CSP induced in ^1H - ^{15}N HSQC spectra of VcDsbA upon the addition of computationally identified hits dihydronaphthalenone fragments **M2** & **M3**.

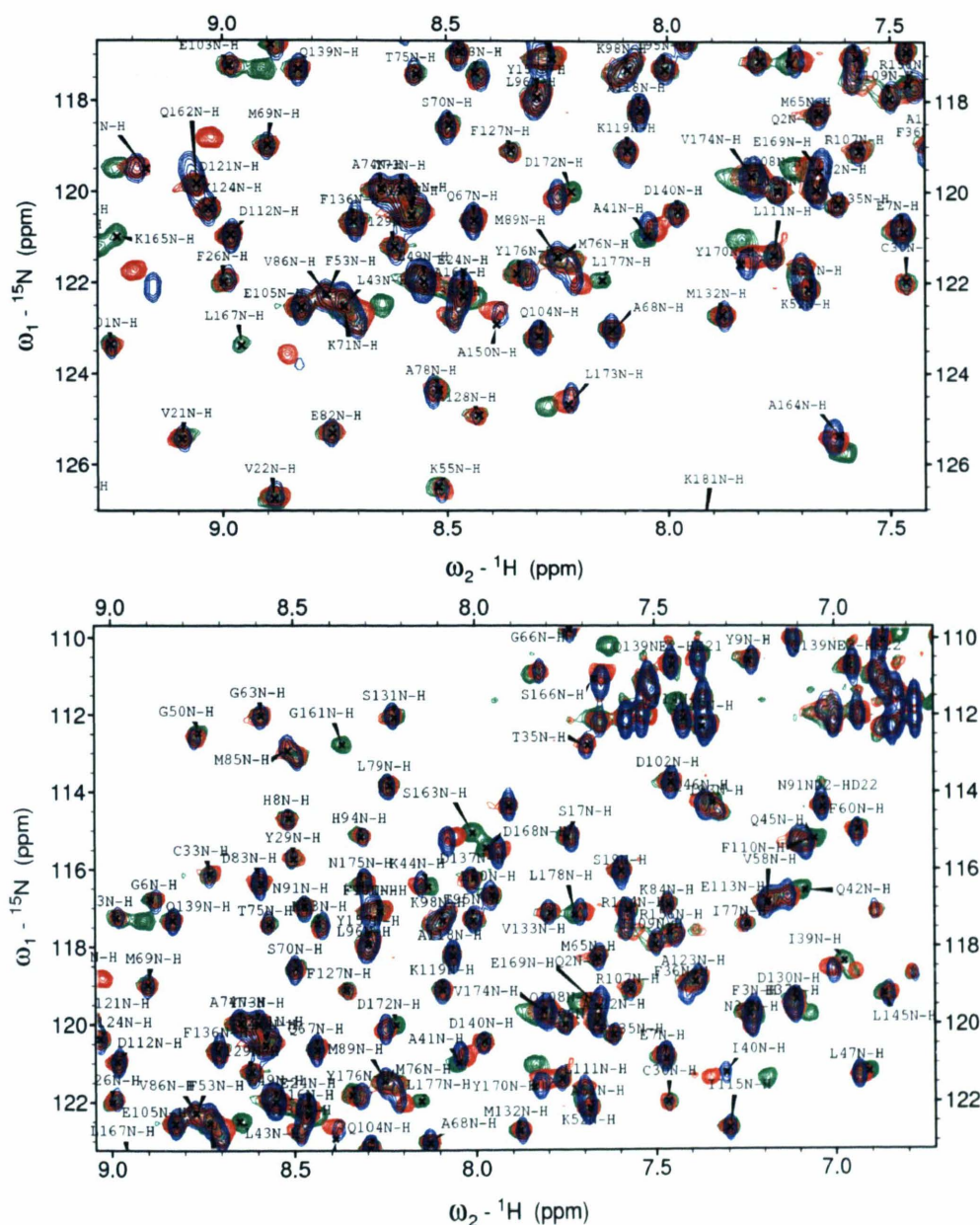


Figure 4-23: Sections of the ^1H - ^{15}N HSQC spectrum of VcDsbA (100 μM) (blue) overlaid with the spectra of VcDsbA in complex (individually) with dihydronaphthalene fragments **M2** (green) and **M3** (red) at 1 mM concentration. Circled resonances show a set of additional or larger CSP in binding site residues of VcDsbA spectra in presence of **M2** as compared to shifts induced by **M3**.

4.2.9.2 Example 2

Design and synthesis of analogues of amino benzophenone fragment: The concept of ligand efficiency (LE) can be used to assess the quality of initial screening hits and also the quality of the leads as they are optimized. LE values higher than $-1.25 \text{ kJmol}^{-1} \text{ HAC}^{-1}$ are considered a good starting point for the hit-to-lead development process based on the desired molecular weight and affinity of the final drug-like compounds.^{169b, 220} Our results show that the benzophenone fragment **N6** binds to the VcDsbA groove with a LE value of $-1.72 \text{ kJmol}^{-1} \text{ HAC}^{-1}$, suggesting this fragment is a good starting point to develop more potent ligands. Furthermore, analysis of the NMR-based docked solution of **N6** suggested a potential strategy for optimization of binding to VcDsbA. Figure 4-24 illustrates two areas in which it might be possible to improve the interaction of the benzophenone fragment with VcDsbA. Firstly, there is a significant region of the groove around the fragment in the direction of favourable interaction sites in the hydrophobic and hydrophilic maps that is not occupied. It was rationalised that by functionalising the 2', 3' and 4' positions of the phenyl ring with various groups it may be possible to more completely fill this region and thereby increase the binding affinity. Secondly, reduction of the ketone gives the alcohol which can potentially make a polar contact to the binding site. To test this hypothesis based on NMR-based docked solution and grid maps, a fragment growth approach was explored.

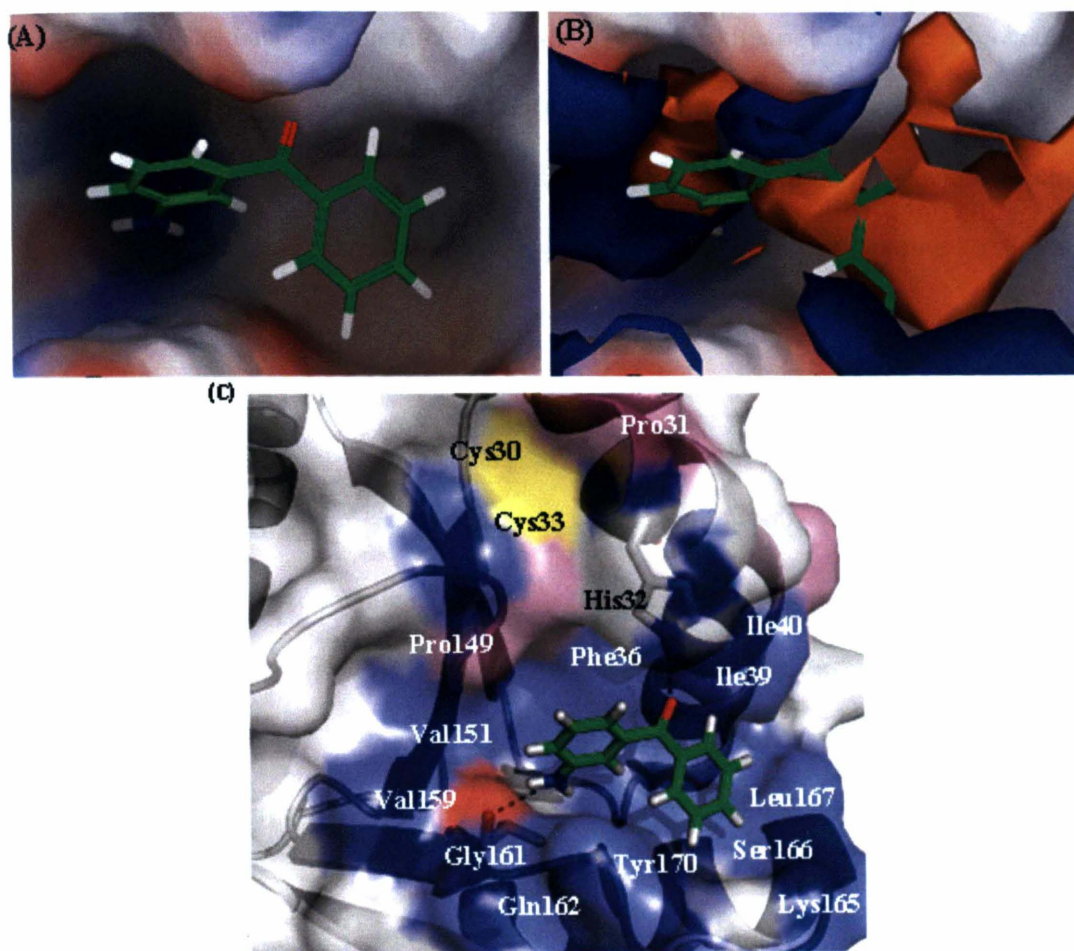


Figure 4-24. (A) NMR constraint based docked pose of benzophenone fragment **N6** (green carbon stick model) in the hydrophobic pocket of VcDsbA (electrostatic surface) and (B) its interactions within surface grid maps. (C) NMR constraint based docked solution of **N6** bound to VcDsbA (PDB code 1BED), mapped with CSPs calculated from HSQC and 3D NMR spectra, induced from the binding of this fragment. Chemical shifts, active site, unassigned residues have been colored as blue, yellow, and pink, respectively, and hydrogen bonds are shown as dotted black lines.

A number of analogues of the amino benzophenone moiety were rationally designed and modelled on the basis of their compatibility with grid maps and ease of chemical synthesis. We introduced mainly hydrophobic groups into the newly designed analogues. These were docked into the VcDsbA with similar settings used to dock the parent compound and many of them showed favourable interactions for binding to VcDsbA surface grid maps.

Consequently the synthesis of a small series of aminobenzophenone analogues was undertaken by Mr Bradley Doak, MIPS. The chemical synthesis was started from the initial hit 3-aminobenzophenone (**N6**) as well as the analogous 2- and 4-substituted aminobenzophenone (**4**, **16**) as parent compounds (Figure 4-25). 2- and 4-aminobenzophenone (**4**, **16**) were found to bind to VcDsbA and also exhibited good binding affinity (Figure 4-25). In total, 18 compounds were synthesized with *N* substituted acetyl-, benzoyl- and *p*-toluenesulfonyl-amino groups in the 2-, 3- and 4-positions as well as the related compounds generated by reduction of the ketone of all derivatives to give the corresponding alcohol.

Binding analysis of the benzophenone series was achieved through HSQC titrations of the compounds using ¹⁵N-labeled VcDsbA in a similar manner to previous experiments. The Table 4-5 shows the chemical structure of the characterized amino-benzophenone analogues. The compound numbers along with substituted functional groups and apparent *K_D* values from NMR-HSQC titrations along with calculated LE values are listed.

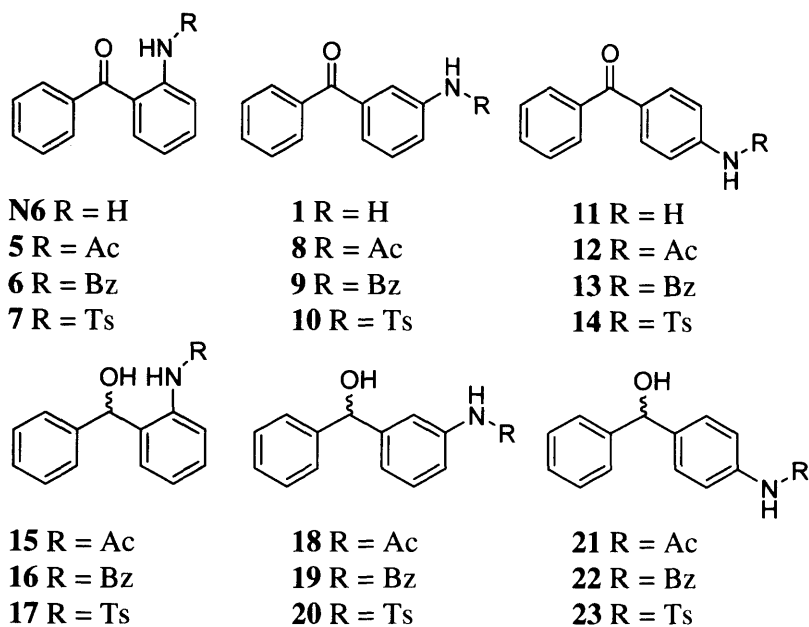
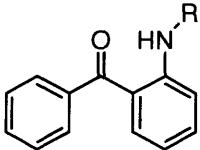
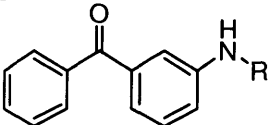
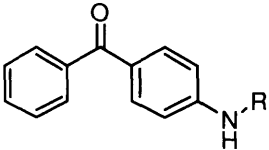
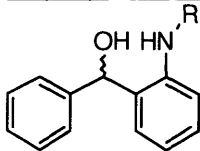
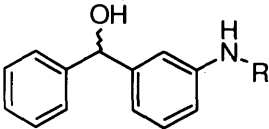
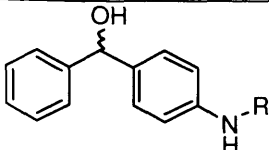


Figure 4-25: Structure of synthesised amino-benzophenone analogues

The observed HSQC indicated a binding site located below the active site in the hydrophobic groove, consistent with that seen for the parent compounds (as shown for fragment **N6** in Figure 4-24 (C)). Most of these compounds were not soluble enough in the previously used buffer conditions for the screening the library for HSQC experiments. Notably, due to poor solubility not all compounds were screened and the buffer system was adjusted to include 5% DMSO to aid solubility. Therefore a direct comparison of these compounds to the other hits, which were characterised in a different buffer system, was not undertaken. Additionally inhibition analysis was not feasible for this series due to poor solubility in assay buffer. Therefore further investigation of the benzophenone fragment was abandoned.

Table 4-5: K_D and ligand efficiency of screened amino-benzophenone analogues

| Compounds | | K_D (μM) | LE ($\text{kJmol}^{-1} \text{HAC}^{-1}$) |
|---|-----------|-------------------------|--|
|  | 4 R = H | 280 ± 38 | 1.35 |
| | 5 R = Ac | 82 ± 13 | 1.30 |
|  | 1 R = H | 128 ± 63 | 1.48 |
| | 8 R = Ac | 330 ± 65 | 1.10 |
|  | 11 R = H | 75 ± 13 | 1.57 |
| | 12 R = Ac | 211 ± 24 | 1.17 |
|  | 15 R = Ac | 197 ± 49 | 1.17 |
| | 16 R = Bz | 129 ± 22 | 0.96 |
| | 17 R = Ts | 79 ± 18 | 0.94 |
|  | 18 R = Ac | 730 ± 290 | 1.00 |
| | 19 R = Bz | 220 ± 79 | 0.91 |
| | 20 R = Ts | 670 ± 110 | 0.73 |
|  | 21 R = Ac | 58 ± 8.0 | 1.34 |
| | 22 R = Bz | 88 ± 20 | 1.00 |
| | 23 R = Ts | 32 ± 8.1 | 1.02 |

4.2.9.3 Example 3

SAR study of Benzimidazole series using HSQC and EC₅₀ data :

A series of benzimidazole analogues was analysed using NMR and EC₅₀ data to study the SAR of rationally designed novel compounds. Fragment **N2** which contains a trifluoromethyl substituted benzimidazole core, induced large HSQC perturbations for the residues in the VcDsbA groove region (Figure 4-12) and was thus selected for further analysis. Fragment **N2** has a measured K_D of approximately 32 μM and LE of 1.5 $\text{kJmol}^{-1} \text{HAC}^{-1}$. This compound also exhibited high inhibition with an EC₅₀ of 8 μM in the motility assay.

As shown in Figures 4-10 and 4-16 (A1 and A2), NMR-based docked solution of fragment **N2** suggested its binding in VcDsbA groove region. Similar to benzophenone fragment **N6**, analysis of this docked pose of **N2** with favourable interactions to surface grid maps suggested that there was space to extend the fragment in an effort to improve its potency.

Subsequently, chemical space around the benzimidazole core was explored through the synthesis of a series of derivatives with a diverse range of substituents at the 2', 5' and 6' position by Mr Bradley Doak (Monash Institute of Pharmaceutical Sciences). Analysis of the binding of this series to VcDsbA was carried out by recording end point HSQC experiments in the presence of the analogs (1 mM). In addition, the effect of the compounds on motility was determined in the presence of a single concentration of the compound (500 μM). For those compounds that displayed 100% inhibition of motility at 500 μM , an EC₅₀ was measured (Figure 4-25 and 4-26).

A preliminary SAR analysis of this series was established using this combination of experimental data, while NMR-based docking was used to study the possible binding mode of selected compounds. HSQC and inhibition data analysis revealed that replacement of the 2-hydroxypropyl side chain with a methyl group (in compound **N2a**) slightly reduced the binding and inhibition. (Figure 4-25) The replacement of the 5' trifluoromethyl group (5'-CF₃) by any other substituent resulted in a significant reduction on VcDsbA binding as illustrated in Figure 4-21. The replacement of the 5'-CF₃ with a single fluorine atom (compound **N2b**), hydroxymethyl or azo group or the removal of the 5'-CF₃ group (in compound **N2e**) resulted in a complete loss of binding. Replacement with a methyl group (compound **N2d**) substantially reduced the extent of CSP in the VcDsbA spectra. Along with a reduction in binding, these compounds without a 5'-CF₃ group also showed reduced inhibition as listed in Figure 4-25.

Overall the data associated with the replacement of 5'-CF₃ suggests the importance of this group for binding and inhibition, possibly due to strong hydrophobic interactions with VcDsbA. It was also observed that the 2-hydroxypropyl side chain contributed to binding. Surface grid maps with NMR-based docked poses support these interactions as shown in Figure 4-11. The 5'-CF₃ group interacts with the hydrophobic maps and makes contact with the hydrophobic groove residues (Figure 4-6 and 4-11). The closest and most highly perturbed residues near the CF₃ group in the docked model include Lys 165, Ser166, Tyr170, Val174, Leu177 and Gln162 in the groove region and Ile39, Ile40 and Leu43 near the active site region (Figure 4-8). Analysis of HSQC data showed the significant

reduction in CSP of the above residues for analogues without the 5'-CF₃ (**N2b**, **N2g** and **N2e**). Additionally, no significant changes were observed in CSP strength upon removal of 2-hydroxypropyl side chain. These observations are consistent with the proposed orientation of fragment **N6** in the groove suggested by NMR-based docking. Additionally, further substitution of the ring with halogen-containing electron-withdrawing groups such as CF₃ (compound **N2k**) and chlorine (compound **N2j**) atom generally enhanced binding and inhibition (Figure 4-20). Alternatively, the addition of a CH₃ (compound **N2i**) reduced the ligand binding.

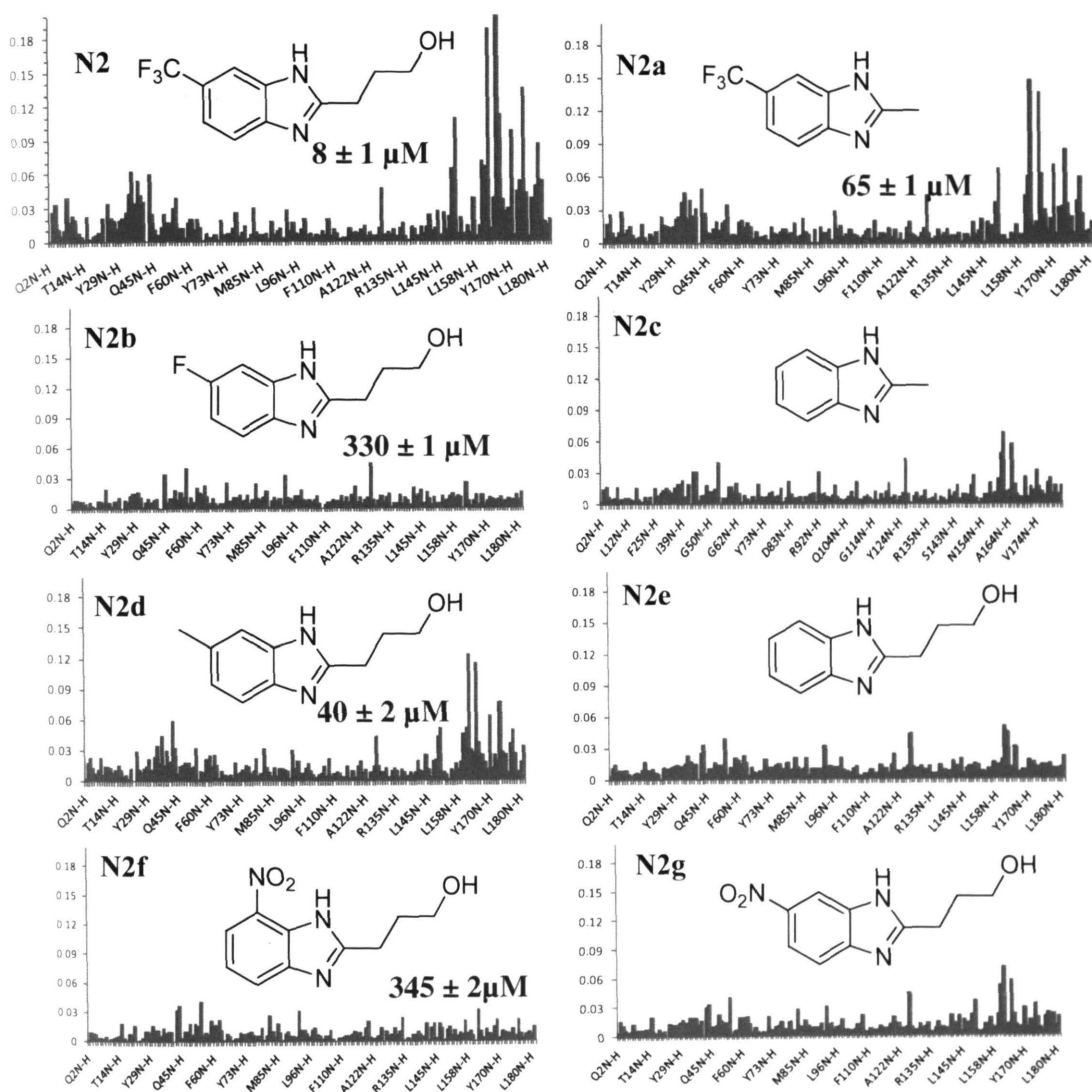


Figure 4-25: Chemical shift changes $\Delta\delta(\text{NH})$ induced in HSQC spectra of ^1H - ^{15}N VcDsbA upon the addition of the benzimidazole analogues **N2a-g** and original hit **N2**. EC_{50} values from the motility assay are also listed for some compounds.

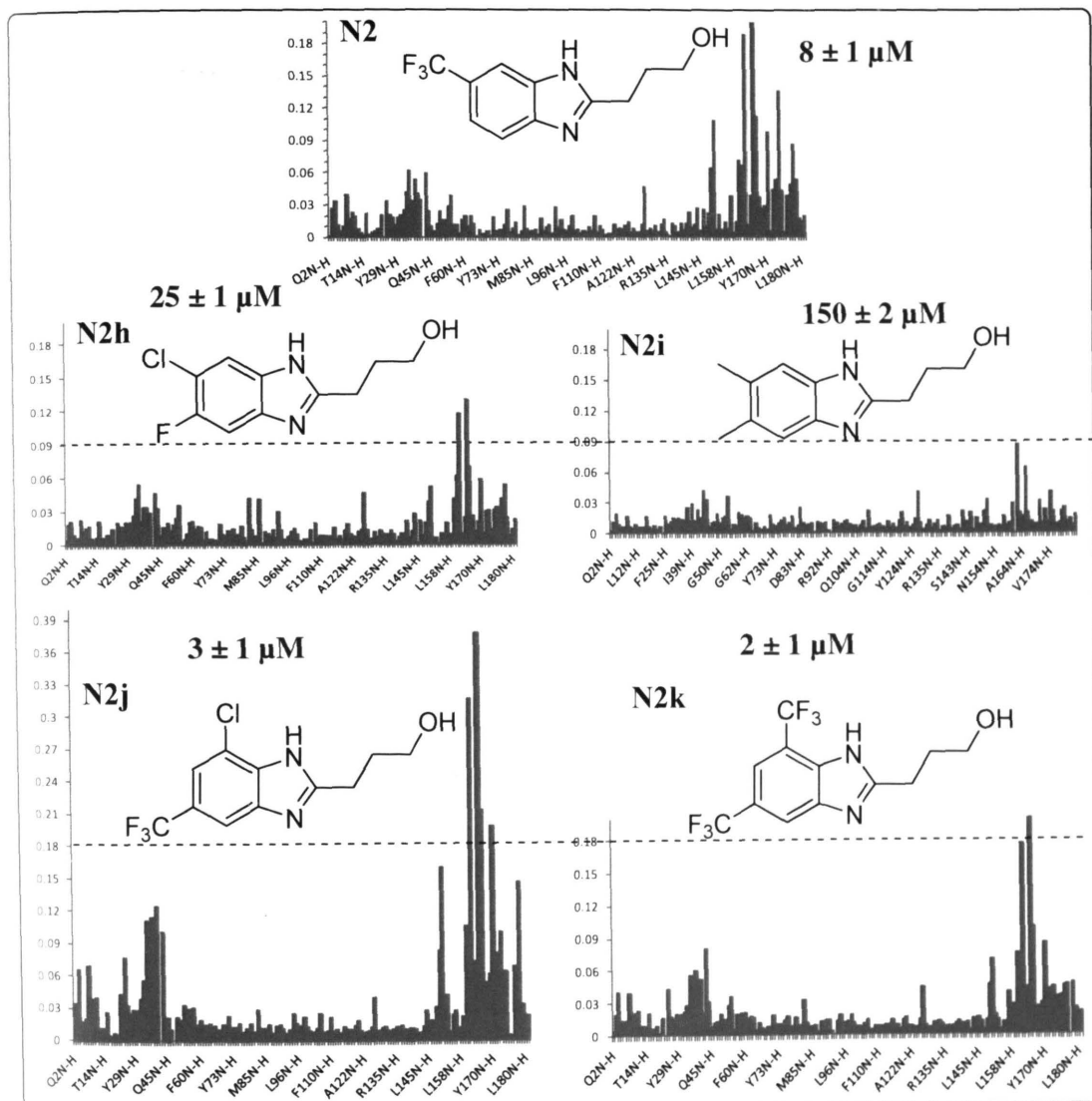


Figure 4-26: Chemical shift changes $\Delta\delta(\text{NH})$ induced in ^1H - ^{15}N HSQC spectra of VcDsbA upon the addition of disubstituted benzimidazole analogues **N2h-k** and the original hit **N2**. EC_{50} values of these compounds in the motility assay are given.

By employing synthetic chemistry and NMR experiments in combination with the use of inhibition data for the benzimidazole analogues, a feasible strategy was established for the further optimization of this class of *VcDsbA* inhibitors. Along with high LE, extensive interactions with the binding site and suitable physiochemical properties, the benzimidazole analogues show good potential for further optimization to give potent small molecule inhibitors of *VcDsbA*. Structural analysis of these compounds suggests that further extension of the benzimidazole core by hydrophobic substituents could provide higher affinity compounds. On the other hand, further optimization of a polar functionality at the 2' position may be able to probe hydrogen bonding interactions towards the open end of the groove. Subsequently, further development of these compounds will be carried out using a combination of structural NMR, docking and medicinal chemistry efforts; however, these studies are beyond the scope of this thesis.

4.2.9.2 Selectivity over *EcDsbA* inhibition

As mentioned before, *VcDsbA* shares a low sequence similarity with *EcDsbA*, but was able to restore motility in *E.coli* $\Delta dsbA$ mutants. To show that the observed effects of DsbA binding fragments were DsbA-specific, wt (*EcDsbA*) *E.coli* and *pVcDsbA* mutants were tested against a set of *EcDsbA* binders that do not bind to *VcDsbA* (Figure 4-4, Table 4-1 and 4-5). These tested compounds inhibited motility in wt (*EcDsbA*) *E.coli*, but showed no inhibition of motility in *VcDsbA* complemented mutants, demonstrating the *EcDsbA* specificity of these compounds and the inhibition of only DsbA-dependant functions in this cell-based assay. Similar cross-binding analysis was conducted for

eighteen compounds which bound to either *Ec*DsbA or *Vc*DsbA by HSQC experiments. As summarized in (Table 4-5) most of the original *Vc*DsbA hits were found not to bind to *Ec*DsbA; however a subset of the *Ec*DsbA hits (**B6**, **C4**, **E3** and **10**) bound to *Vc*DsbA with medium extent shifts.

Table 4-6: Compound number, intensity of CSP observed at 1mM, and % inhibition of motility at 500 μ M of *Vc*DsbA and *Ec*DsbA binding fragments.

| Comp. | CSP strength <i>Ec</i> DsbA and % inhibition of wt <i>Ec</i> | | CSP strength <i>Vc</i> DsbA and % inhibition of - p <i>Vc</i> DsBA | | Comp. | CSP strength <i>Ec</i> DsbA and % inhibition of wt <i>Ec</i> | | CSP strength <i>Vc</i> DsbA and % inhibition of - p <i>Vc</i> DsBA | |
|-------------|---|------|---|------|-----------|---|---|---|------|
| 1 | +++ | 100% | - | 0% | N1 | - | - | ++ | 50% |
| 2 | +++ | 100% | - | 0% | N2 | - | - | +++ | 100% |
| 3 | +++ | 100% | - | 0% | N3 | - | - | ++ | 80% |
| B1 | ++ | 100% | + | 50% | N4 | - | - | ++ | 0% |
| B6 | +++ | n.a | ++ | n.a | N5 | - | - | ++ | 0% |
| C4 | +++ | 100% | ++ | 0% | N6 | - | - | ++ | 0% |
| E1 | +++ | 100% | - | 0% | N7 | - | - | ++ | |
| E2 | +++ | 100% | - | 0% | | | | | |
| E3 | ++ | 100% | ++ | 70% | | | | | |
| 10 | +++ | 100% | ++ | 100% | | | | | |
| SE12 | +++ | 100% | - | 0% | | | | | |

This observation indicates the potential for selectivity of several of the identified hits for their respective DsbA enzymes and is further validated by the motility assay inhibition results (Table 4-5). Notably, in each case tested the compounds bind to the hydrophobic region near the active site of *VcDsbA* or *EcDsbA* enzymes and these compounds selectively inhibit motility in the corresponding whole cell system. The restoration of wt *E.coli* motility by *pVcDsbA* mutant suggests that despite low sequence similarity among *EcDsbA* and *VcDsbA*, these enzymes contains are capable of binding to and oxidising the same protein substrates. The structural differences between hydrophobic grooves in these enzymes are probably responsible for small molecule binding selectivity. This suggests that small molecules that bind selectively to DsbA enzymes can be designed that may target a particular bacterial species to counteract virulence.

4.3 Conclusion

A fragment-based approach was used to identify binders and potential inhibitors of *VcDsbA* by combining multidimensional NMR, HSQC-guided docking and surface grid map-based screening. Fourteen scaffolds exhibiting high to moderate affinity were discovered with a >90 % agreement between experimental and computational results for virtual selection using *VcDsbA* site grid maps. The accuracy and agreement between the computational results and experimental data suggests that the *VcDsbA* binding site model and our overall protocol is reasonable. Indeed we envisage that these methods can be used for the extension of fragments to generate larger molecules to potentially increase

their binding. Benzophenone **N6** and benzimidazole fragments **N2** were further explored by the design and synthesis of an analogue series. These proved to be useful for analysis of SAR however this was limited by the lack of solubility in the case of the benzophenone analogues. Significantly, reliable docked solutions were generated using NMR constraint-based docking for the selected compounds to study small molecule binding to *VcDsbA*. Moreover, the bacterial motility assay was optimized to study the inhibition and selectivity profile of the identified *VcDsbA* hits. The overall results show that *VcDsbA* is a highly tractable target and the compounds we have described should provide a firm platform for further optimisation.

Chapter 5

**Elaboration of phenylthiophene and the phenoxybenzene fragments as
potential inhibitors of *EcDsbA***

5.1 Introduction

Development of a high affinity lead - like compound from a low affinity and low molecular weight fragment presents challenges such as synthetic feasibility, careful monitoring of affinity with respect to size, choice of elaboration strategy and monitoring of physicochemical properties.²²¹ Conventional medicinal chemistry approaches of synthesising analogues of a known 'active' are the most commonly used methods for improving compound affinity and are widely applicable to the fragment elaboration process.²²² The iteration of similarity searching, analogue synthesis, screening and design enables growth of the fragment into a 'hit' compound, which can be facilitated by binding site information from biophysical methods such as NMR and/or X-ray crystallography.²²³ In the FBDD process, solution state NMR is one of the most versatile tools for the study of weak binding interactions between small ligands and protein targets.^{113, 117b, 125a}

Chapter 3 described screening by NMR of a fragment library, which led to the identification of several diverse compounds that bound to *Ec*DsbA with μ M potency and good LE. Some of these compounds inhibited the activity of *Ec*DsbA in a cell-based motility assay, and structures of a number of complexes were determined using X-ray crystallography. Two series of compounds that displayed good LE, activity in the motility assay^{19a, 35, 55b} and for which high-resolution crystal structures were available were the phenylthiophene compounds **1-3** and the phenoxybenzene compound **4**. (Please note that

the compound numbers mentioned in the paper format chapter 3, are different from the compound number discussed in this chapter. Compound **2** discussed in chapter 3 will be numbered as compound **4** in current chapter)

Notably these compounds were readily soluble in NMR sample buffers and the media used for inhibition assays (Figure 5-1). Consequently a number of *in vitro* tests were performed in order to demonstrate that the observed activity was due to inhibition of *EcDsbA* and not toxic effects of these compounds on bacterial cell growth. These tests confirmed that phenylthiophenes and the phenoxybenzene compound have no inherent toxicity for bacterial growth in standard conditions hence can be further analysed by cell-based assays.

LE, which relates the free energy of binding to the number of non-hydrogen atoms in a fragment, is a better parameter than affinity in order to select a lead compound.^{93, 169c} Although the NMR hits **1-3** and **4** have low affinities because of their low molecular weights, they show relatively high LE (up to 1.5-1.8 kJ mol⁻¹HAC⁻¹). Therefore, these *EcDsbA* scaffolds were considered excellent starting points for further chemical optimization to potent leads. This chapter describes attempts to optimize each of these cores, by testing analogues that were either available from commercial sources or could be readily obtained by in-house synthesis. Analogues were restricted to compounds with M.W ≤ 300 to focus on optimization of the core rather than extension of the fragment^{85, 99, 224} and were initially tested for binding to *EcDsbA* by recording "end-point" ¹H-¹⁵N HSQC spectra. The analogues that gave the largest CSP in end-point ¹H-

^{15}N HSQC spectra were characterised further by measuring K_D from ^1H - ^{15}N HSQC titrations and potency by *in vitro* bacterial motility assay. This resulted in the identification of core scaffolds with improved LE values that represent better starting points for further design of novel and more potent inhibitors of *EcDsbA*.

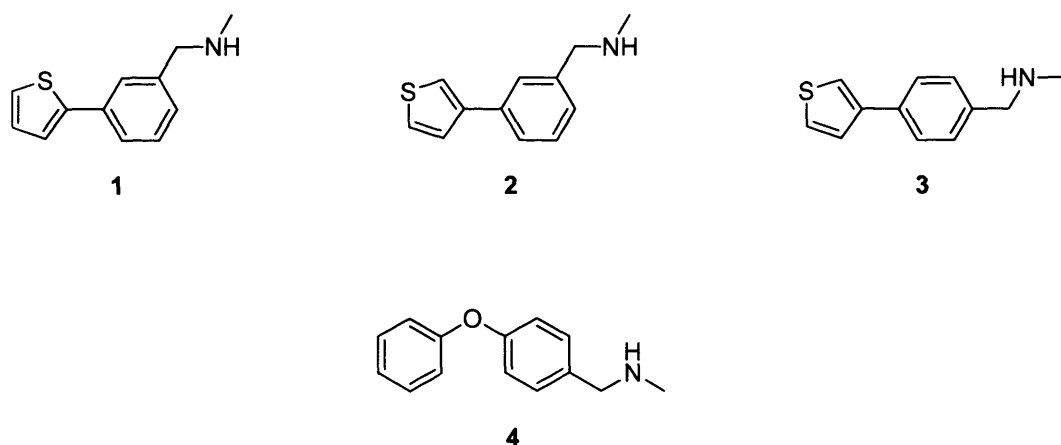


Figure 5-1: Structures of the lead fragments for optimisation. Three phenylthiophene cores comprising the 2-phenylthiophene fragment **1** and the 3-phenylthiophene fragments **2** and **3** as well as the phenoxybenzene fragment **4** were chosen for further optimisation based on their potency, LE, activity and availability of crystal structures.

5.2 Results and Discussion

5.2.1 Characterization of phenylthiophene hits 1-3

In Chapter 3, the identification of three phenylthiophene fragments, which inhibit *EcDsbA* was described. Each had an *N*-methyl-methanamine substituent at either the 3' or 4' position of the phenyl ring. Two of the fragments were 3-phenylthiophenes, whilst the third was a 2-phenylthiophene (Figure 5-1). These fragments induced large CSP (up to 0.17 ppm) in 2D ^1H - ^{15}N HSQC spectra of oxidized *EcDsbA*. (Please refer Appendices section 8.5 and 8.6 for the STD screen result analysis and HSQC chemical shifts of fragments 1-3 upon binding with *EcDsbA*). The perturbed residues were similar in each case although the CSP were of slightly different magnitude (Figure 5-2). The CSP of multiple residues as a function of compound concentrations were used to plot binding isotherm for K_D determination.^{84, 113} The binding affinities reported in this chapter are dissociation constants (K_D) determined by ^1H - ^{15}N HSQC NMR based titration (see Chapter 2 Section 2.3.2). All ligand efficiencies are calculated directly from the NMR dissociation constants and have units of kJ per mole per heavy atom.⁹³

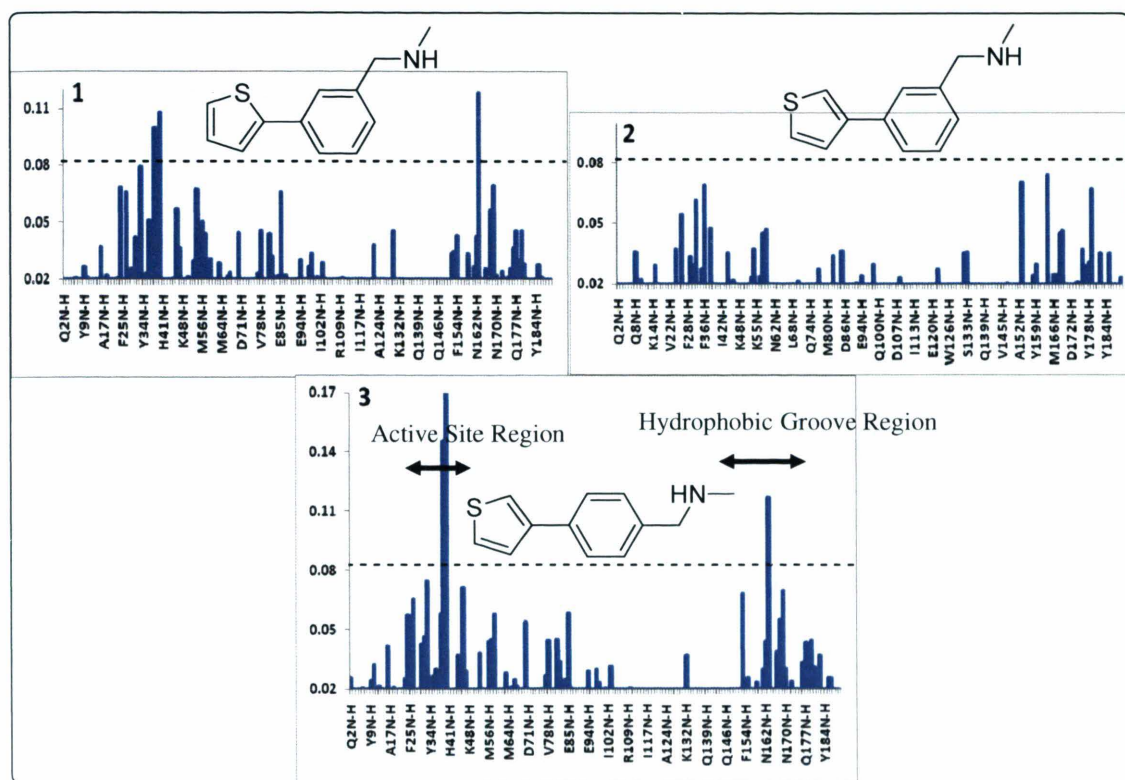


Figure 5-2: Phenylthiophene class of *EcDsbA* hits identified from NMR-based fragment screening. Significant CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of fragments (**1-3**) are shown. A dashed line has been placed on the ppm scale to highlight the extent of CSP in each case. Important binding regions of the protein have been shown as arrows on the lower histogram.

Fragments **1-3** have measured K_D values of 77 μM , 25 μM and 74 μM and LE values of 1.66, 1.80 and 1.67 $\text{kJ mol}^{-1} \text{HAC}^{-1}$ respectively. They showed inhibition in the motility assay with measured EC_{50} values of 52, 36 and 4 μM , respectively. Whilst there is good agreement between the K_D and EC_{50} values for fragments **1** and **2**, the discrepancy observed with fragment **3** may indicate that it is inhibiting motility in a non-*EcDsbA*

dependent fashion or due to properties associated with absorption or membrane permeability and metabolism in the cell-based assay.

The CSP observed in ^1H - ^{15}N HSQC spectra suggested that the fragments were binding in the hydrophobic groove region of the *EcDsbA* structure adjacent to the active site (Figure 5-3 and 5-4). This was confirmed in the case of fragments **1** and **2** by obtaining high resolution structures of their complexes with *EcDsbA*. The X-ray structures confirmed that the fragments bound in the hydrophobic groove of the thioredoxin domain of *EcDsbA*.^{20a} Despite several attempts of co-crystallisation and to soak crystals of *EcDsbA* with fragment **3** no structure was obtained, although based on the similarity of the CSP observed for all three fragments a similar binding mode was inferred. The structure of *EcDsbA* consists of a thioredoxin domain (residues 1-62 and 139-189) with an α helical domain inserted (residues 63-138) at the active-site Cys30-Pro31-His32-Cys33.^{20a 207, 225} As shown in Figure 5-3 the surface of *EcDsbA* contains a number of unique and functionally important features including a peptide binding groove, hydrophobic pocket and hydrophobic patch and an acidic patch on the opposite face to the active site disulfide.³⁴ Interestingly the residues interacted with compounds **1-2** are similar residues involved in the *EcDsbA*-*EcDsbB* binding hence they are binding at a functionally important region.^{42b, 226}

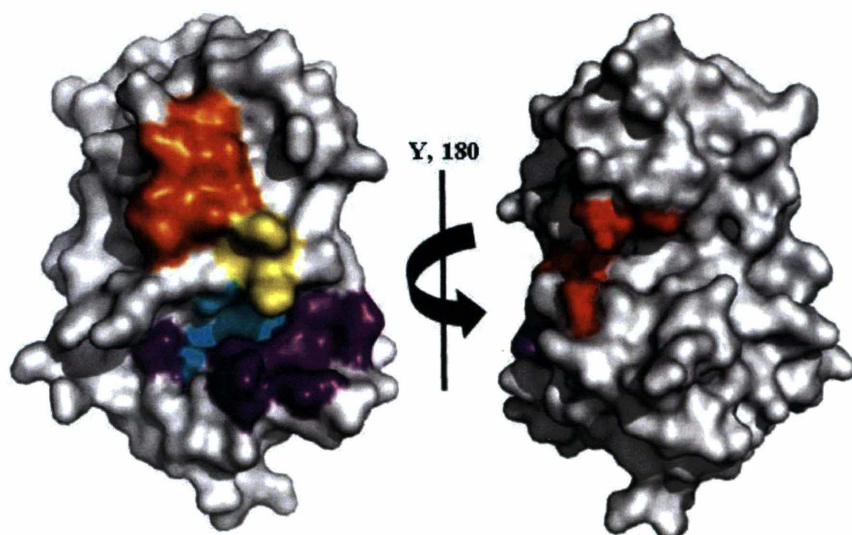


Figure 5-3: The surface features of *EcDsbA* highlighting the active site (yellow), hydrophobic patch (orange), hydrophobic pocket (cyan), hydrophobic groove (violet) and acidic patch (red) which is on the opposite side of the active site.

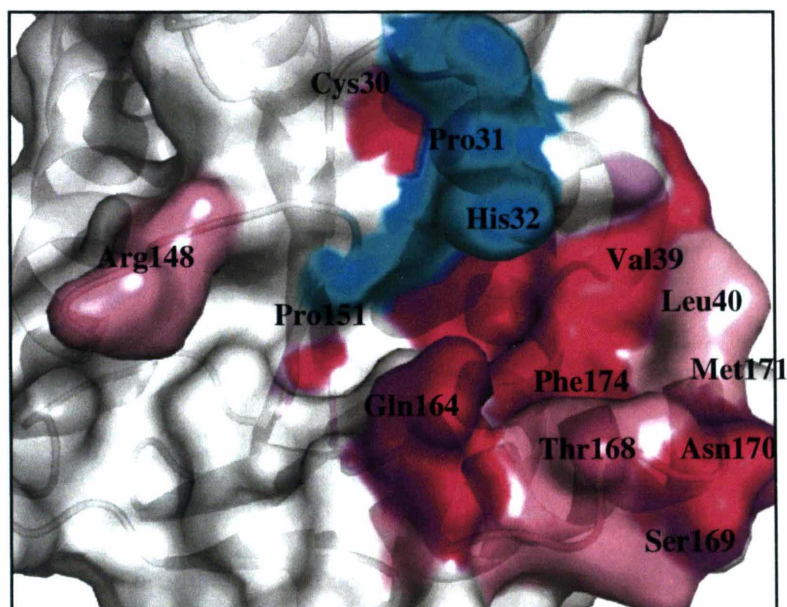


Figure 5-4: *EcDsbA* surface (white) mapped with CSP ($\Delta\delta(\text{NH}) > 0.03$ ppm) induced and calculated from ^1H - ^{15}N HSQC spectra, upon binding of phenylthiophene hit **1** at 1200 μM . Higher, lower shifts and unassigned residues have been colored as magenta, pink and cyan, respectively.

Phenylthiophene fragments **1** and **2** bound in similar and overlapping orientations directly below the active-site disulfide of *EcDsbA* at the midpoint of the hydrophobic groove (Figure 5-5). This binding was dominated by nonpolar hydrophobic and van der Waals (vdW) interactions with the surrounding residues. Fragments **1** and **2** bound only to one of the two molecules of *EcDsbA* in the asymmetric unit of the crystal lattice, and the crystals diffracted to a resolution of 1.95 Å and 1.99 Å, respectively. In both cases, the thiophene ring is surrounded by Phe36 and Leu40 of *EcDsbA* whilst the phenyl ring is located mainly over the hydrophobic pocket where it partially stacks with Phe174. The methanamine side chain lies in close proximity (~3.8 Å) and makes vdW interactions with His32, *cis*-Pro151 and Pro163 that are key residues involved in the interaction of *EcDsbA* with its substrates.^{19c, 42b, 49} An uncharged residue Thr168 along with hydrophobic residue Met171 interacts with the phenyl and thiophene rings respectively via hydrophobic contacts. (Figure 5-5). Altogether crystal structures of fragment **1** and **2** revealed up to 60 vdW interactions to the aliphatic amino acids of the hydrophobic groove as calculated within CCP4 software.^{175b} A notable feature of the complexes formed with these fragments is that no polar interactions were observed in the crystal structures, suggesting that the interactions are largely non-polar and driven by shape-complementarity between the active site and the fragment. This is not unusual for fragment binding to proteins,²²⁷ and despite the lack of polar interactions their high ligand efficiency of binding suggests that these compounds are reasonable starting points for further optimisation.⁹⁸

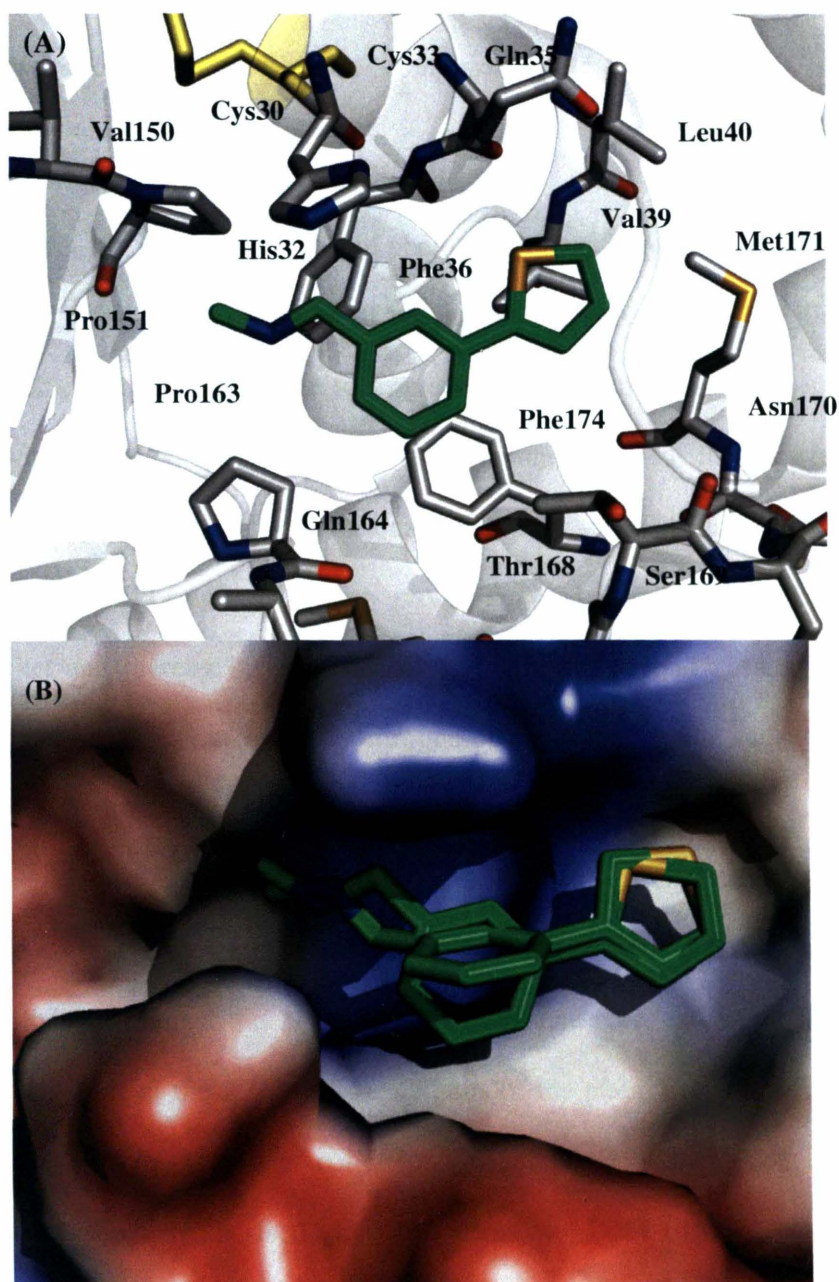


Figure 5-5: (A) Orientation of phenylthiophene fragment **1** (green stick representation) binding at the crystal structure of *EcDsbA* (white stick and ribbon representation). (B) Crystallographic positions of phenylthiophene fragments **1** and **2** (green stick representation) showing their overlapping binding sites in the hydrophobic groove of *EcDsbA* (electrostatic surface representation).

5.2.2 *Elaboration of phenylthiophene compounds 1-3*

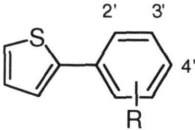
Analysis of the X-ray crystal structures of phenylthiophene fragments **1** and **2** suggested a potential strategy for the optimization of binding to *EcDsbA*. Figure 5-5 illustrates two areas in which it might be possible to improve the interaction of the phenylthiophene fragments with *EcDsbA*. Firstly, there is a significant region of the groove below the fragment that is not occupied. It was rationalised that by functionalising the 2', 3' and 4' positions of the phenyl ring with various groups it may be possible to more completely fill this pocket and thereby increase the binding affinity. Secondly, there is additional space to extend the fragment from the nitrogen of the methanamine substituent towards the open end of hydrophobic groove near the *cis*-Pro151 loop.^{20a, 34}

To test these binding hypotheses a number of approaches were undertaken. Firstly, new compounds that maintained the 2- or 3-phenylthiophene core, but which had a diverse range of substituents at the 2', 3' and 4' positions of the phenyl ring were sourced from either commercial or in-house collections. Secondly, a range of *N*-substituted methanamines were synthesised by Mr Bradley Doak (Monash Institute of Pharmaceutical Sciences). Finally, alternate ring systems consisting of linear 5-5 and 5-6 heterocycles, that were similar to the phenylthiophenes as assessed by their Tanimoto score^{193, 228} were purchased. A Tanimoto similarity threshold of 0.35 was set as the cut-off based on the literature.²²⁹ Substitution of the phenyl ring with pyridine or thiophene was also attempted to explore interaction of this ring as well as enhance polarity and

solubility. Elaboration of the amine via amide couplings as well as replacement of the thiophene ring with a phenol ring were also attempted to increase affinity based on potential interactions of fragments **1** and **2**. Forty six new compounds were chosen or synthesised via this process for evaluation and were first tested for solubility under the conditions employed in the ^1H - ^{15}N HSQC and motility assays. Seven of the selected analogues showed poor solubility in minimal media used for the motility assay and another three were insoluble in the NMR sample buffer and these were discarded. For the remaining analogues end-point ^1H - ^{15}N HSQC spectra were recorded and ranked as (strong, +++), (medium, ++), (weak, +) or (no binding -) on the basis of a threshold applied for overall weighted average ($\Delta\delta$ NH) CSP ≥ 0.03 ppm. This threshold was set according to the extent of the strongest and weakest CSP observed in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon binding of ligands in each series. In addition, the effect of the compounds on motility was determined by undertaking the motility assay^{19a} in the presence of a single concentration of the compound (500 μM) and measuring the % inhibition of motility.^{35, 55b} For those compounds that gave the strongest CSP (+++), the K_D was determined from the ^1H - ^{15}N HSQC titration experiment, and for compounds that displayed 100% inhibition of motility at 500 μM , an EC_{50} was measured.

5.2.2.1 Analogues of the 2-phenylthiophene series

The structures of the 2-phenylthiophene analogues that were purchased and tested are shown in Table 5-1.










| Substituent | 2' | | 3' | | 4' | |
|---|----|-------------|-----|-------------|-----|-------------|
|  | A1 | +++ 100% | 1 | +++ 100% | A11 | +++ 100% |
|  | A2 | ++ 100% | A7 | +++ 100% | | |
|  | A3 | + 0% | A8 | + 30% | A12 | + 0% |
|  | A4 | + 10% | A9 | +++ 100% | A13 | +++ 100% |
|  | A5 | ++ 100% | A10 | ++ 100% | A14 | ++ 100% |
|  | A6 | ++ 100% | | | | |
|  | | | | | A15 | - 0% |

Table 5-1. Analogues of the 2-phenylthiophene fragment **1**. The table shows the chemical structure of the substituent and the position (2', 3' or 4') of attachment to the phenyl ring. The compound number, the magnitude of CSP observed in the ¹H-¹⁵N HSQC end point experiment along with the % inhibition of motility caused by the fragment at a concentration of 500 μM are listed. Fragments highlighted in grey were selected for further characterisation.

Table 5-1 reveals that several different functional groups are tolerated at each position on the ring. A notable finding is that loss of the *N*-methyl substituent from the parent fragment results in a marked reduction in binding (Figure 5-6). As noted above, the methyl group of the parent fragment **1** is involved in a hydrophobic interaction with His32, Pro151 & Pro163, which suggests that these interactions are important for binding (Figure 5-5). This observation is reminiscent of the phenomenon of ‘magic methyls’ which suggests that the addition of a single methyl group can increase potency up to 10-fold by optimally burying of the surface area of a single ‘heavy’ atom.²²⁰ In the present case we were unable to measure K_D values for aminomethyl compounds **A3**, **A8** and **A12** due to their weak CSP, but this in itself is suggestive of a significant drop in K_D . Of the remaining analogues of the 2-phenylthiophene fragment, ten were investigated further by recording ^1H - ^{15}N HSQC titrations to measure their affinity and/or titrations in the motility assay to measure their potency. The results of these experiments are summarised in Table 5-2.

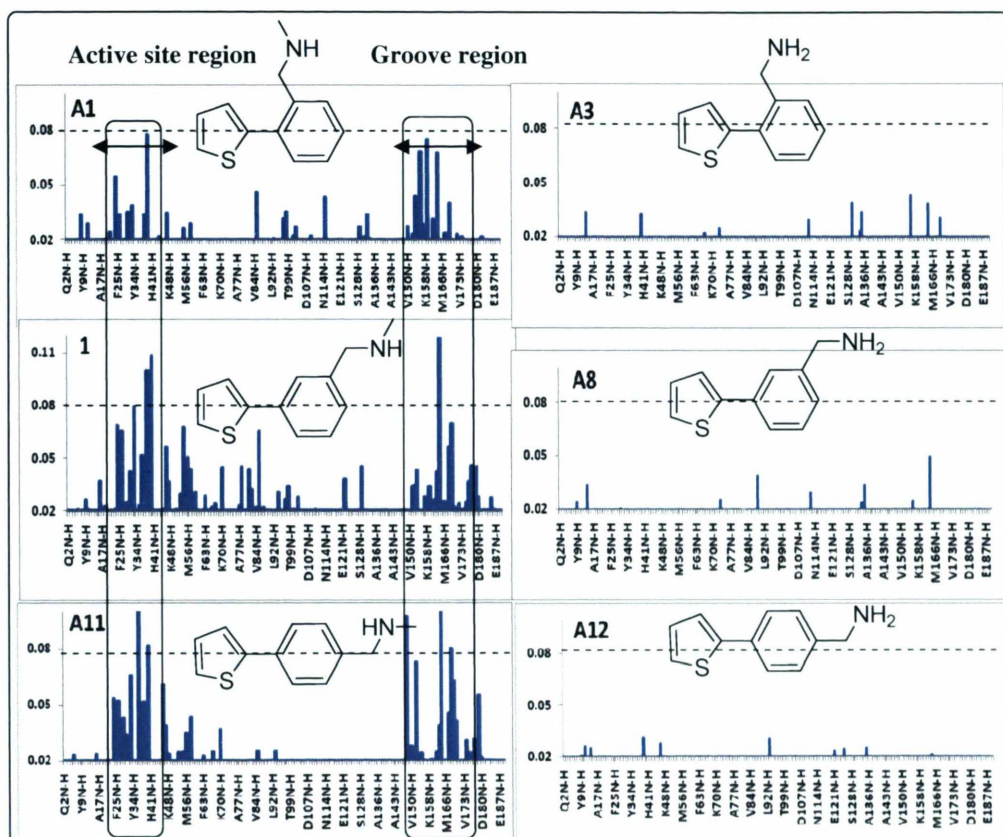


Figure 5-6: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of 2-phenylthiophene analogue with *N*-substituted methanamines **A1**, **1**, **A11** and aminomethyl substitution **A3**, **A8**, **A12**.



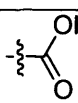

| substituents | 2' | K _D (μM) | EC ₅₀ (μM) | 3' | K _D (μM) | EC ₅₀ (μM) | 4' | K _D (μM) | EC ₅₀ (μM) |
|---|-----------|------------------------|--------------------------|------------|------------------------|--------------------------|------------|------------------------|--------------------------|
|  | A1 | 89 ± 18 | 87 ± 1 | 1 | 77 ± 18 | 51 ± 1 | A11 | 78 ± 18 | 110 ± 1 |
|  | A2 | n.d. | 68 ± 1 | A7 | 37 ± 16 | 38 ± 1 | | | |
|  | A4 | n.d. | n.d. | A9 | 120 ± 24 | 179 ± 1 | A13 | 132 ± 25 | 150 ± 1 |
|  | A5 | n.d. | 154 ± 1 | A10 | n.d. | 162 ± 1 | A14 | n.d. | 125 ± 1 |

Table 5-2. Affinities and potencies of the 2-phenylthiophene analogues.

Across the series of 2-phenylthiophene analogues, it appears that there is generally no clear preference for the substitution pattern of the phenyl ring. One exception to this is the carboxylate substituent, where the 2'-carboxyl substituted compound **A4** showed a marked reduction in potency (as measured in the motility assay) and significantly lower CSP than either the 3'- or 4'- carboxyl compounds **A9** and **A13** respectively (Figure 5-7).

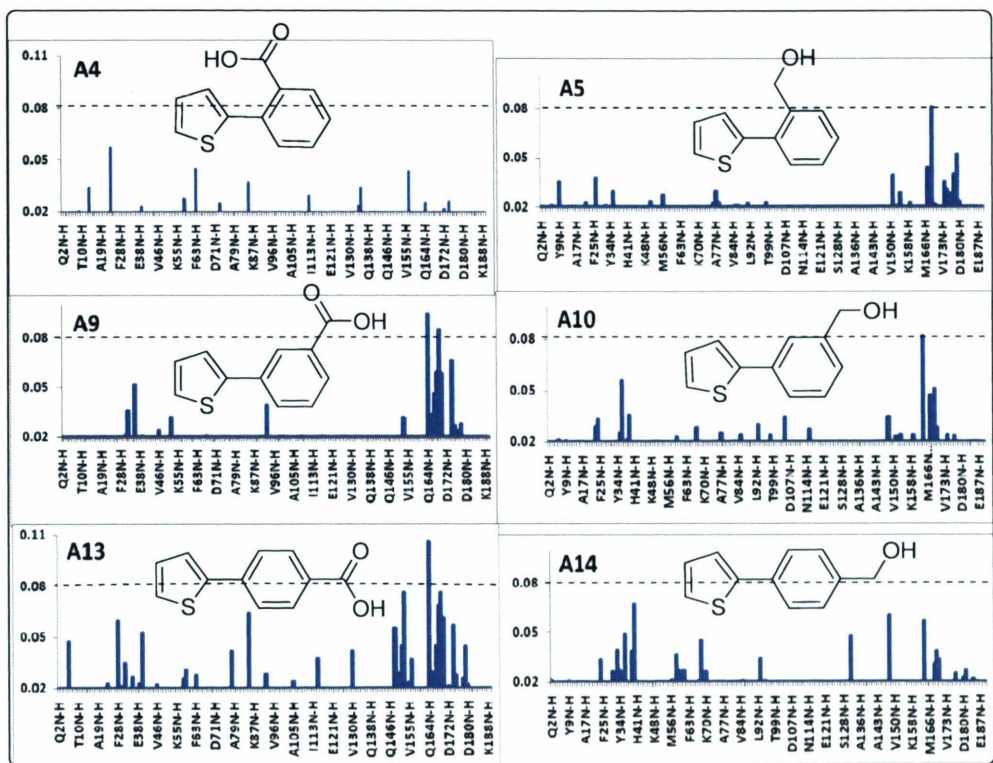


Figure 5-7: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of 2-phenylthiophene analogues with carboxyl substitution **A4**, **A9**, **A13** and hydroxymethyl substitution **A5**, **A10**, **A14**.

Analysis of the patterns of CSP observed across the 2-phenylthiophene analogues revealed that the CSP data were similar with binding in the hydrophobic groove of *EcDsbA* as had been observed with the parent fragment. However, upon closer inspection, some subtle differences in the CSP were observed. Across the *N*-methyl methanamine compounds **A1**, **1**, **A11**, CSP were induced for comparatively similar resonances. Slightly larger CSP were observed for the 3' and 4' substituted compounds (**1**, **A11**). Mapping of these shifts was located as one cluster of large CSP at residues around the active site (Cys30-Cys33) and a second cluster centred on the loop residues that define the bottom edge of the hydrophobic groove of *EcDsbA* (Leu161-Asn170) (Figure 5-6). For the carboxylate-substituted compounds (**A4**, **A9** and **A13**) the CSP were markedly reduced in case of 2'-substituted compound **A4**, (Figure 5-7) corresponding with its apparently lower affinity as suggested from the motility data. The 3'- and 4' substituted compounds on the other hand, induced CSP in the groove albeit that the patterns of perturbation were again subtly different. (Figure 5-7) The hydroxymethyl-substituted compounds **A5**, **A10** and **A14** each gave similar patterns of perturbation although the extent of CSP around the active site for the 2'-substituted compound was somewhat smaller (Figure 5-7). Similar to **A5**, analogue **A6** containing a bulky substitution at 2' with bromomethyl group (Figure 5-8) induced small CSP. With the amine substituted compounds **A2** and **A7** the binding location was consistent with CSP in the hydrophobic groove of *EcDsbA*, although the extent of perturbations observed with the 2'- substituted compound **A2** was generally reduced and there were notably fewer perturbations observed around the active site residues. 2'-Amino substituted compound

A2 showed a marked reduction in potency (as measured in the motility assay) and significantly lower CSP than the 4'- amino substituent (Figure 5-8).

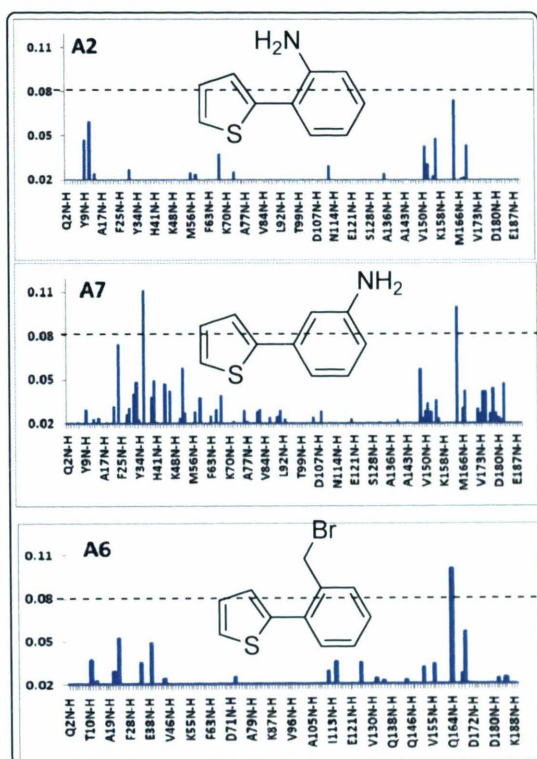


Figure 5-8: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of 2-phenylthiophenes analogues **A2**, **A7** and **A6**.

3' amino substituted compound **A7** induced high CSP up to 0.1 ppm in the active site and groove region with a pattern relatively similar to the original *N*-methyl methanamine compound **1**. Highly perturbed residues in this case were positioned around active site (residues Glu38, Val39, Leu40, Gln35, Phe36, Cys30, Cys33). This compound showed about a 2-fold increase in affinity and inhibition along with the highest LE of 2.2 kJ mol⁻¹ HAC⁻¹ among all the identified *EcDsbA* hits in this study.

5.2.3 *NMR guided molecular docking*

In an effort to characterise the binding poses of these 2-phenylthiophene analogues a number of them were soaked into crystals of *EcDsbA*. Of the compounds tested only the 3-carboxylate compound **A9** showed any additional density in the electron density map obtained from the crystallographic data. However, the data were not of sufficiently high quality to generate a high-resolution model of the complex. Nonetheless, the additional density is consistent with this fragment binding in the hydrophobic groove of *EcDsbA* as suggested by the NMR CSP data. (Data not shown)

X-ray crystallographic complexes of *N*-methyl methanamine phenylthiophene hits **1** and **2** represent the binding mode for this class however analogues of these compounds might interact in a different mode with *EcDsbA* groove. It has been demonstrated in recent fragment elaboration studies that the binding position of the parent core is often not constant but can be either slightly altered or in some cases completely different in analogues.²³⁰ Hence considering the crystallographic binding mode of a parent core to interpret interactions of an analogue might create a biased outcome for the understanding

of binding across a series. As described in chapter 4 NMR-based molecular docking provides an alternative to identify binding locations and orientations in the absence of X-ray crystal structures.^{187, 231} To identify the binding mode and positions of the newer *EcDsbA* hits, we coupled our experimental NMR data with docking calculations. (NMR-based docking protocol and validation results were described in chapter 2, section 2.7 and chapter 4, section 4.2.3 respectively).

Figure 5-9 illustrates the comparison between NMR-based docked solution of fragments **1**, **3** and **4** and the crystallographic coordinates of *EcDsbA*-fragment complexes that we had obtained for these compounds. Hence this approach was able to generate a good agreement between most of the crystallographic pose and docked solutions.

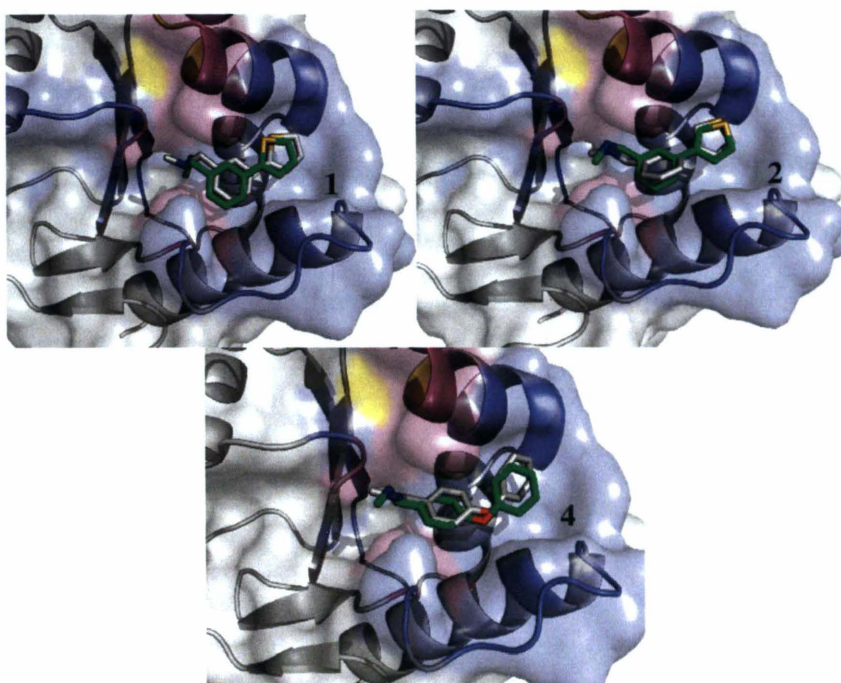


Figure 5-9: Illustration of validation of NMR-based docking. Superposition of the crystallographic coordinates of fragments (white stick representation) bound to *EcDsbA*, with the predicted binding mode of the same fragments (green stick representation). Surface of *EcDsbA*, mapped with the end point ^1H - ^{15}N HSQC CSP (blue) of the corresponding fragment and Cys30-Cys33 and unassigned residues are coloured in yellow and pink respectively.

Subsequently, NMR-based docking with CSP data was used to dock selected analogues into *EcDsbA*. For each studied complex docking grids were generated on the *EcDsbA* NMR binding region of the corresponding fragment. Within this grid area hydrophobic constraints were added in those regions of *EcDsbA* that showed significant CSP in ^1H - ^{15}N HSQC data. A simple relationship was considered between the magnitude of the CSP and the distance from the amino acid's NH atoms and the nearest ligand atom according to the

protocol described by Stark *et al.*¹⁸⁷ A number of docked solutions were optimized to obtain the best scoring pose by using Glide selection criteria.^{185, 188} Information obtained from the selected complexes was compared with ¹H-¹⁵N HSQC data where distances between the docked ligand were measured to surrounding residues within 5 Å. In most cases the distances from the fragment to the -NH of highly perturbed residues of each ligand were found to match well with their experimental CSP values. The highest scoring pose from the NMR-constrained docking of some selected fragments is shown in Figure 5-10.

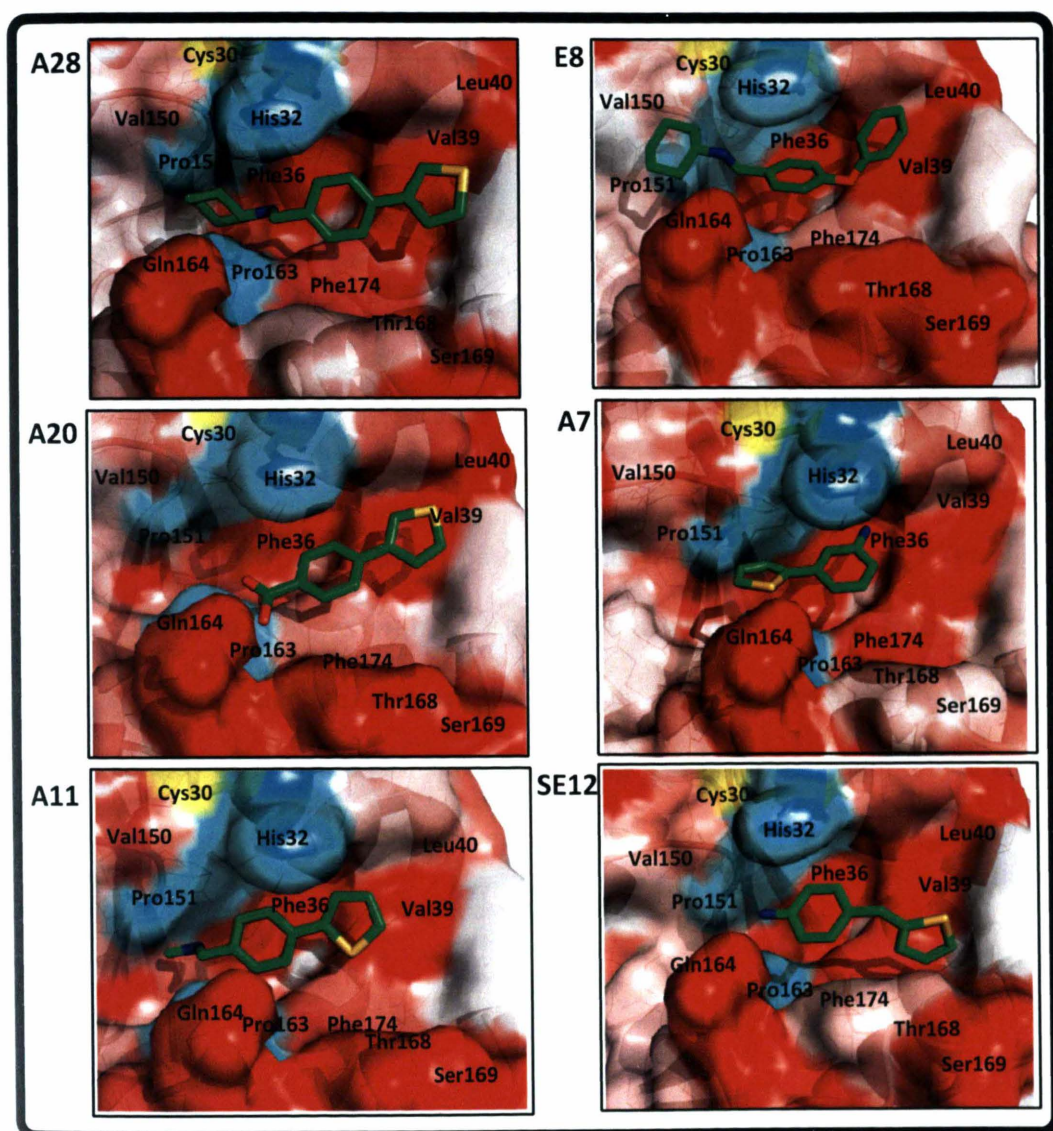


Figure 5-10: CSP maps of *EcDsbA* bound to selected fragments. Residues are coloured according to the extent of CSP with red being the largest CSP observed. Unassigned residues (His32 and prolines are coloured in cyan). Analogues of fragments 1-3 and 4, compounds A28, A20, A11, E8, A7 and SE12 (green stick representation) binding in hydrophobic groove.

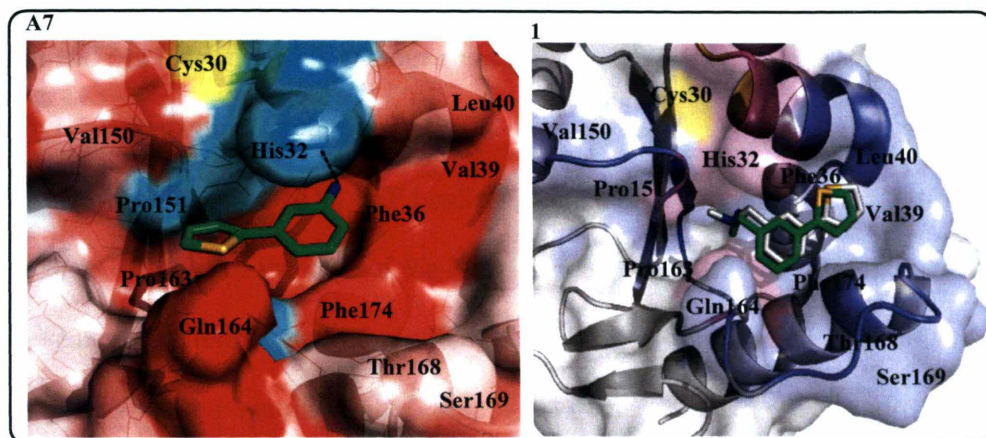


Figure 5-11: NMR-based docked poses of phenylthiophenes **A7** and **1** (green stick representation) in complex with *EcDsbA*. Additionally, the crystal bound pose of **1** (white stick representation) is superimposed with its docked pose. For compounds **A7** and **1**, surface of *EcDsbA* is mapped according to the CSP strength where residues having the largest CSP are coloured red and blue respectively and unassigned residues are coloured in cyan and pink respectively.

The docked solution of the 4-methylaminomethyl analogue **A11** revealed that binding location and orientation of this analogue is similar to the X-ray structure of the original hit *N*-methyl methanamine compound **1**. As shown in Figure 5-10 this position also corresponds with ^1H - ^{15}N HSQC CSP binding data mapped on to the *EcDsbA* structure with docked pose of **A11**. In the case of the 3-amino compound **A7**, the NMR constraint based pose revealed a different binding mode, which is in a flipped orientation relative to the original hit *N*-methyl methanamine compound **1** and also largely supports the ^1H - ^{15}N HSQC binding (Figure 5-11). The docked solution shows that **A7** binds in the hydrophobic pocket and makes vdW contact with the flexible groove and the residues surrounding the active site. Given the enhancement in K_D and activity, we were encouraged to see a hydrogen bond (N—H.....O) between the N of the amino group and the amide backbone carbonyl of His32 in this docked solution. This analogue therefore suggests the possibility of a new interaction to *EcDsbA*. Importantly this is a polar interaction that could be exploited in the future design.

The structures of the available or purchased molecules that were based on structural similarity (Tanimoto > 0.35) to the phenylthiophene fragments are shown in Table 5-3.

Most of the similar molecules (**A32-A36**) that contains an altered thiophene ring such as heteroatom change or substitution of polar or bulky groups showed either weak or no binding (Table 5-3).

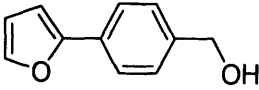
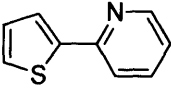
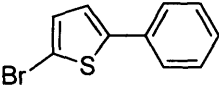
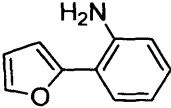
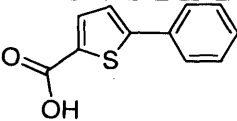
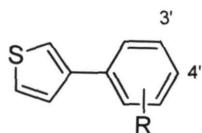
| Compounds | | CSP $\Delta\delta(\text{NH})$ strength | % inhibition |
|---|-----|---|--------------|
|  | A32 | + | 50% |
|  | A33 | + | 20% |
|  | A34 | + | 90% |
|  | A35 | - | 0% |
|  | A36 | + | 10% |

Table 5-3. Chemical structures of the compounds similar to the 2-phenylthiophene fragment. The compound number, magnitude of CSP observed in the end point ^1H - ^{15}N HSQC along with the % inhibition of motility caused by the fragment at a concentration of 500 μM are listed.

5.2.4 Analogues of the 3-phenylthiophene series

The structures of the synthesized or purchased 3-phenylthiophene analogues that were tested are shown in Table 5-4.



| Substituent | 3' | | 4' | | Substituent | 4' | |
|-------------|-----|------------|-----|-------------|-------------|-----|-------------|
| | 2 | ++ 100% | 3 | +++ 100% | | A25 | ++ 30% |
| | A18 | + 100% | | | | A26 | +++ 100% |
| | A17 | ++ 50% | A20 | +++ 100% | | A27 | ++ 50% |
| | A16 | ++ 100% | A19 | + 100% | | A28 | +++ 100% |
| | | | A21 | ++ 100% | | A29 | ++ 100% |
| | | | A22 | + 40% | | A30 | - 0% |
| | | | A23 | ++ 0% | | A31 | ++ 100% |
| | | | A24 | + | | | |

| | | | | | | |
|--|--|--|--|-----|--|--|
| | | | | 50% | | |
|--|--|--|--|-----|--|--|

Table 5-4. Analogues of the 3-phenylthiophene fragment. The table shows the chemical structure of the substituent and the position (3' or 4') at which the substituent was attached to the phenyl ring. The compound number, the magnitude of CSP observed in the end point ¹H-¹⁵N HSQC experiment along with the % inhibition of motility caused by the fragment at a concentration of 500 μM are listed. Fragments highlighted in grey were selected for further characterisation.

3-phenylthiophene analogues showed somewhat similar preferences or tolerances for polar functionality in the 3' or 4' position as observed for the 2-phenylthiophene analogues. In this core 4' position, carboxylate substitution in compound **A20** had improved binding as compared to the 3' carboxylate substitution in compound **A17**. (Table 5-4) As shown in Figure 5-12, **A20** induced higher shifts (up to 0.11 ppm) in the groove region. NMR constraint -based docking suggested that its binding location was similar to the parent fragment, albeit with the carboxyl group oriented down towards the groove residues with a polar contact with Gln164 as shown in Figure 5-9 and Figure 5-13.

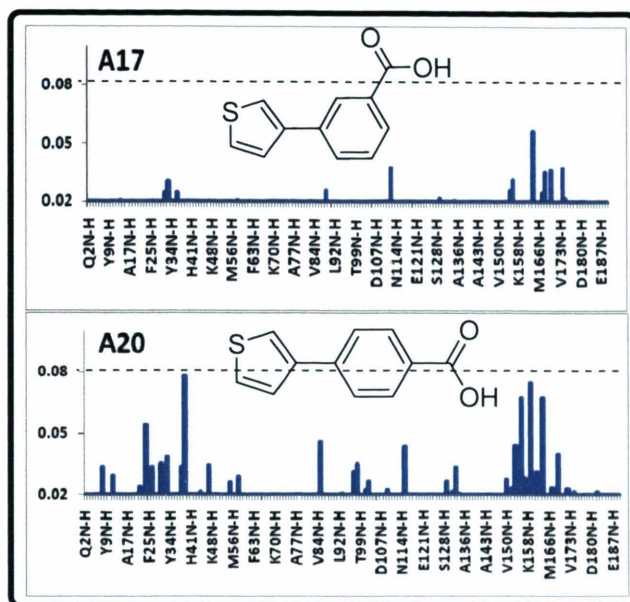


Figure 5-12: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of the 3-phenylthiophenes analogues **A17** and **A20**.

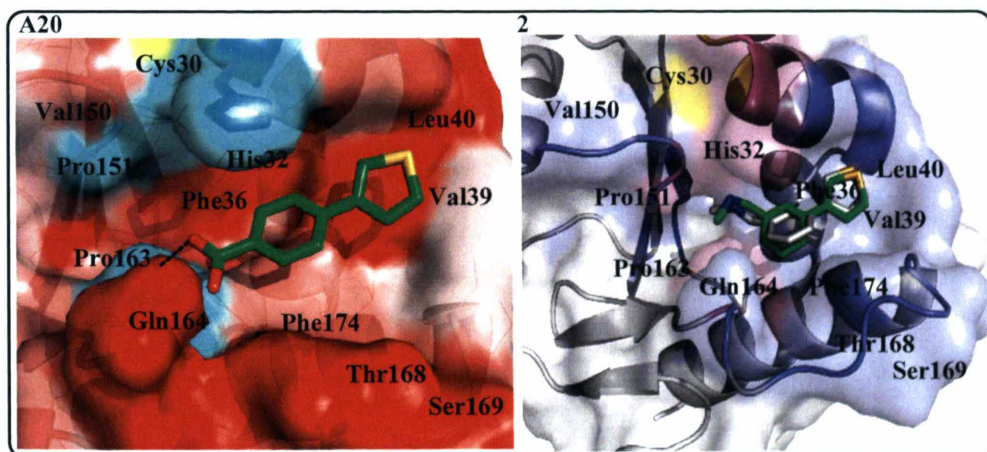



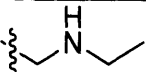
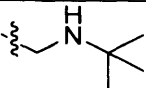
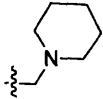
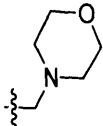
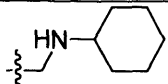


Figure 5-13: NMR-based docked poses of phenylthiophenes **A20** and **2** (green stick representation) in complex with *EcDsbA*. Additionally, the crystal bound pose of **2** (white stick representation) is superimposed with its docked pose. For compounds **A20** and **2**, surface of *EcDsbA* is mapped according to the CSP strength where residues having the largest CSP are coloured red and blue respectively and unassigned residues are coloured in cyan and pink respectively.

Analogues of the 3-phenylthiophene fragment that showed significant CSP and good inhibition of motility were investigated further by recording ^1H - ^{15}N HSQC titrations to measure their affinity and/or titrations in the motility assay to measure their potency. The results of these experiments are summarised in Table 5-5.

| Substituent | 3' | K_D (μM) | EC_{50} (μM) | 4' | K_D (μM) | EC_{50} (μM) |
|---|------------|-------------------------|------------------------------------|------------|-------------------------|------------------------------------|
|  | 2 | 25 ± 9 | 37 ± 1 | 3 | 74 ± 20 | 4 ± 1 |
|  | A16 | n.d. | 183 ± 1 | A19 | n.d. | 180 ± 2 |
|  | A17 | n.d. | n.d. | A20 | 68 ± 25 | 176 ± 1 |
|  | | | | A21 | 124 ± 80 | 172 |
|  | | | | A23 | 116 ± 58 | n.d. |
|  | | | | A25 | 66 ± 40 | n.d. |
|  | | | | A26 | 20 ± 18 | 197 ± 1 |
|  | | | | A28 | 8 ± 14 | 64 ± 2 |

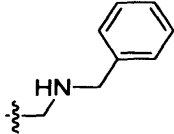
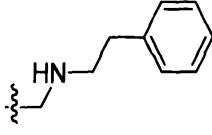
| | | | | | | |
|---|--|--|--|------------|----------|------|
|  | | | | A29 | 106 ± 60 | n.d. |
|  | | | | A31 | 82 ± 47 | |

Table 5-5. Affinities and potencies of the 3-phenylthiophene analogues.

Despite the observation of strong CSP, no enhancement was seen in K_D for a number of polar substituted analogues like the carboxylate or hydroxymethyl in 2 or 3 phenylthiophenes as discussed above. (Table 5-1 and 5-4) Nonetheless, the structural data suggests that in addition to hydrophobic side chains, these polar functionalities might play the role of a molecular anchor that binds the ligand to the *EcDsbA* protein.

The next step was an attempt to explore the steric effects of substituents on the phenyl ring to investigate, firstly whether a substituent is tolerated, and secondly, how large that substituent can be, with the future goal of functionalizing that substituent to make further interactions. The topology of the hydrophobic groove suggests enough space is available to fit larger bulky or hydrophobic groups on the side chain, which may provide favourable interactions. To elaborate hits by further synthesis we selected 3-phenylthiophene fragments **2** and **3** as these hits were found to be the most active *EcDsbA* inhibitors in the motility assay among the hits in this series. X-ray crystal structures of similar fragments **1** and **2** (Figure 5-14) and ^1H - ^{15}N HSQC data helped us to predict the binding mode of fragment **3**. Taken together all this data strongly suggested

that fragment **3** which contains a 4' position side chain and high LE to the *Ec*DsbA has a binding mode very similar to that of fragments **1** and **2**. In house chemical synthesis was carried out by Mr Bradley Doak in order to obtain the desired compounds that were not commercially available and included compounds **A21-A31** with 4' position substitutions (Table 5-4).

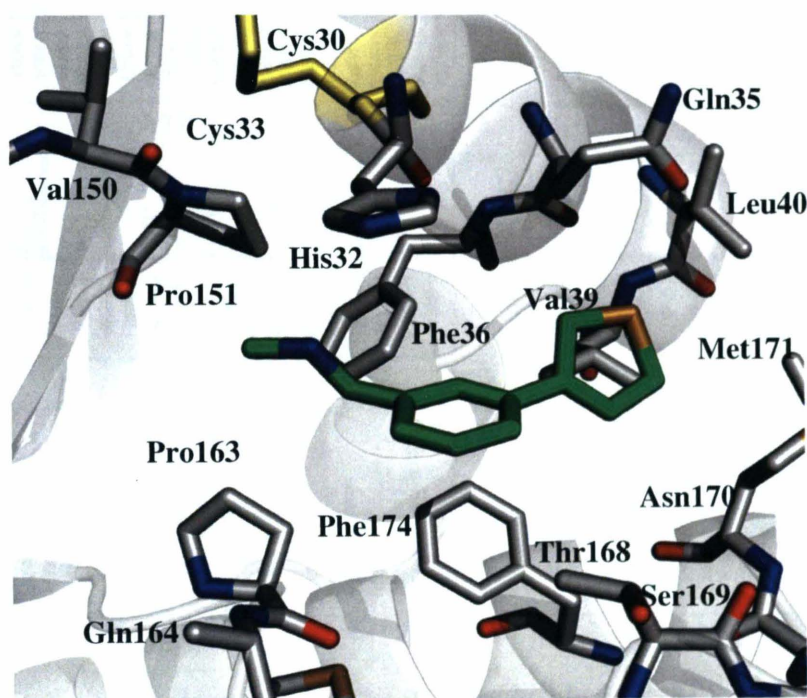


Figure 5-14. Crystal structure of *Ec*DsbA bound to phenylthiophene fragment **2** (green stick representation). Binding was dominated by nonpolar hydrophobic and van der Waals interactions with the surrounding residues (grey stick and ribbon representation) of hydrophobic groove region.

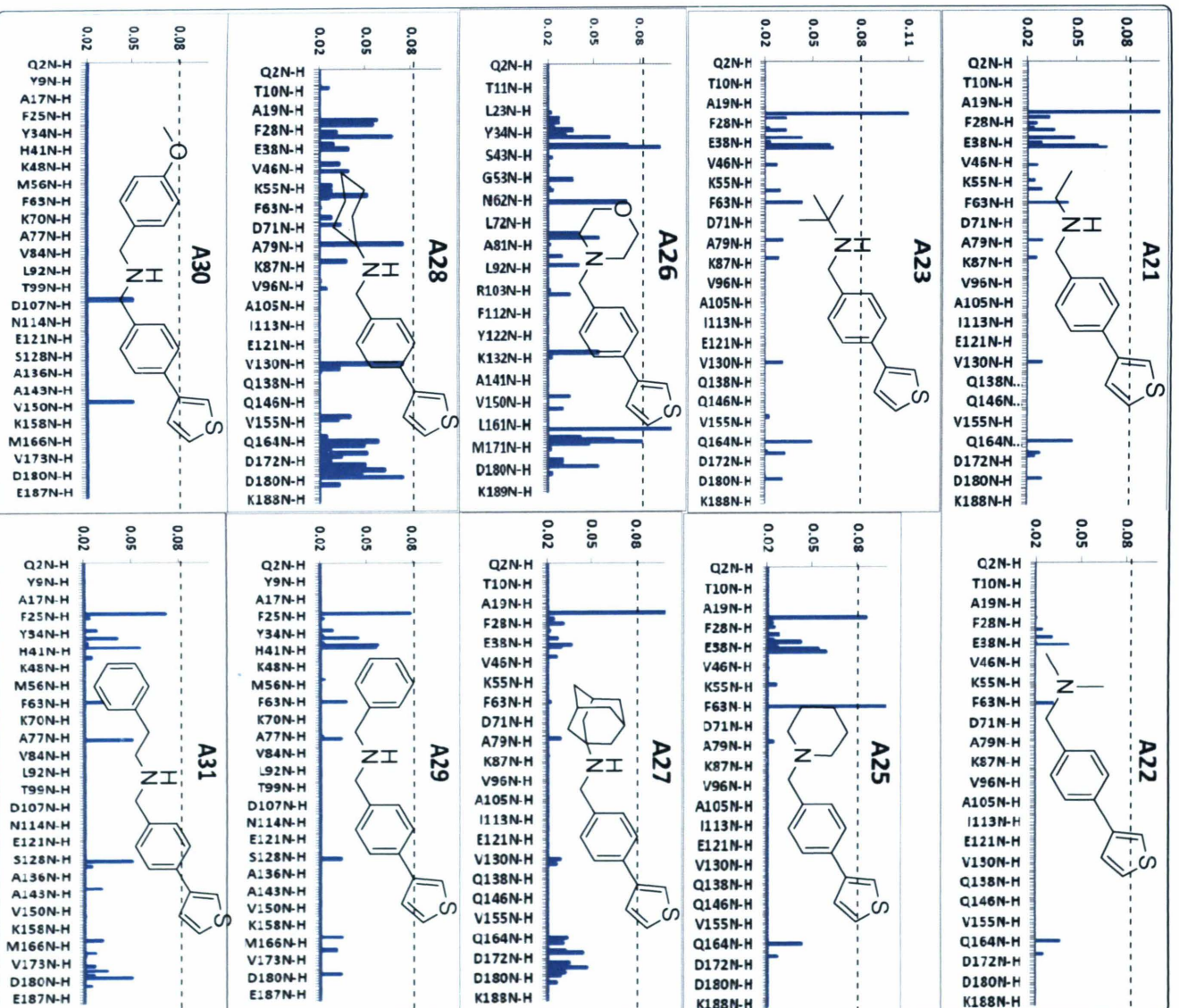


Figure 5-15. CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA*

upon the addition of 4' position sterically substituted analogues of 3-phenylthiophenes; **A21-**

A23 and **A25-A31**.

Substitution with aromatic rings in compounds **A29-A31** resulted in binding affinities similar to hit **3** or worse; however they showed no motility inhibition (Figure 5-15, Table 5-4). Substitution with aliphatic groups in compounds **A21-A23** and **A25-A28** resulted in a range of different binding affinities, from considerably weaker; to approximately a 10-fold improvement. Analysis of NMR binding data suggested that 4' bulkiness is favourable for binding and tolerated up to a cyclohexyl group (**A28**), however replacement with the bulkier adamantyl group (**A27**) showed a reduction in binding. This may suggest the size of adamantane ring is too large to fit in the binding site although poor solubility may also contribute to the loss of binding of compound **A27**. The replacement of the secondary amine with a tertiary amine in the methyl, piperidinyl or morpholino substituted compounds **A22**, **A25** and **A26** resulted in dramatically reduced binding for the methyl **A22**, while the piperidinyl **A25** retained potency. An approximately 4-fold improvement in *EcDsbA* binding affinity was observed in the case of the morpholino analogue **A26** with a K_D of $20 \pm 18 \mu\text{M}$. The pattern of CSP induced by morpholine **A26** along with the NMR-based docking results suggested that its binding location is similar to the original hit, but with a greater extent of CSP both in the groove and active site region. In the case of morpholine **A26** we were successful in increasing the bulkiness of the side chain to gain further binding affinity, while maintaining the solubility by introducing the polar morpholine functionality. Within the amine substituted analogues, cyclohexane **A28** induced the largest CSP and exhibited the high affinity to *EcDsbA*, with a K_D of $8 \mu\text{M} \pm 18 \mu\text{M}$, and also possessed adequate solubility in the NMR sample buffer. Binding of cyclohexyl phenylthiophene **A28** to *EcDsbA* resulted in ^1H - ^{15}N HSQC peaks broadening or disappearance at higher concentrations. This, along with an

intensity reduction for several resonances suggested fast-intermediate exchange binding. **A28** showed 100% inhibition of motility at a concentration of 500 μM and an EC_{50} of 64 μM was determined for this analogue using the motility assay. The discrepancy observed between the K_D and EC_{50} values for cyclohexyl analogue **A28** may be associated with solubility in the different media, adsorption or metabolism in the cell based assay. A comparative analysis of CSP along with NMR-based docking of cyclohexane **A28** and the original hit *N*-methyl methanamine fragment **2** showed binding at similar position although with more hydrophobic contacts in the groove region causing a greater number of shifts in this region. (Figure 5-15 and 5-16)

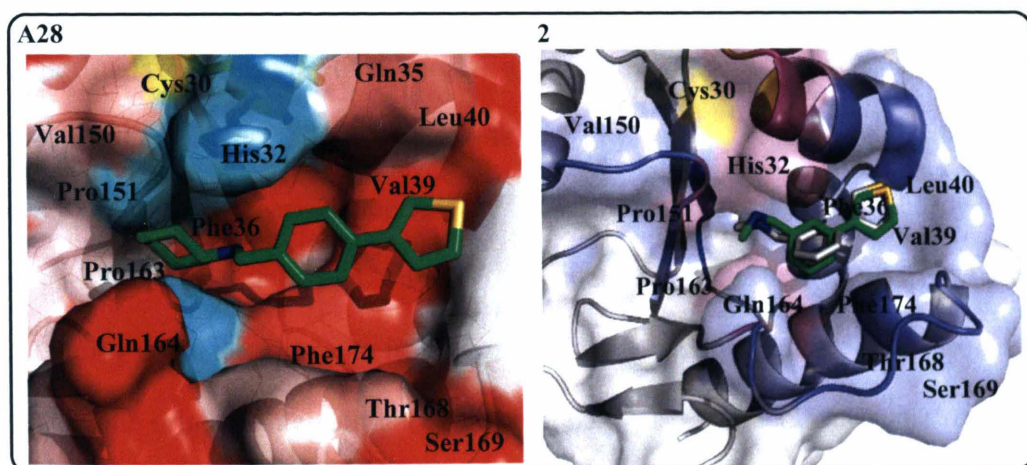


Figure 5-16: NMR-based docked poses of phenylthiophenes **A28** and **2** (green stick representation) in complex with *EcDsbA*. Additionally, the crystal bound pose of **2** (white stick representation) is superimposed with its docked pose. For compounds **A28** and **1**, surface of *EcDsbA* is mapped according to the CSP strength where residues having the largest CSP are coloured red and blue respectively and unassigned residues are coloured in cyan and pink respectively.

This binding hypothesis was confirmed by solving the X-ray crystal structure of cyclohexane **A28** in complex with *EcDsbA*. The structure of this complex was solved to 2.28 Å resolution by employing a crystal soaking method. The omit electron density maps (Fo-Fc) showed clear binding for this ligand with the flexible cyclohexyl ring at an occupancy of 1.0 (Figure 5-17). This crystallographic data revealed the position of the phenylthiophene moiety for compound **A28** was overlapping and but slightly shifted towards the open end of groove as compared to parent phenylthiophene fragment **2**. (Figure 5-18) The docked solution was also found in good agreement with crystallographic data further suggesting the reliability of NMR-based docking. Comparative analysis of the crystal structures of phenylthiophenes **A28**, **1** and **2** revealed that the imidazole ring of His32 moved upon the binding of **A28** (Figure 5-17). The shortest distance from the His32 imidazole ring to the phenyl ring in parent fragment **2** was 4.93 Å. In comparison the shifted imidazole and the phenyl of the cyclohexyl analogue **A28** was 3.39 Å apart. The His32 orientation changes may result from analogue **A28** binding down the groove to form better vdW interactions. Figure 5-18 illustrates that similar to the original fragment hit **2** this compound also interacts with most residues of the hydrophobic groove such as those that participate in the binding of *EcDsbA* with membrane protein *EcDsbB*. Cyclohexyl analogue **A28** binding to *EcDsbA* caused the displacement of 3 additional water molecules (compared to fragment **1-2**) along with significant changes in side chain conformations. This structure revealed that the cyclohexyl component of this compound interacted well with surrounding hydrophobic residues Pro163, Pro151, Val150 in the groove, thus contributing to an increase in affinity compared to fragment **3**.

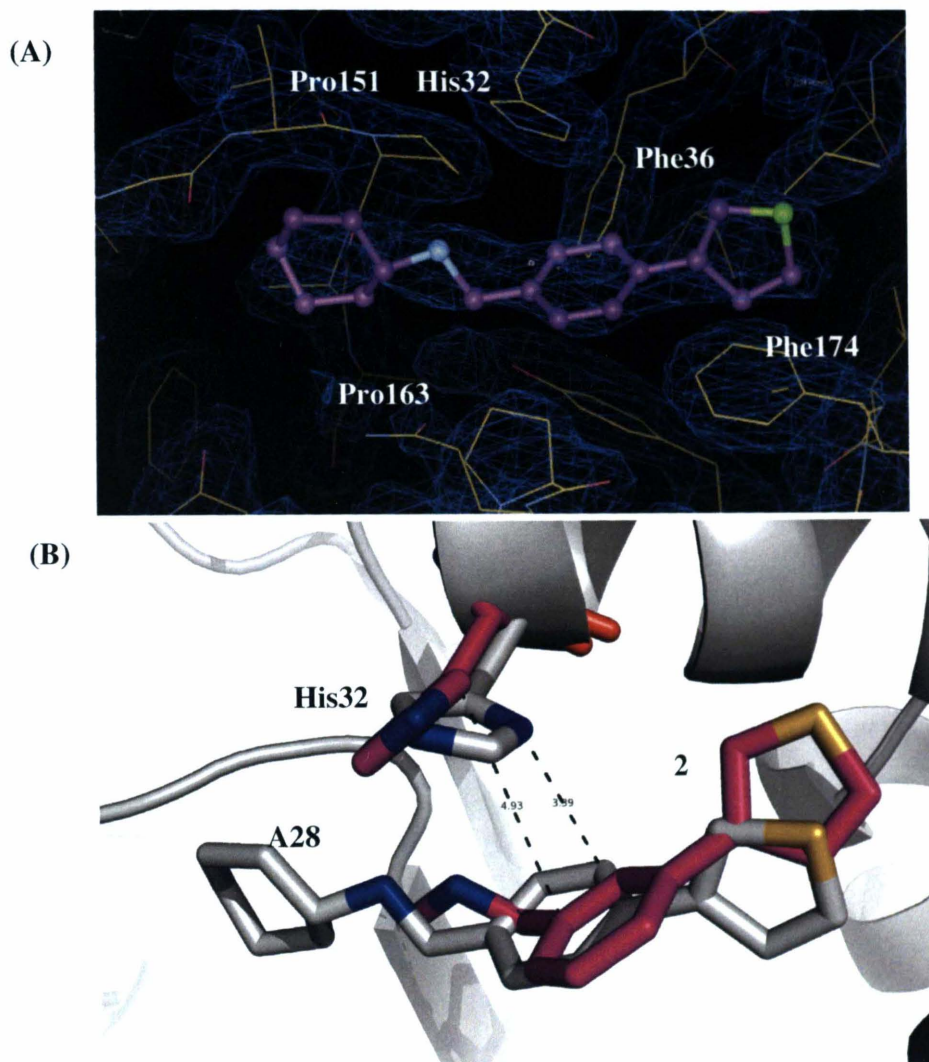


Figure 5- 17: (A): (2Fo - Fc) electron density maps for the *EcDsbA*- cyclohexane analogue **A28** complex (contoured at σ level 1.0) (B) Overlay view of cyclohexane **A28** and original hit *N*-methyl methanamine fragment **2** with active site residue His32 which undergoes a conformational change.

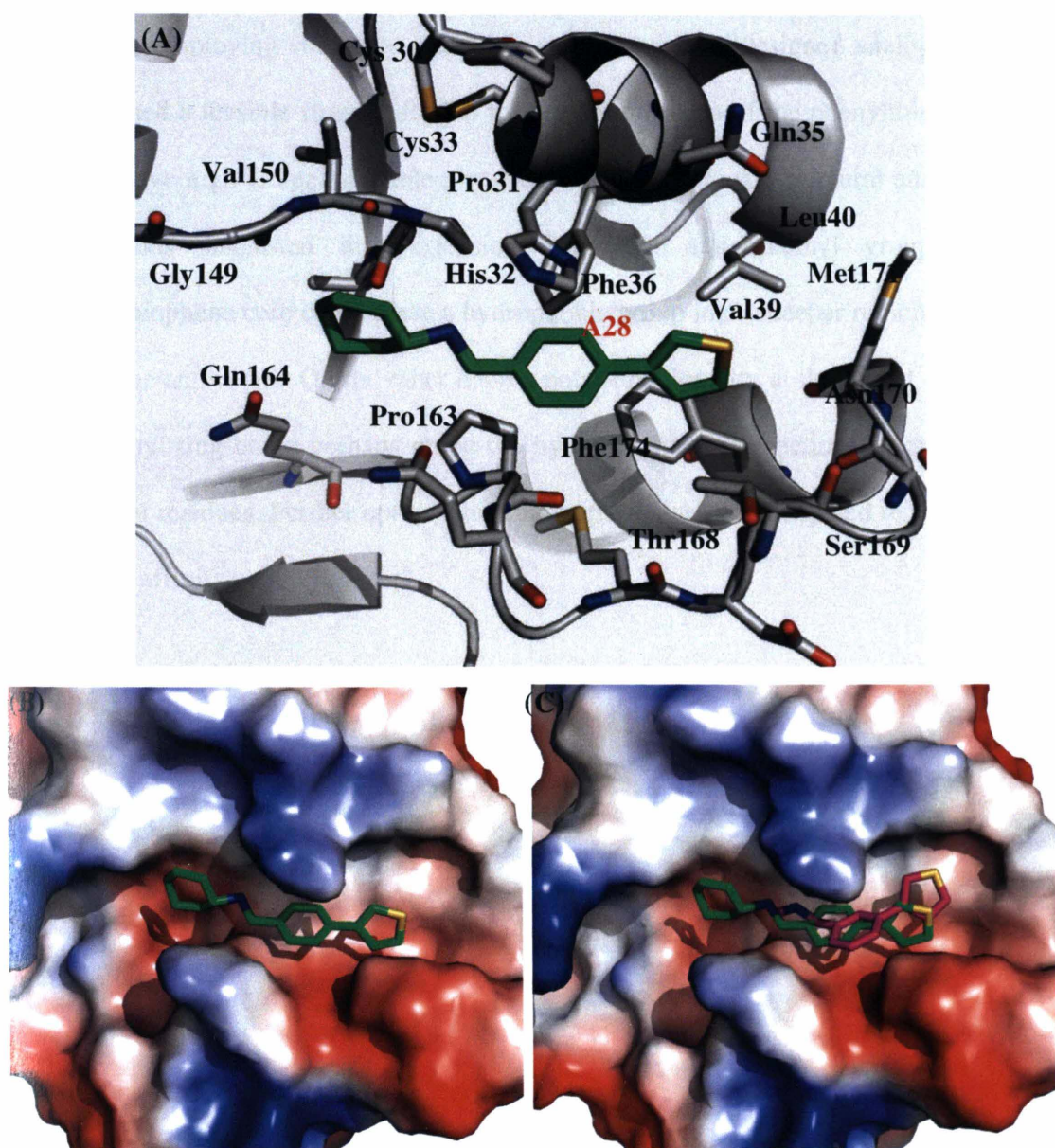


Figure 5-18: (A) Crystallographic binding mode of phenylthiophene cyclohexyl analogue A28 (green stick representation) with *EcDsbA* (grey stick and ribbon representation) (B) Crystallographic position of compound A28 on *EcDsbA* electrostatic surface (C) Comparative view of the crystallographic positions of each compound in the hydrophobic groove of *EcDsbA* shown as an electrostatic surface.

By employing similarity searching and synthesis of designed analogues, we have established a feasible strategy for the further optimization of the phenylthiophenes series which have high LE and suitable physiochemical properties. Structural analysis of these compounds suggested that extension from the aminomethyl group of the 3-phenylthiophene core could place a hydrophobic group into a deeper region of the groove below the active site. On the other hand a polar functionality at the 3' and 4' positions of the phenyl ring could perhaps probe the hydrogen bond interaction to active site His32 and other residues. Further optimisation along these lines is anticipated to result in further gains in affinity.

5.2.5 Analogues of the phenoxybenzene series

5.2.5.1 Characterization of the phenoxybenzene hit

In Chapter 3, the identification of another interesting *EcDsbA* inhibiting fragment containing a phenoxybenzene core was described. Structurally, these compounds are more flexible than the phenylthiophenes and interestingly, the most potent fragment hit (phenoxybenzene **4**) (Figure 5-1) of this class has a similar *N*-methyl methanamine substituent at the 4-position of one phenyl ring whilst the second hit fragment **5** contains a 4-position methanol substituent (Figure 5-19). Phenoxybenzene compounds **4** and **5** showed a very similar binding profile by inducing large CSP (up to 0.15 ppm) in the ^1H - ^{15}N HSQC spectra of oxidized *EcDsbA* within the hydrophobic groove and active site region (Figure 5-19). Fragments **4** and **5** have measured K_D 's of 52 μM and 58 μM and LE's of 1.5 and 1.67 $\text{kJ mol}^{-1} \text{HAC}^{-1}$ respectively. They showed inhibition in the motility assay with measured EC_{50} values of 21 μM and 60 μM , respectively. Whilst there is good agreement between the K_D and EC_{50} values for fragment **5**, the subtle discrepancy observed with fragment **4** may indicate that it is inhibiting motility in a non-*EcDsbA*-dependent manner.

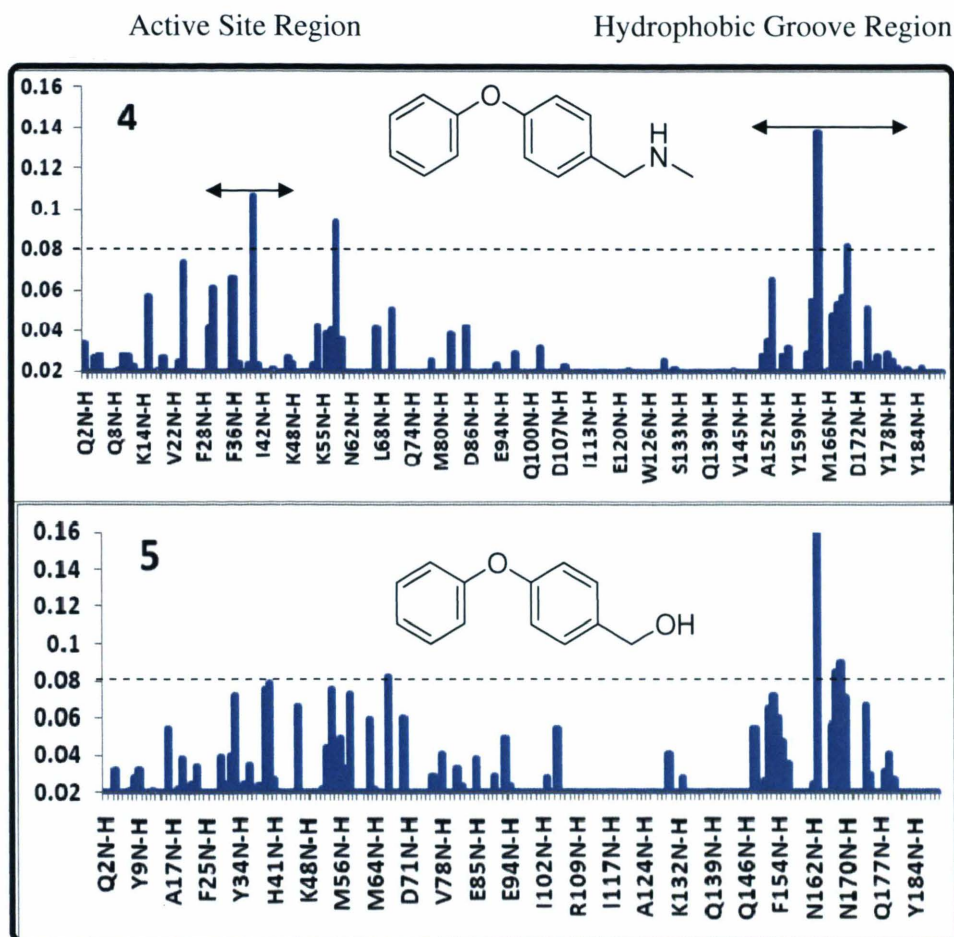


Figure 5-19: Phenoxybenzene class of *EcDsbA* hits identified by NMR-based fragment screening. Significant $\Delta\delta(\text{NH})$ CSP > 0.02 ppm induced in assigned ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of fragments **4** and **5**.

Binding information obtained from ^1H - ^{15}N HSQC experiments for this class was confirmed by solving a 1.9 Å crystal structure of *N*-methyl methanamine compound **4** in complex with *Ec*DsbA. The X-ray structure confirmed that this compound bound in the hydrophobic groove of the thioredoxin domain of *Ec*DsbA in a very similar position to the previously described phenylthiophene compounds **1** and **2**.

Despite several attempts to soak crystals of *Ec*DsbA with fragment **5**, no structures could be obtained. Based on the similarity of the CSP observed for these two fragments a similar binding mode was inferred. Fragment **4** bound directly below the active-site disulfide of *Ec*DsbA at the midpoint of the hydrophobic groove as illustrated in Figure 5-20. Similar to the phenylthiophenes this binding was dominated by nonpolar hydrophobic and vdW interactions with the surrounding residues.

Interestingly, a second crystal structure of fragment **4** was obtained using a shorter soaking time. This structure revealed an alternate orientation of binding with molecule flipped in the opposite orientation, however with the molecule overall occupying a similar position within the groove.. In this case the side chain makes vdW contacts with Gln35, Val39 and Leu40 surrounding the active site. Overall these data revealed two conformations for binding of compound **4** with *Ec*DsbA and indicated the possibility of structure-based elaboration of this core in either direction. Therefore, fragments **4** and **5**, which possess a high LE of 1.5 and 1.67 kJ mol⁻¹ HAC⁻¹, respectively were also considered to be good starting points for further chemical optimization to potent leads.

(A) Crystallographic binding mode of phenylthiophene cyclohexyl analogue **A28** (green stick representation) with *Ec*DsbA (grey stick and ribbon representation) (B)

Crystallographic position of compound **A28** on *EcDsbA* electrostatic surface (C)

Comparative view of the crystallographic positions of each compound in the hydrophobic groove of *EcDsbA* shown as an electrostatic surface.

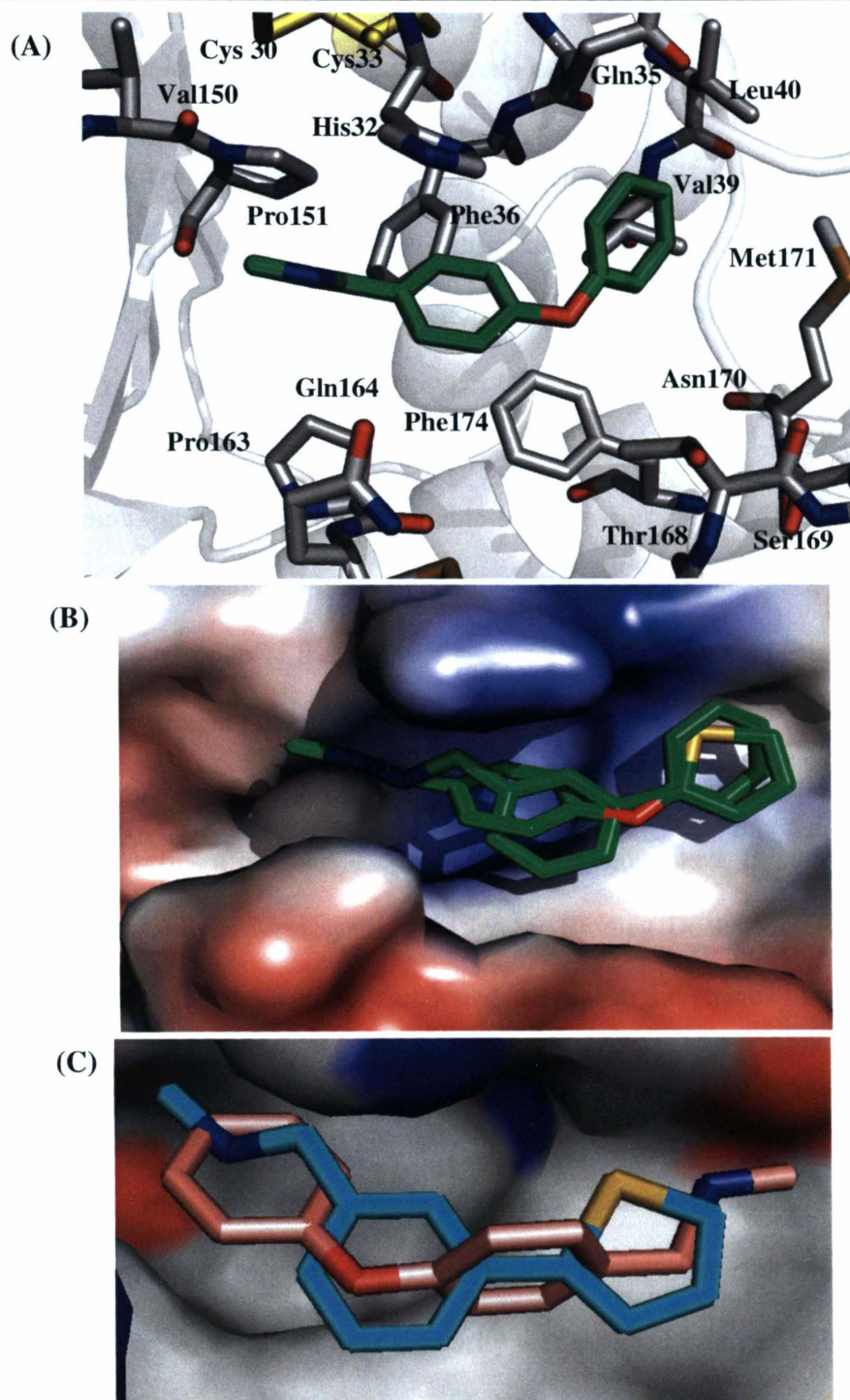


Figure 5-20: (A) Crystallographic binding of phenoxybenzene **4** (green stick) with *EcDsbA* (grey stick and ribbon) (B) Phenylthiophene **1** and **4** (green stick) shown overlapped in the *EcDsbA* (electrostatic surface) (C) Alternative crystallographic orientation of fragment **4** (pink stick) shown overlapped with fragment **1** (cyan stick).

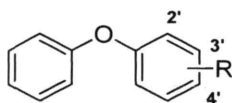
5.2.5.2 **Elaboration of phenoxybenzene hits 4 and 5**

The X-ray crystal structure of phenoxybenzene **4** suggested a potential strategy for optimization of binding to *EcDsbA* for this class of compounds. Figure 5-20 shows two areas in which it might be possible to improve the interaction of the fragments with *EcDsbA*. Firstly, there is a significant region of the groove below the fragment that is not occupied. It was rationalised that by functionalising the 2' and 4' positions of phenyl rings with various groups it may be possible to more completely fill this pocket and thereby increase the binding affinity. Secondly, there is additional space to extend the fragment from the nitrogen of the methanamine, which may pick up further interactions in one of the two different binding modes observed.

To test these binding possibilities and in order to optimize the affinity of the phenoxybenzene series a number of approaches were undertaken. Firstly, a range of *N* substituted methanamines were synthesised by Mr Bradley Doak (Monash Institute of Pharmaceutical Sciences). Secondly, compounds similar to the phenoxybenzene hits that maintained the diphenyl ether core possessing a diverse range of substituents at various positions of the phenyl ring were sourced from either commercial or in-house collections. Finally, alternate ring systems consisting of linear or non-linear 6-6 or 5-6 linked ring systems were purchased for screening. Often for these compounds the oxygen atom was replaced with a carbon atom and their similarity was assessed using a Tanimoto score. A Tanimoto similarity threshold of 0.35 was set as the cut-off based on the literature.²²⁹

Twenty five new compounds were chosen via this process for evaluation and were first tested for solubility under the conditions employed in the ¹H-¹⁵N HSQC and motility

assays. Most of these selected analogues showed good solubility in the minimal media used for the motility assay and were soluble in the NMR sample buffer. Hence end-point ^1H - ^{15}N HSQC spectra were recorded for all compounds, with the resulting spectra being ranked as (strong, +++), (medium, ++), (weak, +) or (no binding -) on the basis of a threshold applied for $\text{CSP} \geq 0.03$ ppm as described above. The effect of the compounds on motility was determined by undertaking the motility assay in the presence of a single concentration of the compound (500 μM), and measuring the percentage inhibition of motility. For those compounds that gave the strongest CSP (+++), the K_D was determined from the ^1H - ^{15}N HSQC titration experiment, and for compounds that displayed 100% inhibition of motility at 500 μM , an EC_{50} was measured. The structures of the phenoxybenzene analogues that were purchased (**E1-E3**) and synthesized (**E4-E8**) are shown in Table 5-6. The structures of similarity based molecules (**SE1-SE17**) are shown in Table 5-8.



| Substituent | 2' | | 4' | | Substituent | 4' | |
|-------------|-----------|-------------|-----------|-------------|-------------|-----------|-------------|
| | E1 | ++ 100% | 4 | +++ 100% | | E4 | ++ 50% |
| | E2 | +++ 100% | 5 | +++ 100% | | E5 | ++ 40% |
| | | | E3 | +++ 100% | | E6 | ++ 60% |
| | | | | | | E7 | + 20% |
| | | | | | | E8 | +++ 100% |

Table 5-6. Analogues of phenoxybenzene fragment **4**. The table shows the chemical structure of the substituent and the position (2' or 4') at which the substituent was attached to the phenyl ring. The compound number, the magnitude of CSP observed in the end point ^1H - ^{15}N HSQC along with the % inhibition of motility caused by the fragment at a concentration of 500 μM are listed. Fragments highlighted in grey were selected for further characterisation.



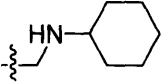
| Substituent | 2' | K _D (μM) | EC ₅₀ (μM) | 4' | K _D (μM) | EC ₅₀ (μM) |
|---|-----------|---------------------|-----------------------|-----------|---------------------|-----------------------|
|  | E1 | n.d | 160 ± 1 | 4 | 52 ± 14 | 20 ± 1 |
|  | E2 | 716 ± 65 | 250 ± 1 | 5 | 60 ± 8 | 60 ± 1 |
|  | | | | E8 | 3 ± 3 | 72 ± 2 |

Table 5-7. Affinities and potencies of selected phenoxybenzene analogues.

Table 5-6 reveals that several different functional groups are tolerated at each position 2' or 4' position on the phenyl ring. Initially *N*-substitution in the 4-position was investigated by NMR. Analysis of the patterns of perturbations observed across the phenoxybenzene analogues **E4-E8** revealed that the CSP data were consistent with binding in the hydrophobic groove of *EcDsbA* as had been observed with the parent diphenyl ether **4**. These compounds were synthesized with increasingly bulky substituents from methyl to cyclohexyl. Aromatic ring substitution in compound **E7** resulted in almost complete loss of binding, while dimethyl and ethyl substituted compounds (**E4** and **E6**, respectively) showed decreased binding. The cyclohexyl substituted compound **E8** showed stronger perturbations (Figure 5-21). This binding profile and preference of

substitution is somewhat similar to the pattern observed with the 4-position of the substituted phenylthiophene series.

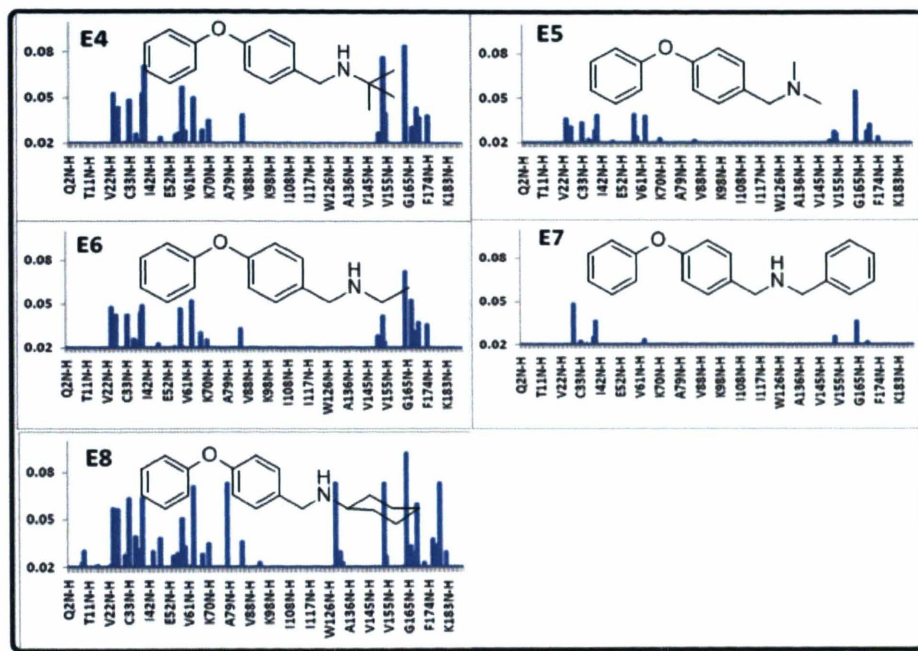


Figure 5-21: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of 4' substituted phenoxybenzene analogues **E4-E8**.

Across this series, cyclohexyl analogue **E8** induced the largest CSP with and approximately 27-fold increase in binding affinity ($K_D \sim 3 \mu\text{M}$) and was identified as the most potent analogue in this study. Disappearing peaks and peak broadening in the ^1H - ^{15}N HSQC spectrum suggested fast-intermediate exchange binding. This compound showed inhibition with an EC_{50} of $72 \mu\text{M}$ in bacterial motility assay. The overall binding profile of analogue **E8** is quite similar to the most potent cyclohexyl compound **A28** in the phenylthiophene series. Analysis of the induced CSP suggested that the binding location of **E8** was similar to initial phenoxybenzene fragment **4**, although possibly with

more hydrophobic contacts in the hydrophobic groove due to the larger steric size of the cyclohexyl group. We were able to solve a 2.2 Å X-ray crystal structure of the *EcDsbA* bound cyclohexyl substituted phenoxybenzene compound **E8** (Figure 5-22). Consistent with the NMR experiments, extra electron density indicated crystallographic binding into the midpoint of hydrophobic groove of *EcDsbA* similar to phenylthiophene cyclohexane **A28**. The electron density maps ($2Fo-Fc$) showed a clearly defined binding mode for this flexible analogue at occupancy 1.0 (Figure 5-23). Binding of analogue **E8** is comparable to the first crystallographic orientation of parent phenoxybenzene fragment **4** and thiophene **A28** due to similar orientation of core and 4-position substitution in all these cases. As shown in Figure 5-22 and 5-23 the orientation of the 4-cyclohexyl group of **E8** is at the end of the hydrophobic groove near active site *cis*-Pro151. Prior to obtaining this crystal complex of **E8** with *EcDsbA* NMR binding data was used to dock the compound **E8** in the binding pocket of *EcDsbA* to study its binding mode as shown in Figure 5-10. The docked solution was found to be in good agreement with the crystallographic data further suggesting the reliability of NMR-based docking to study other analogues where the crystal complexes are not available.

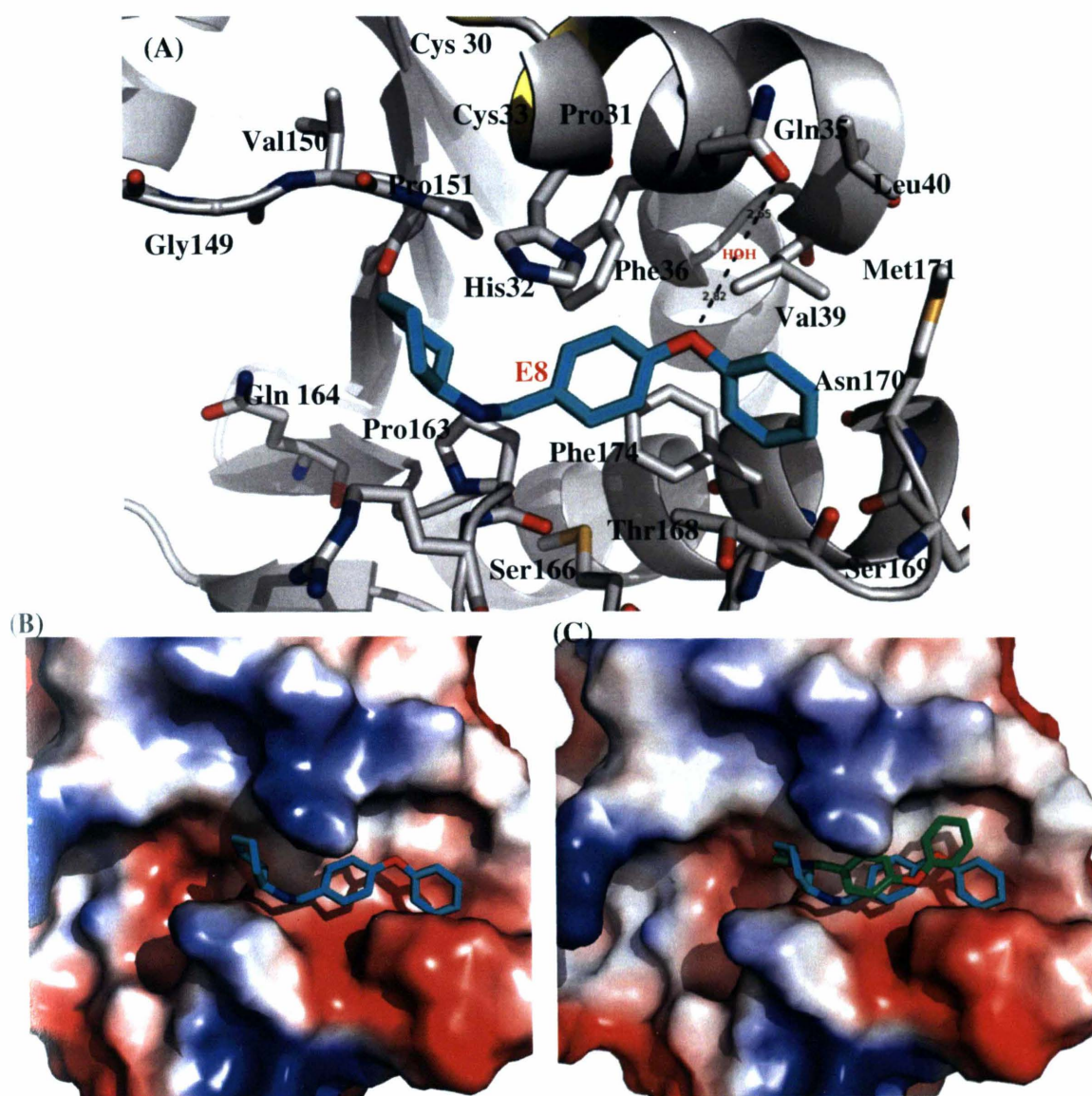


Figure 5-22: (A) Crystallographic binding mode of phenoxyphenyl analogue **E8** (cyan stick representation) with *EcDsbA* (grey stick and ribbon representation) (B) Crystallographic position of compound **E8** (cyan stick representation) on an electrostatic surface of the *EcDsbA* (C) Comparative view of crystallographic positions of compound **E8** with parent fragment **4** (green stick representation) in overlapping and adjacent binding sites on the hydrophobic groove of the *EcDsbA* surface.

The binding mode of the phenoxybenzene analogue **E8** can be outlined using the X-ray crystal structure and the results of the docking and NMR binding data are in agreement with this structure. Figure 5-22 illustrates the position of the phenoxybenzene core of **E8** compared to that of fragment **4** where it is positioned towards the open end of groove. The cyclohexyl ring was located towards the active site His32 and resulted in the imidazole ring moving upto 2.9 Å away compared to apo *EcDsbA*. Without this induced fit, the cyclohexyl ring would ordinarily be placed in a position that would clash with His32. As a result the electron density of His32 was observed to move resulting in an alternate conformation.

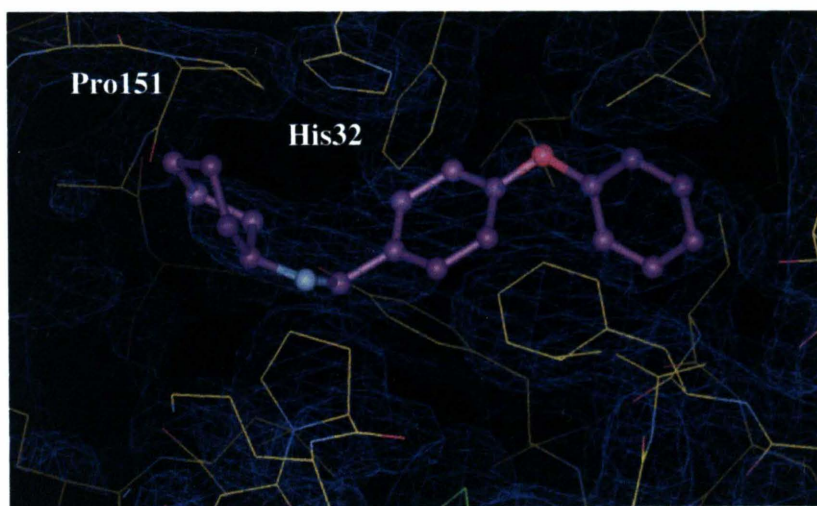


Figure 5- 23: Electron density ($2F_o - F_c$) maps for the *EcDsbA*-fragment **E8** complex (contoured at σ level 1.0).

E8 binding to *EcDsbA* caused displacement of 3 additional water molecules (compared to fragment **1-2**) along with significant changes in side chain conformations. Additionally, this crystal structure revealed a water mediated hydrogen bond between the oxygen of the phenoxybenzene core of **E8** (2.82 Å H₂O 2.65 Å) to Gln35 (Figure 5-22). The vast bulk of the interactions between *EcDsbA* and compound **E8** are apparently hydrophobic interactions. Similar to the original fragment **4**, this compound also interacts with many of the residues of the hydrophobic groove that participate in binding of *EcDsbA* with membrane protein *EcDsbB*. The structure revealed that the cyclohexyl component of this compound interacted with surrounding hydrophobic residues Pro163, Pro151, Val150 in the groove. This contributed to an increase in affinity relative to fragment **4** and made a number of interactions across the hydrophobic groove (Figure 5-22).

On the other hand, analysis of the pattern of CSP of the phenoxybenzene analogues **E1 - E3**, **4** and **5** suggested that generally there is no clear preference for the polar substitution of the phenyl ring as all of these analogues induced strong CSP (Figure 5-24). Alcohol **E2** consists of a 2'-position methanol substitution on the original hit **5** and showed a similar binding pattern with CSP up to 0.15 ppm. This compound induced large CSP in residues positioned around the active site along with higher shifts in groove region Gln164, Thr168. Additionally many shifts changes were observed in the surrounding region possibly due to large conformational changes in the flexible groove region. However the overall decrease in K_D and LE in this case is indicative of the fact

that ^1H - ^{15}N HSQC CSP strength and binding affinity of the ligand do not necessarily correlate.

Compared to 4'- *N*-methanamine compound **4** the 2' substitution in analogue **E1** resulted in the loss of interactions in the groove region, with fewer interactions around the active site (Figure 5-19 and 5-24). Interestingly, the CSP mapping of **E1** supports the second conformation of fragment **4** in the crystal structure where the side chain interacts with residues Gln35 and Leu40 surrounding the active site. In this orientation the side chain at the 2'-position is likely to make fewer interactions with residues in the groove. Removal of the second phenyl ring in compound **SE3** causes the loss of CSP in the groove, again supporting the second conformation of fragment **4** as illustrated in Figure 5-24.

A subtle improvement was observed in binding affinity for the similarity-based compound **SE1**, where the second phenyl ring in parent fragment **5** is replaced by a pyridine ring. (Figure 5-25) Pyridine **SE1** induced up to 0.05 ppm higher CSP in the active site region, however lesser shifts in the groove region as compared to parent fragment **5**. This compound showed good solubility and also inhibited bacterial motility with an EC_{50} value of 71 μM , suggesting it could be a useful starting point in designing better compounds.

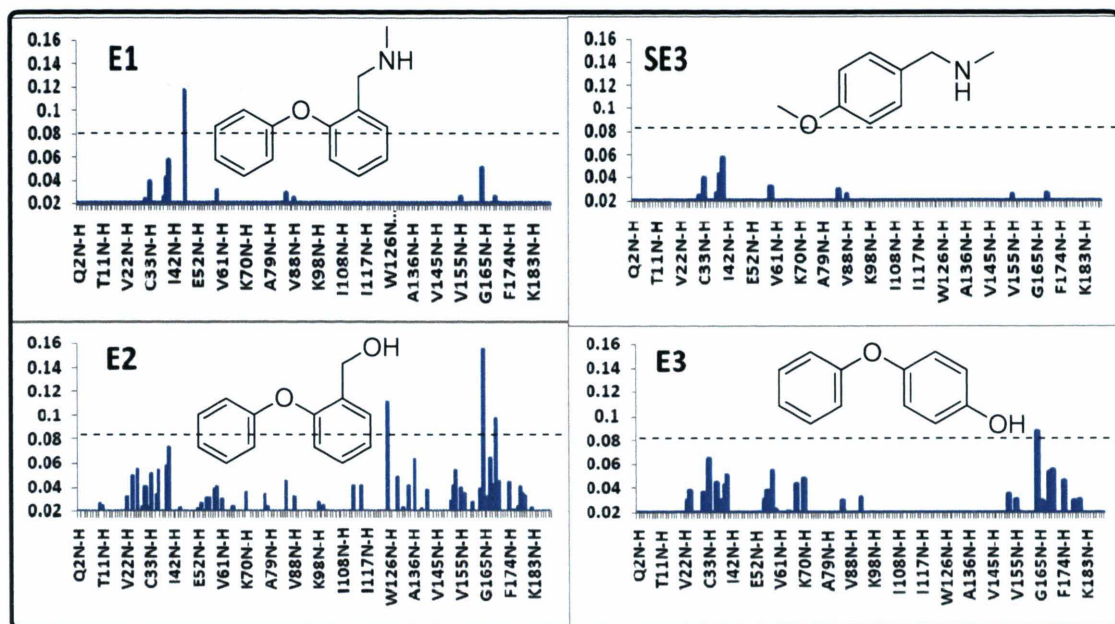


Figure 5-24: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of analogues **E1-E3** and **SE3**.

Following the similarity search process to identify ligand for screening, an interesting molecule was identified which had a methylene linked phenylthiophene **SE12**. (Table 5-8) Compound **SE12** can be considered to be a combination of the features of both the phenylthiophene and phenoxybenzene series where one phenyl ring is replaced by a thiophene ring and the oxygen linker is replaced by a carbon atom (Tanimoto similarity score 0.5 with phenoxybenzene core and 0.63 with phenylthiophene). Additionally, the 4'-position side chain was replaced by an amino group. Moderately strong CSP were observed in the hydrophobic groove and active site regions upon **SE12**-*EcDsbA* complex formation in solution (Figure 5-23).

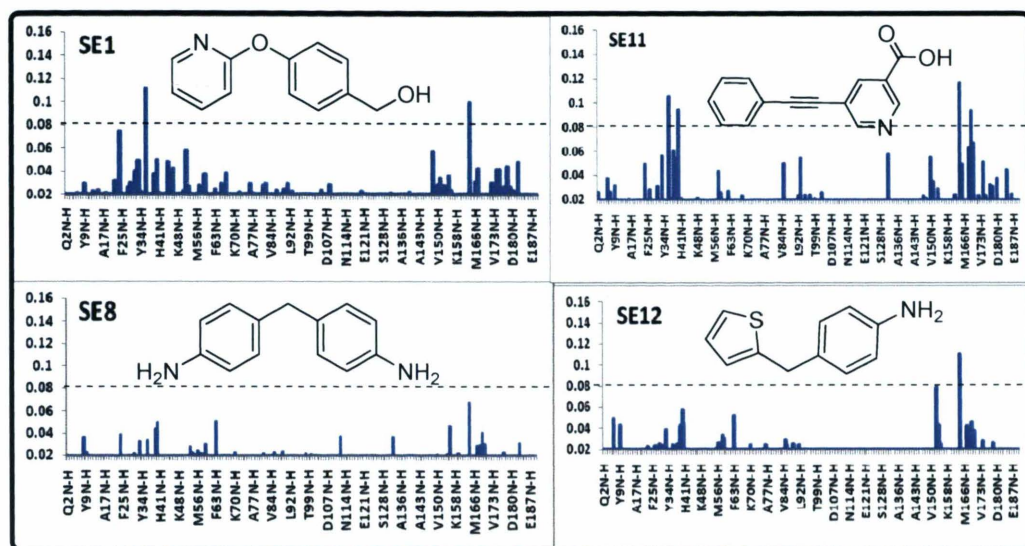


Figure 5-25: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of phenoxybenzene similarity molecules **SE1**, **SE8**, **SE11** and **SE12**.

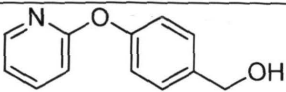
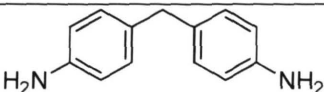
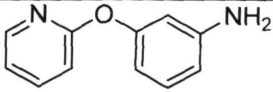
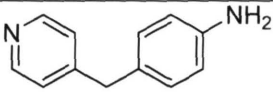
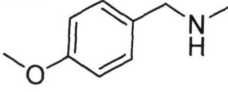
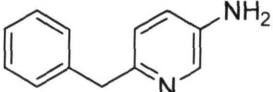
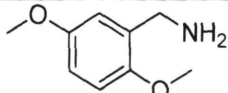
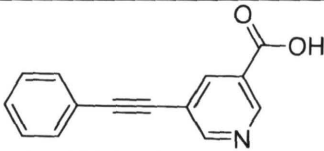
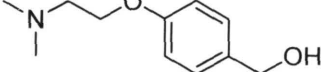
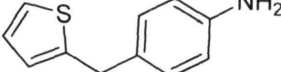
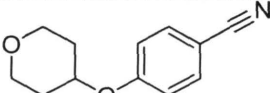
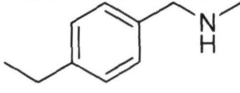
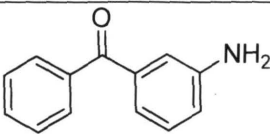
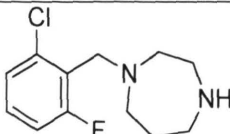
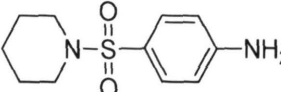
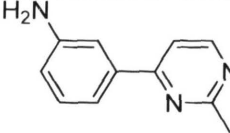
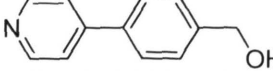
| Compound | Profile | | Compound | Profile | |
|---|---------|-------------|---|---------|-------------|
|  | SE1 | +++ 100% |  | SE8 | +++ 100% |
|  | SE2 | + 20% |  | SE9 | - 0% |
|  | SE3 | + 40% |  | SE10 | ++ 70% |
|  | SE4 | + 70% |  | SE11 | +++ 0% |
|  | SE5 | + 0% |  | SE12 | +++ 100% |
|  | SE6 | + 40% |  | SE13 | + 20% |
|  | SE7 | - 0% |  | SE14 | - - |
| | | |  | SE15 | + 0% |
| | | |  | SE16 | + 10% |
| | | |  | SE17 | ++ 40% |

Table 5-8. Molecules selected on the basis of similarity to phenoxybenzene fragment.

This table shows the compounds with minimum Tanimoto similarity scores of 0.35 with fragment **4** and contains various heterocyclic rings or linkers in 5-6 ring systems. The

compound number, the magnitude of CSP observed in the end point ^1H - ^{15}N HSQC along with the % inhibition of motility caused by the fragment at a concentration of 500 μM are listed. Compounds highlighted in grey were selected for further characterisation.

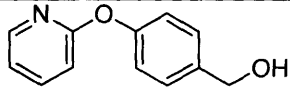
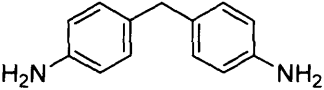
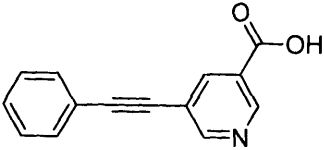
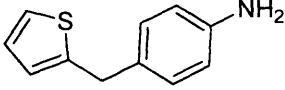
| Compound | | K_D (μM) | EC_{50} (μM) |
|---|------|-------------------------|------------------------------------|
|  | SE1 | 40 ± 15 | 70 ± 1 |
|  | SE8 | 70 ± 20 | 180 ± 1 |
|  | SE11 | 60 ± 5 | n.d |
|  | SE12 | 8 ± 8 | 18 ± 2 |

Table 5-9. Affinities and potencies of the similarity molecules of the phenoxybenzene fragment.

This compound **SE12** bound to *EcDsbA* with about a 10-fold higher affinity and also higher LE $2.2 \text{ kJmol}^{-1}\text{HAC}^{-1}$ as compared to the original phenylthiophene hit fragments **1-3** and the phenoxybenzene fragment **4**. Additionally, **SE12** shows strong inhibition with an EC_{50} value of $18 \text{ }\mu\text{M}$ along with good solubility in all tested media/buffers. Overall **SE12**, which combines features of the most potent original *EcDsbA* binding fragments **1-3** and **4** showed a comparatively better binding, LE and inhibition profile. NMR based docking suggested its binding in *EcDsbA* hydrophobic groove in a similar orientation to the original hits (Figure 5-10).

NMR binding and docking studies were later confirmed by a crystal structure of $2.2 \text{ }\text{\AA}$ which revealed density for compound **SE12** in a similar area across the hydrophobic groove as observed in crystallographic binding of the original hit phenylthiophene fragments **1**, **2** and phenoxybenzene **4** (Figure 5-26). This complex revealed two overlapping conformations of binding across the active site and groove region. As illustrated in Figure 5-26 these two conformations contains almost similar and overlapping position of 4' phenyl position amino group of **SE12** which is in between Phe36 and Phe174 and points towards open end of the groove. The thiophene ring in these conformations adopts two opposite orientations where the first is bent towards the lower end of groove near Asn170 and second is bent towards the upper end closer to Gln35 (Figure 5-26). Additionally, NMR-based docking closely supports the former orientation of **SE12**. The possibility remains that these two conformations of **SE12** are a consequence of the molecule's flexibility and two conformations of this ligand are possible during binding in the hydrophobic groove.

Overall the crystallographic complex showed a complimentary match between the core scaffolds and the hydrophobic pocket located in groove region of *EcDsbA* (Figure 5-26). Importantly, in addition to vdW and hydrophobic contacts, the amino group of **SE12** made a water mediated hydrogen bond to a structural water molecule HOH 307 (Figure 5-27). HOH 307 is located within hydrogen bonding distances to Pro151, Pro163 and His32 along with two other water molecules. This polar interaction by **SE12** in the *EcDsbA* groove might be an important factor for high affinity and inhibition capability of this compound. Therefore this inhibitor is an excellent starting point for further lead optimization as it contains binding and structural features of two of the most promising *EcDsbA* binding scaffolds along with demonstrating an improvement in polar contacts with the binding site.

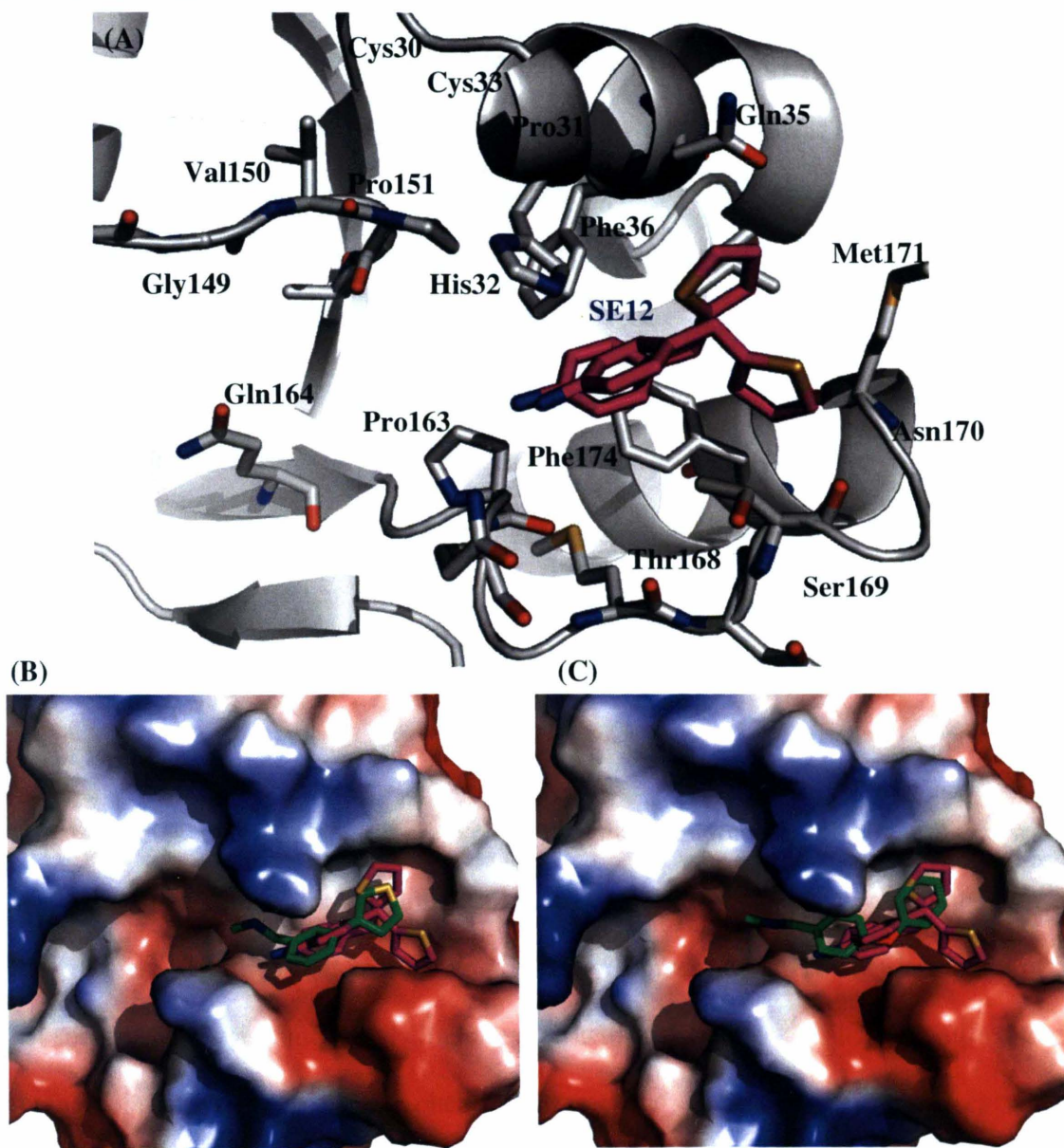


Figure 5-26: (A) Crystallographic binding mode of **SE12** (pink stick representation) with *EcDsbA* (grey stick and ribbon representation). Crystallographic position of compound **SE12** on the *EcDsbA* electrostatic surface in overlapping positions of (B) phenylthiophene fragment **2** (green stick) and (C) the phenoxyphenyl fragment **4** (green stick) on the hydrophobic groove of the *EcDsbA* surface.

Overall we have identified the phenoxybenzene chemical class of *EcDsbA* inhibitors and characterized analogues with low micro molar dissociation constants. Compounds **SE1** and **E8** show promise for further optimization to produce potent small molecule inhibitors of *EcDsbA*. The crystallographic binding mode suggested that a 4' phenyl cyclohexyl group extension off the phenoxybenzene core could add steric bulk and place the molecule into a deeper region of groove. On the other hand, replacement of phenyl with pyridine could provide a hydrogen bond interaction to active site residues. Additionally, identification of **SE12** that showed good LE, may also represent a useful starting point for drug design.

5.3 Conclusion

In these studies, the hydrophobic groove and active site region of *EcDsbA* structure was thoroughly characterised for small molecule binding. Preliminary structure-activity based features were identified that will facilitate the development of novel *EcDsbA* inhibitors that specifically target functionally important protein surface sites. We have identified the optimal binding features of the phenoxybenzene and phenylthiophenes classes with versatile scaffolds **A7**, **A26**, **A28**, **SE1**, **E8** and **SE12**. These were shown to possess a high ligand efficiency and low μM affinity and to exhibit inhibitory activity against an *EcDsbA*-dependent phenotype. Important structural and binding features such as solvent replacement and key polar and hydrophobic contacts made by these compounds indicates these ligands have provided excellent insights in the development of potential *EcDsbA* inhibitors.

Chapter 6

Structural Characterisation of Diverse Scaffolds Binding to Adjacent Sites of *Ec*DsbA

6.1 Introduction

In chapter 5, it was shown that *EcDsbA* binds analogues of phenylthiophenes and phenoxybenzenes with binding affinities varying from 2 μ M - 1 mM. The majority of these molecules appeared to bind in a broadly similar mode to *EcDsbA*, possessing hydrophobic moieties and polar functionalities that make interactions in the hydrophobic groove region of the protein. This small molecule binding region of *EcDsbA* is the same region that interacts with the membrane based protein *EcDsbB*¹⁷ during the oxidoreductase catalytic cycle (Figure 6-1A). A subsequent study found that a peptide that was designed as a substrate mimic bound in a location that partially overlapped the DsbB binding site, but did not make interactions in the hydrophobic groove.^{19c} It was suggested that the groove might be necessary in the context of a substrate protein (rather than a peptide mimetic) to bind the region of the protein containing the second cysteine that would attack the mixed disulfide and allow release of the oxidised substrate from DsbA. Thus, small molecules binding in the hydrophobic groove could potentially disrupt the catalytic cycle in two ways. The DsbA-substrate complex also suggests other there are regions of the DsbA surface other than the hydrophobic groove that may be suitable binding sites for small molecule inhibitors - notably the interface between the thioredoxin and α -helical domains.

In addition to the scaffolds discussed in Chapter 5, five other classes of chemically distinct compounds that displayed medium to strong CSP in ¹H-¹⁵N HSQC based NMR screening, were identified: benzofuran 6 and B1, benzothiophene 7 and C1, chroman 8, phenylthiazole 9 and phenol 10 (Figure 6-1B). Although these fragments showed weaker binding and inhibition activity than those discussed in Chapter 5, further investigation of their binding site may prove of value.

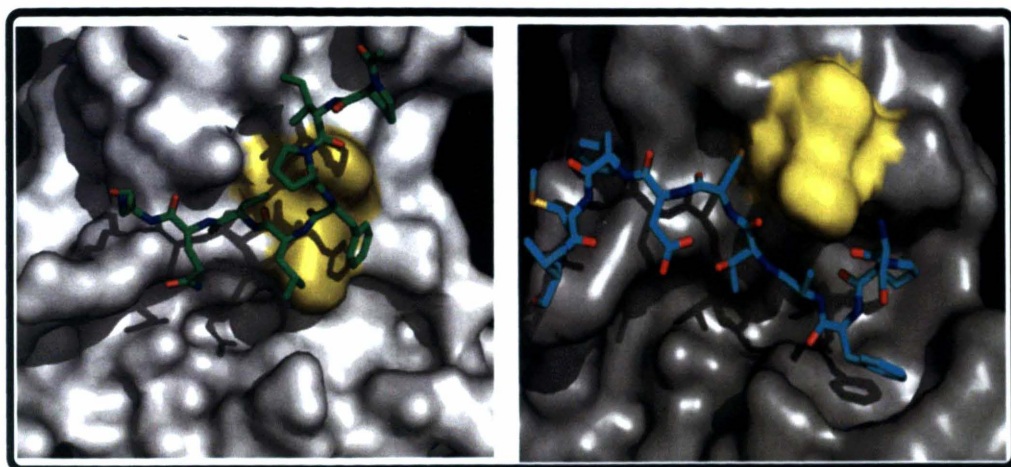


Figure 6-1A: The left panel shows the *EcDsbA*-substrate complex where the protein structure and bound peptide are represented as white surface and green sticks respectively. The panel on the right shows the *EcDsbA*-*EcDsbB* complex where the *EcDsbA* shown as grey surface and the interacting loop of *EcDsbB* is shown as cyan sticks.

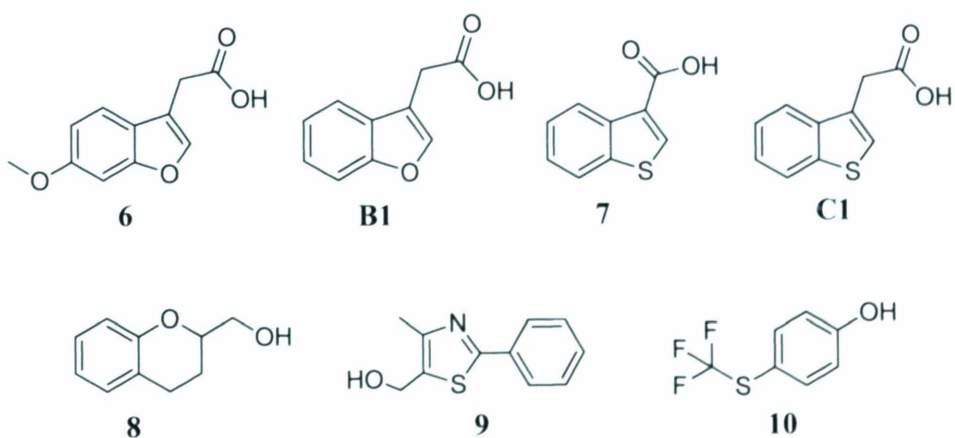


Figure 6-1B: Structures of additional fragment hits identified through NMR screening.

This chapter describes attempts to optimize each of these cores, by testing analogues that were either available from commercial sources or could be readily obtained by in-house synthesis. A range of analogues were identified on the basis of substructure searching, and similarity based searching to identify similar structures that may have diverse core scaffolds. A limited number of analogs or molecules selected based on similarity were tested (except the phenol compound 10). Analogs were restricted to compounds with MW \leq 300 in order to focus on optimization of the core rather than extension of these initially weak binding fragments.^{89, 110, 120a}

The current chapter also presents evidence for the binding of ligands at alternate and adjacent sites in the hydrophobic groove of *EcDsbA*. In order to identify these distinct small molecule binding sites, the characterisation of the structure of ligands bound to *EcDsbA* has been undertaken using X-ray crystallography²³² and NMR spectroscopy.²³³

6.2 Results and Discussion

6.2.1 *Characterisation and elaboration of EcDsbA hits containing 6,5-fused rings*

The second round of ^1H - ^{15}N HSQC NMR-based screening identified 14 fragment hits comprised of 6,5-fused heterocyclic ring systems with distinct functional groups attached at various positions. All of the original hits within this broad category of fragments demonstrated weak to medium binding against *EcDsbA* by inducing CSP up to 0.1 ppm in the ^1H - ^{15}N HSQC spectra of oxidised *EcDsbA*. These hits included compounds with such

diverse cores as benzofuran, benzothiophene, thienopyridine, imidazole, indoline, benzoxadiazole, chroman and benzimidazole. The majority of these hits showed CSP in the hydrophobic groove region of *EcDsbA* however they produced somewhat different patterns of CSP in the ^1H - ^{15}N HSQC. The 6-methoxy benzofuran **6** and benzothiophene **C1** showed the strongest CSP in this class and therefore were selected for further analysis by NMR, X-ray crystallography and inhibition assay (Figure 6-1B).

6.2.2.1 Characterization of benzofuran hit 6

Benzofuran fragment **6** has a measured K_D of $340 \pm 30 \mu\text{M}$ and a LE of 1.3 kJ. It showed inhibition in the motility assay with a measured EC_{50} value of $220 \pm 1 \mu\text{M}$. The CSP observed in the ^1H - ^{15}N HSQC spectra suggested that this fragment binds in the hydrophobic groove region of *EcDsbA* adjacent to the active site (Figure 6-2). This was established by solving a preliminary crystal structure (resolution 2.1 Å) of fragment **6** in complex with *EcDsbA* (Figure 6-2 and 6-3). ‘Preliminary’ crystallographic analysis will be described at several sections of this chapter and this refers to the data where extra electron density was observed in (Fo-Fc) maps, although in some cases there was density only for part of the ligand and in others the ligand occupancy was less than 100%. These preliminary crystallographic studies provide evidence of the binding site of the fragments and may be improved through optimisation of the soaking protocols as described in method section 2.5.

This structure revealed that benzofuran **6** binds in a slightly different but adjacent position compared to the previously identified sites for phenylthiophenes **1** and **2** at the edge of the groove near the Arg148-Gly149-Val150 loop (Figure 6-2 and 6-3).

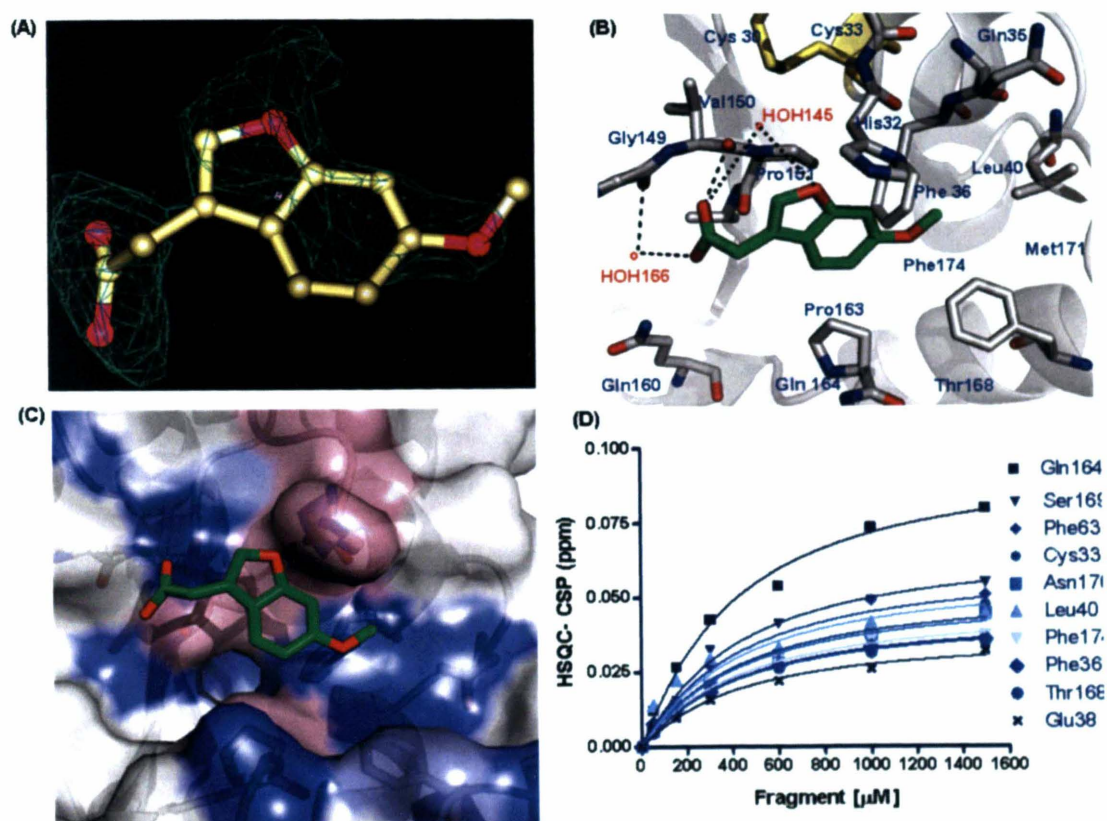


Figure 6-2: (A) Omit electron density maps of benzofuran hit **6** (shown as a yellow carbon stick model) at contour level σ 2.5 (B) Preliminary crystallographic *EcDsbA* binding mode. (C) NMR-based heat map (blue: CSP > 0.03 ppm, pink: unassigned residues) with crystallographic coordinates of hit **6** bound to *EcDsbA* (D) ^1H - ^{15}N HSQC titration binding curve (K_D measurement) of hit **6**.

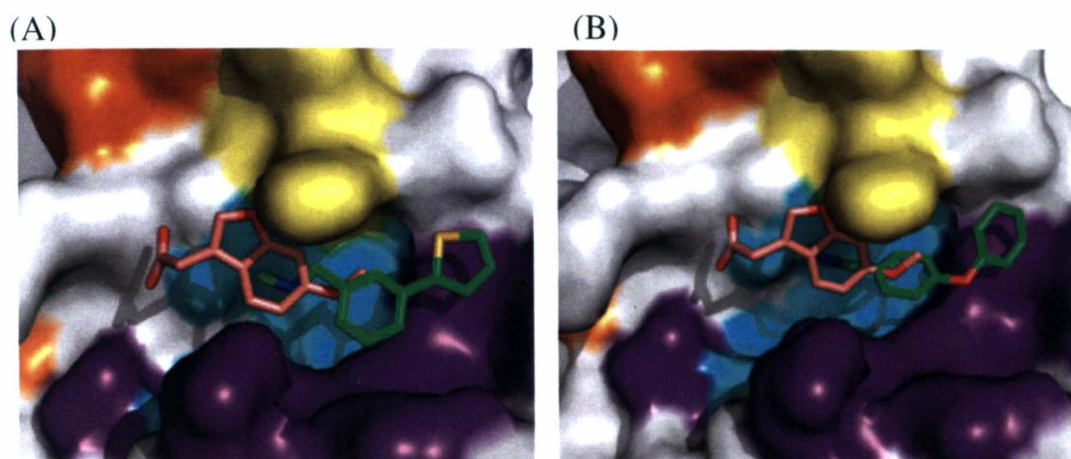


Figure 6-3: Crystallographic binding position of benzofuran 6 with (A) phenylthiophene 1 and (B) phenoxybenzene 4 which binds in overlapping and adjacent binding sites in the hydrophobic groove (purple) and pocket (cyan) near the active site (yellow) and hydrophobic patch (orange) of *EcDsbA* (white solid surface).

As illustrated in Figure 6-3, the methoxy side chain of fragment **6** overlaps with the position of fragments **1**, **2** and **4**. The 5- position phenyl atom of bezofuran ring is about 2.8 Å away from the nearest edge of the phenyl ring of the phenylthiophenes **1** and **2** and 1.95 Å away from the nearest edge of the central phenyl ring of compound **4** in the hydrophobic groove. Although the resolution of the crystal structure along with shape of the ligand's extra density was not of a very high quality, a preliminary analysis of the binding mode was possible and was assisted by the NMR data. This analysis suggested two possible hydrogen bonds between the carboxyl group of benzofuran **6** and the backbone oxygen atoms of Val150 and Gly149, one of which (Val150) is water mediated. The furan-ring oxygen also may make a water mediated hydrogen bond with Val150 (Figure 6-2). However, like other

hits, most of the interactions between *EcDsbA* and fragment **6** are hydrophobic in nature.²³⁴ The benzofuran ring is surrounded by the hydrophobic residues Pro163 and Pro151 and points towards the active site residue His32. The 6-methoxy substituent interacts with His32 and Phe36 and the benzene ring interacts with Pro163 and Gln164 via vdW interactions. At the edge of the groove, the Arg148-Gly149-Val150 loop lies in close proximity to the 3'-acetic acid group (Figure 6-2). These interactions probably contribute to the chemical shift changes in the flexible groove region that can be seen in the NMR-based heat map in Figure 6-2. A reasonable agreement was found between the crystal structure and the NMR location of this fragment and CSP of up to 0.08 ppm were observed in the groove region. The binding location of the benzofuran hit suggested that this adjacent binding pocket on the *EcDsbA* hydrophobic groove can be explored further and may be of interest for fragment elaboration by a linking strategy. Secondly this structurally diverse chemical class provides new polar interactions to *EcDsbA* as compared to the phenylthiophene and phenoxybenzene fragments. Hence benzofuran **6** was considered a good starting point for further chemical optimization to more potent leads.

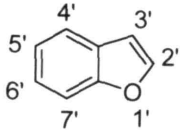
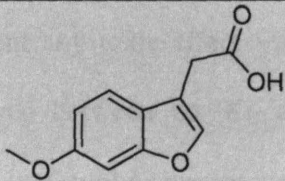
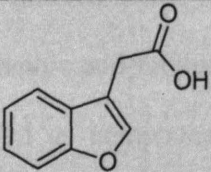
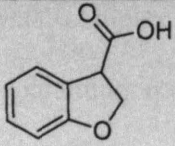
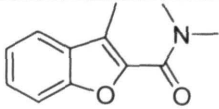
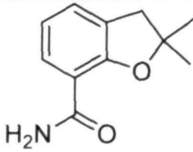
6.3.2 *Elaboration of benzofuran fragment 6*

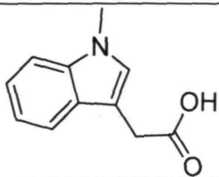
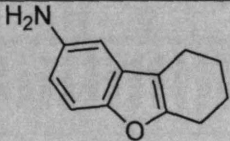
Analysis of the X-ray crystal structure of fragment **6** suggested a potential strategy for optimization of binding to *EcDsbA*. Figure 6-2 and 6-3 illustrates those portions of fragment **6** where it might be possible to improve the interaction with *EcDsbA*. Firstly there is a large unoccupied region to one side of the benzofuran ring system. Functionalizing the 4', 5' and 6' positions of the benzene ring with various groups may fill this pocket and

thereby increase the binding affinity. Secondly, a chemically simple place to attach various functionalities is at the 3'-acetic acid. The reaction of the acid group with amines or alcohols to form the corresponding esters or amides may allow for new interactions to form with the binding site. These linkages connecting the benzofuran core to another ring system or small hydrophobic moiety would perhaps pick up additional binding interactions in or around the *EcDsbA* active site.

To test these binding hypotheses a number of approaches were undertaken. Firstly, new compounds that maintained the benzofuran or benzothiophene core, but which had a diverse range of substituents on the phenyl ring were sourced from either commercial or in-house collections (Table 6-1). In all cases the 7-position remained unsubstituted based on the crystal structure which suggested the possibility of a steric clash with the active site in that case. End-point ^1H - ^{15}N HSQC spectra were recorded for the initially selected five compounds and another original hit B1 of this class and ranked as previously described. Additionally for analogues that induced large CSP, K_D , motility assay and crystallographic trials were undertaken.

Table 6-1: Analogues of the benzofuran **6**. The magnitude of CSP observed in the end point ^1H - ^{15}N HSQC along with the % inhibition of motility caused by the fragment at a concentration of 500 μM are listed. Fragments highlighted in grey were selected for further characterisation and for these compounds the K_D and EC_{50} are listed.

| Structure  | Compound | CSP strength and/or K_D (μM) | Motility % Inhibition and/or EC_{50} (μM) |
|--|-----------|---|--|
|  | 6 | ++ 340 ± 30 | 100% 220 ± 2 |
|  | B1 | ++ 310 ± 27 | 100% 56 ± 2 |
|  | B2 | +++ 500 ± 37 | 50% 615 ± 1 |
|  | B3 | ++ | 20% |
|  | B4 | ++ | 50% |

| | | | |
|---|-----------|---------------|------|
|  | B5 | ++ | 30% |
|  | B6 | +++ 46 ± 5 | 100% |

Another original benzofuran hit **B1**, which lacks the 6'-methoxy substituent of fragment **6**, showed a similar extent of CSP with a slight increase in binding affinity (Figure 6-4). Surprisingly, this compound was 4-fold more potent in inhibiting bacterial motility (without any toxic effects on cell growth) as compared to fragment **6**. The discrepancy observed between the K_D and EC_{50} values for benzofuran **B1** may indicate that this compound inhibits motility in a non-*EcDsbA* dependent fashion. The NMR binding, X-ray crystallographic and inhibition profile of this compound was discussed in Chapter 3 (and was referred as compound **3**, for the publication purposes). Briefly, a high resolution *EcDsbA* crystal complex of benzofuran **B1** revealed a quite similar binding mode to the same core containing fragment **6**, supporting the adjacent binding site on the *EcDsbA* surface close to the Arg148-Gly149-Val150 loop (Figure 6-5). Comparatively higher CSP in the hydrophobic groove region were induced by fragment **6** (with the 6'-methoxy substituent) compared to **B1**, corresponding to the observed crystallographic binding modes where **B1** makes fewer interactions with groove residues such as Gln164-Asn170 (Figure 6-4 and 6-5).

Active site region Hydrophobic groove region

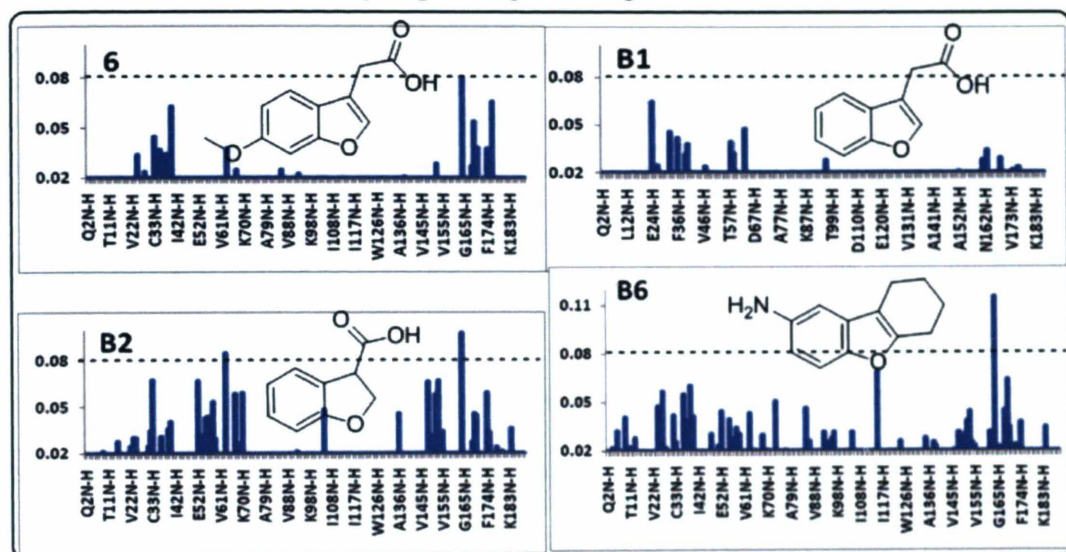


Figure 6-4: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of the analogues of fragment 6.

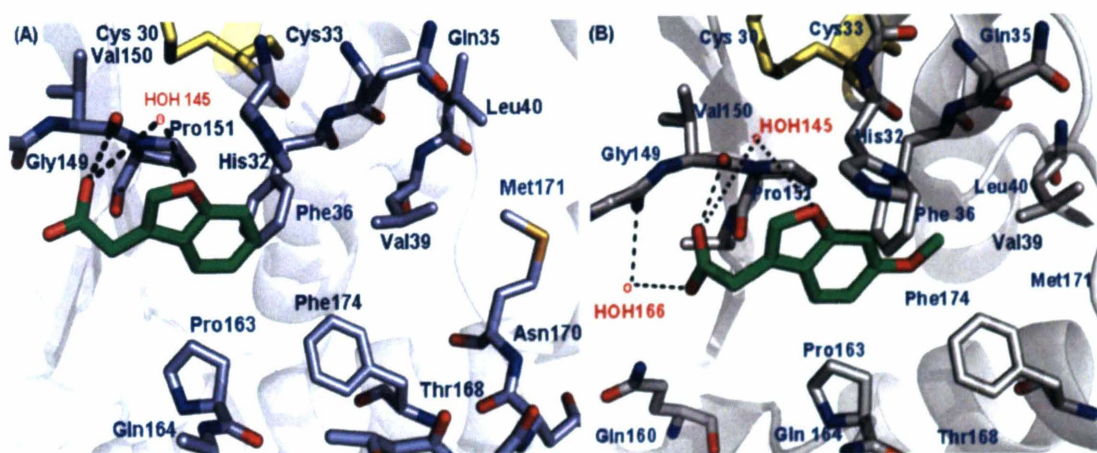


Figure 6-5: Crystallographic binding modes of benzofuran fragments 6 and analogue B1 (green stick representation) to *EcDsbA* (white stick and ribbon representation).

Dihydrobenzofuran **B2**, containing a 3'-carboxylic acid substituent showed strong CSP, however data revealed weaker values for the K_D and EC_{50} (Figure 6-4, Table 6-1). Additionally three analogues **B3** - **B5** showed only moderate CSP and appeared to bind in the hydrophobic groove. Unfortunately, since compounds **B2** - **B6** contain more than one structural change from the parent fragment it is difficult to interpret the NMR binding data to derive useful SAR information (Table 6-1). Therefore further elaboration of this class was carried out using structure-guided synthesis and medicinal chemistry approaches by Mr. Christopher Schreurs (MIPS), however further details and results are beyond the scope of this thesis.

Fragment **B6**, which contains a tetrahydrodibenzofuran core, was also a strong binder. This compound induced CSP up to 0.11 ppm in the *EcDsbA* ^1H - ^{15}N HSQC spectra, however it presented a different pattern compared to benzofuran fragment **6**. NMR-based mapping suggested a binding location in the hydrophobic groove, although there were comparatively smaller CSP in the active site region (Figure 6-6). Additionally a high resolution (1.9 Å) crystal structure was obtained where **B6** fits very well into the *Fo-Fc* electron density,²³⁵ and was located in accord with the approximate NMR binding site of this compound. However, the crystallographic position of this ligand is very close to the interface between monomer A and B (Figure 6-6). The possible interactions with two molecules in the crystal structure are potentially due to crystal packing interactions, hence this structure was considered to have dubious biological relevance. Secondly, this compound showed poor solubility in assay minimal media and NMR sample buffers so we did not investigate **B6** further.

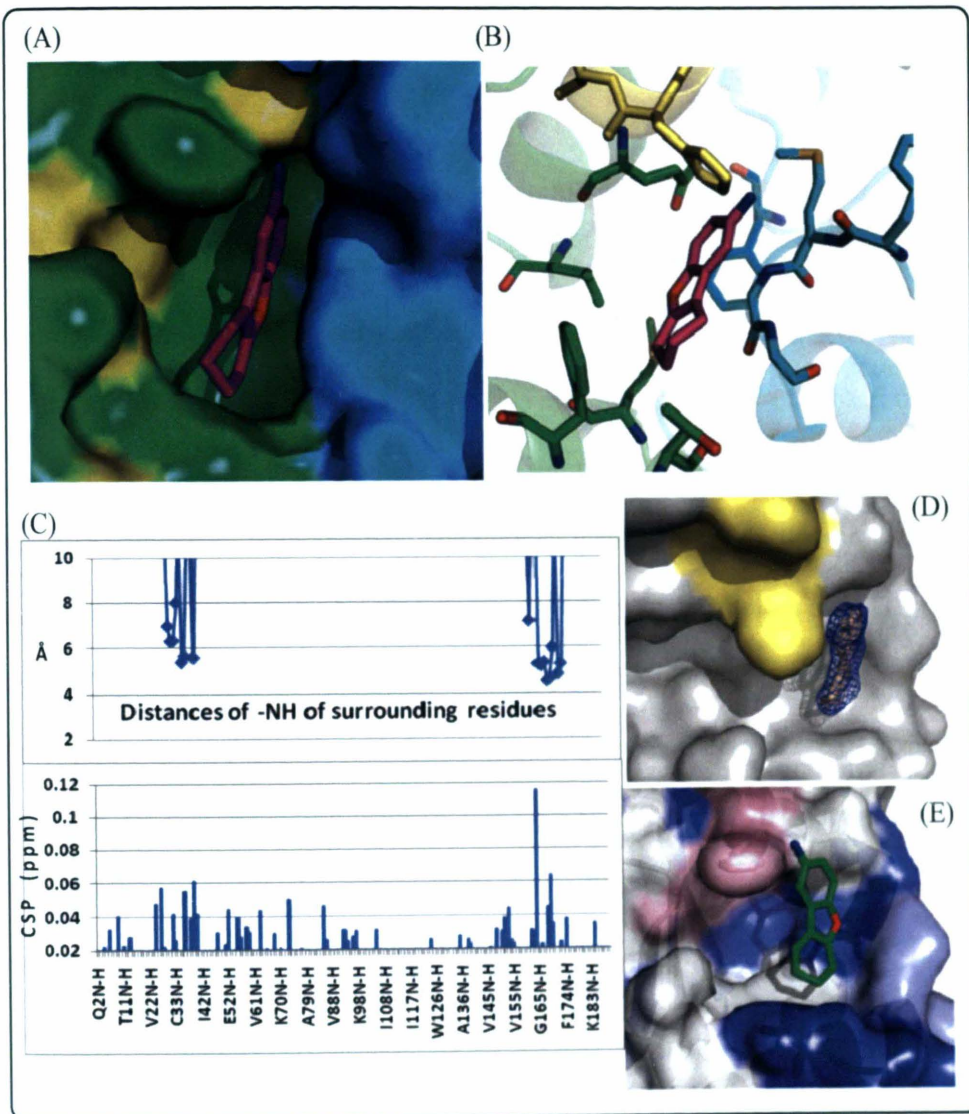
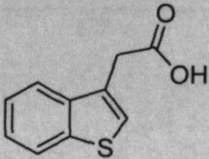
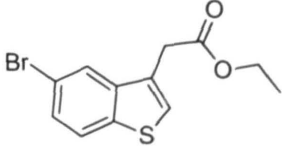
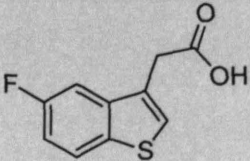
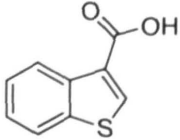
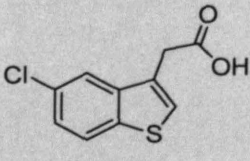
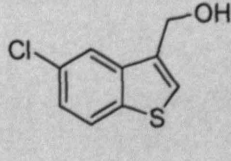


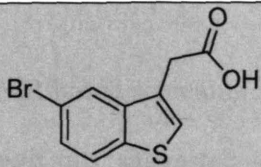
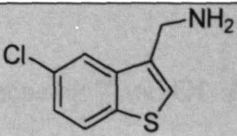
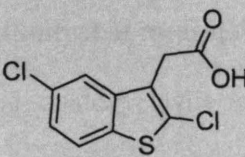
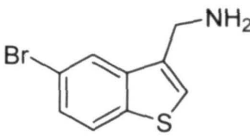
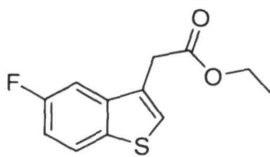
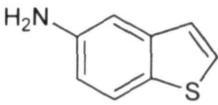
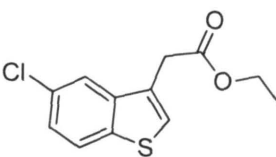
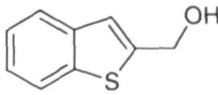
Figure 6-6: (A) Binding of fragment **B6** with *EcDsbA*. The crystals have two molecules of DsbA in the asymmetric unit. Chain A (green) and B (cyan) are shown as (A) surface and (B) stick respectively. (C) Comparison of HSQC-CSP and distances from amides (surrounding 5 Å) to the *EcDsbA* bound fragment **B6**. (D) Omit electron density maps of **B6** at contour level σ 2.5 (E) NMR based heat map (CSP > 0.02 ppm, pink: unassigned residues) with coordinates of fragment **B6** (green stick representation).

6.3.3. *Elaboration of benzothiophene hit*

For the initial exploration trials of the benzothiophene class, analogues where the core of the original hit **C1** was kept unchanged but with a diverse range of substituents at different positions of the ring were selected. Additionally, molecules that were assessed by their Tanimoto similarity score (threshold of 0.35 was set as the cut-off based on the literature) were selected.²²⁹ These compounds were either purchased or synthesized and then tested as shown in Table 6-2. Eighteen new compounds were chosen via this process for evaluation and were first tested for solubility under the conditions employed in NMR and motility assays. Five analogues showed poor solubility in the minimal media used for the motility assay and three of these were insoluble in the NMR sample buffer and these were discarded. For the remaining analogues end-point ^1H - ^{15}N HSQC spectra were recorded and compounds were ranked as described above (Table 6-2). The analogues with largest perturbations were characterised further by K_D , motility assay and crystallographic studies.^{111b}

Table 6-2. Analogs of the benzothiophene hit **C1**. The table shows the chemical structure of the substituent and the position (3' and 5' or 2') at which the substituent was attached to the benzothiophene core. The compound number, the magnitude of CSP observed in the end-point ^1H - ^{15}N HSQC spectra along with the % inhibition of motility caused by the fragment at a concentration of 500 μM are listed. Fragments highlighted in grey were selected for further characterisation.

| Compound Structure and CSP Strength | % inhibition | Compound Structure and CSP Strength | % inhibition |
|--|-----------------|---|-----------------|
| C1 ++  | 50% | C8 ++  | 20% |
| C2 +++  | 100% | 7 +  | 20% |
| C3 +++  | 100% | C9 ++  | 40% |
| C4 +++ | 100% | C10 +++ | 100% |

| | | | |
|---|------|---|-----|
|  | |  | |
| C5 +++ | 100% | C11 + | 0% |
|  | |  | |
| C6 + | 10% | C12 ++ | 10% |
|  | |  | |
| C7 + | 10% | C13 ++ | 10% |
|  | |  | |

The NMR based screening identified benzothiophene hit **7** and **C1**. A comparison among these two hits suggested that the presence of the 3' acetic acid in **C1** instead of 3' carboxylic acid in **7** improved the CSP strength greatly and allowed the determination of a low affinity, but measurable, K_D of $880 \pm 77 \mu\text{M}$ (Figure 6-7, Table 6-3). In light of the chemical analogy, with a similar shape and sidechain, it might be expected that benzothiophene acetic acid **C1** binds in a similar position to benzofuran acetic acid **B1**. However, CSP based mapping indicated a binding location near the active site in the hydrophobic pocket, showing comparatively more CSP towards the opposite side of the groove, similar to benzofurans **6** and **B1**. This observation suggested that changing the nature of the heteroatom (S or O) changes the binding significantly due to their varying hydrophobic, steric, or electronic effects on the ring system.

In order to further characterise the interaction of benzothiophene fragments with *Ec*DsbA a series of closely related analogues of **C1** were purchased. Additionally, a number of ester and halogen derivatives containing the benzothiophene core were synthesized by Ms Joan Ho (MIPS) (**C2**, **C4** and **C6 - C8**, Table 6-2). Initially, analogue **C3** was identified with a 5'-chlorine substitution that showed ^1H - ^{15}N HSQC binding with moderate CSP of up to 0.09 ppm (Figure 6-7). This hit was further analysed by measuring the K_D ($86 \pm 7 \mu\text{M}$) by NMR, and a determining an EC_{50} of $128 \pm 2 \mu\text{M}$ in the bacterial motility assay (Table 6-3).

The 5-position chlorine was replaced with fluorine and bromine atoms in analogues **C2** and **C4**. Additionally, the acetic acid functionality of all the halogen

substituted analogues (Table 6-2) was esterified as the ethyl ester. Based on their CSP mapping, all of these analogues appeared to bind in the hydrophobic groove of *EcDsbA*, however, there were subtle differences in the pattern of CSP according to the nature of substituent halogen and 3' substitution. All of the ester analogues showed little or no binding, compared to the corresponding acid functionalised analogues, highlighting the importance of this polar group. Additionally, the 5-bromo benzothiophene **C4** was identified as the most potent hit of this series, showing a 64 fold increase in affinity ($K_D = 14 \pm 13 \mu\text{M}$) as compared to **C3**. The increased affinity of **C4** may be due to the steric size or electronic effects or a combination of both, of the 5-bromo substituent.

Replacement of the 3'-acetate group with a hydroxymethyl substituent in compound **C9** or the aminomethyl in analogue **C10** resulted in a reduction in binding affinity. Although methylamine **C10** induced slightly greater CSP than chlorobenzothiophene **C3**, the binding affinity was weaker; again illustrating that ^1H - ^{15}N HSQC CSP strength and binding affinity do not always correlate.

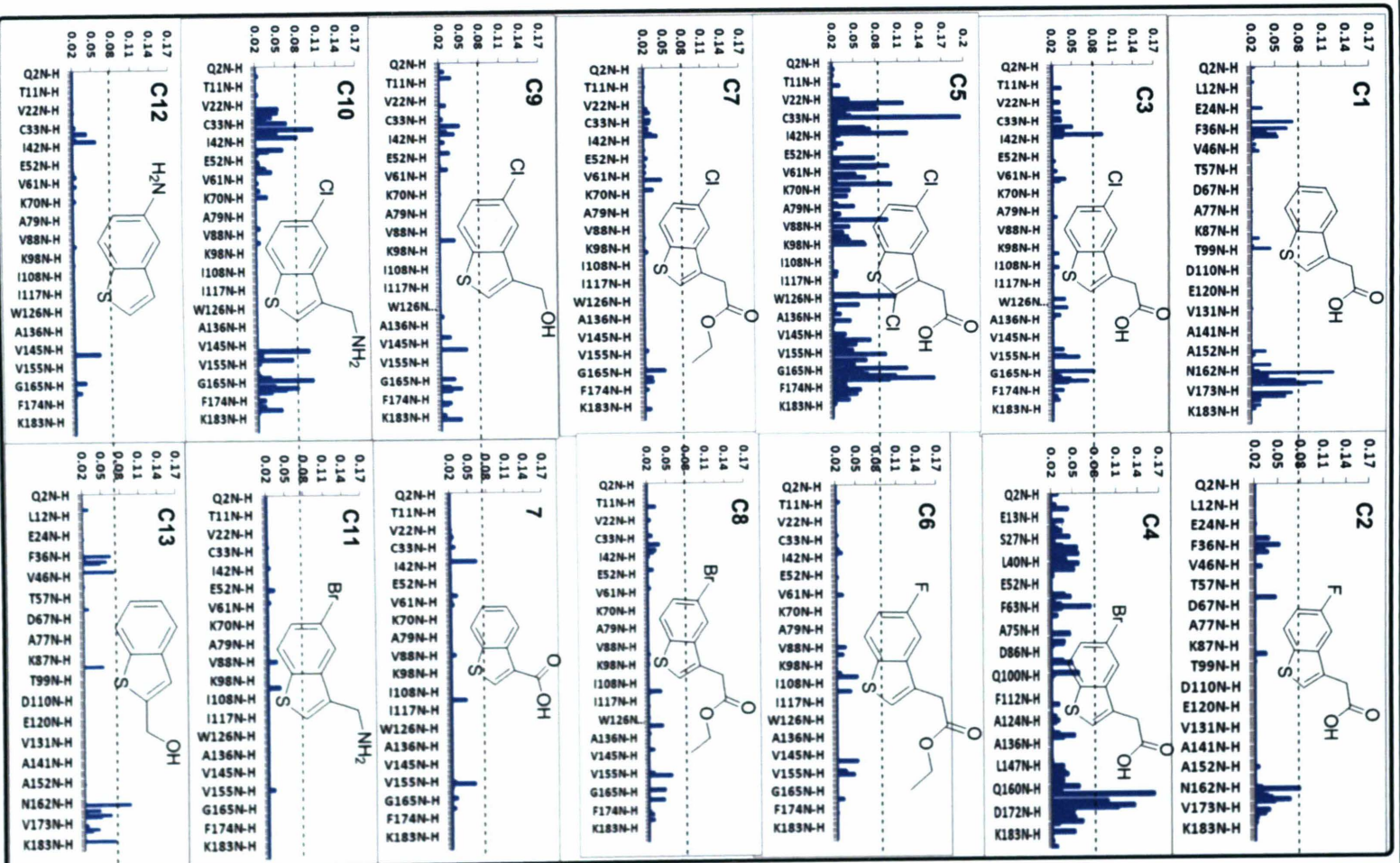
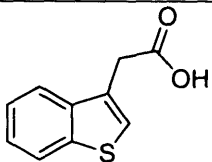
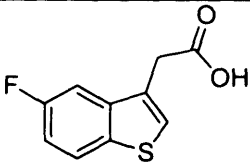
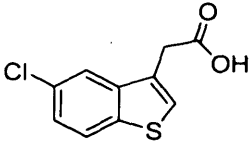
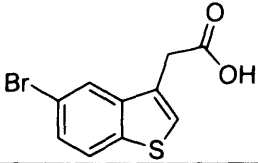
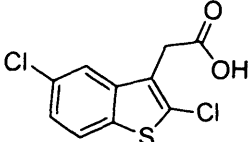
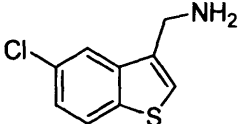


Figure 6-7: CSP ($\Delta\delta$ (NH) > 0.02 ppm) of ^1H - ^{15}N HSQC spectra of *EcDsbA* induced by benzothioephene analogues.

Interestingly the extent of CSP from the dichlorobenzothiophene **C5** were significantly greater (Table 6-2, Figure 6-7). This compound induced 3-fold higher CSP of up to 0.19 ppm with a K_D determined by NMR of $113 \pm 6 \mu\text{M}$. This analogue also showed inhibition in the motility assay with measured EC_{50} value of $150 \pm 2 \mu\text{M}$. CSP mapping suggested a binding location in the hydrophobic groove with significant changes in the hydrophobic patch region as well (Figure 6-9).

Table 6-3: K_D (μM) and EC_{50} (μM) of benzothiophenes analogues

| Structure | Compound | K_D (μM) | EC_{50} (μM) |
|---|-----------|-------------------------|------------------------------------|
|  | C1 | 880 ± 77 | 735 ± 2 |
|  | C2 | 30 ± 9 | 125 ± 1 |
|  | C3 | 86 ± 7 | 128 ± 2 |
|  | C4 | 14 ± 13 | 170 ± 2 |
|  | C5 | 113 ± 6 | 150 ± 2 |

| | | | |
|---|------------|--------------|-------------|
|  | C10 | 540 ± 60 | 110 ± 1 |
|---|------------|--------------|-------------|

6.3.3.1 Preliminary crystallographic trials for benzothiophenes

In order to structurally characterise this class of *EcDsbA* ligands, crystallographic soaking experiments were performed on compounds **C1**, **C3** - **C5** and **C10**, initially selected due to the extent of their CSP (Figure 6-7, Table 6-2). Somewhat surprisingly, where electron density was observed for this class of ligands, it was weak for all or part of the fragments. For instance, the density of the 5'-Cl substituted compound **C3**, the 5'-Br substituted compound **C4** and the unsubstituted original hit **C1** was weak, while for the 2, 5-dichloro substituted compound **C5** portions of the electron density were missing. Despite this missing density, on the basis of this preliminary data **C5** was considered a reasonable hit and its crystallographic binding mode is described in more detail below. The methylamine **C10**, which exhibited moderate binding affinity, showed well-defined electron density, however its location appeared to be biased by crystal contacts (Figure 6-8). We considered the binding location of methylamine **C10**, to be more likely a crystallographic artifact Figure (6-8).

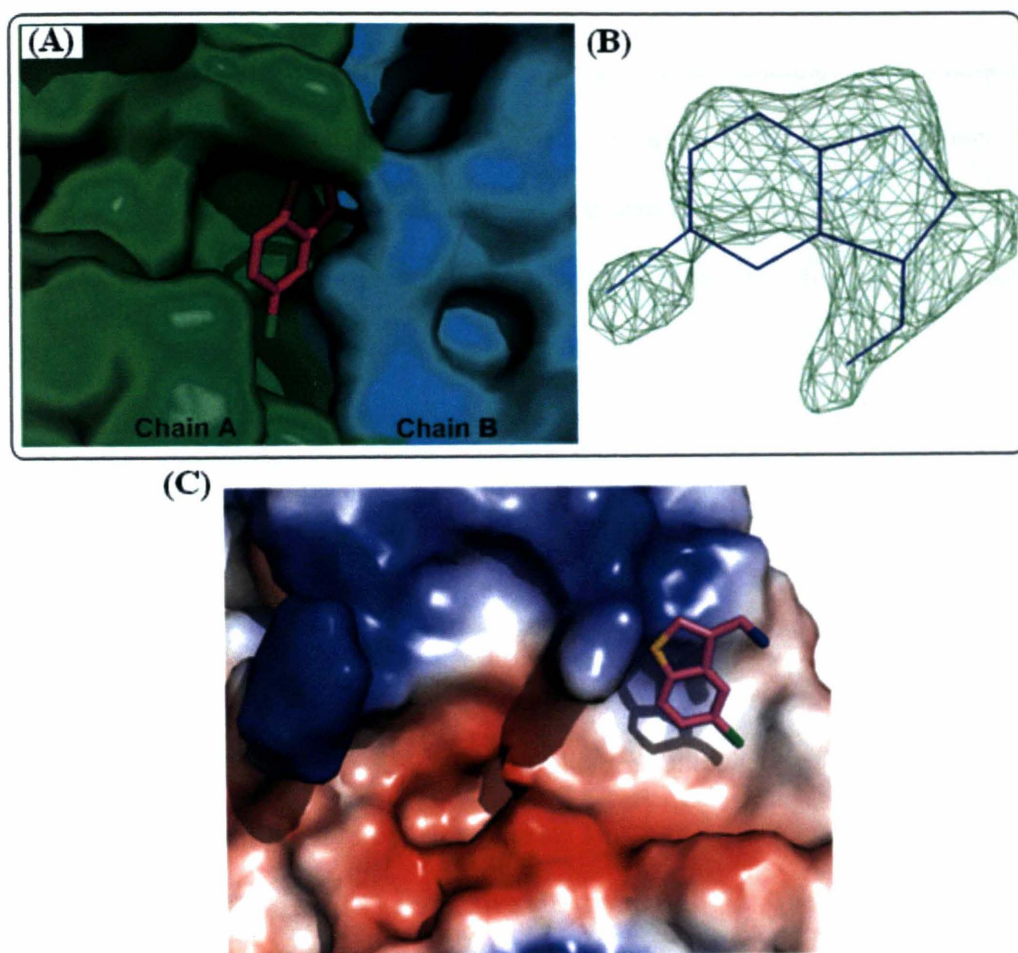


Figure 6-8: (A) Preliminary crystallographic binding of the benzothiophene analogue **C10** at the crystallographic interface in *EcDsbA*, Chain A (green) and Chain B (cyan) shown as a solid surface model (B) Omit electron density maps at contour level σ 2.5 (C) binding of **C10** with *EcDsbA* chain a shown as electrostatic surface model.

For 2,5-dichlorobenzothiophene **C5**, structural information was obtained by solving a 2.1 Å crystal structure of the compound in complex with *Ec*DsbA. A large area of extra (Fo-Fc) electron density was observed in monomer A of the asymmetric unit near the active site, however a small amount of density was missing for some parts of the molecule (Figure 6-9). This may be due to the lower affinity and correspondingly lower occupancy of this ligand compared to previously studied, more potent ligands. Refinement and minimization of the complex suggested that **C5** binds at the midpoint of the groove and in a slightly overlapping position compared to the most potent initial fragment hits, phenylthiophenes 1-3 (Chapter 5). The closest edge of the phenyl ring of **C5** is 1.8 Å away from the closest edge of the phenyl of the analogous benzofurans **B1** and **6** in their crystal complexes (Figure 6-9). The benzothiophene ring points towards the active site and lies in the hydrophobic pocket. This ring is surrounded by hydrophobic residues Phe36, Pro151, Pro163 and active site residue His32. The benzothiophene ring also stacks via hydrophobic contacts with Phe174. The 3-acetic acid substituent interacts via a hydrogen bond (O—H \cdots O, 3.49 Å) with Thr168 and introduces a new key interaction for this class (Figure 6-9). ITC data for the binding of several of the fragments are presented in Appendix. Benzothiophene **C5** was found to have a higher enthalpy of binding than was measured for several of the other fragments. The observed high enthalpy changes on binding of dichlorobenzothiophene **C5** to *Ec*DsbA are consistent with the polar interactions seen in the crystal structure, which may also contribute to large CSP in the flexible groove region.²³⁶ It is also conceivable that the two hydrophobic

chloro substituents provide extra vdW interactions with groove residues Val39, Leu40 and Pro163 (Figure 6-9), which contribute to the affinity.

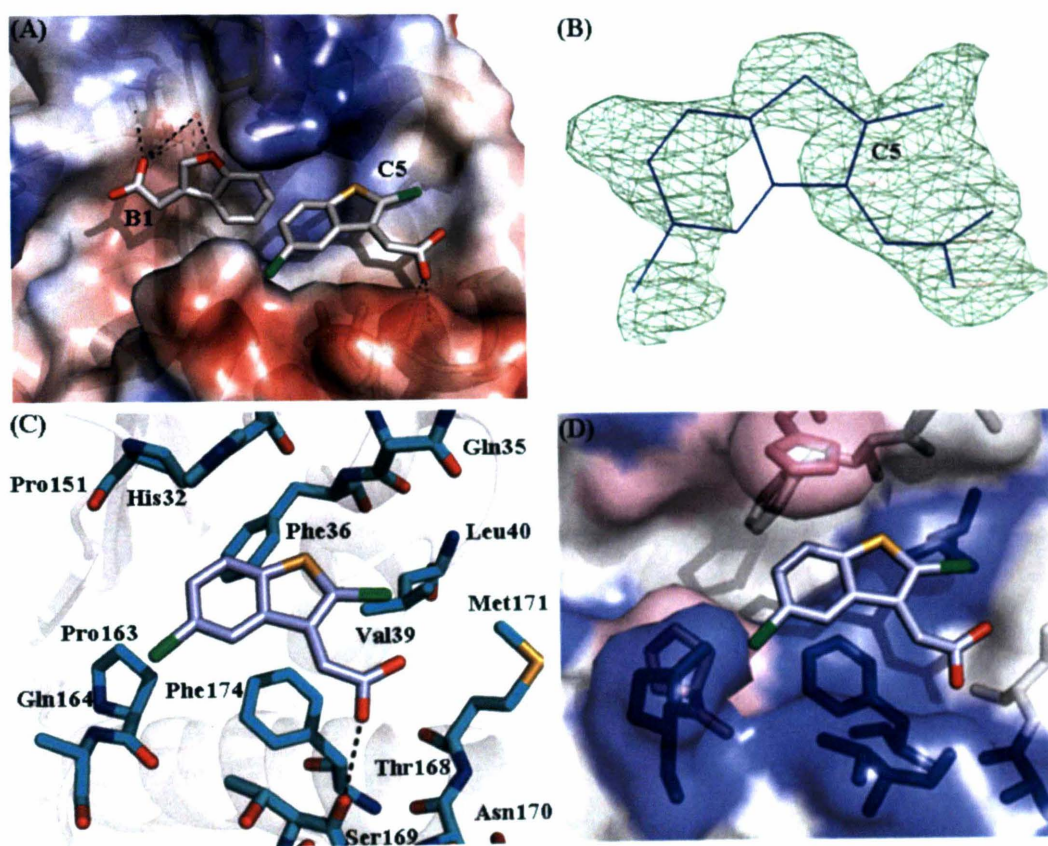


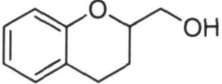
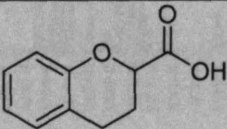
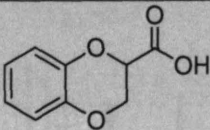
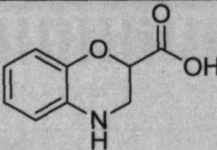
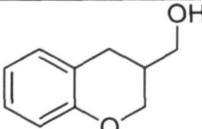
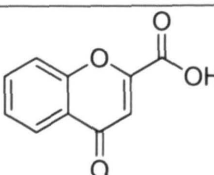
Figure 6-9: (A) Crystallographic orientation of benzofuran fragment **B1** shown along with coordinates of the benzothiophene fragment **C5** (white stick representation) in the groove of *EcDsbA* (represented as an electrostatic surface). (B) Omit electron density maps (green) of **C5** (blue line) at contour level σ 2.5. (C) Binding mode of **C5** with *EcDsbA* where protein and ligand are shown as cyan and white stick respectively. (D) NMR-HSQC based heat map (blue: CSP > 0.03 ppm, pink: unassigned residues) from the binding of **C5** at 1 mM with crystallographic coordinates of fragment **C5** bound to *EcDsbA* (white surface).

The crystallographic and NMR-derived location of fragment C5 were consistent, (as shown in Figure 6-9), with additional CSP of up to 0.1 ppm observed in the hydrophobic patch region (Figure 6-7). Although the site of binding is distant from the hydrophobic patch (comprised of residues Phe29, Phe63-Gly66, Leu68 and Val150) it is possible that these CSP arise due to relayed conformational changes in the β 2 strand due to an interaction with the active site residues. Although the electron density observed for this fragment is less than ideal, the proposed binding location and novel key interactions suggest this class may be of future interest for fragment elaboration. In particular, its binding location next to benzofuran 6 suggests the possibility of fragment linking.²³⁷

6.3.4. *Elaboration of Chroman hit*

The last scaffold identified in the category of fused ring systems was the 6,6-fused chroman fragment **8**. NMR screening showed this fragment bound to *Ec*DsbA, exhibiting moderate CSP of up to 0.07 ppm (Table 6-4, Figure 6-10). The perturbations were observed in a similar place in the hydrophobic groove to the initial benzofuran hit **6**. A limited exploration via a similarity search was initiated for this class and six analogues selected and further investigated. These molecules contain the 6,6-fused heteroatom ring system with the modifications to the pyran ring portion at the 4-position, or the 2-position sidechain functionality (Table 6-4).

Table 6-4: Structures and activity of analogues of the chroman fragment **8**. Fragments highlighted in grey were selected for further characterization, resulting in the K_D (μM) and EC_{50} (μM) values shown.

| Structure | Compound | Binding & %Inhibition | K_D (μM) | EC_{50} (μM) |
|---|-----------|--------------------------|-------------------------|-----------------------------|
|  | 8 | ++ 0% | | |
|  | D1 | +++ 80% | 114 ± 16 | 180 ± 2 |
|  | D2 | +++ 60% | 160 ± 7 | 330 ± 1 |
|  | D3 | +++ 0% | 90 ± 8 | n.d |
|  | D4 | ++ 0% | | |
|  | D5 | + 0% | | |

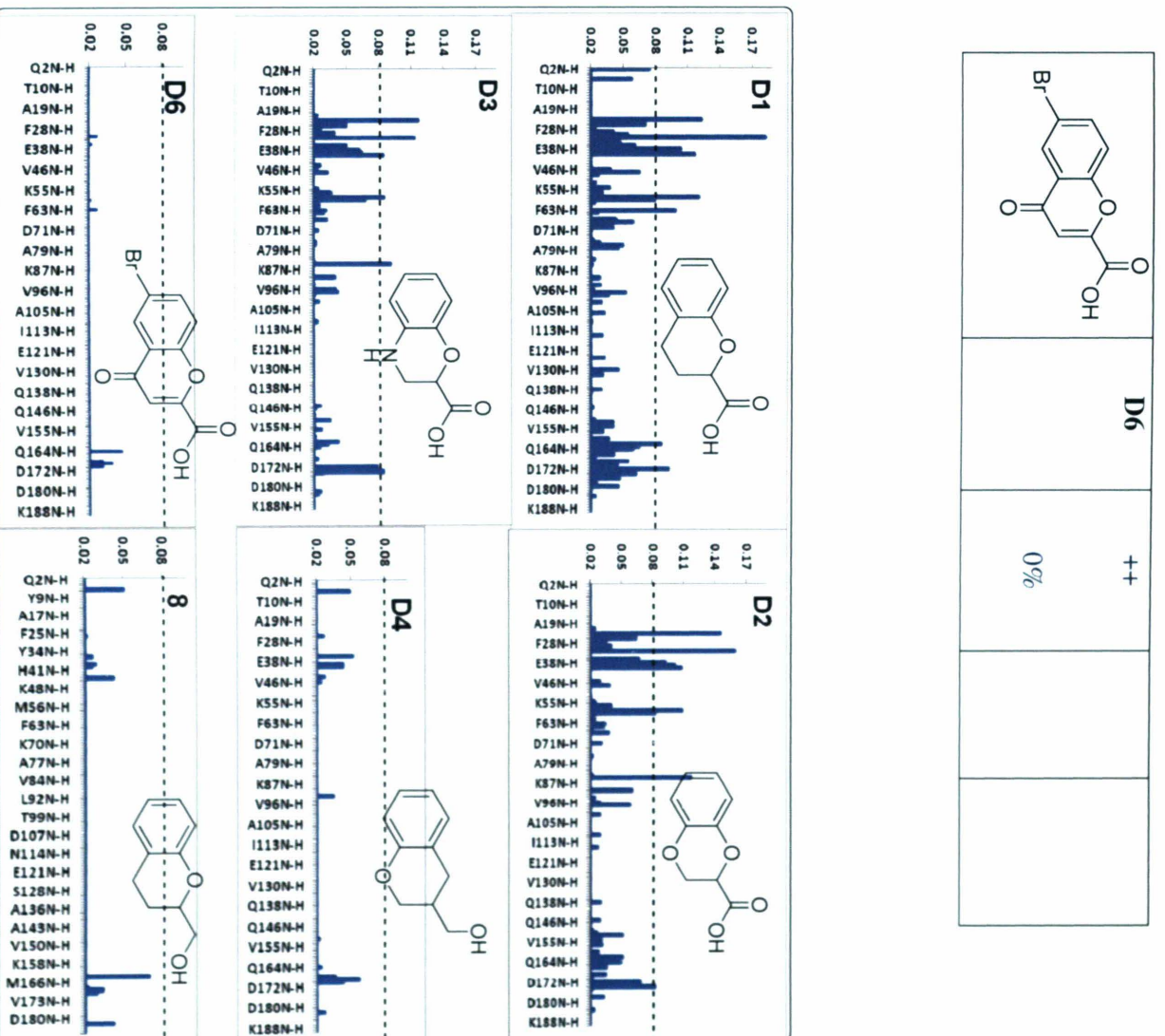


Figure 6-10: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in the ^1H - ^{15}N HSQC spectra of *EcDsbA* upon the addition of chroman analogues **D1** - **D4**, **D6** and **8**.

Of the tested analogues, three (**D1-D3**) showed CSP to a greater extent than the hit fragment **8**, with CSP in ^1H - ^{15}N HSQC spectra of up to 0.2 ppm. NMR titrations revealed K_D values ranging 90-160 μM for these compounds (Table 6-4, Figure 6-10). The pattern of CSP for these hits suggested their binding close to the active site and hydrophobic patch although also with some moderate CSP in the hydrophobic groove, that may result from conformational changes on binding.

Crystallographic complexes of all three of the strongest chroman binders **D1**, **D2** and **D3** (Table 6-4) were refined to give their structures bound to *EcDsbA* (resolution ranging 2.0-2.1 Å). The extra density observed for **D1** was not very clear and suggested the possibility of two orientations below the active site. Unfortunately, these modeled positions show clashes with active site residues and inter-molecular interactions with monomer B, which suggest this extra density may not give a reliable indication of the solution binding position.

In compounds **D2** and **D3** the quality of extra density was slightly better and both of these fragments showed two binding poses surrounding the active site region. As shown in Figure 6-11 one position clashing with His32 was common for both **D2** and **D3** but the proximity to Chain B again suggested this pose may be a crystallographic artifact. The second positions for these hits were slightly different, and within the hydrophobic groove (Figure 6-11), however, the overall quality of these preliminary data sets was too poor to give detailed interaction information, but did confirm crystallographic binding. Overall, these three carboxylic group containing compounds **D1**, **D2** and **D3** induced greater changes in the binding site environment. Additionally, due to their small size these fragments exhibited high ligand efficiency in the range of 1.65-1.76 $\text{kJ mol}^{-1} \text{HAC}^{-1}$.

Therefore these preliminary results might suggest these scaffolds are suitable for further exploration in the process of developing potent ligands for *EcDsbA*.

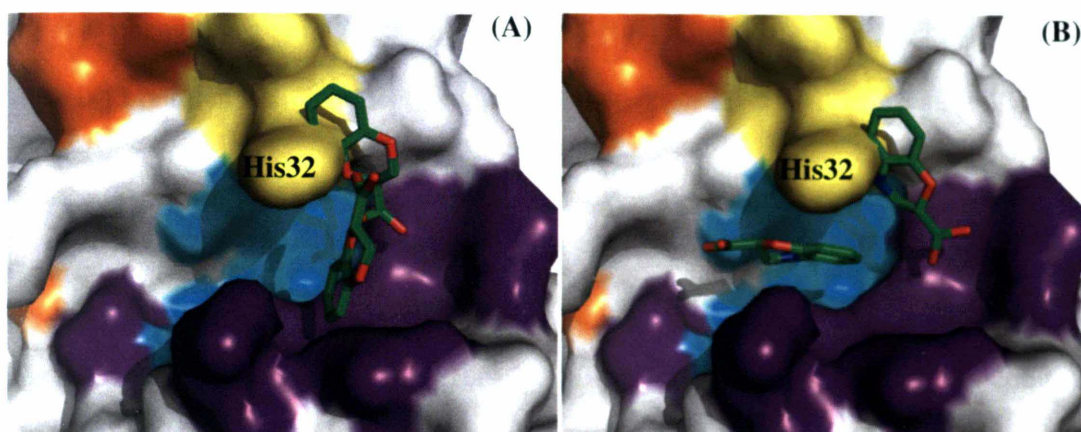


Figure 6-11: Preliminary crystallographic binding positions of chroman analogues (green stick representation) (A) **D2** and (B) **D3**. Both are shown in different binding sites in the hydrophobic groove and the pocket close to the active site and hydrophobic patch of the *EcDsbA* (shown as white solid surface), which are mapped in purple, cyan, yellow and orange respectively.

In spite of the promising results outlined above, as well as good solubility, these compounds showed poor or no inhibition in the bacterial motility. This is possibly due to the compound having insufficient membrane permeability to reach the site of action. As a result, this class of compound was not pursued further.

6.3.5. Characterization and elaboration of phenylthiazoles

Another hit fragment in the 5,6-linked biaryl class is the phenylthiazole **9**, which may be considered structurally related to the previously explored phenylthiophene hits. The original phenylthiazole hit fragment **9** induced strong CSP in the ^1H - ^{15}N HSQC spectra of up to 0.13 ppm, in a similar location to the phenylthiophene hits **1-3**, however it displayed a different pattern of CSP (Figure 6-12). Mapping of the ^1H - ^{15}N HSQC CSP onto the EcDsbA surface suggested a binding location close to the active site and the helical domain $\alpha 1$, but also with large changes in the groove. An NMR-based dissociation constant (K_D) of $378 \pm 86 \mu\text{M}$ (corresponding to a LE of $1.4 \text{ kJ mol}^{-1} \text{ HAC}^{-1}$) was measured for compound **9**, which corresponds reasonably well with the EC_{50} value of $340 \pm 2 \mu\text{M}$ determined in the bacterial motility assay. A similarity search was performed on commercial collections in order to identify available analogues in the phenylthiazole class.

Analogues were selected and tested where the phenylthiazole core was kept unchanged, but with differing substituents at various positions. This process identified phenylthiazole **9b** which contains a 5-carboxylic acid substitution in place of the hydroxymethyl side chain in the original hit **9**. This modification improved the K_D and EC_{50} values. Alternatively, substitution with the hydrophobic trifluoro methyl in the *p*-phenyl position, as in compound **9a**, resulted in weaker CSP as illustrated in Figure 6-12. Additionally three similarity molecules **9c** - **9e** showed moderate CSP in the end point ^1H - ^{15}N HSQC analysis and appeared to bind in the hydrophobic groove (Figure 6-12). Since these compounds contain more than one structural change from the original hit, and given

their apparently moderate CSP on the ^1H - ^{15}N HSQC spectrum, they were not investigated further (Table 6-5).

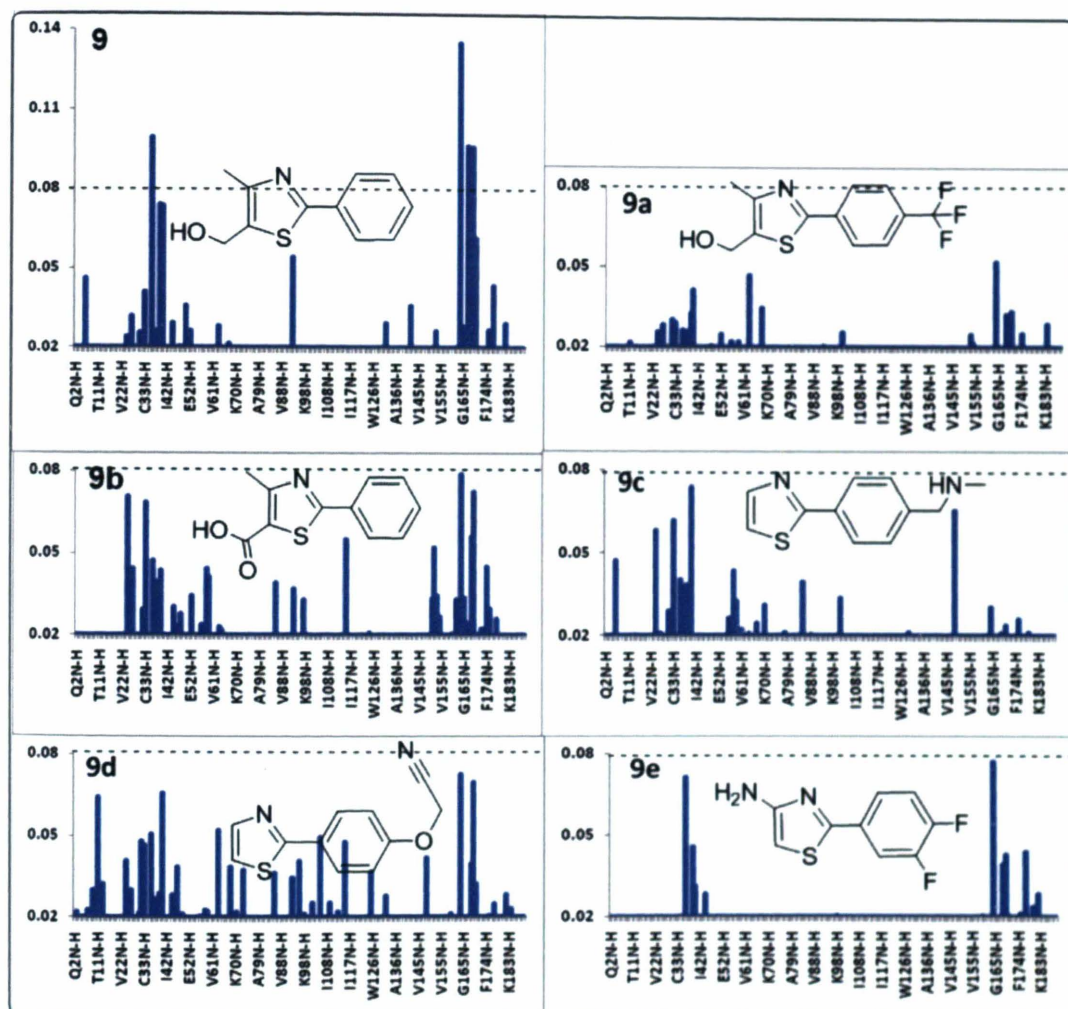
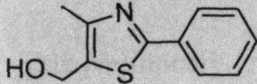
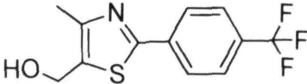
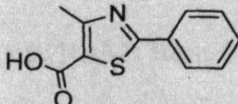
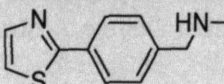
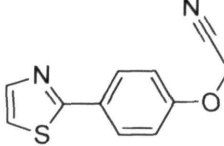
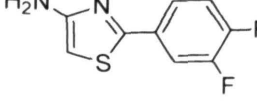


Figure 6-12: CSP ($\Delta\delta$ (NH) > 0.02 ppm) induced in the ^1H - ^{15}N HSQC spectra of EcDsbA upon the addition of phenyl thiazole analogues (**9** and **9a** - **9e**).

Table 6-5: Analogues of the phenylthiazole hit **9**. The compound number, the magnitude of CSP observed in the ^1H - ^{15}N HSQC end point titration along with the % inhibition of motility caused by the fragment at a concentration of 500 μM are listed. Fragments highlighted in grey were selected for further characterisation and for those K_D (μM) and EC_{50} (μM) are listed.

| Structure | Compound | Binding & % Inhibition | K_D (μM) | EC_{50} (μM) |
|---|-----------|---------------------------|-------------------------|------------------------------------|
|  | 9 | +++ 80% | 378 ± 86 | 340 ± 2 |
|  | 9a | ++ 70% | | |
|  | 9b | +++ 90% | 75 ± 8 | 250 ± 1 |
|  | 9c | ++ 100% | n.d. | 88 ± 2 |
|  | 9d | ++ 10% | | |
|  | 9e | ++ 60% | | |

Subsequently, in order to obtain detailed structural insights for phenylthiazoles as *EcDsbA* ligands, **9** and **9b** were selected for crystallographic soaking experiments due to their different structures and promising binding data. The binding mode of this class was determined by solving the X-ray crystal structure of phenylthiazoles **9** and **9b** bound to *EcDsbA*. The preliminary structures of these complexes were solved to 2.2 Å and 2.0 Å respectively. The omit electron density maps (*F_o-F_c*) showed clearly defined density for these hits at an occupancy 1.0, although compound **9** was missing some density from the edges of the phenyl ring (Figure 6-13).

This preliminary crystallographic data suggested the position of the phenylthiazole moiety was overlapping, but shifted towards Pro151, when compared to the phenylthiophene fragments **1** and **2** (Figure 6-13). Carboxylic acid containing compound **9b** showed binding in two conformations with a large area of electron density around the active site His32. Both of these conformations revealed hydrogen bonding involving the imidazole ring of active site His32 and the carboxylate oxygen. One conformation of **9b** also made a water-mediated hydrogen bond with active site residues Pro31 and Gln35 (Figure 6-13). Two possible explanations for the observed two binding conformations of compound **9b** could be envisaged: (1) the ligand possibly interacts with *EcDsbA* in the two different orientations, which is also supported by the agreement observed between NMR generated CSP and the crystallographic binding site (Figure 6-13 (B1)). (2) The influence of the other crystallographic monomer, producing an artifact hence second conformation was observed towards monomer B.

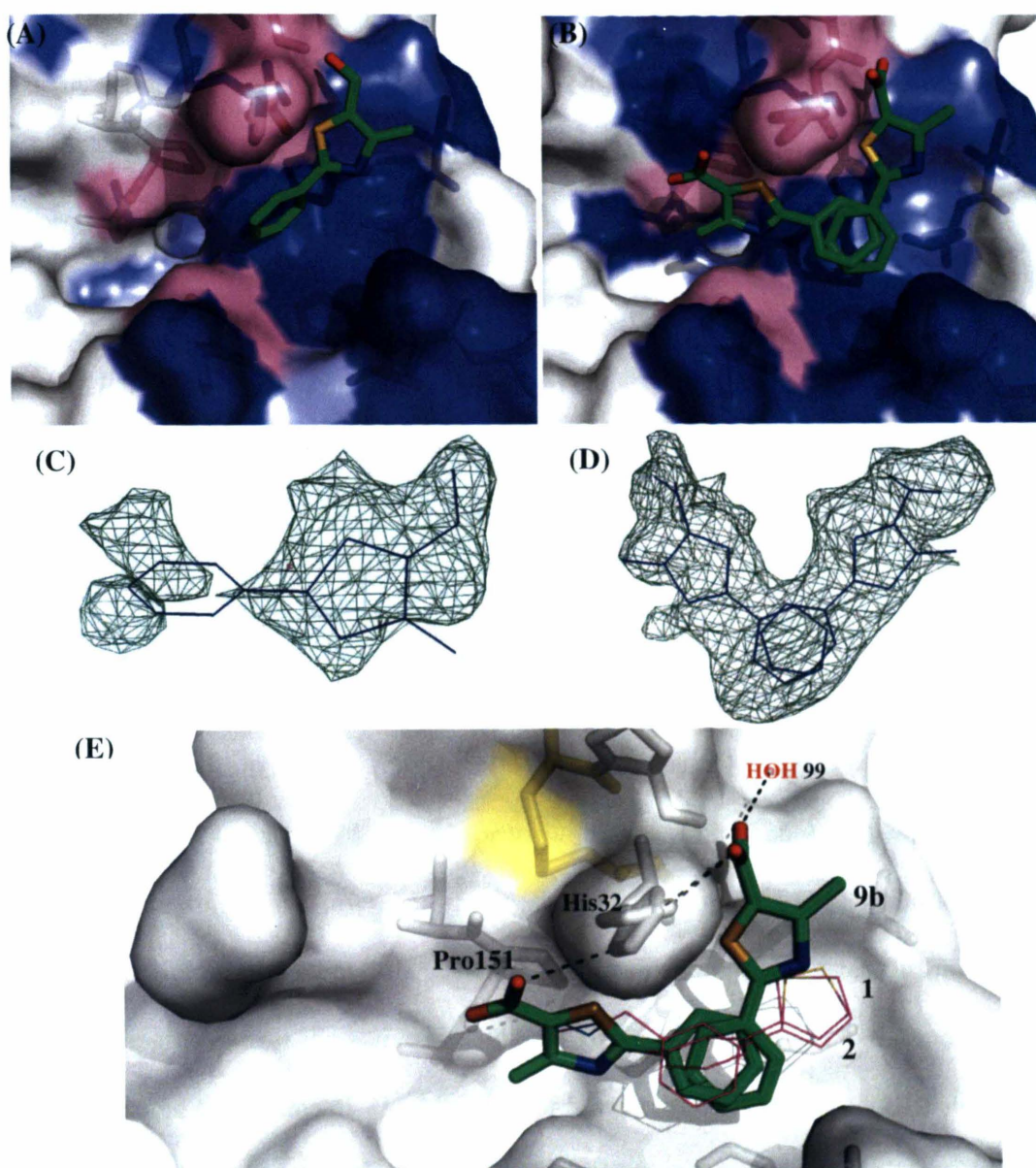


Figure 6-13 (A) and (B): NMR based heat maps (blue CSP > 0.03 ppm, pink: unassigned residues) with crystal bound coordinates of phenylthiazole fragments **9** and **9b** respectively (green stick representation). (C) and (D): Omit electron density maps of fragments **9** and **9b** respectively at contour level σ 2.5 obtained from crystallographic binding with *EcDsbA*. (E) *EcDsbA* (white surface representation) binding of **9b** compared with phenylthiophenes **1** and **2** (as pink wire representation).

The phenylthiazole class of ligands provides novel hydrogen bonding networks around the *EcDsbA* active site which have not been observed in any other explored fragments in this study, suggesting these hits may also represent suitable starting points for further elaboration.

6.3.6. Characterization of phenol hit fragment 10

Among the remaining *EcDsbA* fragment hits, about 10 fragments were identified which contained only one aromatic ring substituted with 1-3 simple functionalities. These hits also appeared to bind around the active site with varying patterns of ^1H - ^{15}N HSQC CSP.

The phenol fragment **10** (referred to as compound **4** in Chapter 3) was identified as the strongest of these hits, based on ^1H - ^{15}N HSQC analysis (Figure 6-14). A dissociation constant of $33 \pm 24 \mu\text{M}$ was calculated for trifluoromethylthiophenol (**10**), which due to its low molecular weight and low μM affinity showed the highest LE ($2.1 \text{ kJ mol}^{-1} \text{ HAC}^{-1}$) amongst the initial fragment hits. Chemical shift mapping suggested it binds in the hydrophobic groove near the active site. Phenol **10** showed inhibition in the motility assay with a measured EC_{50} value of $73 \mu\text{M}$. It is worth mentioning that fragment **10** also showed the highest inhibition in several other *in vitro* assays to measure direct inhibition of DsbA. These assays were optimized by Dr Kieran Rimmer (MIPS) and further details of these studies are given in appendix section 4.5.

Crystallographic soaking experiments resulted in the determination of the structure of the *EcDsbA*-fragment **10** complex at 2 \AA resolution after refinement. Interestingly, in this case a novel fragment binding site was observed on the hydrophobic groove near Gln35 and

Met171 (Figure 6-14). X-ray crystallography revealed that phenol **10** binds in a site 6 Å away from the position of the benzofuran hit **6**, whereas it overlaps with the position of the thiophene rings of phenylthiophenes **1-2** and the phenyl of phenoxybenzene **4** (Figure 6-14 and 6-15). The trifluoromethylthio-group fits in a sub-pocket on the hydrophobic groove completely and interacts via hydrophobic contacts with Phe174 (Figure 6-14). The structure of this fragment bound to *EcDsbA* showed that its aromatic ring lies on an essentially hydrophobic surface in the vicinity of non-polar and uncharged residues Met171, Leu40 and Val39. Additionally, a hydrogen bond (1.94 Å, $N-H\cdots O$) was observed between the phenol oxygen atom and the side chain amide of residue Glu35 from the $\alpha 1$ helix of the thioredoxin domain (Figure 6-14). The location of the CSP in the 1H - ^{15}N HSQC spectrum is consistent with the crystallographic binding position of fragment **10**. The hydrogen bond observed with this fragment is a new polar interaction to *EcDsbA*. As such, fragment **10** provides valuable new information for further chemical optimization to potent leads. Additionally, the binding location of hit **10** raises interesting possibilities for fragment elaboration by linking or merging (Figure 6-15) strategies.²³⁷⁻²³⁸

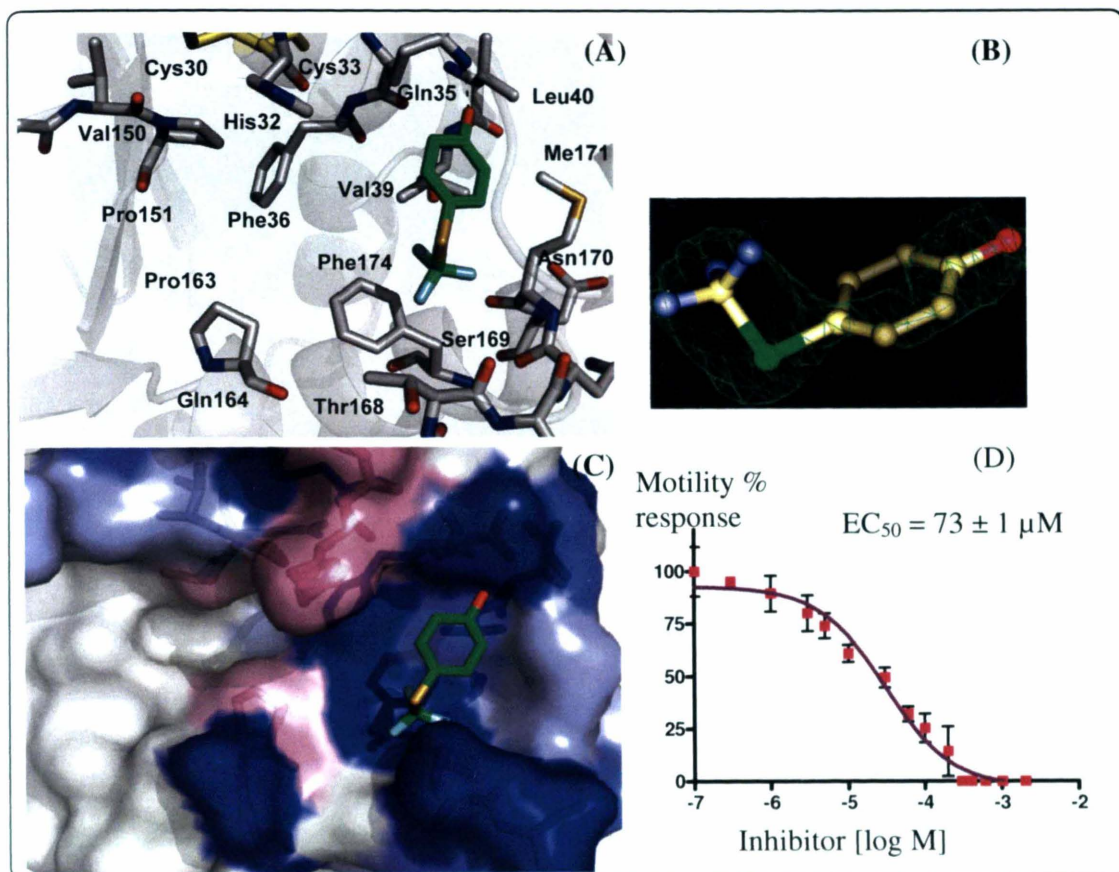


Figure 6-14: Characterization of phenol fragment hit **10**. (A) Crystallographic binding mode in hydrophobic groove of *EcDsbA* (B) Omit electron density maps of **10** (shown as yellow carbon stick models) at contour level σ 2.5 (C) NMR based heat map (blue: (NH) > 0.03 ppm, pink: unassigned residues) with crystallographic position of **10** (D) Motility inhibition of wt *E. coli* shown as sigmoidal dose response plot obtained from titration of fragment **10** identified by NMR screening.

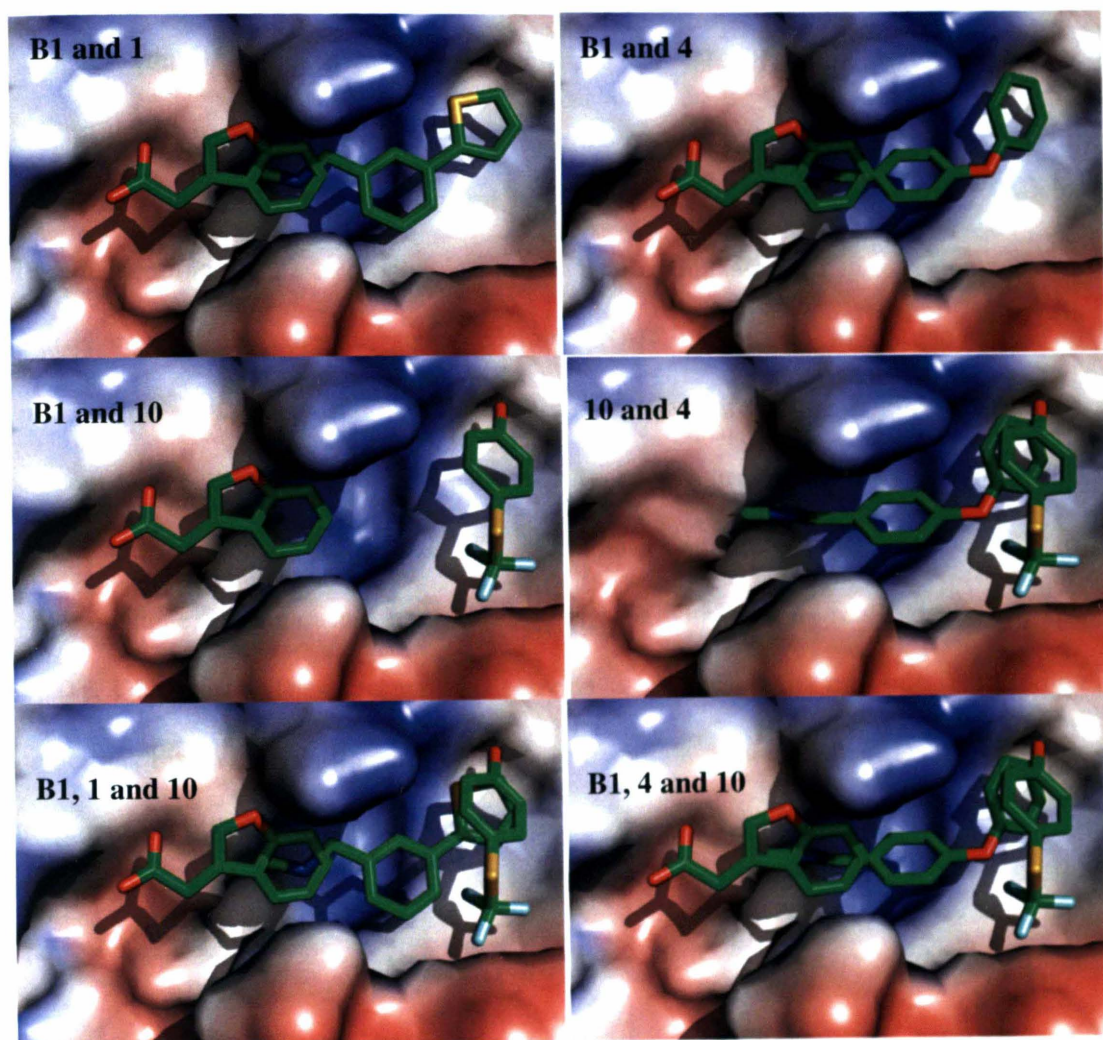


Figure 6-15: Overlay views of crystallographic binding positions of phenylthiophene fragments **1**, phenoxyphenyl fragment **4**, benzofuran fragment **B1** and phenol fragment **10** (all shown as green carbon stick models) in overlapping and adjacent binding sites on the hydrophobic groove of *EcDsbA* (which is shown as an electrostatic surface model)

In summary, NMR-based fragment screening identified several promising hits and their analogues which were shown to possess strong ^1H - ^{15}N HSQC and crystallographic binding and/or to exhibit inhibitory activity against *Ec*DsbA. The resulting structural information yielded insight into several novel series of *Ec*DsbA binders that bind in adjacent locations in the hydrophobic groove. The optimization of these fragments via small changes to the core has generated initial SAR information which can be used to guide further elaboration of *Ec*DsbA ligands. A strategy involving linking or merging fragments crystallized in nearby locations may be a more viable way to probe this binding site. We have identified the preliminary binding features of benzofuran, benzothiophene, chroman and benzothiazole classes with various scaffolds **B1**, **6**, **C4**, **C5**, **D1**, **D2** and **9b**. These compounds provide valuable insights into possible binding interactions that will assist in further efforts to elaborate these fragment DsbA inhibitors into potent lead-like molecules.

Chapter 7

Summary and Future Directions

DsbA is a periplasmic oxidoreductase enzyme and had been found to be centrally responsible for the correct oxidative folding of numerous virulence determinants in Gram-negative bacteria. Hence DsbA represents a potential target for developing inhibitors that could counteract the virulence of Gram-negative pathogens. This thesis describes a structure assisted-fragment based lead discovery program carried out to validate DsbA as an antibacterial drug target by identifying small molecule inhibitors. This project has utilized a broad range of complementary techniques, from NMR based fragment library screening, validation and biophysical characterization of identified fragment hits, DsbA inhibition activity assessment, structural studies and fragment elaboration to identify potential inhibitors of DsbA.

The FBDD studies conducted in this project employed biophysical screening and binding characterization with approaches that required relatively large quantities of pure protein. For the experiments conducted in chapters 3-6, bacterial protein expression procedures for labeled and unlabeled proteins were employed as outlined in chapter 2. Overexpression utilizing *E. coli* (BL21) strains containing *EcDsbA* and *VcDsbA* expression vectors allowed for milligram quantities of DsbA per litre of media to be produced. The purification procedure outlined in chapter 2 involved two steps of FPLC (PHP and MonoQ columns) to produce pure protein as shown by the single band on SDS-PAGE (Figure 2-3).

The study conducted in chapter 3 outlines the power of a fragment-based strategy in the field of DsbA-ligand discovery, providing potential lead compounds for the development

of antivirulence drugs. A fragment library comprised of chemically diverse compounds was screened in two rounds of NMR experiments (STD and HSQC) in order to find potential small molecule inhibitors of the bacterial oxidoreductase protein *EcDsbA*. This screening led to the identification of distinct fragment classes, which were demonstrated to bind near the active site of *EcDsbA* by NMR spectroscopy coupled with X-ray crystallography. Several of these hits also inhibited DsbA-dependant activity as assessed by measuring the DsbA-dependent phenotype of motility. Structural studies revealed that the strongest fragment hits bind in the hydrophobic groove of *EcDsbA* in adjacent or overlapping positions. These hits were shown to possess high LE and low μM EC_{50} values against DsbA, providing excellent starting points for the development of more potent *EcDsbA* inhibitors.

In chapter 4, a fragment-based approach was used to identify and characterize binders and potential inhibitors of *VcDsbA*, which is a homologue of *EcDsbA* and also represents a suitable antibacterial target. To accomplish this aim for target enzyme *VcDsbA*, a combination of experimental and computation approaches were employed, including multidimensional NMR, HSQC-guided docking, and surface grid map based screening along with the bacterial motility assay. Reliable docked solutions were generated using NMR constraint-based docking for selected compounds to study their binding mode to *VcDsbA*. Fourteen scaffolds exhibiting high to moderate affinity were discovered by employing a new approach to virtual selection using *VcDsbA* site grid maps. The close correlation between computational results and experimental data suggested that this protocol is robust for *VcDsbA*-small molecule interactions, hence can be used for the

extension of fragments into larger molecules with greater potency. Previously identified benzophenone and benzimidazole fragments were further characterized and elaborated and an initial SAR analysis of these classes of *VcDsbA* inhibitors was developed. Moreover, a bacterial motility assay was optimized to study inhibition and the selectivity profile of identified *VcDsbA* hits. Overall results shows that *VcDsbA* is a tractable target, and the compounds discovered should provide a firm platform for a further lead generation process.

Chapters 5 and 6 represent the fragment elaboration studies for the most promising hits identified from the NMR-based fragment screen of *EcDsbA*, as described in chapter 3. These hits were further characterized by NMR-based K_D , cell-based inhibition assay and high resolution or preliminary X-ray crystallography. After initial characterization a diverse set of analogues were screened to optimize the potency and binding affinity of each fragment. The resulting structural information yielded insights into binding features of several novel series of *EcDsbA* ligands. Chapter 5 described the further elaboration of those original hits for which high resolution crystal structures could be obtained. Additionally, NMR-based docking studies complemented the crystal structures for newer ligands and proved useful for SAR analysis and analogue design.

A systemic strategy was established for the further optimization of phenyl thiophene and phenoxy benzene chemical classes using a combination of experimental approaches. These ligands have a high LE with suitable physiochemical properties for further development. Structural analysis of phenyl thiophene analogues containing diverse

functional groups suggested that extension from the aminomethyl group of the 3-phenyl thiophene core could place a hydrophobic group into the groove to significantly enhance the affinity. Additionally a number of analogues were investigated containing a variety of functional groups at different positions to observe the preference for binding to DsbA. Novel phenoxybenzene analogues with improved binding to *Ec*DsbA were identified which represent suitable points for further optimization. The crystallographic binding mode of the strongest-binding analogue suggested that substitution with a larger hydrophobic group from the 4-position of the phenoxybenzene core would create favourable interactions. Additionally, further elaboration of compounds which contain merged features of phenylthiophenes and phenoxybenzene classes can provide useful insights to develop a novel series for potential inhibitors of DsbA.

Chapter 6 described the exploration and characterization of original hit fragments from the benzofuran, benzothiophene, chroman, benzothiazole and phenol classes. For these classes preliminary crystallographic data was used in combination with NMR data to propose binding models, which were supported by inhibition data. This chapter also describes experimental evidence for the binding of ligands at alternate sites in the hydrophobic groove of *Ec*DsbA. This binding diversity is potentially valuable in combination with the other inhibitors described in Chapter 5, in proposing avenues to explore further optimization of these ligands through linking or merging approaches. The growth of these fragments via small substitutions to the core has generated some initial SAR information, however further exploration is required to enable development of these series. A strategy involving linking or merging fragments crystallized in nearby locations

may be a more viable way to probe these adjacent binding sites. These compounds were also shown to possess strong binding and to exhibit inhibitory activity against *EcDsbA*. The polar and hydrophobic contacts made by these compounds will prove valuable in the development of potential *EcDsbA* inhibitors.

In conclusion, this thesis opens up a range of possibilities of study for future investigation. The discovery of small molecule inhibitors with nM potency for DsbA remains to be achieved, however these ongoing efforts are likely to benefit from the SAR and protein-ligand structures of *EcDsbA* in complex with a variety of small drug-like molecules, such as those described in this study. Additionally, the cell-based phenotypic assay results described for the molecules in this thesis provide convincing evidence that the inhibition of DsbA is worth pursuing in the drive to develop new, more effective antimicrobial therapies.

Chapter 8

Appendices

8.1 Full papers

8.1.1 E.coli DsbA-Peptide Complex

8.1.2 Pseudomonas aeruginosa DsbA – Structure and Characterization

The Structure of the Bacterial Oxidoreductase Enzyme DsbA in Complex with a Peptide Reveals a Basis for Substrate Specificity in the Catalytic Cycle of DsbA Enzymes^{*,[5]}

Received for publication, March 4, 2009, and in revised form, April 22, 2009. Published, JBC Papers in Press, April 22, 2009, DOI 10.1074/jbc.M109.011502

Jason J. Paxman^a, Natalie A. Borg^{5,1}, James Horne^a, Philip E. Thompson^a, Yanni Chin^a, Pooja Sharma^a, Jamie S. Simpson², Jerome Wielens², Susannah Piek³, Charlene M. Kahler^{3,2}, Harry Sakellaris⁴, Mary Pearce¹, Stephen P. Bottomley¹, Jamie Rossjohn^{5,3}, and Martin J. Scanlon^{1,4}

From ^aMedicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Monash University (Parkville Campus), 381 Royal Parade, Parkville, Victoria 3052, the ^bProtein Crystallography Unit, Australian Research Council Centre of Excellence in Structural and Functional Microbial Genomics, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, the ^cSchool of Biomedical, Biomolecular and Chemical Sciences, QEII Medical Centre, University of Western Australia, Crawley, Western Australia 6009, and the ^dDepartment of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia

Oxidative protein folding in Gram-negative bacteria results in the formation of disulfide bonds between pairs of cysteine residues. This is a multistep process in which the dithiol-disulfide oxidoreductase enzyme, DsbA, plays a central role. The structure of DsbA comprises an all helical domain of unknown function and a thioredoxin domain, where active site cysteines shuttle between an oxidized, substrate-bound, reduced form and a DsbB-bound form, where DsbB is a membrane protein that reoxidizes DsbA. Most DsbA enzymes interact with a wide variety of reduced substrates and show little specificity. However, a number of DsbA enzymes have now been identified that have narrow substrate repertoires and appear to interact specifically with a smaller number of substrates. The transient nature of the DsbA-substrate complex has hampered our understanding of the factors that govern the interaction of DsbA enzymes with their substrates. Here we report the crystal structure of a complex between *Escherichia coli* DsbA and a peptide with a sequence derived from a substrate. The binding site identified in the DsbA-peptide complex was distinct from that observed for DsbB in the DsbA-DsbB complex. The structure revealed details of the DsbA-peptide interaction and suggested a mechanism by which DsbA can simultaneously show broad specificity for substrates yet exhibit specificity for DsbB. This mode of binding was supported by solution nuclear magnetic resonance data as well as functional data,

which demonstrated that the substrate specificity of DsbA could be modified via changes at the binding interface identified in the structure of the complex.

The formation of disulfide bonds is a critical step in the correct folding and stability of many secreted proteins. In Gram-negative bacteria, disulfide bond formation occurs in the periplasm and is catalyzed by enzymes of the Dsb family. The Dsb family contains several members, which mediate different aspects of disulfide bond formation and isomerization (1). DsbA is the enzyme that is primarily responsible for the formation of disulfide bonds in newly synthesized substrate proteins (Fig. 1). In this reaction, oxidized DsbA reacts with a substrate protein to generate a mixed disulfide intermediate. This covalent reaction intermediate is rapidly resolved to release the oxidized substrate and reduced DsbA. Reduced DsbA is in turn reoxidized by the inner membrane protein DsbB (2).

Bacteria lacking a functional DsbA display pleiotropic phenotypes, since the folding of a large number of disulfide bond-containing proteins is disrupted. Many of these are secreted proteins, such as toxins and surface proteins, that contribute to bacterial virulence (3). For example, DsbA is required for the formation of a functional type III secretion system in many bacteria, including *Pseudomonas aeruginosa* (4, 5), *Shigella flexneri* (6), *Salmonella enterica* serovar Typhimurium (7), and *Yersinia pestis* (8); *dsbA* *Vibrio cholerae* are unable to secrete cholera toxin (9); *dsbA* strains of *E. coli* exhibit reduced levels of β -lactamase activity (10) and are hypersensitive to benzyl penicillin, dithiothreitol (11), and some divalent metal cations (12). Furthermore, DsbA has been shown to be necessary for intracellular survival of *S. flexneri* (13) and *P. aeruginosa* (4), and DsbA is required for virulence of *S. enterica* in a mouse infection model (7). Each of these phenotypes has been attributed to the lack of disulfide formation in protein substrates of DsbA. Thus, there has been considerable interest in the structural basis of DsbA activity and selectivity and its role in bacterial virulence.

* This work was supported in part by Australian Research Council (ARC) Grant LP0455508 and National Health and Medical Research Council (NHMRC) Grant 455860.

The atomic coordinates and structure factors (code 3DMS) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

⁵ The on-line version of this article (available at <http://www.jbc.org/>) contains supplemental Figs. 1–4 and Tables 1 and 2.

¹ Supported by an NHMRC Career Development Award.

² Supported by the Ada Bartholomew Trust.

³ Supported by an ARC Federation Fellowship. To whom correspondence may be addressed. Tel.: 61-3-99029236; Fax: 61-3-99054699; E-mail: jamie.rossjohn@med.monash.edu.au.

⁴ To whom correspondence may be addressed. Tel.: 61-3-99039540; Fax: 61-3-99039582; E-mail: martin.scanlon@pharm.monash.edu.au.

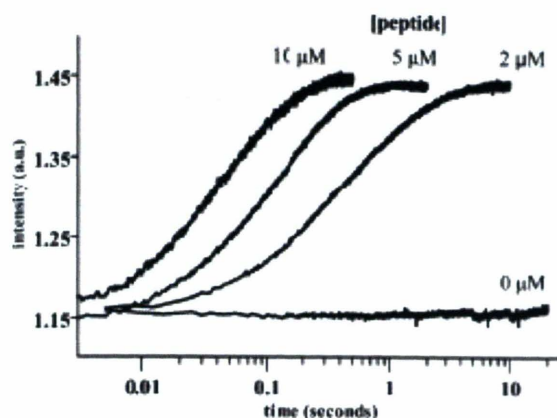


FIGURE 2. Oxidation of the SigA peptide by DsbA. Oxidase activity was determined by monitoring the oxidation of SigA peptide substrate. The kinetics of the oxidation reactions were determined using stopped-flow fluorescence by monitoring the increase in fluorescence upon reduction of EcDsbA (1 μ M) in the assay. Progress curves for oxidation of the peptide substrate are shown at increasing peptide concentrations (0, 2, 5, and 10 μ M). The rate of oxidation increases with increasing peptide concentration.

peptide was determined. The covalent complex was stabilized by replacing the intermolecular disulfide with a more stable isosteric thioether bond (37) to prevent product release. To form the complex, a nine-residue peptide encompassing SigA residues Pro⁵⁷⁰–Asp⁵⁷⁸ was synthesized with a homoserine residue (Hse) in place of the native cysteine (Ac-P¹IPFL-Hse-QKD⁹-NH₂). The hydroxyl of the homoserine was substituted with bromine to generate the homobromoalanine analogue, which was reacted with Cys³⁰ of EcDsbA to form a covalent thioether complex between the peptide and EcDsbA. The isosteric thioether bond is more stable than the corresponding disulfide, and hence the thioether-linked peptide is essentially bound irreversibly to EcDsbA. The peptide was synthesized with an acetylated N terminus and an amidated C terminus to avoid introduction of charges at the termini of the peptide, which would not be present in the context of the intact SigA protein. Diffraction quality crystals of the complex were formed by soaking the homobromoalanine-SigA-peptide into pre-formed crystals of reduced EcDsbA. Analysis of the crystals by SDS-PAGE revealed that the peptide had formed a complex with ~50% of the DsbA in the crystal, as estimated from the intensity of the respective bands on SDS-polyacrylamide gels (Fig. 3a). The structure of the complex was solved by molecular replacement and refined to a resolution of 1.9 Å ($R_{\text{work}} = 21.7\%$, $R_{\text{free}} = 25.2\%$) (Table 1).

Description of the Structure—Each of the four EcDsbA molecules within the asymmetric unit adopted a typical DsbA fold (Fig. 3b), wherein the active site C₃₀PHC₃₃ sequence of EcDsbA is positioned at the end of helix α 1 within the TRX domain of the protein. The TRX domain contains an inserted α -helical domain. The two domains are linked via a loop between strand β 3 of the TRX domain and helix α 2 in the helical domain (residues 62–66) and through a loop at the end of the long helix α 6 (residues 129–144) (14). The four EcDsbA molecules within the asymmetric unit were highly similar to one another and

The Structure of a DsbA-Peptide Complex

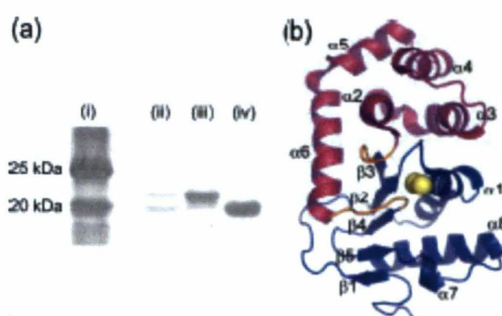


FIGURE 3. Analysis of the DsbA peptide crystal. a, analysis by SDS-PAGE (lane ii) revealed that approximately half of the EcDsbA in the crystal had formed a covalent complex with the peptide. The peptide-DsbA complex (lane ii) was clearly resolved from free EcDsbA (lane iv) under the conditions used to run the gel. (Molecular weight markers are shown in lane i.) b, ribbon diagram of EcDsbA (Protein Data Bank code 3DKS, chain C). Each of the four molecules in the asymmetric unit adopted a typical DsbA fold comprising a thioredoxin (blue) and α -helical (magenta) domains and the insertion points (orange). The sulfur atoms of the active site cysteine residues are shown in yellow CPK representation.

TABLE 1

Data collection and refinement statistics

Values in parenthesis are for the highest resolution shell.

| DsbA-peptide complex | |
|------------------------------------|-------------------|
| Data collection | |
| Space group | $P2_12_12_1$ |
| Cell dimensions | |
| a, b, c (Å) | 85.4, 87.9, 112.5 |
| α, β, γ (degrees) | 90.0, 90.0, 90.0 |
| Resolution (Å) | 1.9 |
| R_{merge} | 0.109 (0.793) |
| $I/\sigma(I)$ | 22.1 (2.5) |
| Completeness (%) | 99.8 (99.7) |
| Redundancy | 7.1 (6.6) |
| Refinement | |
| Resolution (Å) | 1.9 |
| No. of reflections | 66,227 |
| $R_{\text{merge}}/R_{\text{free}}$ | 21.7/25.2 |
| No. of atoms | |
| Protein | 5805 |
| Ligand/ion | 117 |
| Water | 488 |
| B-factors | |
| Protein | 24.48 |
| Ligand/ion | 37.41 |
| Water | 28.84 |
| r.m.s. deviations | |
| Bond lengths (Å) | 0.009 |
| Bond angles (degrees) | 1.02 |

superimposed over the backbone atoms for residues 2–184 with overall root mean square (r.m.s.) deviation of <0.5 Å. They are also similar to previously reported structures of reduced (Protein Data Bank code 1A2L) and oxidized (Protein Data Bank code 1FVK) EcDsbA and superimpose over the backbone atoms for residues 2–184 with overall r.m.s. deviation of <0.5 Å. Of the four EcDsbA molecules in the asymmetric unit, two (chains C and D) were covalently linked to a SigA peptide (chains E and F, respectively), whereas the other two DsbA molecules (chains A and B) were present in their reduced form. The covalent complex was formed via a thioether bond between the C₃₀ in the homoserine of the SigA peptide and S₃₀ in Cys³⁰ of EcDsbA. Analysis of the structures revealed that binding of the peptide to the first EcDsbA molecule blocked the equivalent binding site on the adjacent EcDsbA, accounting for the obser-

The Structure of a DsbA-Peptide Complex

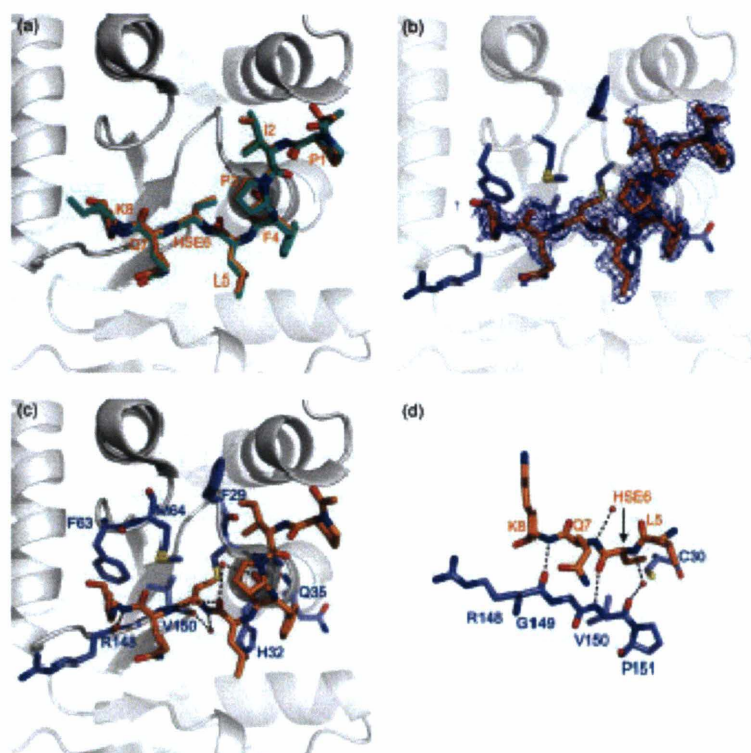


FIGURE 4. The substrate binding site of DsbA. *a*, comparison of the peptide conformation in the two models of the EcDsbA-peptide complex (Protein Data Bank code 3DKS). Peptide chains E and F are shown in orange and green stick representation, respectively, and the residues are labeled in orange. The EcDsbA molecule to which peptide E is covalently attached (chain C) is shown in a schematic diagram. *b*, omit ($2F_o - F_c$) electron density map (contoured at 0.8 σ) for the peptide (chain E). EcDsbA residues in contact with the peptide are shown in a blue stick representation. *c*, polar interactions observed in the complex are shown as black dotted lines between the interacting residues. Water atoms are shown as red spheres. The EcDsbA-peptide complex (chains C and E) is shown in the same orientation as above, and EcDsbA residues are labeled in blue. *d*, the peptide forms an antiparallel interaction with residues in the cis-proline loop of EcDsbA. Polar interactions are shown as black dotted lines between the peptide (in orange) and EcDsbA (in blue).

vation that only half of the EcDsbA molecules in the crystal were present as a complex. The peptides in each of the two EcDsbA-peptide complexes were highly similar and had an r.m.s. deviation of 0.21 Å over eight C α atoms (Fig. 4*a*). The side chains of the glutamine at position 7 and lysine at position 8 were mobile in both peptide complexes, as determined by the lack of prominent electron density in the $2F_o - F_c$ map (Fig. 4*b*). The aspartate at position 9 was disordered and not modeled in either complex.

DsbA-Peptide Interaction—Analysis of the structure of the covalent EcDsbA-SigA-peptide complex revealed that the peptide was bound at the interface of the TRX and α -helical domains of DsbA, with contacts being made to both domains. Details of the DsbA-peptide complex are presented in Fig. 4. The intermolecular thioether bond displayed an architecture similar to that of a right-handed disulfide bond (47) with dihedral angles defined by $C_{\alpha C30}-C_{\beta C30}-S_{\gamma C30}-C_{\gamma H100}$, $C_{\beta C30}-S_{\gamma C30}-C_{\gamma H100}-C_{\beta H100}$, and $S_{\gamma C30}-C_{\gamma H100}-C_{\beta H100}-C_{\alpha H100}$ of 76°/81°, 92°/100°, and 147°/154° in the two intermolecular complexes of the asymmetric unit, respectively. This is a conformation similar to

that of the mixed disulfide that is observed in a TRX-substrate complex (48), indicating that the thioether provides a suitable isosteric replacement of the disulfide.

In addition to the covalent linkage between the peptide and EcDsbA, substrate binding was stabilized by hydrogen bonding and van der Waals interactions (Fig. 4, *c* and *d*), and the binding interface was almost identical in each complex. Hse⁶ and Lys⁸ from SigA formed hydrogen bonds with Val¹⁵⁰ and Arg¹⁴⁸ in the cis-proline loop of EcDsbA, respectively, such that the loop and substrate peptide were arranged in an antiparallel fashion, and a water-mediated hydrogen bond was formed between the amide proton of Hse⁶ and the carbonyl oxygen of Val¹⁵⁰ (Fig. 4*d*). In addition to the hydrogen-bonding interactions, Phe²⁹ of EcDsbA packed against Ile² of the peptide, whereas Pro²¹, His³², and Gln³⁵ of DsbA made contacts to Phe⁴ of SigA. Hse⁶ of the peptide contacted Met⁶⁴ of DsbA, and Gln⁷ and Lys⁸ made contacts with Phe⁶³, Arg¹⁴⁸, and Gly¹⁴⁹. The binding interface in the DsbA-peptide complex has a total buried surface area of 760 and 795 Å² in the two structures, respectively, and a surface complementarity of 74% as calculated in CCP4 (49). There were very slight differences in the packing interactions in the two

DsbA-substrate complexes present in the asymmetric unit of the unit cell as a result of small changes in the orientation of the side chain of Ile² and the mobility of Lys⁸ (Fig. 4*a*). In addition, there were slight differences in the side chain orientations of Phe²⁹ and Met⁶⁴ in the two molecules of EcDsbA that were covalently attached to the peptide. Notwithstanding these small differences, the 10 residues of EcDsbA that contact the peptide in the complex could be superimposed over their C α atoms with an r.m.s. deviation of 0.12 Å, and the H-bonding interactions were conserved in the two complexes, indicating that the mode of interaction is similar. In both cases, residues around the active site (²⁹FCPH³²) together with those in the loop connecting the TRX and α -helical domains (Phe⁶³ and Met⁶⁴) and residues in the loop preceding the cis-Pro¹⁵¹ (Arg¹⁴⁸, Gly¹⁴⁹, and Val¹⁵⁰) formed >95% of the binding surface area.

Characterization of DsbA-Peptide Binding—Given that only two of the EcDsbA molecules out of the four present in the asymmetric unit were observed to bind the peptide, there was concern that the mode of peptide binding was a crystallographic artifact. In order to address this concern, the binding

The Structure of a DsbA-Peptide Complex

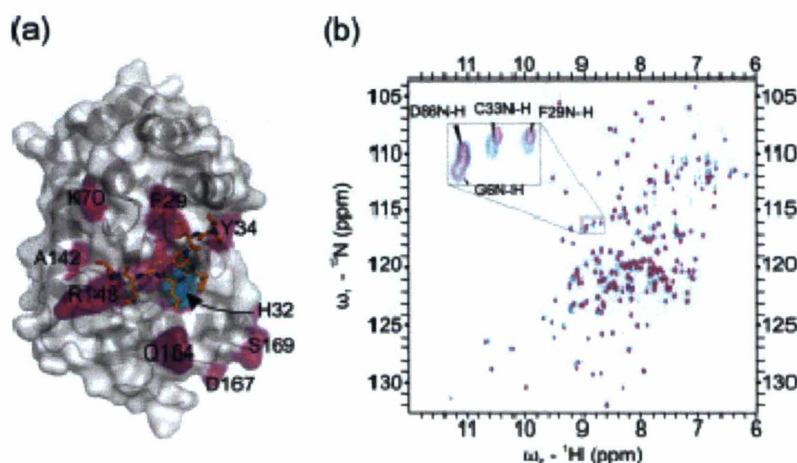


FIGURE 5. Analysis of EcDsbA-SigA peptide interaction by NMR spectroscopy. *a*, EcDsbA residues for which chemical shift perturbations are observed upon the addition of the peptide are colored in magenta on the crystal structure of the EcDsbA-peptide complex (Protein Data Bank code 3DKS, chain C; the peptide 3DKS, chain E, is in orange stick format). His³², which is not observed in either HSQC spectrum is colored in cyan. A continuous surface is formed by residues that form the EcDsbA-peptide interface in the crystal. *b*, ¹⁵N HSQC spectrum of EcDsbA in the absence (magenta) and presence (cyan) of SigA peptide. The expansion demonstrates perturbations that are observed for some residues (e.g. Phe¹⁶⁸ and Cys¹⁰⁴), whereas other residues (e.g. Gly⁶ and Asp⁹⁶) are unaffected.

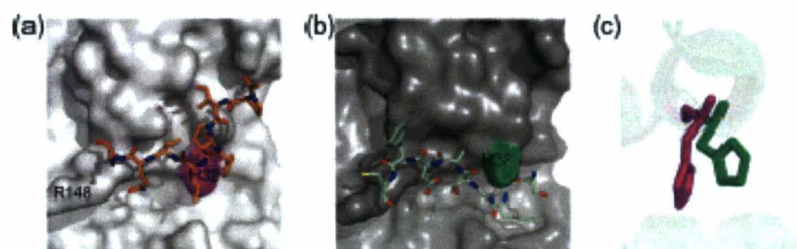


FIGURE 6. Comparison of the complexes of substrate-DsbA and DsbB-DsbA (Protein Data Bank code 2HI7). *a*, surface view of the interaction between EcDsbA and SigA peptide. The surface of EcDsbA is shown with the SigA peptide in an orange stick representation. The position of His³² on the surface is highlighted in magenta. *b*, structure of the DsbA-DsbB complex. The surface of EcDsbA is shown with the bound Cys¹⁰⁴ loop in a green stick representation, and the position of His³² is highlighted. *c*, superposition of the two complexes reveals that His³² in the complex with SigA peptide (magenta) adopts a different rotameric state compared with that observed in the complex with DsbB.

location of the SigA peptide was determined by measuring chemical shift perturbations in ¹H-¹⁵N HSQC NMR experiments on uniformly ¹⁵N-labeled EcDsbA in the absence and presence of the peptide (Fig. 5). The chemical shifts of resonances in HSQC spectra are exquisitely sensitive to changes in the environment of the amide groups, and measurement of chemical shift perturbation is widely used as a means to identify the location of ligand binding. A peptide containing homoserine in place of cysteine was employed to prevent formation of a covalent complex and to allow identification of residues whose resonances were perturbed in the context of a noncovalent complex between the peptide and EcDsbA. Analysis of the spectra revealed that many of the most significant perturbations were observed for residues that formed a continuous surface at the interface between the α -helical and TRX domains (Fig. 5*a*). The location of the perturbed chemical shifts in the NMR spectra was consistent with the mode of binding observed in the x-ray structure and suggests that the structure represents a relevant model of the biological complex. In contrast, many of

the residues in the hydrophobic groove that is the binding site for the periplasmic loop of DsbB were observed in the ¹⁵N HSQC spectrum of EcDsbA but were not significantly perturbed upon the addition of the peptide. This suggests that the SigA peptide binds in a location that is distinct from the binding site observed in the crystal structure of the DsbA-DsbB complex (21). An interesting feature of NMR data was that no peak was observed in the ¹⁵N HSQC spectrum for His³² either in the presence or absence of peptide, which is consistent with this residue undergoing conformational exchange, which has previously been suggested to be important in substrate recognition and catalysis in DsbA enzymes (50).

Comparison of DsbA-Substrate and DsbA-DsbB Complex.—Comparison of our structure with the EcDsbA-EcDsbB complex structure revealed that the conformation of EcDsbA was similar in both cases (Fig. 6). The two EcDsbA molecules (Protein Data Bank codes 3DKSD and 2HI7A) superimposed with an r.m.s. deviation of 0.59 Å over 179 C α atoms. Both the SigA peptide and the periplasmic loop of EcDsbB interacted with ¹⁴⁸RGV¹⁵⁰ in the *cis*-proline loop of EcDsbA; however, the side chains of residues ⁹⁹PFA¹⁰¹ from the DsbB loop were buried in the hydrophobic groove, where they interacted with Pro¹⁵¹,

Pro¹⁶³, Gln¹⁶⁴, Thr¹⁶⁸, Met¹⁷¹, and Phe¹⁷⁴ of EcDsbA (Fig. 6*b*). Thus, although the general mode of interaction with DsbA is maintained, DsbB appears to make additional interactions within the hydrophobic groove on the surface of DsbA. These result in the higher buried surface area observed in the DsbA-DsbB complex (1340 Å²) (21), which may also account for the greater specificity of the DsbA-DsbB interaction.

Despite their general similarity, there were some differences observed between the DsbA molecules in the DsbA-substrate and DsbA-DsbB complexes. A notable difference was the orientation of the side chain of His³² (Fig. 6*c*). In the complex with peptide (Protein Data Bank code 3DKS chain C), His₃₂ adopted a *gauche*⁺ conformation ($\chi_1 = 82^\circ$), where its side chain lay across the face of the hydrophobic groove and blocked access to the peptide substrate. In the complex with DsbB (Protein Data Bank code 2HI7), His³² adopted a *trans* conformation ($\chi_1 = -177^\circ$), which allowed access to the hydrophobic groove of DsbA and where its side chain made van der Waals interactions with Ala¹⁰²-Thr¹⁰³ in the DsbB loop. In the original crystal

The Structure of a DsbA-Peptide Complex

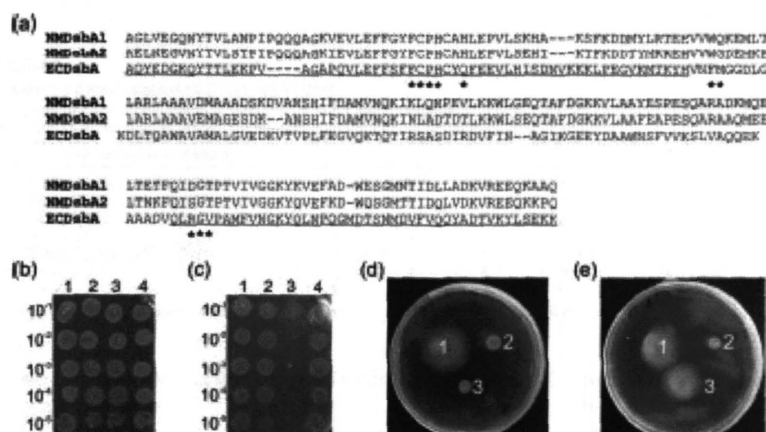


FIGURE 7. Phenotypic analysis of chimeric DsbA. *a*, primary structure of EcDsbA, NmDsbA1, and NmDsbA2. Residues comprising the thioredoxin domain of EcDsbA are underlined. Residues that make contacts to the peptide substrate in the crystal structure of the complex are highlighted (*). *b–e*, phenotypic assays on *dsbA*[−] *E. coli* (JCB571) complemented with different DsbA proteins. *b*, JCB571 alone (lane 3), JCB571 complemented with EcDsbA in pTrc99A as a positive control (lane 4), and JCB571 complemented with EcTDNmDsbA1α (lane 1) or EcTDNmDsbA2α (lane 2) were able to grow on LB agar without DTT, indicating that each of the cell lines was viable. *c*, the same cell lines were tested for growth on LB agar containing 15 mM DTT and 1 mM isopropyl 1-thio-β-D-galactopyranoside to induce DsbA expression. Each of the cell lines that had been complemented with the DsbA proteins grew, whereas JCB571 did not, indicating that each construct expressed an active DsbA. *d*, neither JCB571 (zone 2) nor JCB571 complemented with EcTDNmDsbA1α (zone 3) was motile. In contrast, the positive control (JCB571 complemented with EcDsbA) (zone 1) was motile. *e*, both the positive control strain (zone 1) and JCB571 complemented with EcTDNmDsbA2α (zone 3) were motile, whereas JCB571 (zone 2) was not.

TABLE 2
Phenotypic analysis of DsbA chimeras

All *dsbA* genes were cloned into the low copy vector pHS676 and assessed for their ability to restore *E. coli dsbA* mutant strain JCB571 resistance to DTT and motility.

| <i>dsbA</i> | DTT resistance | Motility |
|-------------------------------------|----------------|----------|
| JCB570 (wild type) | Yes | Yes |
| JCB571 (<i>dsbA</i> [−]) | No | No |
| JCB571 transformed with | | |
| EcDsbA | Yes | Yes |
| EcTDNmDsbA1α | Yes | No |
| EcTDNmDsbA1α (QLRGV) | Yes | Yes |
| EcTDNmDsbA2α | Yes | Yes |
| EcTDNmDsbA1α (QLRGV) | Yes | Yes |

structure of oxidized EcDsbA (14), there were two molecules in the asymmetric unit. These two molecules of oxidized EcDsbA differed in the orientation of the His³² side chain (in the structure of oxidized EcDsbA, Protein Data Bank code 1FVK, χ_1 = 160 and 72°, respectively). His³² has been shown to destabilize the oxidized form of EcDsbA (50), and it has been suggested previously that movement of this residue may be required to allow substrate access to the active site (50).

Functional Analysis of DsbA—To determine the role of the substrate-binding residues of DsbA in defining specificity, we constructed chimeric proteins comprising the TRX domain of EcDsbA and the α-domain of either NmDsbA1 or NmDsbA2. These are referred to as EcTDNmDsbA1α and EcTDNmDsbA2α, respectively. Alignment of the sequences of EcDsbA, NmDsbA1, and NmDsbA2 (Fig. 7*a*) revealed that within the chimeric DsbA enzymes, the loop ¹⁴⁸RGV¹⁵⁰ in EcDsbA is replaced by DGT in EcTDNmDsbA1α and EcTDNmDsbA2α, respectively. In addition, residues Phe⁶³-Met⁶⁴ of EcDsbA are replaced with WQ and WG in

EcTDNmDsbA1α and EcTDNmDsbA2α, respectively. In contrast, the residues that form the hydrophobic groove in EcDsbA are retained in the chimeras.

In order to test the effect of these mutations, phenotypic analysis was performed using the *dsbA*[−] *E. coli* strain JCB571 complemented with EcDsbA, EcTDNmDsbA1α, or EcTDNmDsbA2α. First, the sensitivity to DTT of each of the transformants was determined. Strains of *dsbA*[−] *E. coli*, such as JCB571, are sensitive to DTT and are unable to grow on LB agar containing DTT (Fig. 7, *b* and *c*). Transformation of JCB571 with either of the chimeric proteins or EcDsbA restored resistance to DTT. This suggested that each construct was capable of expressing a functional DsbA enzyme and that the expressed DsbA was able to be reoxidized in the periplasm.

DsbA is also required for the correct folding and stability of a component of the flagellar motor of *E. coli* (FlgI) (51). In the absence of DsbA, the P-ring component of the flagellar motor is not formed, which renders *dsbA*[−] *E. coli*, such as JCB571, nonmotile. The phenotypic assays revealed that cells transformed with either EcDsbA or EcTDNmDsbA2α were motile, whereas those transformed with EcTDNmDsbA1α were not (Fig. 7, *d* and *e*). This is in contrast to observations with the intact *Neisseria* DsbA enzymes, wherein NmDsbA1 restored motility to *dsbA*[−] *E. coli*, whereas NmDsbA2 did not (29). However, when the five residues immediately preceding the *cis*-proline in EcTDNmDsbA1α (QIDGT) were replaced with the corresponding sequence from EcDsbA (QLRGV), the resulting chimeric protein was capable of both conferring resistance to DTT and restoring motility in JCB571. Similarly, replacing the five residues preceding the *cis*-proline in EcTDNmDsbA2α (QISGT) with the corresponding sequence from EcDsbA (QLRGV), the resulting chimeric protein was capable of both conferring resistance to DTT and restoring motility in JCB571.

Despite the relatively high conservation within the sequences of NmDsbA1 and NmDsbA2, our analysis indicates that changes introduced by constructing the chimeric proteins are sufficient to alter the substrate specificity, but they do not destroy the oxidoreductase activity of the proteins or, apparently their ability to be reoxidized by DsbB. These data suggest that neither the TRX domain nor the α-domain is solely responsible for conferring substrate specificity and are consistent with the structural details, which reveal that binding takes place at a surface that encompasses the interface formed between the two domains. Further, they suggest that substrate recognition by oxidized DsbA is different in some respects from the interaction of reduced DsbA with DsbB, such that changes

The Structure of a DsbA-Peptide Complex

TABLE 3

Sequence conservation in DsbA substrates

Forty-five sequences of confirmed substrates of DsbA were compared to assess sequence conservation around the cysteine residue. Analysis of the sequences revealed that there was no conservation beyond the cysteine residue itself, as determined using ConScore

| Conservative mutations | | | | | | | | | | | | | |
|-----------------------------------|------|------|------|------|------|-------|------|------|------|------|------|--|--|
| Hydrophobic, small (AVLI) | 17 | 9 | 13 | 8 | 13 | 0 | 15 | 15 | 11 | 5 | 7 | | |
| Hydrophobic, aromatic (FYW) | 2 | 8 | 3 | 7 | 3 | 0 | 3 | 4 | 4 | 5 | 2 | | |
| Acidic (DE) | 6 | 5 | 6 | 5 | 7 | 0 | 5 | 4 | 2 | 7 | 9 | | |
| Basic (RHK) | 7 | 4 | 5 | 6 | 1 | 0 | 5 | 6 | 6 | 8 | 4 | | |
| Neutral polar (SMTCNQ) | 10 | 15 | 9 | 14 | 18 | 45 | 14 | 9 | 13 | 18 | 15 | | |
| Glycine | 2 | 3 | 6 | 4 | 2 | 0 | 2 | 1 | 5 | 1 | 2 | | |
| Proline | 1 | 1 | 3 | 1 | 1 | 0 | 1 | 6 | 4 | 1 | 5 | | |
| Percentage of conserved mutations | 40.5 | 36.6 | 36.1 | 35.0 | 42.9 | 100.0 | 35.7 | 39.5 | 36.1 | 41.9 | 40.5 | | |

in substrate specificity can be achieved without apparent loss of recognition of DsbB.

DISCUSSION

We have determined a high resolution crystal structure of EcDsbA in complex with a nine-residue peptide derived from the autotransporter protein SigA of *S. flexneri*. This structure revealed that the peptide bound outside the hydrophobic groove of EcDsbA and contacted residues at the interface between the TRX and α -domains. The complex was stabilized by main-chain hydrogen bonds between substrate and residues in the *cis*-proline loop of the TRX domain, such that the two peptide chains were arranged in an antiparallel fashion. The mode of binding observed in the peptide-DsbA complex displayed some similarity to that of a trapped substrate-thioredoxin complex (Protein Data Bank entry 2IWT) (48). In both cases, the complex was stabilized by backbone-to-backbone hydrogen bonds between the substrate and residues in the *cis*-proline loop of the TRX domain, and the loop and peptide were arranged in an antiparallel fashion. The heavy atoms of the residues of the active site (XCXX(C/S)) and the residues in the loop immediately preceding the conserved *cis*-proline residue for the two complexes (R¹⁴⁰GVP in EcDsbA and E⁸⁸AMP in thioredoxin) superimposed with an r.m.s. deviation value of 0.52 Å. Furthermore, the buried surface areas observed for the DsbA-substrate complex (760 and 795 Å²) were similar to that in the thioredoxin-substrate complex (760 Å²) (48). The presence of mostly main-chain interactions as well as the relatively low interacting surface in the EcDsbA-SigA complex is consistent with the broad substrate specificity observed with most DsbA enzymes and the diverse sequences of potential DsbA substrates (31). Calorimetric studies have previously demonstrated that the interactions between DsbA and its substrate proteins are relatively weak (37), which suggested that small changes in the active site of the enzyme may be sufficient to alter substrate specificity. This is supported by previous mutational data, whereby it was demonstrated that a V150G mutation in EcDsbA was defective in complex formation with substrate proteins (31). A similar mode of binding has now been observed for a number of TRX domain-containing enzymes, including thioredoxin (48, 52, 53), glutaredoxin (54), glutathione transferase (55) as well as the mixed disulfide complex of DsbC-DsbD (56). This mode of interaction is emerging as a general feature of substrate recognition in TRX domain-substrate complexes (48).

However, comparison of the current structure and that of the EcDsbA-EcDsbB complex, suggests that there are subtle differ-

ences in the manner in which reduced and oxidized EcDsbA bind their respective substrates. Thus, EcDsbB bound within the hydrophobic groove of reduced EcDsbA, whereas the SigA peptide substrate did not. Comparison of the primary structure of proteins that have been identified as substrate proteins for EcDsbA (Table 3) reveals that there is no conserved sequence of hydrophobic residues in the substrate proteins that would presumably be required for favorable interaction within the hydrophobic groove. In fact, for those proteins that have been confirmed to be substrates of EcDsbA, there is no sequence conservation in the residues flanking the cysteine (Table 3).

As a consequence of the different location of binding, the SigA peptide did not contact the conserved *cis*-Pro¹⁵¹ within the hydrophobic groove. It has previously been demonstrated that the mutation P151T in EcDsbA results in the accumulation of DsbA-substrate complexes (31) in the periplasm of *E. coli*. This suggests that the P151T mutant is able to both recognize and form a mixed disulfide with a range of substrate proteins, but is defective in the step of the reaction pathway where the complex is resolved with the release of oxidized substrate and reduced DsbA (Fig. 1). The structure and activity of EcDsbA in which the *cis*-proline was mutated to alanine has previously been reported (57). The activity of EcDsbA_{P151A}, as determined from its ability to oxidize alkaline phosphatase in a cell-based assay, was significantly lower than that of wild type EcDsbA. However, EcDsbB-mediated reoxidation of EcDsbA_{P151A} was not impaired. In the structure of EcDsbA_{P151A}, the conformation of the active site residues ³⁰CPHC³³ was unchanged, whereas the positions of residues in the *cis*-proline loop were significantly perturbed. These findings suggest that the residues in the loops surrounding the active site may be of greater importance in substrate recognition and binding by DsbA than the hydrophobic groove that is the binding site for DsbB.

Structures have recently been reported for SaDsbA (24) and NmDsbA3 (23), both of which are oxidoreductase enzymes that, in contrast to EcDsbA, display a limited substrate repertoire as well as *E. coli* DsbL, which is thought to have specificity for the enzyme arylsulfate sulfotransferase (25). Comparison of each of these structures with that of EcDsbA reveals significant differences around the regions identified as the peptide binding surface in the current study. Notably, SaDsbA and NmDsbA3 have a similar TcP sequence in the *cis*-proline loop. In the case of SaDsbA, mutation of this sequence to VcP, which is more common in Gram-negative DsbA enzymes, enhanced the oxidoreductase activity as measured in an insulin reduction assay (24). Thus, it

The Structure of a DsbA-Peptide Complex

appears that the sequence of the *cis*-proline loop may be important in dictating the substrate binding of DsbA, as has previously been demonstrated with other TRX domain proteins (27, 28). Although DsbL retains the VcP sequence in the *cis*-proline loop, it lacks the hydrophobic surface features that surround the active site in EcDsbA and VcDsbA, which supports the notion that these are important for substrate binding and specificity.

Although the structures of EcDsbA are very similar throughout the catalytic cycle, differences have been observed in the relative orientation of the TRX and α -helical domains in several structures of DsbA (17, 18). These are present both for different DsbA molecules in the same crystal structure, between DsbA structures solved by different groups, and between the different forms of DsbA. This has led to the suggestion that dynamics may play a role in the catalytic activity of DsbA (20). Analysis of the dynamics in reduced and oxidized forms of VcDsbA has revealed the presence of interdomain motions in the reduced form of the protein (18), which result in an opening of the hydrophobic groove. Redox-dependent conformational changes have previously been demonstrated to result in the opening of a cavity in the TRX domain-containing protein ResA (58), which contributes to substrate specificity. In the case of DsbA, the conformations of the different redox states are similar, but changes in the dynamics appear to allow the reduced form of DsbA to access a more open conformation, which may facilitate the interaction with the DsbB loop and enable the formation of a mixed disulfide between Cys¹⁰⁴ of EcDsbB and Cys³⁰ of EcDsbA.

The structure of the EcDsbA-peptide complex presented herein provides an insight into the specificity observed within the catalytic cycle of this enzyme. In conjunction with the description of the structure of a EcDsbA-EcDsbB complex (21) and differences in the dynamics of the different redox states of the VcDsbA enzyme (18), a clearer picture is beginning to emerge of the factors regulating the catalytic cycle of DsbA. Thus, oxidized DsbA appears to bind its substrates through a mechanism that is common to many TRX family oxidoreductases via the formation of backbone hydrogen bonds to a loop immediately prior to the conserved *cis*-proline residue of the TRX fold. Subtle changes in the composition of this loop and the residues that surround it appear to influence substrate specificity. The relatively higher specificity of reduced DsbA for DsbB results from additional interactions of DsbB within the hydrophobic groove on the surface of DsbA, which may become more accessible as a result of interdomain motions that are present in the reduced DsbA.

Acknowledgments—We greatly appreciate the gift of strains JCB570 and JCB571 from Professor James Bardwell (University of Michigan, Ann Arbor, MI). We thank the BIOCARS staff at the Advanced Photon Source for assistance with data collection. NMR data were analyzed using the program SPARKY.⁶

REFERENCES

- Collet, J. F., and Bardwell, J. C. (2002) *Mol. Microbiol.* 44, 1–8
- Bardwell, J. C., Lee, J. O., Jander, G., Martin, N., Belin, D., and Beckwith, J.

- (1995) *Proc. Natl. Acad. Sci. U. S. A.* 90, 1038–1042
- Heras, B., Shouldice, S. R., Totsika, M., Scanlon, M. J., Schembri, M. A., and Martin, J. L. (2009) *Nat. Rev. Microbiol.* 7, 215–225
- Ha, U. H., Wang, Y., and Jin, S. (2005) *Infect. Immun.* 71, 1590–1595
- Dacheux, D., Epaulard, O., de Groot, A., Gaery, B., Leberre, R., Aitree, I., Polack, B., and Toussaint, R. (2002) *Infect. Immun.* 70, 3973–3977
- Watarai, M., Tobe, T., Yoshikawa, M., and Sasaki, C. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 4977–4981
- Miki, T., Okada, N., and Dambra, H. (2004) *J. Biol. Chem.* 279, 34631–34642
- Jackson, M. W., and Plano, G. V. (1999) *J. Bacteriol.* 181, 5126–5130
- Peck, J. A., and Taylor, R. K. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 6210–6214
- Bardwell, J. C., McGovern, K., and Beckwith, J. (1991) *Cell* 67, 581–589
- Mitsiades, D., Georgopoulos, C., and Raina, S. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 7084–7088
- Stafford, S. J., Humphreys, D. P., and Lund, P. A. (1999) *FEMS Microbiol. Lett.* 174, 179–184
- Yu, J. (1998) *Infect. Immun.* 66, 3909–3917
- Martin, J. L., Bardwell, J. C., and Kuriyan, J. (1995) *Nature* 365, 464–468
- Hu, S. H., Peck, J. A., Rattigan, E., Taylor, R. K., and Martin, J. L. (1997) *J. Mol. Biol.* 268, 137–146
- Schirra, H. J., Renner, C., Cziisch, M., Huber-Wunderlich, M., Holak, T. A., and Glockshuber, R. (1998) *Biochemistry* 37, 6263–6276
- Guddat, L. W., Bardwell, J. C., and Martin, J. L. (1998) *Structure* 6, 757–767
- Horne, J., d'Auvergne, E. J., Coles, M., Velkov, T., Chin, Y., Charman, W. N., Prankerd, R., Gooley, P. R., and Scanlon, M. J. (2007) *J. Mol. Biol.* 371, 703–716
- Martin, J. L. (1995) *Structure* 3, 245–250
- Guddat, L. W., Bardwell, J. C., Zander, T., and Martin, J. L. (1997) *Protein Sci.* 6, 1148–1156
- Inaba, K., Murakami, S., Suzuki, M., Nakagawa, A., Yamashita, F., Okada, K., and Ito, K. (2006) *Cell* 127, 789–801
- Dutton, R. J., Boyd, D., Berkmen, M., and Beckwith, J. (2008) *Proc. Natl. Acad. Sci. U. S. A.* 105, 11933–11938
- Vivian, J. P., Scoullar, J., Robertson, A. L., Bottomley, S. P., Horne, J., Chin, Y., Wielens, J., Thompson, P. F., Velkov, T., Peck, S., Byrnes, E., Reddick, T., Wilce, M. C., Kahler, C. M., Romjohn, J., and Scanlon, M. J. (2008) *J. Biol. Chem.* 283, 32452–32461
- Heras, B., Kurz, M., Jarrott, R., Shouldice, S. R., Frei, P., Robin, G., Cemazar, M., Thöny-Meyer, L., Glockshuber, R., and Martin, J. L. (2008) *J. Biol. Chem.* 283, 4261–4271
- Grimshaw, J. P., Stürmann, C. U., Brozzo, M. S., Malojcic, G., Grütter, M. G., Capitani, G., and Glockshuber, R. (2008) *J. Mol. Biol.* 380, 667–680
- Dumoulin, A., Grauschopf, U., Bischoff, M., Thöny-Meyer, L., and Berger-Bächi, B. (2005) *Arch. Microbiol.* 184, 117–128
- Hiniker, A., Ren, G., Heras, B., Zheng, Y., Laurinec, S., Jobson, R. W., Stuckey, J. A., Martin, J. L., and Bardwell, J. C. (2007) *Proc. Natl. Acad. Sci. U. S. A.* 104, 11670–11675
- Ren, G., Stephan, D., Xu, Z., Zheng, Y., Tang, D., Harrison, R. S., Kurz, M., Jarrott, R., Shouldice, S. R., Hiniker, A., Martin, J. L., Heras, B., and Bardwell, J. C. (2009) *J. Biol. Chem.* 284, 10150–10159
- Sinha, S., Langford, P. R., and Kroll, J. S. (2004) *Microbiology* 150, 2993–3000
- Tinsley, C. R., Voulhoux, R., Beretti, J. L., Tommaseu, J., and Nassif, X. (2004) *J. Biol. Chem.* 279, 27078–27087
- Kadokura, H., Tian, H., Zander, T., Bardwell, J. C., and Beckwith, J. (2004) *Science* 303, 534–537
- Darby, N. J., and Creighton, T. E. (1995) *Biochemistry* 34, 3576–3587
- Ruddock, L. W., Hirst, T. R., and Freedman, R. B. (1996) *Biochem. J.* 315, 1001–1005
- Skórko-Glonek, J., Sobiecka-Szakutala, A., and Lipińska, B. (2006) *Acta Biochim. Pol.* 53, 585–589
- Amann, E., Ochs, B., and Abel, K. J. (1988) *Gene* 69, 301–315
- Marley, J., Lu, M., and Bracken, C. (2001) *J. Biomol. NMR* 20, 71–75
- Coupric, J., Vinci, F., Dugave, C., Quémeur, F., and Moustiez, M. (2000) *Biochemistry* 39, 6732–6742

⁶ T. D. Goddard and D. G. Kneller, unpublished data.

The Structure of a DsbA-Peptide Complex

38. Collaborative Crystallography Project 4 (1994) *Acta Crystallogr. Sect. D* 50, 760–763
39. Delaglio, F., Girzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* 6, 277–293
40. Ayed, A., Mulder, F. A., Yi, G. S., Lu, Y., Kay, L. E., and Arrowsmith, C. H. (2001) *Nat. Struct. Biol.* 8, 756–760
41. Warrens, A. M., Jones, M. D., and Lechler, R. I. (1997) *Gene* 186, 29–35
42. Chung, C. T., and Miller, R. H. (1988) *Nucleic Acids Res.* 16, 3580
43. Sardesai, A. A., Genevans, P., Schwager, F., Ang, D., and Georgopoulos, C. (2003) *EMBO J.* 22, 1461–1466
44. Macrae, R. M. (1986) *Methods Enzymol.* 125, 563–581
45. Wielen, J., Crosby, I. T., and Chalmers, D. K. (2005) *J. Comput. Aided Mol. Des.* 19, 301–317
46. Bader, M., Mease, W., Zander, T., and Bardwell, J. (1998) *J. Biol. Chem.* 273, 10302–10307
47. Petersen, M. T., Jonson, P. H., and Petersen, S. B. (1999) *Protein Eng.* 12, 535–548
48. Maeda, K., Häggblom, P., Finnie, C., Svensson, B., and Henriksen, A. (2006) *Structure* 14, 1701–1710
49. Lawrence, M. C., and Coleman, P. M. (1993) *J. Mol. Biol.* 234, 946–950
50. Goddard, L. W., Bardwell, J. C., Glockshuber, R., Huber, W., Wunderlich, M., Zander, T., and Martin, J. L. (1997) *Protein Sci.* 6, 1893–1900
51. Mizukuri, Y., Yakushi, T., Kawagishi, I., and Hosima, M. (2006) *J. Bacteriol.* 188, 4190–4197
52. Qin, J., Clore, G. M., Kennedy, W. M., Huth, J. R., and Gronenborn, A. M. (1995) *Structure* 3, 289–297
53. Qin, J., Clore, G. M., Kennedy, W. P., Kuszewski, J., and Gronenborn, A. M. (1996) *Structure* 4, 613–620
54. Nordstrand, K., Åslund, F., Holmgren, A., Otting, G., and Berndt, K. D. (1999) *J. Mol. Biol.* 286, 541–552
55. Ladner, J. E., Parsons, J. F., Rife, C. L., Gilliland, G. L., and Armstrong, R. N. (2004) *Biochemistry* 43, 352–361
56. Haeber, P. W., Goldstone, D., Katzen, F., Beckwith, J., and Metcalf, P. (2002) *EMBO J.* 21, 4774–4784
57. Charbonnier, J. B., Belin, P., Moutiez, M., Stura, E. A., and Quémener, F. (1999) *Protein Sci.* 8, 96–105
58. Colbert, C. L., Wu, Q., Erbel, P. J., Gardner, K. H., and Deisenhofer, J. (2006) *Proc. Natl. Acad. Sci. U. S. A.* 103, 4410–4415

Characterization of the DsbA Oxidative Folding Catalyst from *Pseudomonas aeruginosa* Reveals a Highly Oxidizing Protein that Binds Small Molecules

Stephen R. Shouldice,¹ Begonia Heras,¹ Russell Jarrott,¹ Pooja Sharma,²
Martin J. Scanlon,² and Jennifer L. Martin¹

Abstract

Bacterial antibiotic resistance is an emerging global crisis, and treatment of multidrug-resistant gram-negative infections, particularly those caused by the opportunistic human pathogen *Pseudomonas aeruginosa*, remains a major challenge. This problem is compounded by a lack of new antibiotics in the development pipeline only two new classes have been developed since the 1960s, and both are indicated for multidrug-resistant gram-positive infections. A promising new approach to combat antibiotic resistance is by targeting bacterial virulence, rather than bacterial viability. The bacterial periplasmic protein DsbA represents a central point for antivirulence intervention because its oxidoreductase activity is essential for the folding and function of almost all exported virulence factors. Here we describe the three-dimensional structure of this DsbA target from *P. aeruginosa*, and we establish for the first time that a member of this enzyme family is capable of binding small molecules. We also describe biochemical assays that validate the redox activity of PaDsbA. Together, the structural and functional characterization of PaDsbA provides the basis for future studies aimed at designing a new class of antivirulence compounds to combat antibiotic-resistant *P. aeruginosa* infection. *Antioxid. Redox Signal.* 12, 921–931.

Introduction

THE GRAM-NEGATIVE organism *Pseudomonas aeruginosa* is an opportunistic human pathogen and a leading cause of hospital-acquired infections (2). Such nosocomial infections are associated with wounds, surgery, invasive devices, or mechanical ventilation and *P. aeruginosa* infection is particularly prevalent in the immunocompromised and critically ill (2). In hospitals, the widespread use of broad-spectrum antibiotics suppresses patients' normal flora and encourages cross-infection between patients. Surveillance programs that monitor the incidence of antimicrobial resistance have revealed increases in both the number of *P. aeruginosa* infections and the frequency of multidrug-resistant isolates (25).

Pseudomonas aeruginosa possesses a diverse array of virulence factors that enable it to evade host defences. Many of the processes that confer virulence or antibiotic resistance are mediated by secreted proteins, including flagellar proteins, type III secretion factors, pili, and enzymes such as proteases and β -lactamases. To function correctly, these secreted pro-

teins must be folded and remain intact in the harsh extracellular environment; for this reason, most incorporate disulfide bonds. Incorporation of disulfide bonds takes place in the periplasm and is catalyzed by enzymes of the Dsb family. The DsbA enzyme from *P. aeruginosa* (PaDsbA) plays a pivotal role; it is required for the expression of elastase, exotoxin A, protease IV, lipase, and alkaline phosphatase (7, 11, 15, 31, 32, 49). PaDsbA is also required for the formation of a functional type III secretion system, and *P. aeruginosa* organisms lacking PaDsbA are defective in twitching motility and in the expression of the pilin subunit pilA. Taken together, these findings highlight the intimate relationship between PaDsbA activity and *P. aeruginosa* pathogenicity and identify PaDsbA as a potential target for the development of a new class of antibacterial agents that target virulence rather than viability (18). This approach has been advocated as a means of generating "more effective drugs with a lower propensity for inducing bacterial resistance" (39).

The DsbA of *Escherichia coli* (EcDsbA) is the best-characterized member of the DsbA family. However, recent

¹The University of Queensland, Institute for Molecular Bioscience, Division of Chemistry and Structural Biology, Brisbane, Queensland, Australia.

²Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Monash University (Parkville Campus), Parkville, Victoria, Australia.

structural and biochemical evidence illustrates that significant differences exist among DsbA homologues (17, 22, 50). For example, the peptide-binding groove identified in the EcDsbA structure (34) and subsequently shown to be used by its redox partner protein DsbB (24) is virtually absent in the structure of *Staphylococcus aureus* DsbA (SaDsbA) (17). In addition, a hydrophobic patch believed to be important for substrate binding in EcDsbA is not present in SaDsbA.

To investigate the role of PaDsbA further, we determined the 1.5-Å resolution crystal structure of this protein and characterized its redox properties. Our data show that PaDsbA is one of the most oxidizing proteins yet characterized, and the structure of PaDsbA reveals important differences compared with other structurally characterized DsbA proteins, including the unexpected finding that PaDsbA binds small molecules. Collectively, these studies reveal the biochemical function of PaDsbA and suggest a basis for future structure-based drug-design studies targeting this important virulence determinant.

Materials and Methods

Cloning, expression and purification

The codon-optimized *Pseudomonas aeruginosa* dsbA gene lacking the signal sequence and inserted into the PCR-Blunt vector was obtained from GeneArt (Regensburg, Germany) and subsequently amplified by using forward and reverse primers compatible with ligation-independent cloning (LIC). The amplified gene was then inserted into a modified pET21a vector. This modified vector encodes a leader sequence consisting of an N-terminal maltose-binding protein (MBP) followed by the tobacco-vein mottling virus (TMV) protease recognition sequence. This plasmid also encodes for TMV so that the MBP fusion is cleaved during protein folding. Immediately after the TMV site is an His₆-tag, followed by a spacer containing the tobacco etch virus (TEV) protease recognition sequence and a LIC sequence based on a central SspI restriction site. The cloning protocol introduced three additional amino acid residues (SNA) at the N-terminus of the mature protein when the His₆-tag is cleaved with TEV protease. The sequence of the final construct was confirmed to be that of *Pseudomonas aeruginosa* dsbA (GeneID:877731). The plasmid was transformed into *Escherichia coli* BL21(DE3)/pLysS for expression by autoinduction (47). In brief, cells were grown with agitation at 200 rev/min for 24 h at 303 K in autoinduction media supplemented with 50 µg/ml ampicillin and 34 µg/ml chloramphenicol. After expression, the cells were harvested with centrifugation, flash frozen in liquid nitrogen, and stored at 193 K. Cells were lysed in 25 mM HEPES, pH 7.5, 150 mM NaCl, 0.5% TritonX-100, protease inhibitor cocktail III (A.G. Scientific, Inc., San Diego, CA) and DNase. The sample was sonicated, and cell debris was removed by centrifugation (41,399 g for 30 min). His₆-tagged PaDsbA was purified by using PrePac High Yield Resin (USB-Millennium Science, Surrey Hills, Victoria, Australia) and eluted with 25 mM HEPES, pH 7.5, 150 mM NaCl, and 500 mM imidazole. The peak fractions were pooled, and the protein concentration was measured at 280 nm by using a NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE). Purified recombinant His₆-tagged TEV (generated from pRK793) (5) was added to a molar ratio of 1:100 to cleave the His tag from PaDsbA. This mixture was dialyzed extensively against

25 mM HEPES, pH 7.5, 150 mM NaCl, at 277 K for 48 h. The mixture was then passed over PrePac High Yield Resin once more so that the flowthrough contained the cleaved PaDsbA, and the resin retained His₆-tagged TEV. The cleaved PaDsbA was further purified by gel-filtration chromatography by using a Superdex S-75 column (AKTA; GE Healthcare, Piscataway, NJ).

Peak fractions containing PaDsbA were pooled and concentrated to 100 mg/ml in 10 mM HEPES, pH 7.5, by using Amicon Ultra centrifugal filter devices with a 10-kDa cutoff (Millipore, Billerica, MA). The protein concentration was again measured at 280 nm, and the purity was assessed with SDS-PAGE analysis.

Disulfide isomerase assay

An *in vitro* assay monitoring the refolding of scrambled RNaseA was used to measure the isomerase activity of PaDsbA, EcDsbC, and EcDsbA (19). Disulfide-scrambled RNaseA was produced as previously described (17). The isomerization of scrambled RNaseA (40 µM) disulfide bonds to their native conformation was measured by incubation in 100 mM sodium phosphate buffer, 1 mM ethylenediamine tetraacetic acid (EDTA), pH 7.0, 10 µM dithiothreitol (DTT), with 10 µM protein sample. Positive and negative controls consisted of two additional reactions with folded native RNaseA or buffer, respectively. All assays were performed at 298 K, and RNaseA activity was determined by monitoring cAMP hydrolysis at 296 nm at several time points.

Chemical unfolding and refolding

Oxidized PaDsbA was diluted to 1.5 µM in 20 mM Hepes-NaOH, 170 mM NaCl, 0.1 mM EDTA, pH 7.0 buffer containing guanidinium chloride at various concentrations and incubated at room temperature for 24 h. Reduced PaDsbA was unfolded under identical conditions but in the presence of 0.75 mM DTT. Refolding experiments were performed by first unfolding oxidized and reduced PaDsbA stock solutions (37.5 µM) in 20 mM Hepes-NaOH, 170 mM NaCl, 0.1 mM EDTA (pH 7.0), 6 M guanidinium chloride (containing 20 mM DTT for the reduced protein) for 20 h at room temperature, and then taking 20 µl of the unfolded protein solution and mixing it with 480 µl of buffer containing guanidinium chloride at various concentrations. These solutions were then incubated for 24 h at room temperature. Three independent experiments were performed for each transition, which were measured fluorimetrically at 353 nm, by using an excitation wavelength of 280 nm. Data were fitted according to a two-state model (38, 44).

Thermal stability

The thermal-denaturation equilibrium was measured with methods we described previously (17). We confirmed that PaDsbA disulfide oxidation was complete by using the Elman assay before measuring the temperature-induced unfolding curves. The reduced sample was prepared by adding 0.75 mM DTT (final concentration) to the oxidized sample. Temperature-induced unfolding was measured at 226.5 nm for the oxidized sample and 225 nm for the reduced sample on a Jasco J-810 spectropolarimeter. The experiment was conducted by heating the samples from 25°C to 95°C at

PSEUDOMONAS AERUGINOSA DsbA

929

1°C/min (unfolding). Once 95°C was reached, samples were then cooled to 25°C (refolding).

PaDsbA redox potential

The fractions of reduced and oxidized PaDsbA (2 µM) in 100 mM sodium phosphate, 1 mM EDTA, pH 7.0, at 298 K, and different concentrations of excess reduced glutathione (GSH) and oxidized glutathione (GSOG) (51) were determined from the redox-dependent PaDsbA fluorescence at 340 nm (excitation, 280 nm) after equilibrium for 24 h. Three independent experiments were performed to give the estimate of the equilibrium constant. The redox potential was then calculated as described previously (51).

Insulin-reduction assay

We measured the ability of PaDsbA, EcDsbC, and EcDsbA to catalyze insulin reduction in the presence of DTT by using methods described previously (21). In brief, reaction mixtures were prepared in cuvettes containing 130 µM insulin, and protein catalyst (5 to 10 µM) in 100 mM sodium phosphate buffer, 1 mM EDTA, pH 7.0. Reactions were started with the addition of DTT to a final concentration of 0.33 mM. After thorough mixing, the rate of precipitation was monitored by recording the increased turbidity of the reaction mixture at 650 nm every 30 s. Absorbance measurements were made by using a 1-cm cuvette and a Cary50 spectrophotometer. The noncatalyzed reduction of insulin by DTT was monitored in a control reaction without catalyst.

EcDsbA complementation

For these experiments, we cloned the *psdA* gene into the arabinose-inducible pBAD33 vector (14). The constructs were then chemically transformed into the nonmotile *E. coli* *dsbA* mutant and *dsbA*/*dsbB* double-mutant strains [JC8817 and JC8818, respectively (4)] and plated on LB chloramphenicol plates. As positive controls, we used JC8817 and JC8818 cells containing pBAD33::EcDsbA. Triplicates of fresh colonies were stabbed into minimal agar (M63) containing 40 mg/ml of each amino acid, and 1 mg/ml arabinose to induce expression. Additionally, the experiment was performed by using agar without arabinose to monitor background complementation. The experiment also was performed in the absence and presence of L-cysteine to determine whether L-cysteine could act as a general oxidant for PaDsbA. Plates were incubated for 7 to 10 h at 310 K before analysis of the swarming of *E. coli* cells.

Crystallization and diffractor data measurement

The hanging-drop vapor-diffusion method was used for crystallization of PaDsbA with the commercially available screens Crystal Screens 1 and 2, polyethylene glycol (PEG)/Ion Screen, index (Hampton Research, San Diego, CA), Wizard Crystal screens 1 and 2 (Emerald BioSystems, Bainbridge Island, WA), the JCSG1 Suite (Qingen Pty Ltd, Doncaster, Victoria, Australia) and the PACT premier screen (Molecular Dimensions Limited, Newmarket, Suffolk, England). Crystallization trials were set up in 96-well plates by using a Mosquito crystallization robot (TTP Labtech, Melbourne, Cambridge, England). Each drop consisted of 100 nL protein solution and 100 nL well solution. The crystallization

plates were incubated at 293 K and imaged by using a temperature-controlled RockImager (Formulatrix, Waltham, MA). Crystals of PaDsbA appeared within 2 days in a condition from the JCSG1 Suite screen comprising 24% (wt/vol) PEG of average molecular mass of 1,500 Da (PEG 1500), and 20% (vol/vol) glycerol. A fine screen around this condition was then set up by using VDXm 24-well hanging-drop plates and 18-mm silicized coverslips (both from Hampton Research, San Diego, CA). To increase the crystal size, each coverslip held one 2-µL drop containing 1 µL protein solution and 1 µL well solution. The optimized condition consisted of 24% (wt/vol) PEG 1500, and 22% (vol/vol) glycerol. We attempted to solve the structure of PaDsbA by molecular replacement with Phaser (36) and the CCP4 program CHAINSAW, by using as models the coordinates of DsbA structures available at the time: DsbA from *Escherichia coli* (PDB code 1DSB, Seq ID 27%), *Staphylococcus aureus* (3BC1 Seq ID 21%), *Wallabies papilloma* (3PAR Seq ID 8%), *Vibrio cholerae* (1RED Seq ID 28%), and *Nitrosomonas europaea* (2IN0 Seq ID 12%). However, these trials were unsuccessful, and we therefore pursued single-wavelength anomalous dispersion (SAD) phasing methods. Selenomethionine (SeMet) PaDsbA was produced by using a protocol similar to that previously described for the production of SeMet CcmG (9). In brief, *E. coli* BL21(DE3)pLysS cells containing the PaDsbA expression plasmid described earlier were grown at 310 K in M63 minimal media containing 50 µg/ml SeMet (D/L mixture). Cultures were induced at the exponential phase of growth ($OD_{600} = 0.6$) with a final concentration of 0.5 mM isopropyl β -D-thiogalactopyranoside (IPTG), and cells were harvested 4 h after induction. SeMet PaDsbA was purified as described earlier for native PaDsbA. SeMet incorporation was confirmed with mass spectrometry. SeMet PaDsbA crystals appeared with the same protein concentration and crystallization conditions as native PaDsbA.

For SeMet PaDsbA data collection, crystals were harvested from hanging drops by using a 0.2- to 0.3-mm CryoLoop Copper Magnetic HT Cryocool (Hampton Research) and placed in a 1-µL drop consisting of 35% (wt/vol) PEG 1500 and 25% (vol/vol) glycerol for ~30 s before flash cooling in liquid nitrogen (our crystal-cooling trials indicated that both PEG and glycerol were necessary to achieve adequate cryoprotection). The PaDsbA crystal was maintained at 100 K throughout data collection. X-ray data were collected at the Australian Synchrotron on PX1 beamline 3BM1 (wavelength, 0.953645 Å). The crystal-to-detector distance was 150 mm, and 1-degree oscillation images were collected for a total of 720 degrees, with an exposure time of 5 s each. This collection strategy resulted in a highly redundant data set ideally suited for SAD phasing. Diffraction data were processed and scaled by using the HKL2000 program package (37).

Structure determination and refinement

The structure of PaDsbA was solved by SAD phasing. PaDsbA contains six Met residues, and SeMet PaDsbA crystals were predicted to have one molecule in the asymmetric unit, according to the Matthews analysis (35). The locations of all six selenium atoms were determined by using Phenix AutoSol (1), and the model was improved with iterative model-building, density modification and refinement by using the Phenix AutoBuild wizard. This was followed by

cycles of interactive manual refitting of the model by using the program COOT (10) and refinement with phenix.refine (1). During the later stages of refinement, glycerol molecules and water molecules were built into ($F_o - F_c$) difference maps. Some surface residues with weak electron density were modeled with reduced occupancies for the side-chain atoms (Asn², Asp⁴, Asp⁵, Glu¹⁴, Ser²⁶, Lys²⁸, Tyr⁴¹, Arg⁴², Glu⁴⁷, Glu⁵⁶, Lys¹⁰⁰, Glu¹⁰¹, His¹¹⁸, Lys¹²⁶, Lys¹²⁸, Glu¹³⁰, Asp¹³³, Lys¹³², Glu¹³², Lys¹⁴⁰, Lys¹⁴³, Glu¹⁷³, Lys¹⁸¹, Lys¹⁸⁶, Lys¹⁹³), and some residues were modeled with alternate conformations (Ser²⁷, Ser²², Cys³⁷, Cys⁶⁰, Ile⁴⁷, Glu⁵², Ser⁶³, Glu¹¹⁹, Ser¹²⁷, Met¹³⁸, Glu¹³⁹, Met¹⁵¹, Lys¹⁶¹, Lys¹⁷⁷). Table 1 provides the statistics for the x-ray data collection and final refined model. The coordinates and structure factors of SeMet PtdAbA have been deposited (PDB code 3H93).

Results

PaDspA is active in the insulin-reduction assay

Before obtaining the crystal structure of PaDsbA, we examined the redox activity of the purified protein to confirm that the protein was active and to characterize it more fully than had previous studies (49). First, we used the insulin-

TABLE 1 X-RAY DATA COLLECTION AND REFINEMENT STATISTICS FOR Pd_2SbA

| | |
|---|--|
| Data collection | |
| Wavelength (Å) | 0.95364 |
| Resolution range (Å) | 50.0–1.50 |
| Space group | P4 ₁ |
| Unit cell dimensions (Å) | a, b = 41.15; c = 98.32 α = β = γ = 90.0 |
| Observed reflections | 732,228 |
| Unique reflections | 26,111 |
| R _{merge} | 0.076 (0.361) |
| Completeness (%) | 100 (100) |
| <I> / <σ(I)> | 46.5 (11.3) |
| Redundancy | 28.0 (17.3) |
| SAD Phasing | |
| Resolution (Å) | 29.1–1.50 |
| Number of selenium sites | 6/6 |
| found/expected | |
| Mean figure of merit | 0.62b |
| Refinement | |
| Resolution (Å) | 50.1–1.50 |
| Completeness for range (%) | 99.92 |
| R _{factor} | 11.0 (9.5) |
| R _{free} | 15.2 (17.8) |
| Number of non-H protein atoms | 1,567 |
| Number of waters | 322 |
| Number of non-H glycerol atoms | 4 × 6 |
| Wilson B | 11.65 |
| Average B factor (Å ²): All atoms | 16.47 |
| Average B factor (Å ²): Water | 32.08 |
| Average B factor (Å ²): Non-solvent | 13.32 |
| R.m.s.d. from ideal geometry | |
| Bonds (Å) | 0.005 |
| Angles (°) | 0.969 |
| Ramachandran plot | |
| Residues in most-favored/ additionally-allowed regions (%) | 98.5/1.5 |

*Values in parentheses refer to the highest-resolution shell.

reduction assay (21). In this assay, oxidoreductase enzymes reduce the interchain disulfides of insulin, causing precipitation of the insoluble insulin B-chain. Precipitation is monitored as an increase in absorbance at 650 nm. We found that the purified *Pa*DebA we produced was as active as *Escherichia coli* DebA (*Ec*DebA) in this assay, but less active than the disulfide isomerase *E. coli* DebC (*Ec*DebC) (Fig. 1A), in agreement with an earlier report (49). We therefore set about a more-comprehensive analysis of the redox activity of *Pa*DebA.

Perfume has weak disinfectant germicidal activity

The Dab family of proteins vary in their ability to catalyze disulfide isomerization. We measured the ability of PaDabA to isomerize, or shuffle, incorrect disulfides of scrambled RNaseA. We found that, under the conditions of our assay, PaDabA yielded ~50% active RNaseA after 300 min (Fig. 1B). By comparison, the isomerase EcDabC yielded ~90% activity. In this assay, the activity of PaDabA is more similar to that of EcDabA, a strong thiol oxidant with a very oxidizing disulfide. The next step was to determine the redox potential of PaDabA, to identify whether it, too, is highly oxidizing.

PaDsbA is highly oxidizing

The redox potentials of DsbA proteins vary tremendously; the most reducing DsbA characterized to date is *Wolbachia pipiensis* α -DsbA1 (Wpz-DsbA1), with a value of -163 mV (27), and the most oxidizing are the neisserial NmDsbA proteins [*Neisseria meningitidis*, -80 mV (29)] and *E. coli* DsbL (EcDsbL), the specialized DsbA from pathogenic *E. coli* [redox potential, -95 mV (13)]. We used the standard assay for measuring redox potential at pH 7.0 and 298 K, from the equilibrium constant with glutathione. The data yielded an equilibrium constant (K_{eq}) for PaDsbA of $9.96 \pm 0.88 \times 10^{-6}$ M, corresponding to an intrinsic redox potential (E°) of -94 mV (Fig. 1C). PaDsbA is thus 10 times more oxidizing than EcDsbA [$K_{eq} = 12 \times 10^{-5}$ M; $E^\circ = -122$ mV (23)] and as oxidizing as EcDsbL (13).

The active-site disulfide bond of PaCsbA is destabilizing

Disulfides typically stabilize protein structures. However, for DsbAs, the active-site disulfide has been shown to destabilize [e.g., EcDsbA (52); Wpa-DsbA1 (27)] or to have no effect [e.g., *Staphylococcus aureus* DsbA (17)] on the stability of the protein. We performed guanidinium chloride induced unfolding/refolding equilibria experiments on PaDsbA to determine the ΔG_{fold} . The transitions were fully reversible and cooperative, and all unfolding curves fitted a two-state model (Fig. 1D) (38, 44). The oxidized form of PaDsbA destabilizes the protein compared with the reduced form ($\Delta\Delta G_{\text{ox/red}}$ 13.1 kJ/mol; reduced PaDsbA $\Delta G_{\text{fold}} = -60.7 \pm 5.5$ kJ/mol; cooperativity, 26.6 ± 2.2 kJ/mol/M; midpoint of transition, 2.61 M; oxidized PaDsbA $\Delta G_{\text{fold}} = -56.6 \pm 3.9$ kJ/mol; cooperativity, 30.9 ± 2.1 kJ/mol/M, midpoint of transition, 1.83 M). On the basis of $\Delta\Delta G$ values, the disulfide form of the PaDsbA active site is as destabilizing as that of EcDsbA ($\Delta\Delta G_{\text{ox/red}}$ 14.8 kJ/mol; oxidized EcDsbA $\Delta G_{\text{fold}} = -33.5 \pm 1.2$ kJ/mol vs. reduced EcDsbA $\Delta G_{\text{fold}} = -48.3 \pm 2.5$ kJ/mol (23)).

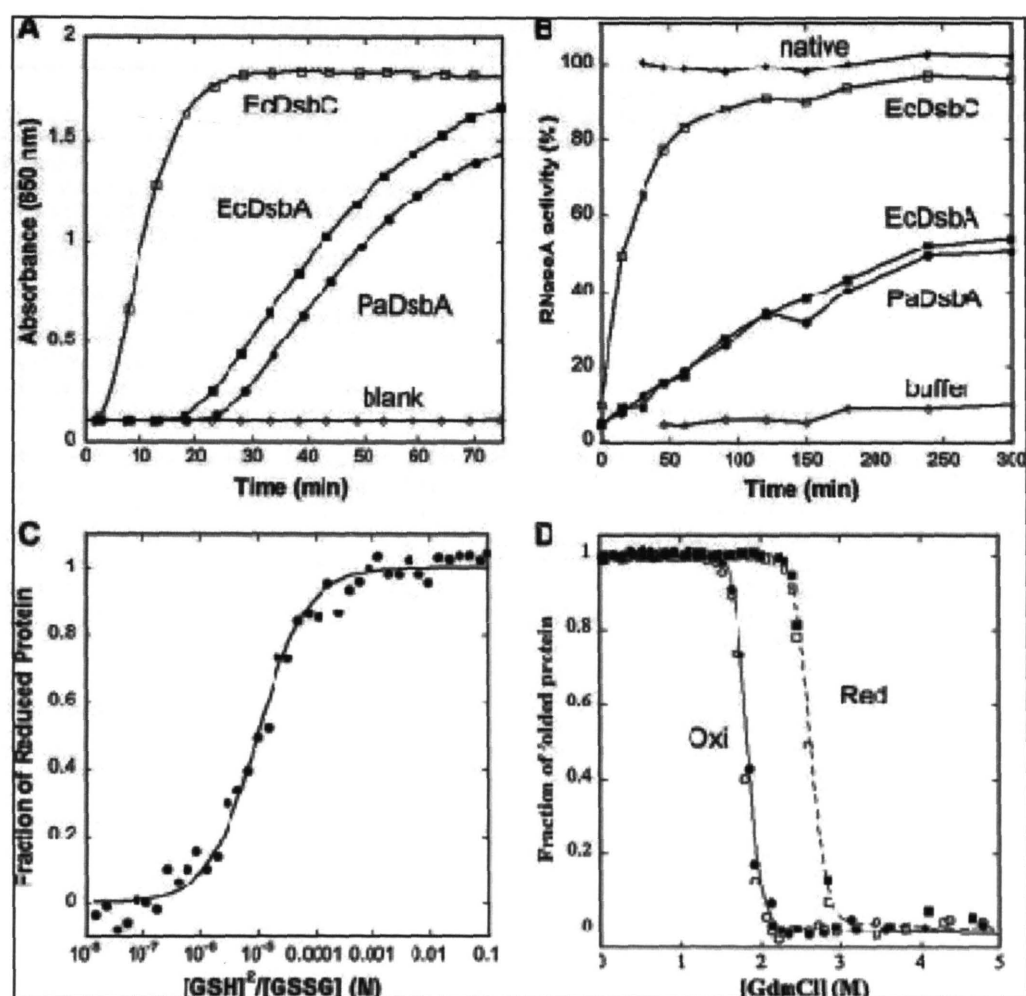


FIG. 1. Characterization of PaDsbA. (A) Disulfide reductase activity of PaDsbA was determined by using the insulin-reduction assay. The assay was performed with 10 μ M PaDsbA (●) or no protein (Blank ○). For comparison, we also measured the reductase activity of 5 μ M EcDsbC (□) and 10 μ M EcDsbA (■). A: time zero, DTT was added to each reaction, and the catalyzed reduction of insulin was measured as an increase in absorbance at 650 nm. (B) Disulfide isomerase activity of PaDsbA was measured by its ability to shuffle disulfide bonds of scrambled ribonuclease A (scRNaseA). scRNaseA was incubated with 10 μ M of either PaDsbA (●), EcDsbC (□) or EcDsbA (■). Native RNaseA (◆) was used as positive control, and buffer, as blank (○). The cleavage of cCMP by correctly folded RNaseA was followed spectroscopically. (C) The fraction of oxidized and reduced PaDsbA after equilibration in redox buffers containing different GSH/GSSG ratios was determined from the redox state dependent fluorescence of PaDsbA. These data were used to calculate K_{eq} and redox potential. (D) Guanidinium induced unfolding/refolding equilibria of PaDsbA. Transitions were measured fluorimetrically at 353 nm. Solid symbols, unfolding transitions; open symbols, refolding transitions. The normalized fluorescence data for oxidized (●, ○) and reduced (■, □) PaDsbA were fitted to a two-state model of folding (solid and dashed lines, respectively).

We also performed temperature-induced unfolding experiments on PaDsbA to determine T_m , the transition midpoint of folded to unfolded protein, as monitored by the far-UV CD signal. These results also show that oxidized PaDsbA destabilizes the protein compared with the reduced form (T_m values, 345.8 ± 0.08 K and 355.8 ± 0.19 K, respectively), but we found that the unfolding was not reversible for PaDsbA under the conditions we used. Conversely, EcDsbA unfolding by using this approach was reversible (T_m values: unfolding oxidized 341.7 ± 0.18 K; unfolding reduced

350.9 ± 0.21 K; refolding oxidized 338.4 ± 0.65 K; refolding reduced 349.1 ± 0.43 K).

PaDsbA complements EcDsbA in vivo

By using an *in vivo* complementation assay with an anisotropy-inducible PaDsbA expression plasmid, we showed that PaDsbA complements EcDsbA in an *E. coli* DsbA-null strain. This agrees with results of others (49) who used a PaDsbA expression plasmid in an *E. coli* DsbA-null

TABLE 2. COMPLEMENTATION OF *E. COLI* DSB MUTATIONS WITH PA_{DSB} A MEASURED BY MOTILITY OF THE BACTERIA ON SOFT AGAR PLATES

| | No arabinose; no expression from plasmid (negative control) | Arabinose added to induce expression from the plasmid |
|---|---|---|
| Minimal media | | |
| EcDsbA-null strain: EcDsbA plasmid (positive control) | — | + |
| EcDsbA/EcDsbB-null strain: EcDsbA plasmid | — | — |
| EcDsbA-null strain: PaDsbA plasmid | — | + |
| EcDsbA/EcDsbB-null strain: PaDsbA plasmid | — | — |
| Cystine-supplemented media | | |
| EcDsbA-null strain: EcDsbA plasmid (positive control) | — | + |
| EcDsbA/EcDsbB-null strain: EcDsbA plasmid | — | + |
| EcDsbA-null strain: PaDsbA plasmid | — | + |
| EcDsbA/EcDsbB-null strain: PaDsbA plasmid | — | + |

strain that did not require arabinose for expression. Furthermore, we used an *E. coli* strain lacking both DsbA and DsbB and found that PaDsbA does not complement this strain unless the media is supplemented with cystine. These results suggest that PaDsbA interacts with EcDsbB, and that this interaction is required *in vivo* for re-oxidation of PaDsbA. A summary of the results is presented in Table 2.

Structure of PaDsbA

Hanging-drop crystallization trials were conducted on purified native PaDsbA by using a Mosquito nanoliter crye-

tization robot. Initial crystallization conditions were optimized, and the resulting crystals were found to diffract to high resolution. SeMet-labeled protein crystals were then used to determine the crystal structure by SAD methods, making use of high-resolution, high-redundancy data collected at the Australian Synchrotron. The resulting 1.5-Å resolution structure reveals one molecule of PaDsbA in the asymmetric unit. The refined model matches the data very well, with R_{factor} and R_{free} values of 11.0 and 15.2% respectively (Table 1). The final crystallographic refinement parameters, electron-density maps, Ramachandran plot (41), and MolProbity analysis (30) indicate that the PaDsbA crystal structure is of very high quality.

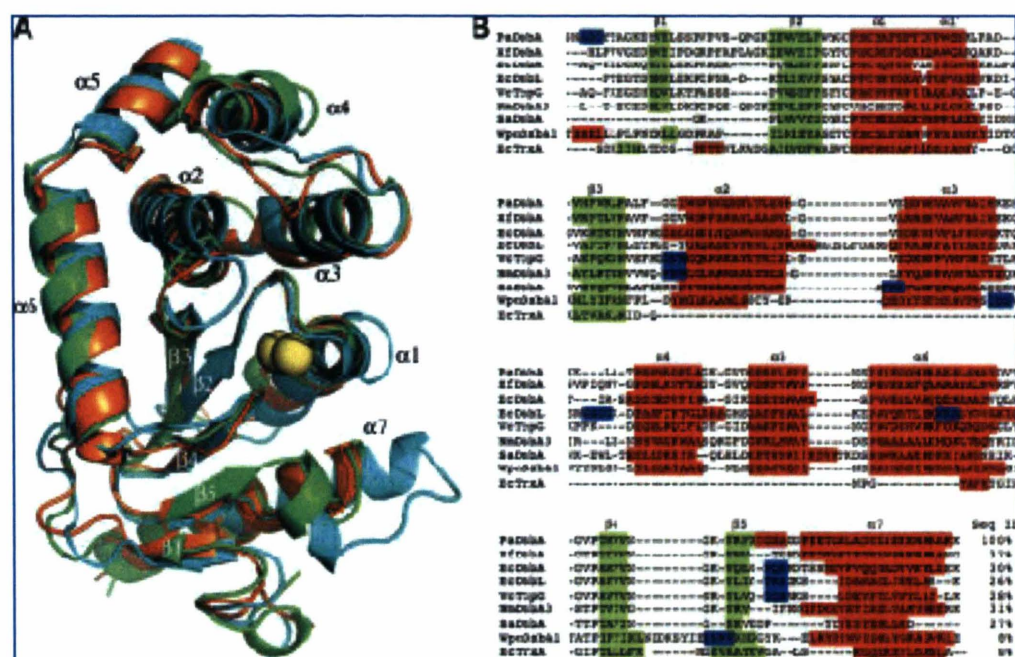


FIG. 2. PaDsbA structure. (A) Overlay of structures of PaDsbA (green), its closest structural homologue XfDsbA (orange), and EcDsbA (cyan). Sulfur atoms of the catalytic cysteines of PaDsbA are shown as yellow spheres. (B) Structure-based sequence alignment performed by using STRAP, with secondary structural elements labeled. Cyan denotes 3_{10} -helices, green shows β -strands, and α -helices are red. The coordinates of each protein were used to make a multiple sequence alignment with TM-align (53). Secondary structural elements were calculated by using DSSP (26). Proteins used and their respective PDB codes are PaDsbA (3H93), XfDsbA (2RFM), EcDsbA (1FVK), EcDsbB (3CTM), VcDsbA (1BFD), NmdsbA3 (2ZNM), SaDsbA (3BCI), WpaDsbA1 (3F4R), and EcTrxA (2TRX).

In agreement with all other structurally characterized DsbA molecules, PaDsbA contains a thionedoxin (TRX) fold (53) with an inserted α -helical domain. A Dali search (20) indicates that the closest structural homologue to PaDsbA is a DsbA from the plant pathogen *Xylella fastidiosa* (XfDsbA, 37% sequence identity, PDB code 2REM). A structural superposition of PaDsbA and its homologues XfDsbA and EcDsbA resulted in overall root-mean-squared deviations of 1.6 Å (186 C α aligned) and 2.2 Å (188 C α aligned), respectively (Fig. 2A). Structural differences between DsbA homologues are highlighted in a sequence alignment (Fig. 2B), showing, for example, that long loops present in EcDsbL ((13) between α 2 and α 3) and WpaDsbA1 ((27) between β 4 and β 5) are unique to these molecules. Also evident from Fig. 2 is that the "bulge" in EcDsbL that connects helices α 3 and α 4 is absent in PaDsbA, but present in WpaDsbA1 and XfDsbA. The most obvious region of variability among the structurally characterized

DsbA proteins is in the loop connecting β 5 and α 7 (Fig. 2B), which is involved in binding substrates and the redox-partner protein DsbE. We note that the α s-Pro loop residues of PaDsbA are GVcisPT, whereas those of EcDsbA and EcDsbL are GVcisPA (Fig. 2B).

Glycerol-binding sites

Unexpectedly, the electron-density map revealed the presence of four bound glycerol molecules, located in discrete locations on PaDsbA (Fig. 3). Two glycerol molecules are bound to the active-site face of the enzyme, and two more are bound to the opposite face. Three of the four glycerol molecules bind to the TRX fold, and one of these three (Glyc-1) is located close to the catalytic cysteines. The Glyc-1 binding site lies on the face of β -strands 1 and 5. Glyc-1 is modeled in two subtly different conformations, both held in place by

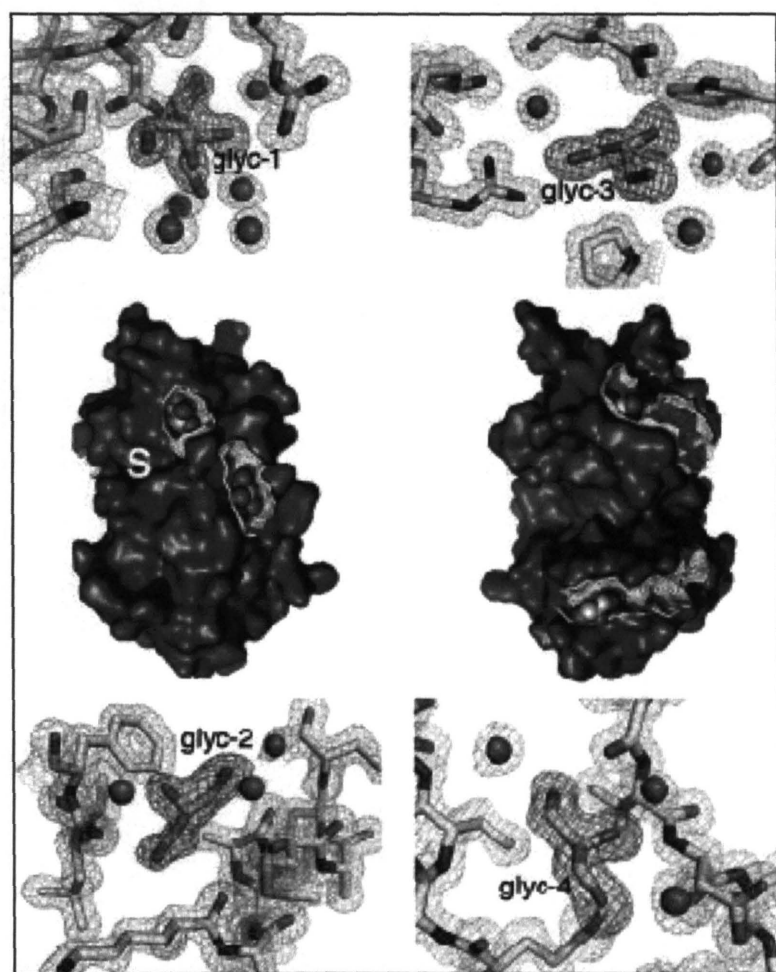


FIG. 3. Binding cavities on PaDsbA. Surface representations of PaDsbA are shown for both faces of the molecule, with the active site denoted by a white S (left orientation is rotated 180 degrees around the Z-axis compared with Fig. 2A, and the right image is rotated 180 degrees around Y, compared with the left image). The four glycerol binding-site cavities/grooves are shown as a mesh [calculated from PDBsum (29)] with the glycerol molecules located inside. OMIT maps are also shown for each of the four bound glycerol molecules (each close-up view displays electron-density maps, light gray for the $2F_o - F_c$ map contoured at 1σ , and dark gray for the $F_o - F_c$ map contoured at 3σ , obtained after refinement of the model in which the glycerol molecules were omitted).

hydrogen bonds to the main-chain carbonyl oxygen of Lys²⁰ (after $\beta 1$) and the side chain of residue Arg¹⁶³ (85). In addition, five ordered water molecules form interactions with Glyc-1. This site is important not only because of its proximity to the active-site cysteines, but also because this region is involved in binding substrates (43) and the essential partner protein DsbB (24) (Fig. 4).

Glyc-2, the second glycerol molecule that binds on the active-site surface of PaDsbA, is located between the C-

terminal end of $\alpha 6$ and the truncated loop joining $\beta 3$ and $\alpha 2$. Glyc-2 forms hydrogen bonds to the carbonyl oxygen of Leu⁶⁶ and to nearby water molecules. It also forms a hydrogen bond to a symmetry-related PaDsbA through the side chain of Lys¹²⁴. Of the two glycerol molecules bound on the opposite face of PaDsbA (Figs. 3 and 4), Glyc-3 is located in an area rich in acidic and basic residues and is within hydrogen-bonding distance of the side chains of Asp¹⁸⁰ and Arg¹⁸⁷ on helix $\alpha 7$ and to three ordered water molecules. Finally, Glyc-4 is located in a relatively hydrophobic pocket between the loop joining helices $\alpha 2$ and $\alpha 3$ and the loop connecting $\alpha 4$ and $\alpha 5$ (Figs. 3 and 4). It forms hydrogen bonds with the main-chain carbonyl oxygen of Gly¹¹⁹ and the amide nitrogen of Asp¹²¹, located in the $\alpha 4$ - $\alpha 5$ loop, and with three ordered water molecules.

Electrostatic surface features of PaDsbA

The surface electrostatics of DsbA proteins vary tremendously. For example, EcDsbA and the DsbA from *Vibrio cholerae*, VcTspG, have a hydrophobic patch and a hydrophobic groove surrounding the active-site cysteine residues (22, 34), whereas EcDsbL (13) and Wpx-DsbA1 (27) are basic, and the DsbA from *Staphylococcus aureus* is polar (18). PaDsbA has a smaller hydrophobic patch than EcDsbA, and its surface is more basic, although not so basic as EcDsbL. The $\beta 5$ - $\alpha 7$ loop that forms the hydrophobic groove of PaDsbA is also truncated compared with EcDsbA (Fig. 2E).

Discussion

Pseudomonas aeruginosa is a particularly nasty opportunistic human pathogen that causes serious nosocomial infections (42) and places significant economic stress on health care systems (6). Increasing reports are appearing of multidrug-resistant strains (40, 46), and *P. aeruginosa* possesses numerous pathways to evade host defense mechanisms and to damage cells (18), including the ability to inject virulence proteins directly into host cells during infection through a type III secretion system (8, 12). PaDsbA plays a pivotal role in these virulence mechanisms by assembling and stabilizing virulence factors, including the secretion system (7, 15, 49). *Pse-*

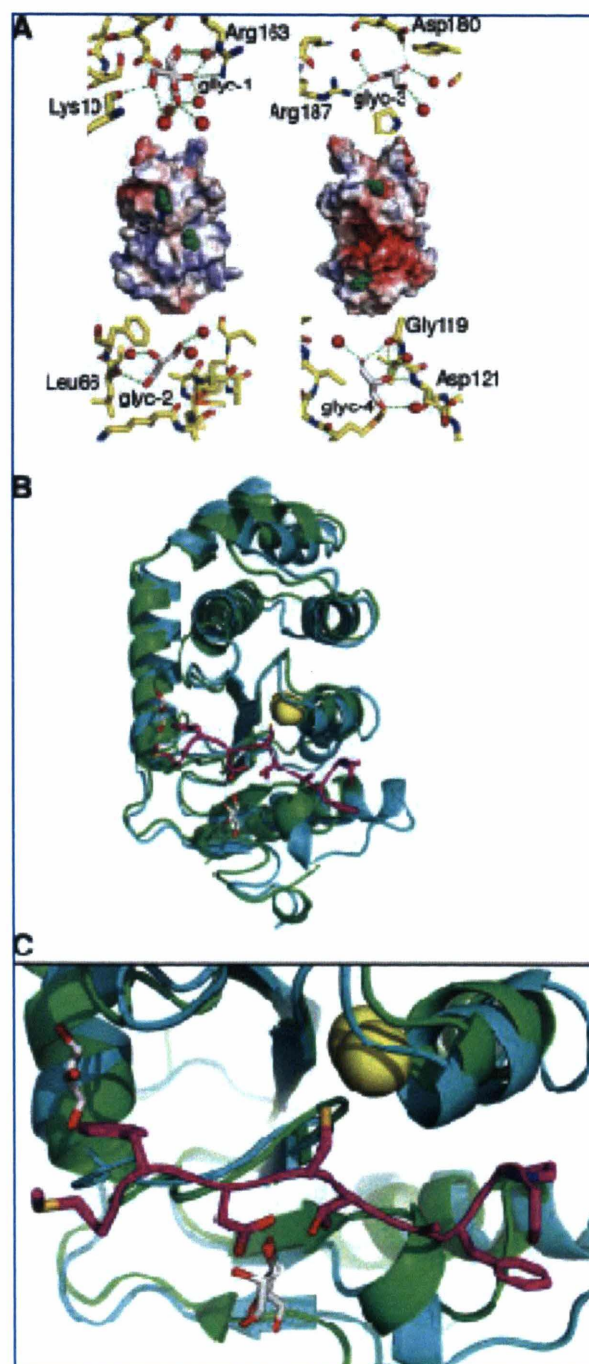


FIG. 4. PaDsbA interactions with glycerol. (A) Electrostatic surface of PaDsbA and the interactions formed with the glycerol molecules. Electrostatic surface representations of PaDsbA are illustrated in the center (in the same orientation as in Fig. 3), with red indicating negatively charged (-5 kT), and the blue, positively charged regions ($+5$ kT). The four bound glycerol molecules are shown in green, and the catalytic active site is indicated by a black S. The four surrounding panels show hydrogen-bonding interactions (green dashed lines) with each of the four glycerol molecules. Glyc-1 is modeled as two alternate conformations. This panel was generated by using APBS (3) and PyMOL (Delano Scientific, Palo Alto, CA). (B) Overlay of PaDsbA (green) and EcDsbA (cyan) structures in the same orientation as in Fig. 2A. The sulfur atoms of the catalytic cysteine residues of PaDsbA are shown as yellow spheres. The loop of EcDsbB (100-107) that interacts with EcDsbA is shown in magenta. (C) Close-up view of the hydrophobic groove showing the binding sites of the EcDsbB loop to EcDsbA and of glycerol molecules to PaDsbA.

8.2 Mass spectrometry for *EcDsbA* and *VcDsbA*

Mass spectrometry was performed on the two proteins used in this study. *EcDsbA* and *VcDsbA*. All of protein samples were produced according to protocol outlined in chapter 2. The mass spectrometry experiments were performed by Dr Kieran Rimmer (MIPS) at the Bio21, Molecular Science and Biotechnology Institute, The University of Melbourne, Australia. (<http://www.bio21.unimelb.edu.au/platform-technologies/mass-spectrometry>).

Overall analysis included trypsin digestion of the protein samples, followed by MALDI-TOF to determine the list of masses in the digested sample. Peptide mass fingerprint was undertaken to confirm the identity of the protein. Figure 8-1 and 8-2 shows the mass spec of DsbA samples.

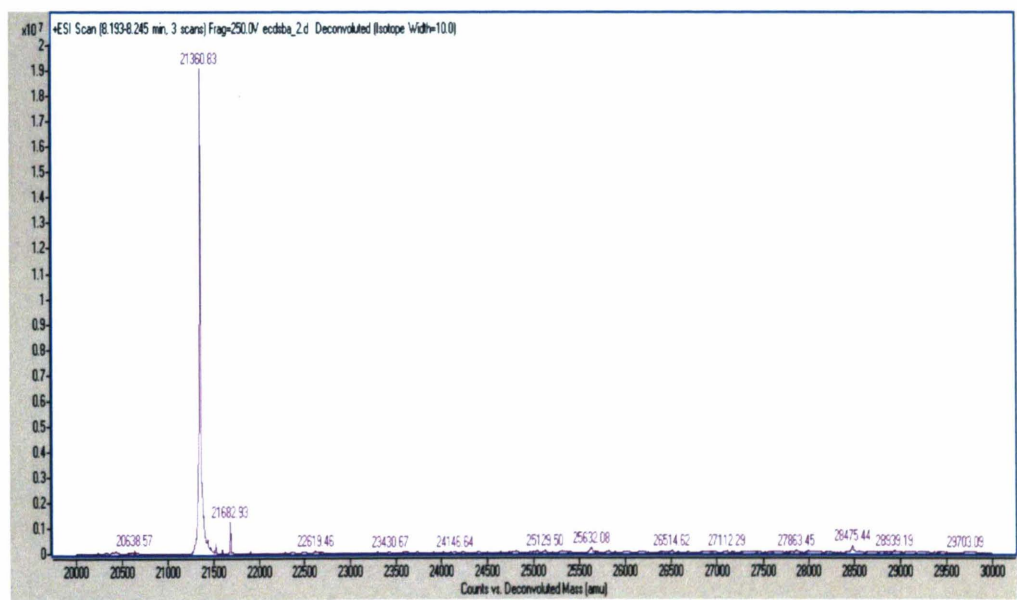


Figure 8-1: Mass spec analysis of unlabeled *EcDsbA*.

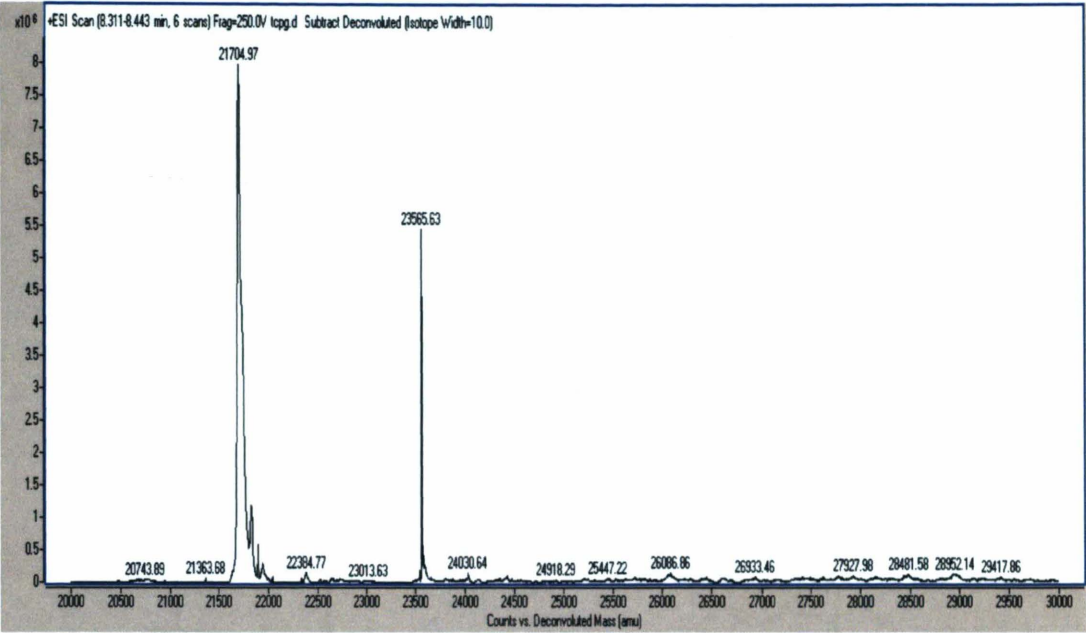
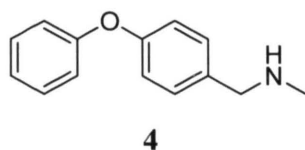
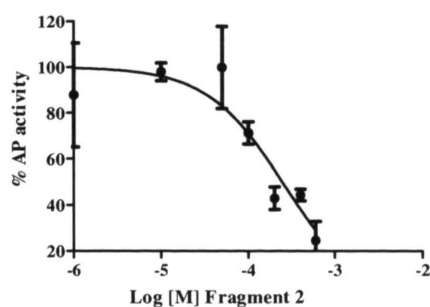
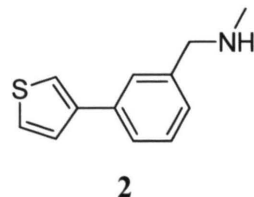
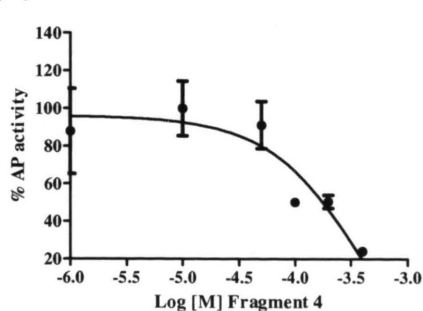


Figure 8-2: Mass spec analysis of unlabeled VcDsbA.

8.3 Preliminary Assay Results

8.3.1. Alkaline Phosphatase Activity Assay

(A)



(B)

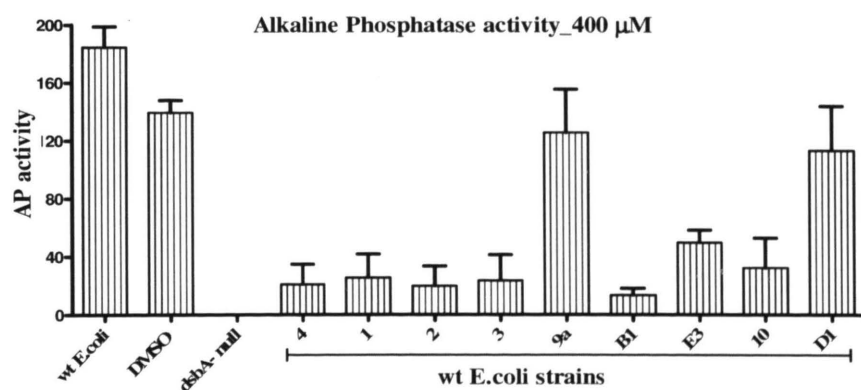


Figure 8-3: (A) Concentration-dependant plots of Alkaline Phosphatase activity in presence of phenylthiophene fragment **2** and phenoxybenzene fragment **4**. (B) AP activity response in presence of fragments (from chapter 5-6), listed on X axis at a concentration of 400 μM.

8.3.2. Fluorescence (FRET) Peptide assay

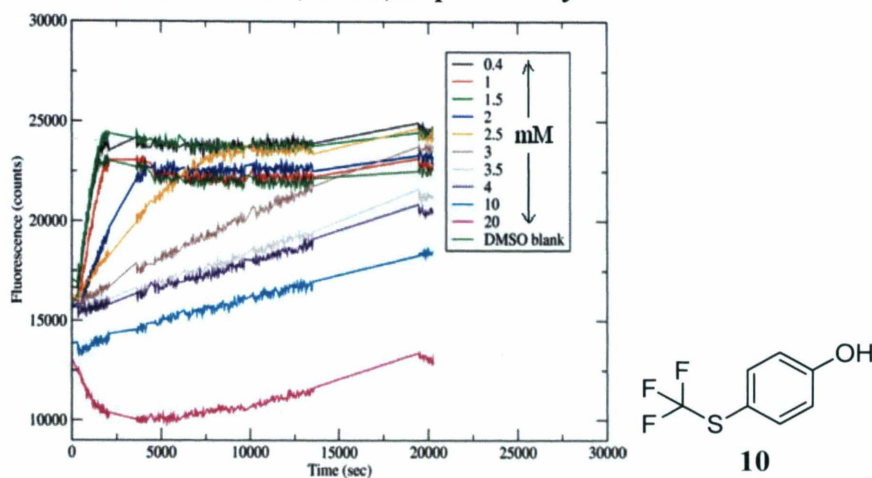


Figure 8-4: Inhibitor concentration-dependant plot of DsbA oxidising activity in FRET-peptide assay with Phenol fragment **10**.

8.3.2. Insulin precipitation assay

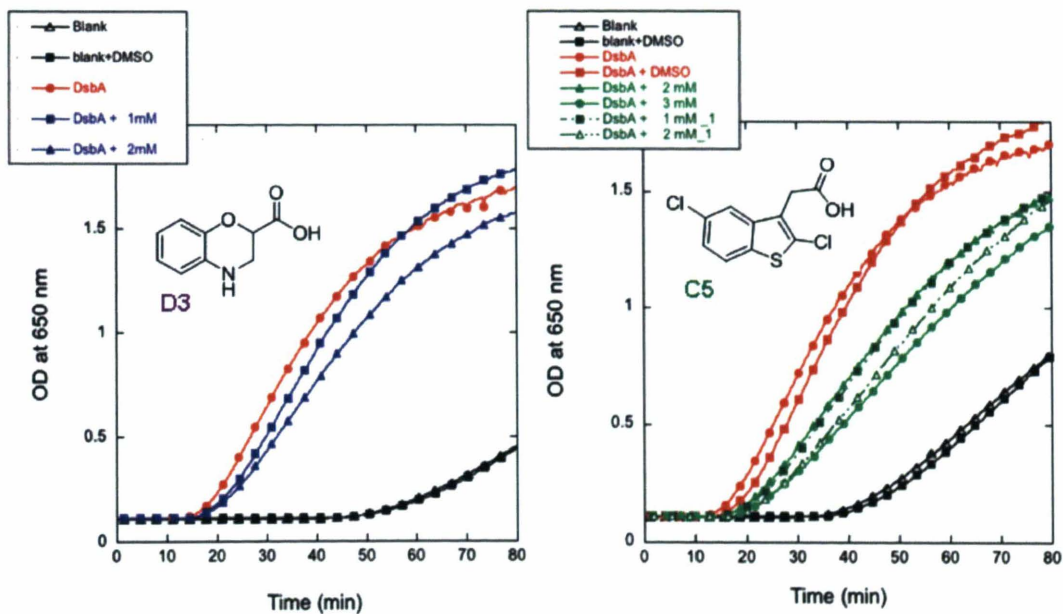


Figure 8-5: UV spectrum obtained from insulin reduction assay in presence two of the *EcDsbA* binding fragments chroman **D3** (left panel) and benzothiophene **C5** (right panel) at different concentrations 1-3 mM.

8.3.3 Motility Assay

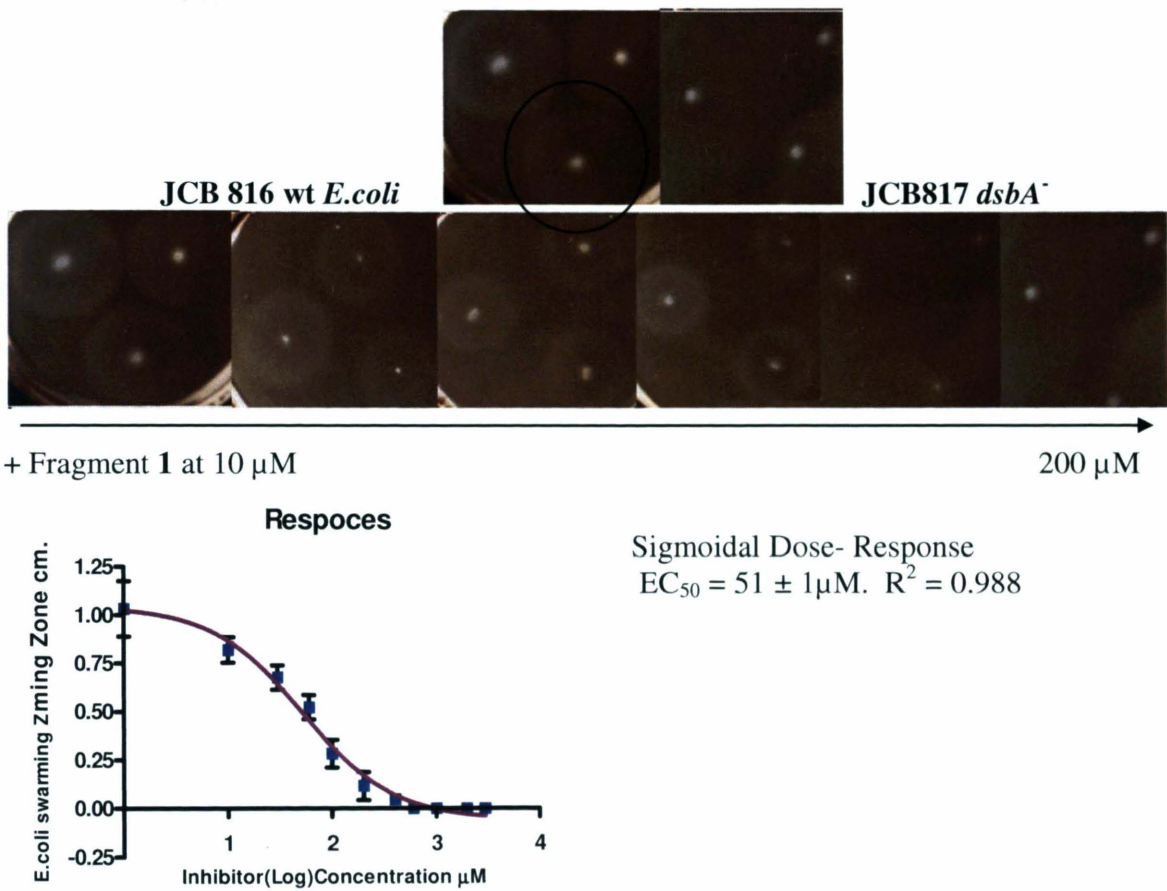


Figure 8-6: Measurement of EC₅₀ of phenylthiophene fragment 1 by bacterial motility assay.

8.4. Isothermal Calorimetry binding studies

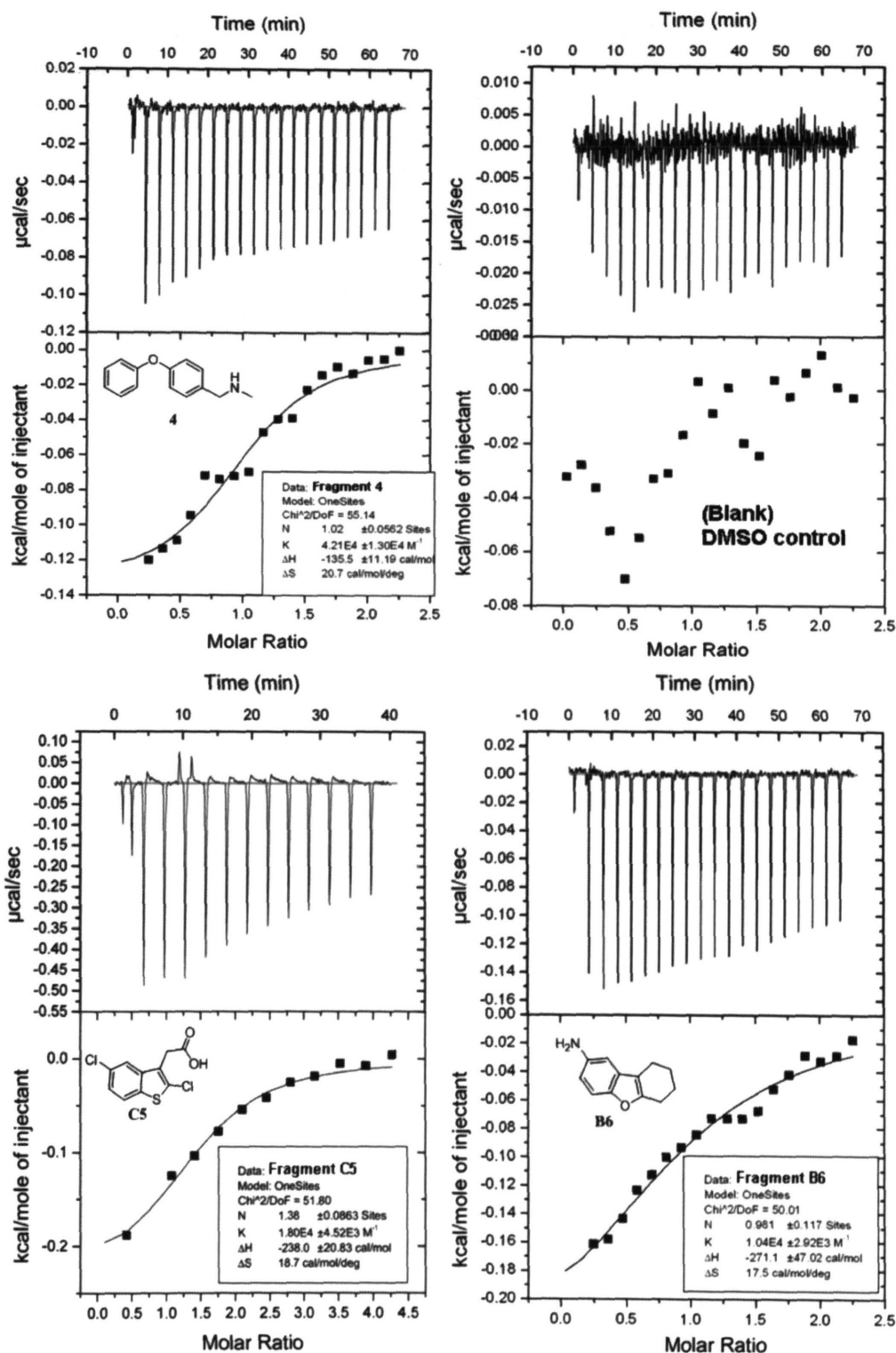


Figure 8-5: *Ec*DsbA –ligand binding curves from ITC experiments for three of the tested fragments (4, C5 and B6) described in chapter 5 and 6.

8.5. STD screening results for phenylthiophene hits

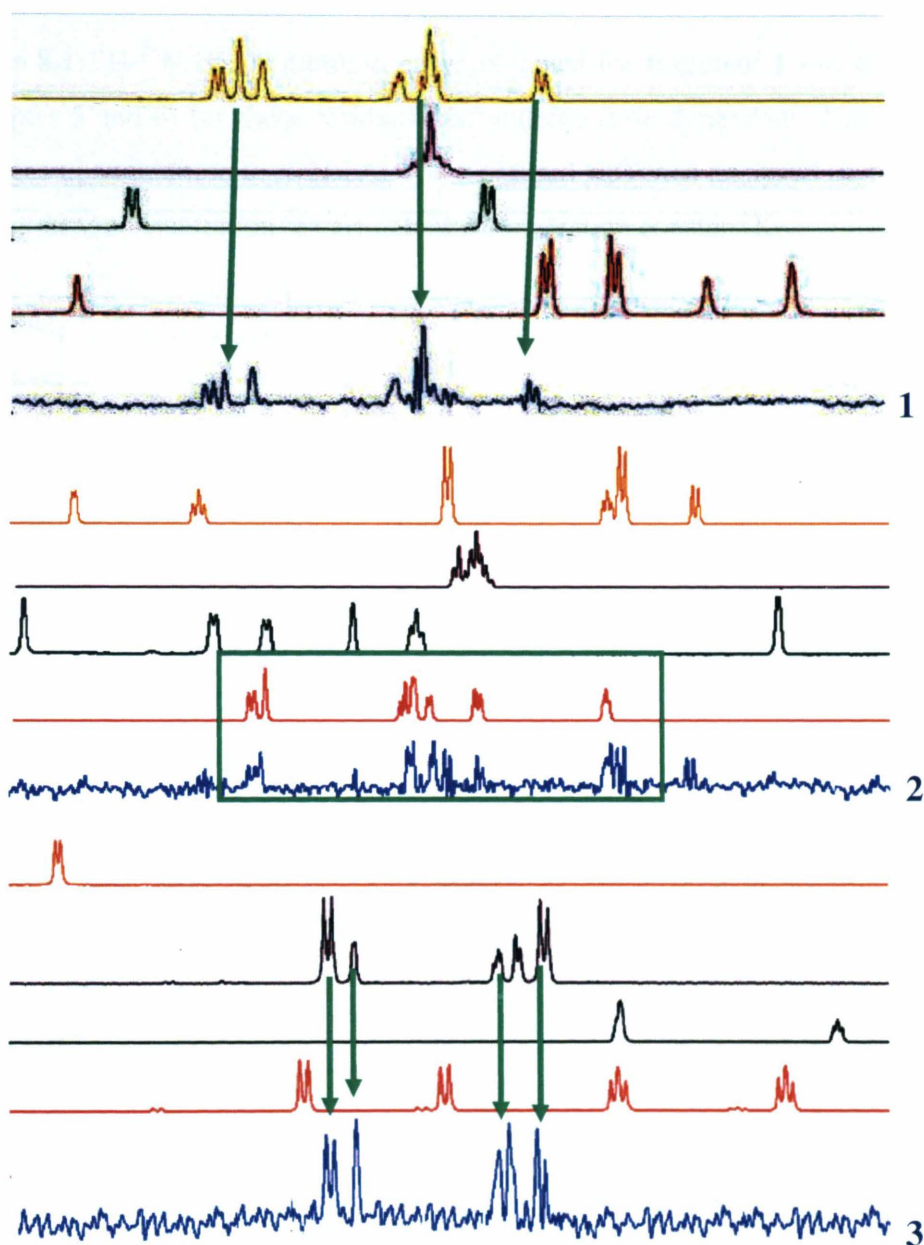
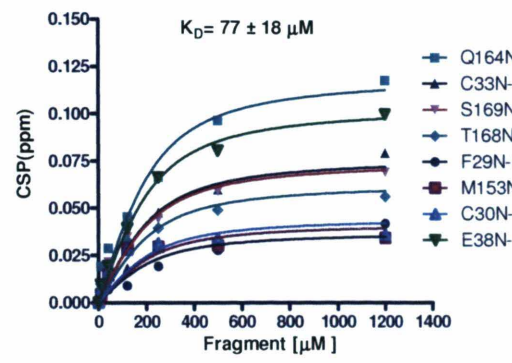
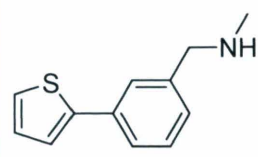
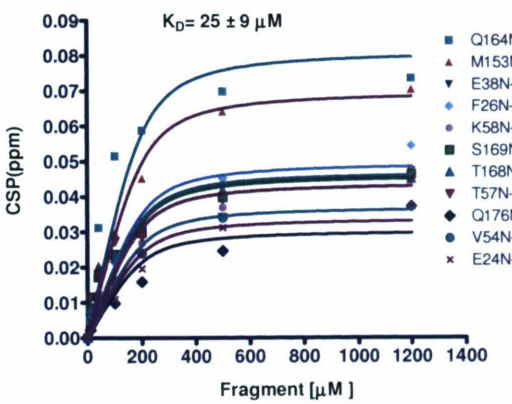
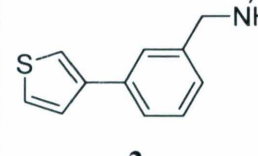
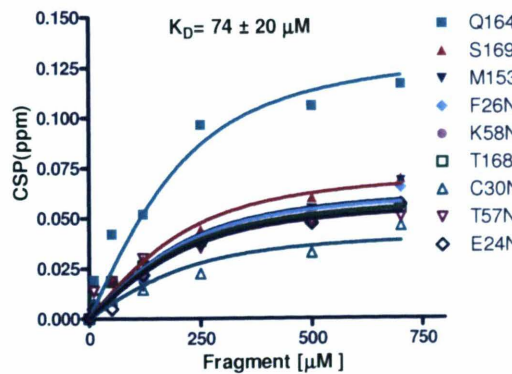
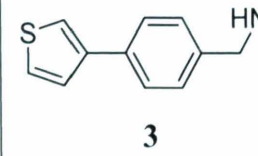
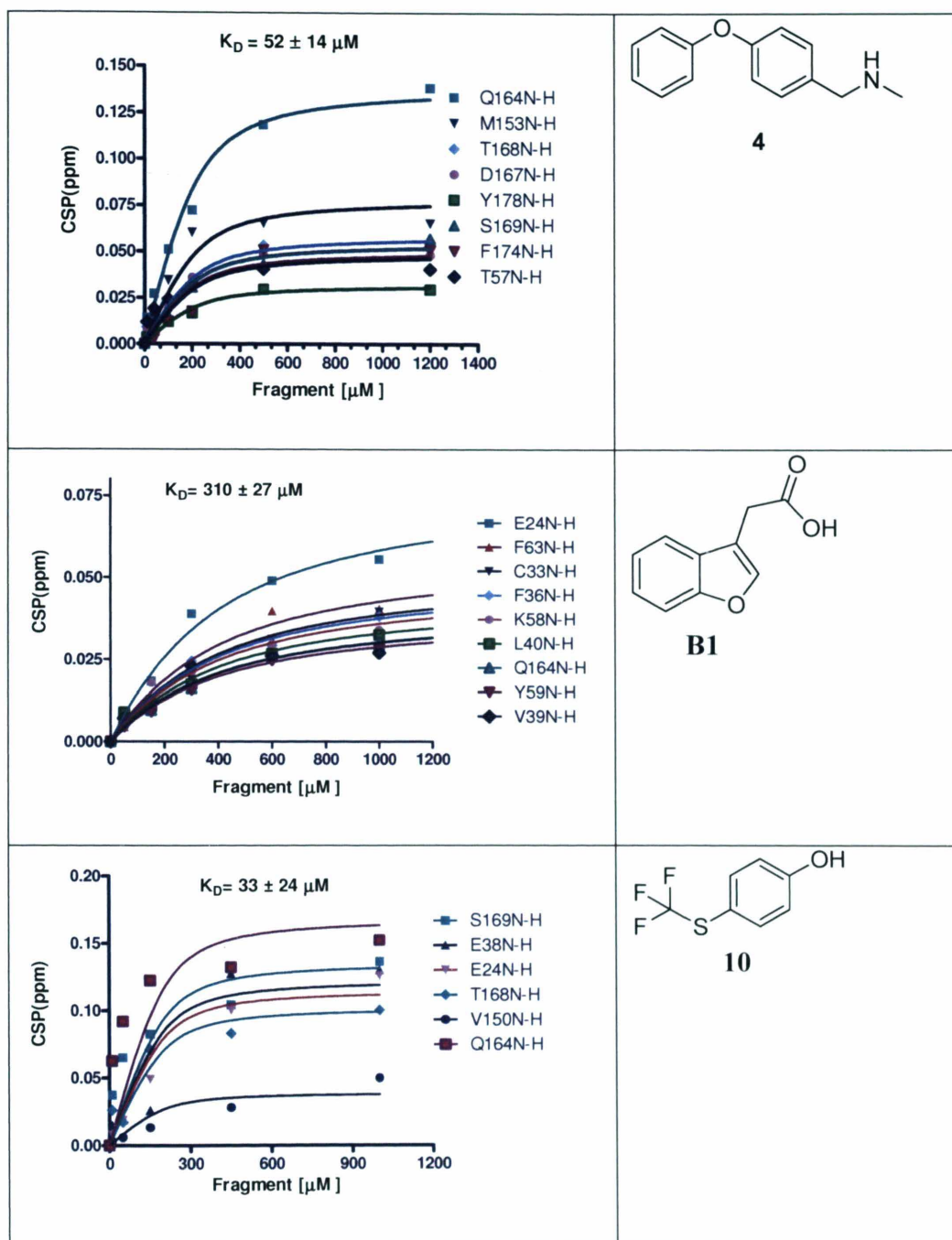


Figure 8-6: Identification of fragment **1-3** (Chapter 5) as *EcDsbA* binders from STD NMR based screening by comparing the STD spectra recorded for a mixture of fragments with the reference 1D ^1H spectra for each fragment. Green arrows and box connect resonances observed in the STD spectrum with their corresponding signals in the 1D ^1H reference spectrum of the ligand.

8.6. *Ec*DsbA-HSQC titration results for phenylthiophene hits

Table 8.1: ¹H-¹⁵N HSQC titration curve obtained for fragment **1** - **4**, **B1** and **10** (Chapter 5 and 6) for those residues that undergo dose dependant chemical shift changes upon addition to *Ec*DsbA. The measured CSP was analysed as a function of fragment concentration to calculate the dissociation constant *K_D*

| | |
|--|--|
|  <p>$K_D = 77 \pm 18 \mu\text{M}$</p> <p>Legend:</p> <ul style="list-style-type: none"> Q164N-H C33N-H S169N-H T168N-H F29N-H M153N-H C30N-H E38N-H |  <p>1</p> |
|  <p>$K_D = 25 \pm 9 \mu\text{M}$</p> <p>Legend:</p> <ul style="list-style-type: none"> Q164N-H M153N-H E38N-H F26N-H K58N-H S169N-H T168N-H T57N-H Q176N-H V54N-H E24N-H |  <p>2</p> |
|  <p>$K_D = 74 \pm 20 \mu\text{M}$</p> <p>Legend:</p> <ul style="list-style-type: none"> Q164N-H S169N-H M153N-H F26N-H K58N-H T168N-H C30N-H T57N-H E24N-H |  <p>3</p> |



8.7. *EcDsbA*-fragment and peptide binding shifts

8.7.1 ^1H - ^{15}N HSQC chemical shifts of *EcDsbA* in solution complexes of phenylthiophene hits. (Chapter 5)

| Residue | <i>EcDsbA</i> -APO | | <i>EcDsbA</i> -Fragment 1 | | <i>EcDsbA</i> -Fragment 2 | | <i>EcDsbA</i> -Fragment 3 | |
|---------|--------------------|--------|---------------------------|--------|---------------------------|--------|---------------------------|--------|
| | N | NH | NH | H | NH | H | NH | H |
| Q2N-H | 119.931 | 8.275 | 119.882 | 8.263 | 119.945 | 8.26 | 119.869 | 8.251 |
| Y3N-H | 120.391 | 7.633 | 120.285 | 7.624 | 120.293 | 7.63 | 120.304 | 7.619 |
| E4N-H | 122.908 | 8.963 | 122.952 | 8.958 | 122.955 | 8.961 | 122.925 | 8.967 |
| D5N-H | 125.68 | 8.507 | 125.772 | 8.5 | 125.741 | 8.498 | 125.769 | 8.5 |
| G6N-H | 116.327 | 8.973 | 116.44 | 8.984 | 116.411 | 8.978 | 116.451 | 8.98 |
| K7N-H | 121.84 | 8.142 | 121.779 | 8.134 | 121.768 | 8.135 | 121.752 | 8.134 |
| Q8N-H | 115.614 | 8.51 | 115.642 | 8.495 | 115.565 | 8.475 | 115.626 | 8.502 |
| Y9N-H | 112.07 | 7.568 | 111.917 | 7.556 | 111.973 | 7.568 | 111.929 | 7.558 |
| T10N-H | 112.13 | 9.532 | 112.218 | 9.516 | 112.24 | 9.518 | 112.299 | 9.513 |
| T11N-H | 121.39 | 9.057 | 121.508 | 9.059 | 121.517 | 9.06 | 121.513 | 9.059 |
| L12N-H | 129.158 | 8.705 | 129.055 | 8.694 | 129.07 | 8.692 | 129.045 | 8.693 |
| E13N-H | 123.279 | 8.651 | 123.287 | 8.647 | 123.294 | 8.643 | 123.291 | 8.645 |
| K14N-H | 117.272 | 7.802 | 117.336 | 7.786 | 117.262 | 7.789 | 117.296 | 7.793 |
| V16N-H | 123.195 | 9.031 | 123.031 | 9.004 | 123.111 | 9.005 | 123.05 | 8.996 |
| A17N-H | 132.033 | 8.603 | 132.047 | 8.603 | 132.047 | 8.603 | 132.047 | 8.603 |
| G18N-H | 110.718 | 8.423 | 110.645 | 8.404 | 110.662 | 8.41 | 110.627 | 8.408 |
| A19N-H | 122.363 | 7.423 | 122.312 | 7.414 | 122.322 | 7.417 | 122.305 | 7.413 |
| Q21N-H | 121.761 | 8.438 | 121.734 | 8.424 | 121.78 | 8.432 | 121.777 | 8.428 |
| V22N-H | 116.064 | 7.702 | 116.007 | 7.705 | 116.05 | 7.703 | 115.978 | 7.705 |
| L23N-H | 129.281 | 8.861 | 129.308 | 8.841 | 129.334 | 8.858 | 129.33 | 8.837 |
| E24N-H | 128.835 | 9.113 | 129.276 | 9.121 | 129.076 | 9.111 | 129.199 | 9.122 |
| F25N-H | 121.524 | 9.239 | 121.496 | 9.23 | 121.489 | 9.228 | 121.473 | 9.227 |
| F26N-H | 119.78 | 8.975 | 120.131 | 9.012 | 120.091 | 9.001 | 120.156 | 9.005 |
| S27N-H | 108.862 | 6.543 | 108.773 | 6.533 | 108.876 | 6.542 | 108.796 | 6.533 |
| F28N-H | 131.395 | 11.327 | 131.35 | 11.351 | 131.307 | 11.335 | 131.458 | 11.322 |
| F29N-H | 115.927 | 8.704 | 115.714 | 8.678 | 115.779 | 8.68 | 115.708 | 8.678 |
| C30N-H | 120.44 | 7.306 | 120.459 | 7.342 | 120.419 | 7.335 | 120.316 | 7.348 |
| C33N-H | 115.874 | 8.826 | 115.554 | 8.764 | 115.591 | 8.783 | 115.565 | 8.769 |

| | | | | | | | | |
|--------|---------|-------|---------|-------|---------|-------|---------|-------|
| Y34N-H | 120.137 | 7.756 | 120.127 | 7.75 | 120.144 | 7.747 | 120.181 | 7.752 |
| Q35N-H | 118.657 | 7.718 | 118.51 | 7.722 | 118.482 | 7.719 | 118.486 | 7.721 |
| F36N-H | 120.612 | 8.432 | 120.372 | 8.467 | 120.461 | 8.367 | 120.441 | 8.446 |
| E37N-H | 115.175 | 7.417 | 115.123 | 7.428 | 115.151 | 7.418 | 115.19 | 7.433 |
| E38N-H | 113.603 | 7.905 | 114.25 | 7.911 | 113.841 | 7.935 | 113.856 | 7.863 |
| V39N-H | 116.744 | 7.141 | 117.081 | 7.183 | 116.813 | 7.145 | 116.2 | 7.26 |
| L40N-H | 115.188 | 7.873 | 114.679 | 7.799 | 115.167 | 7.86 | 115.1 | 7.702 |
| H41N-H | 111.464 | 6.37 | 111.406 | 6.364 | 111.45 | 6.362 | 111.433 | 6.363 |
| I42N-H | 119.745 | 8.346 | 119.741 | 8.337 | 119.719 | 8.342 | 119.727 | 8.34 |
| S43N-H | 118.544 | 8.539 | 118.529 | 8.531 | 118.494 | 8.534 | 118.529 | 8.535 |
| D44N-H | 121.484 | 7.766 | 121.473 | 7.772 | 121.392 | 7.798 | 121.697 | 7.784 |
| N45N-H | 117.186 | 7.616 | 117.198 | 7.596 | 117.195 | 7.601 | 117.19 | 7.596 |
| V46N-H | 120.538 | 8.702 | 120.759 | 8.657 | 120.497 | 8.681 | 120.284 | 8.761 |
| K47N-H | 118.728 | 8.348 | 118.792 | 8.383 | 118.757 | 8.366 | 118.773 | 8.376 |
| K48N-H | 115.175 | 7.417 | 115.123 | 7.428 | 115.161 | 7.424 | 115.19 | 7.433 |
| K49N-H | 116.784 | 7.388 | 116.846 | 7.402 | 116.776 | 7.392 | 116.824 | 7.403 |
| L50N-H | 120.316 | 7.173 | 120.281 | 7.193 | 120.338 | 7.173 | 120.314 | 7.183 |
| E52N-H | 121.427 | 8.482 | 121.512 | 8.486 | 121.466 | 8.481 | 121.47 | 8.486 |
| G53N-H | 111.712 | 8.744 | 111.545 | 8.73 | 111.737 | 8.721 | 111.483 | 8.729 |
| V54N-H | 122.868 | 7.57 | 122.59 | 7.518 | 122.841 | 7.533 | 122.807 | 7.58 |
| K55N-H | 127.298 | 8.207 | 127.401 | 8.219 | 127.322 | 8.206 | 127.377 | 8.215 |
| M56N-H | 122.51 | 8.755 | 122.623 | 8.708 | 122.57 | 8.733 | 122.625 | 8.715 |
| T57N-H | 123.848 | 8.701 | 123.63 | 8.673 | 123.605 | 8.676 | 123.612 | 8.675 |
| K58N-H | 123.514 | 8.763 | 123.52 | 8.745 | 123.264 | 8.736 | 123.203 | 8.731 |
| Y59N-H | 125.398 | 9.33 | 125.583 | 9.319 | 125.319 | 9.333 | 125.425 | 9.324 |
| H60N-H | 123.086 | 9.163 | 123.063 | 9.15 | 123.099 | 9.15 | 123.081 | 9.148 |
| N62N-H | 117.838 | 9.176 | 117.814 | 9.173 | 117.89 | 9.173 | 117.918 | 9.162 |
| F63N-H | 112.061 | 6.476 | 111.909 | 6.492 | 112.102 | 6.477 | 112.06 | 6.504 |
| M64N-H | 117.086 | 6.904 | 117.139 | 6.896 | 117.077 | 6.904 | 117.055 | 6.907 |
| G65N-H | 107.894 | 8.543 | 107.812 | 8.551 | 107.83 | 8.549 | 107.767 | 8.55 |
| G66N-H | 108.916 | 8.124 | 108.864 | 8.104 | 108.904 | 8.106 | 108.876 | 8.1 |
| D67N-H | 128.176 | 9.068 | 128.141 | 9.045 | 128.134 | 9.05 | 128.127 | 9.048 |
| L68N-H | 119.632 | 8.069 | 119.621 | 8.067 | 119.641 | 8.068 | 119.621 | 8.062 |
| G69N-H | 106.991 | 7.624 | 106.969 | 7.626 | 106.969 | 7.626 | 106.969 | 7.626 |
| K70N-H | 121.748 | 7.503 | 121.597 | 7.465 | 121.672 | 7.485 | 121.664 | 7.555 |
| D71N-H | 121.982 | 7.882 | 121.999 | 7.877 | 121.993 | 7.882 | 121.976 | 7.879 |
| L72N-H | 120.409 | 8.809 | 120.426 | 8.8 | 120.362 | 8.797 | 120.396 | 8.799 |
| T73N-H | 120.029 | 7.987 | 120.001 | 8.003 | 119.971 | 7.994 | 120.035 | 7.998 |

| | | | | | | | | |
|---------|---------|--------|---------|-------|---------|--------|---------|--------|
| Q74N-H | 125.021 | 7.881 | 125.048 | 7.878 | 125.05 | 7.883 | 124.989 | 7.873 |
| A75N-H | 123.086 | 9.163 | 123.063 | 9.15 | 123.099 | 9.15 | 123.081 | 9.148 |
| W76N-H | 120.401 | 9.102 | 120.464 | 9.086 | 120.433 | 9.093 | 120.434 | 9.09 |
| A77N-H | 118.657 | 7.718 | 118.51 | 7.722 | 118.482 | 7.719 | 118.486 | 7.721 |
| V78N-H | 119.924 | 8.184 | 119.747 | 8.22 | 119.925 | 8.179 | 119.726 | 8.216 |
| A79N-H | 121.18 | 8.693 | 121.172 | 8.683 | 121.162 | 8.682 | 121.168 | 8.684 |
| M80N-H | 115.939 | 8.214 | 115.9 | 8.211 | 115.953 | 8.214 | 115.895 | 8.21 |
| A81N-H | 123.848 | 8.701 | 123.63 | 8.673 | 123.798 | 8.692 | 123.612 | 8.675 |
| L82N-H | 113.881 | 8.385 | 113.794 | 8.356 | 113.781 | 8.355 | 113.759 | 8.357 |
| G83N-H | 111.103 | 7.761 | 111.047 | 7.754 | 111.026 | 7.758 | 111.007 | 7.755 |
| V84N-H | 108.916 | 8.124 | 108.864 | 8.104 | 108.904 | 8.106 | 108.876 | 8.1 |
| E85N-H | 126.562 | 10.716 | 126.65 | 10.78 | 126.592 | 10.752 | 126.583 | 10.774 |
| D86N-H | 116.327 | 8.973 | 116.44 | 8.984 | 116.411 | 8.978 | 116.451 | 8.98 |
| K87N-H | 117.625 | 7.867 | 117.556 | 7.848 | 117.559 | 7.855 | 117.56 | 7.851 |
| T89N-H | 120.189 | 7.684 | 120.241 | 7.686 | 120.27 | 7.683 | 120.22 | 7.687 |
| V90N-H | 120.401 | 9.102 | 120.464 | 9.086 | 120.433 | 9.093 | 120.434 | 9.09 |
| L92N-H | 118.244 | 7.592 | 118.216 | 7.577 | 118.239 | 7.572 | 118.229 | 7.575 |
| F93N-H | 119.701 | 8.209 | 119.663 | 8.215 | 119.729 | 8.224 | 119.726 | 8.216 |
| E94N-H | 116.553 | 9.213 | 116.723 | 9.198 | 116.685 | 9.2 | 116.702 | 9.195 |
| G95N-H | 108.055 | 8.448 | 108.014 | 8.436 | 108.049 | 8.442 | 108.049 | 8.436 |
| V96N-H | 119.341 | 7.941 | 119.369 | 7.927 | 119.364 | 7.934 | 119.417 | 7.927 |
| Q97N-H | 113.049 | 8.34 | 113.035 | 8.366 | 113.119 | 8.356 | 112.995 | 8.369 |
| K98N-H | 117.367 | 7.943 | 117.218 | 7.967 | 117.186 | 7.953 | 117.344 | 7.966 |
| T99N-H | 106.539 | 8.078 | 106.548 | 8.082 | 106.548 | 8.082 | 106.548 | 8.082 |
| Q100N-H | 114.704 | 6.714 | 114.611 | 6.699 | 114.611 | 6.709 | 114.603 | 6.701 |
| I102N-H | 120.976 | 7.095 | 120.909 | 7.102 | 120.932 | 7.096 | 120.969 | 7.096 |
| R103N-H | 126.153 | 9.508 | 126.032 | 9.487 | 126.123 | 9.498 | 126.031 | 9.483 |
| S104N-H | 115.384 | 8.26 | 115.393 | 8.257 | 115.364 | 8.259 | 115.38 | 8.257 |
| A105N-H | 123.281 | 9.199 | 123.295 | 9.199 | 123.22 | 9.208 | 123.302 | 9.211 |
| S106N-H | 114.918 | 8.284 | 114.847 | 8.287 | 114.876 | 8.283 | 114.841 | 8.288 |
| D107N-H | 120.223 | 7.408 | 120.204 | 7.399 | 120.228 | 7.402 | 120.221 | 7.398 |
| I108N-H | 119.7 | 7.298 | 119.66 | 7.281 | 119.624 | 7.278 | 119.701 | 7.28 |
| R109N-H | 120.976 | 7.095 | 120.909 | 7.102 | 120.932 | 7.096 | 120.969 | 7.096 |
| D110N-H | 117.186 | 7.616 | 117.198 | 7.596 | 117.195 | 7.601 | 117.19 | 7.596 |
| V111N-H | 120.223 | 7.408 | 120.204 | 7.399 | 120.228 | 7.402 | 120.221 | 7.398 |
| F112N-H | 120.391 | 7.633 | 120.285 | 7.624 | 120.293 | 7.63 | 120.304 | 7.619 |
| I113N-H | 121.974 | 8.328 | 121.911 | 8.322 | 121.916 | 8.323 | 121.893 | 8.319 |
| N114N-H | 120.447 | 8.516 | 120.471 | 8.507 | 120.445 | 8.511 | 120.424 | 8.504 |

| | | | | | | | | |
|---------|---------|-------|---------|-------|---------|-------|---------|-------|
| A115N-H | 121.048 | 7.65 | 121.015 | 7.651 | 121.017 | 7.645 | 121.015 | 7.65 |
| I117N-H | 125.074 | 8.054 | 125.036 | 8.047 | 125.052 | 8.046 | 125.062 | 8.044 |
| K118N-H | 125.79 | 8.684 | 125.831 | 8.692 | 125.821 | 8.689 | 125.846 | 8.694 |
| G119N-H | 113.423 | 9.303 | 113.475 | 9.296 | 113.452 | 9.295 | 113.479 | 9.297 |
| E120N-H | 117.545 | 9.356 | 117.491 | 9.359 | 117.479 | 9.368 | 117.498 | 9.361 |
| E121N-H | 120.342 | 7.087 | 120.344 | 7.082 | 120.308 | 7.086 | 120.317 | 7.085 |
| Y122N-H | 122.783 | 8.375 | 122.632 | 8.345 | 122.702 | 8.351 | 122.819 | 8.384 |
| D123N-H | 119.213 | 8.792 | 119.182 | 8.788 | 119.219 | 8.793 | 119.205 | 8.788 |
| A124N-H | 121.048 | 7.65 | 121.015 | 7.651 | 121.017 | 7.645 | 121.015 | 7.65 |
| A125N-H | 120.629 | 7.835 | 120.559 | 7.832 | 120.663 | 7.834 | 120.67 | 7.828 |
| W126N-H | 121.524 | 9.239 | 121.496 | 9.23 | 121.489 | 9.228 | 121.473 | 9.227 |
| N127N-H | 110.42 | 6.98 | 110.405 | 6.979 | 110.384 | 6.975 | 110.38 | 6.973 |
| S128N-H | 116.42 | 7.795 | 116.348 | 7.782 | 116.39 | 7.783 | 116.351 | 7.78 |
| V130N-H | 119.924 | 8.184 | 119.747 | 8.22 | 119.925 | 8.179 | 119.924 | 8.184 |
| V131N-H | 122.072 | 7.194 | 122.051 | 7.208 | 122.086 | 7.207 | 122.066 | 7.212 |
| K132N-H | 121.484 | 7.766 | 121.473 | 7.772 | 121.392 | 7.798 | 121.697 | 7.784 |
| S133N-H | 115.614 | 8.51 | 115.642 | 8.495 | 115.565 | 8.475 | 115.626 | 8.502 |
| L134N-H | 124.094 | 8.171 | 124.02 | 8.17 | 123.994 | 8.171 | 123.987 | 8.17 |
| V135N-H | 123.279 | 8.651 | 123.287 | 8.647 | 123.294 | 8.643 | 123.291 | 8.645 |
| A136N-H | 121.246 | 7.454 | 121.244 | 7.457 | 121.225 | 7.455 | 121.224 | 7.458 |
| Q137N-H | 118.544 | 8.539 | 118.529 | 8.531 | 118.494 | 8.534 | 118.529 | 8.535 |
| Q138N-H | 119.341 | 7.941 | 119.369 | 7.927 | 119.364 | 7.934 | 119.417 | 7.927 |
| Q139N-H | 121.081 | 8.163 | 121.074 | 8.173 | 121.058 | 8.163 | 121.025 | 8.17 |
| K140N-H | 121.917 | 8.602 | 121.899 | 8.596 | 121.874 | 8.6 | 121.867 | 8.598 |
| A141N-H | 119.539 | 8.129 | 119.52 | 8.132 | 119.549 | 8.127 | 119.592 | 8.139 |
| A142N-H | 116.784 | 7.388 | 116.846 | 7.402 | 116.776 | 7.392 | 116.824 | 7.403 |
| A143N-H | 120.981 | 7.584 | 120.954 | 7.573 | 120.95 | 7.577 | 120.954 | 7.571 |
| D144N-H | 120.872 | 9.077 | 120.881 | 9.059 | 120.919 | 9.062 | 120.904 | 9.061 |
| V145N-H | 108.596 | 7.056 | 108.554 | 7.055 | 108.569 | 7.056 | 108.57 | 7.054 |
| Q146N-H | 119.371 | 7.857 | 119.335 | 7.846 | 119.327 | 7.847 | 119.346 | 7.844 |
| L147N-H | 118.244 | 7.592 | 118.216 | 7.577 | 118.239 | 7.572 | 118.229 | 7.575 |
| R148N-H | 123.281 | 9.199 | 123.295 | 9.199 | 123.22 | 9.208 | 123.302 | 9.211 |
| G149N-H | 106.011 | 7.147 | 106.032 | 7.144 | 106.032 | 7.144 | 106.032 | 7.144 |
| V150N-H | 110.11 | 8.23 | 110.092 | 8.247 | 110.008 | 8.237 | 110.055 | 8.248 |
| A152N-H | 122.682 | 8.154 | 122.896 | 8.159 | 122.771 | 8.151 | 122.752 | 8.15 |
| M153N-H | 121.565 | 8.548 | 121.71 | 8.574 | 121.874 | 8.6 | 121.867 | 8.598 |
| F154N-H | 124.401 | 9.714 | 124.677 | 9.72 | 124.311 | 9.721 | 124.5 | 9.723 |
| V155N-H | 122.769 | 9.486 | 122.674 | 9.479 | 122.867 | 9.487 | 122.934 | 9.483 |

| | | | | | | | | |
|---------|---------|--------|---------|--------|---------|--------|---------|--------|
| N156N-H | 127.1 | 10.234 | 127.03 | 10.218 | 127.147 | 10.223 | 127.043 | 10.218 |
| K158N-H | 114.775 | 7.91 | 114.787 | 7.896 | 114.663 | 7.893 | 114.829 | 7.904 |
| Y159N-H | 117.367 | 7.943 | 117.218 | 7.967 | 117.186 | 7.953 | 117.344 | 7.966 |
| Q160N-H | 124.732 | 9.548 | 124.678 | 9.544 | 124.707 | 9.542 | 124.684 | 9.54 |
| L161N-H | 113.049 | 8.34 | 113.035 | 8.366 | 113.119 | 8.356 | 112.995 | 8.369 |
| N162N-H | 119.165 | 8.411 | 119.14 | 8.369 | 119.193 | 8.397 | 119.064 | 8.37 |
| Q164N-H | 115.058 | 8.029 | 114.293 | 8.027 | 114.583 | 8.02 | 114.3 | 8.027 |
| G165N-H | 107.228 | 7.974 | 107.251 | 7.977 | 107.251 | 7.977 | 107.251 | 7.977 |
| M166N-H | 119.376 | 7.374 | 119.445 | 7.351 | 119.34 | 7.35 | 119.379 | 7.357 |
| D167N-H | 121.079 | 8.648 | 120.947 | 8.647 | 120.932 | 8.638 | 121.168 | 8.684 |
| T168N-H | 112.355 | 7.841 | 112.235 | 7.788 | 112.291 | 7.797 | 112.294 | 7.787 |
| S169N-H | 118.523 | 8.699 | 118.081 | 8.711 | 118.238 | 8.714 | 118.072 | 8.696 |
| N170N-H | 117.622 | 7.319 | 117.752 | 7.31 | 117.584 | 7.305 | 117.659 | 7.289 |
| M171N-H | 123.086 | 9.163 | 123.063 | 9.15 | 123.099 | 9.15 | 123.081 | 9.148 |
| D172N-H | 120.73 | 8.229 | 120.728 | 8.253 | 120.734 | 8.245 | 120.69 | 8.252 |
| V173N-H | 121.982 | 7.882 | 121.999 | 7.877 | 121.993 | 7.882 | 121.976 | 7.879 |
| F174N-H | 121.635 | 8.226 | 121.502 | 8.222 | 121.556 | 8.209 | 121.572 | 8.233 |
| V175N-H | 116.766 | 8.03 | 116.621 | 8.019 | 116.747 | 8.031 | 116.643 | 8.029 |
| Q176N-H | 116.983 | 7.354 | 117.179 | 7.374 | 117.107 | 7.386 | 117.178 | 7.368 |
| Q177N-H | 119.502 | 8.742 | 119.314 | 8.707 | 119.377 | 8.721 | 119.28 | 8.715 |
| Y178N-H | 126.237 | 8.939 | 126.049 | 8.931 | 126.058 | 8.926 | 126.039 | 8.929 |
| A179N-H | 119.924 | 8.184 | 119.747 | 8.22 | 119.885 | 8.251 | 119.726 | 8.216 |
| D180N-H | 118.605 | 8.856 | 118.505 | 8.833 | 118.556 | 8.839 | 118.556 | 8.826 |
| T181N-N | 120.137 | 7.756 | 120.127 | 7.75 | 120.144 | 7.747 | 120.181 | 7.752 |
| V182N-H | 121.484 | 7.766 | 121.473 | 7.772 | 121.392 | 7.798 | 121.697 | 7.784 |
| K183N-H | 121.625 | 7.561 | 121.646 | 7.556 | 121.705 | 7.561 | 121.664 | 7.555 |
| Y184N-H | 119.632 | 8.069 | 119.621 | 8.067 | 119.641 | 8.068 | 119.621 | 8.062 |
| L185N-H | 117.852 | 8.134 | 117.686 | 8.124 | 117.829 | 8.099 | 117.719 | 8.119 |
| S186N-H | 115.023 | 8.465 | 114.971 | 8.446 | 114.997 | 8.454 | 114.941 | 8.443 |
| E187N-H | 119.253 | 7.041 | 119.26 | 7.035 | 119.315 | 7.04 | 119.267 | 7.038 |
| K188N-H | 122.844 | 7.194 | 122.788 | 7.194 | 122.79 | 7.199 | 122.776 | 7.195 |
| K189N-H | 128.986 | 7.952 | 128.948 | 7.94 | 128.885 | 7.935 | 128.908 | 7.937 |

8.7.2 ¹H -¹⁵N HSQ C chemical shifts of EcDsbA in complex with SigA (N-terminally acetylated) peptide. (Paper 8.1.1)

| Residue | N | NH |
|---------|---------|--------|
| Q2N-H | 119.926 | 8.282 |
| Y3N-H | 120.431 | 7.631 |
| E4N-H | 122.876 | 8.963 |
| D5N-H | 125.678 | 8.502 |
| G6N-H | 116.579 | 8.985 |
| K7N-H | 121.822 | 8.134 |
| Q8N-H | 115.616 | 8.495 |
| Y9N-H | 111.771 | 7.559 |
| T10N-H | 112.146 | 9.523 |
| T11N-H | 121.373 | 9.055 |
| L12N-H | 129.171 | 8.706 |
| E13N-H | 123.039 | 8.664 |
| K14N-H | 117.258 | 7.803 |
| V16N-H | 123.193 | 9.029 |
| A17N-H | 132.031 | 8.6 |
| G18N-H | 110.721 | 8.422 |
| A19N-H | 122.361 | 7.423 |
| Q21N-H | 121.733 | 8.439 |
| V22N-H | 116.096 | 7.709 |
| L23N-H | 129.239 | 8.858 |
| E24N-H | 128.34 | 9.113 |
| F25N-H | 121.546 | 9.229 |
| F26N-H | 119.652 | 8.963 |
| S27N-H | 108.907 | 6.547 |
| F28N-H | 131.436 | 11.326 |
| F29N-H | 116.063 | 8.707 |
| C30N-H | 120.33 | 7.282 |
| C33N-H | 116.133 | 8.844 |
| Y34N-H | 120.097 | 7.808 |
| Q35N-H | 118.574 | 7.727 |

| | | |
|--------|---------|-------|
| F36N-H | 120.427 | 8.433 |
| E37N-H | 115.226 | 7.406 |
| E38N-H | 113.477 | 7.919 |
| V39N-H | 116.708 | 7.126 |
| L40N-H | 115.205 | 7.884 |
| H41N-H | 111.475 | 6.369 |
| I42N-H | 119.741 | 8.341 |
| S43N-H | 118.546 | 8.537 |
| D44N-H | 121.4 | 7.787 |
| N45N-H | 117.195 | 7.618 |
| V46N-H | 120.481 | 8.697 |
| K47N-H | 118.693 | 8.365 |
| K48N-H | 115.226 | 7.406 |
| K49N-H | 116.834 | 7.382 |
| L50N-H | 120.33 | 7.167 |
| E52N-H | 121.551 | 8.552 |
| G53N-H | 111.786 | 8.742 |
| V54N-H | 121.704 | 7.555 |
| K55N-H | 127.186 | 8.196 |
| M56N-H | 122.474 | 8.775 |
| T57N-H | 123.834 | 8.692 |
| K58N-H | 123.788 | 8.775 |
| Y59N-H | 125.079 | 9.338 |
| H60N-H | 123.273 | 9.199 |
| V61N-H | 111.925 | 6.174 |
| N62N-H | 117.844 | 9.167 |
| F63N-H | 111.736 | 6.446 |
| M64N-H | 117.077 | 6.91 |
| G65N-H | 107.927 | 8.538 |
| G66N-H | 108.935 | 8.128 |
| D67N-H | 128.19 | 9.09 |
| L68N-H | 120.032 | 8.065 |
| G69N-H | 106.978 | 7.617 |
| K70N-H | 121.829 | 7.525 |
| D71N-H | 121.993 | 7.884 |
| L72N-H | 120.418 | 8.809 |
| T73N-H | 120.037 | 7.99 |
| Q74N-H | 124.994 | 7.877 |

| | | |
|---------|---------|-------|
| A75N-H | 123.079 | 9.158 |
| W76N-H | 120.333 | 9.09 |
| A77N-H | 118.574 | 7.727 |
| V78N-H | 119.731 | 8.21 |
| A79N-H | 121.165 | 8.689 |
| M80N-H | 115.952 | 8.205 |
| A81N-H | 123.834 | 8.692 |
| L82N-H | 113.864 | 8.38 |
| G83N-H | 111.048 | 7.754 |
| V84N-H | 108.935 | 8.128 |
| E85N-H | 126.472 | 10.65 |
| D86N-H | 116.324 | 8.962 |
| K87N-H | 117.643 | 7.863 |
| V88N-H | 104.175 | 7.042 |
| T89N-H | 120.178 | 7.752 |
| V90N-H | 120.333 | 9.09 |
| L92N-H | 118.235 | 7.583 |
| F93N-H | 119.84 | 8.163 |
| E94N-H | 116.478 | 9.21 |
| G95N-H | 108.091 | 8.449 |
| V96N-H | 119.434 | 7.977 |
| Q97N-H | 113.025 | 8.319 |
| K98N-H | 117.373 | 7.916 |
| T99N-H | 106.523 | 8.08 |
| Q100N-H | 114.695 | 6.718 |
| I102N-H | 120.91 | 7.122 |
| R103N-H | 126.103 | 9.51 |
| S104N-H | 115.367 | 8.256 |
| A105N-H | 123.592 | 9.225 |
| S106N-H | 114.899 | 8.283 |
| D107N-H | 120.23 | 7.406 |
| I108N-H | 119.734 | 7.295 |
| R109N-H | 120.965 | 7.089 |
| D110N-H | 117.195 | 7.618 |
| V111N-H | 120.23 | 7.406 |
| F112N-H | 120.431 | 7.631 |
| I113N-H | 121.963 | 8.325 |
| N114N-H | 120.434 | 8.511 |

| | | |
|---------|---------|-------|
| A115N-H | 120.997 | 7.713 |
| G116N-H | 105.572 | 7.812 |
| I117N-H | 125.062 | 8.054 |
| K118N-H | 125.794 | 8.68 |
| G119N-H | 113.432 | 9.302 |
| E120N-H | 117.531 | 9.353 |
| E121N-H | 120.367 | 7.088 |
| Y122N-H | 122.781 | 8.368 |
| D123N-H | 119.211 | 8.786 |
| A124N-H | 120.993 | 7.639 |
| A125N-H | 120.645 | 7.832 |
| W126N-H | 121.536 | 9.232 |
| N127N-H | 110.431 | 6.978 |
| S128N-H | 116.429 | 7.792 |
| V130N-H | 119.731 | 8.21 |
| V131N-H | 122.07 | 7.194 |
| K132N-H | 121.809 | 7.779 |
| S133N-H | 115.616 | 8.495 |
| L134N-H | 124.086 | 8.166 |
| V135N-H | 123.039 | 8.664 |
| A136N-H | 121.282 | 7.451 |
| Q137N-H | 118.546 | 8.537 |
| Q138N-H | 119.289 | 7.94 |
| Q139N-H | 121.057 | 8.152 |
| K140N-H | 121.934 | 8.603 |
| A141N-H | 119.512 | 8.125 |
| A142N-H | 116.834 | 7.382 |
| A143N-H | 120.939 | 7.59 |
| D144N-H | 120.86 | 9.071 |
| V145N-H | 108.63 | 7.062 |
| Q146N-H | 119.331 | 7.852 |
| L147N-H | 118.235 | 7.583 |
| R148N-H | 123.592 | 9.225 |
| G149N-H | 106.014 | 7.152 |
| V150N-H | 110.04 | 8.222 |
| A152N-H | 122.628 | 8.162 |
| M153N-H | 121.551 | 8.552 |
| F154N-H | 124.388 | 9.697 |

| | | |
|---------|---------|-------|
| V155N-H | 122.739 | 9.483 |
| N156N-H | 127.127 | 10.24 |
| G157N-H | 105.605 | 9.399 |
| K158N-H | 114.735 | 7.91 |
| Y159N-H | 117.373 | 7.916 |
| Q160N-H | 124.743 | 9.548 |
| L161N-H | 113.025 | 8.319 |
| N162N-H | 118.693 | 8.365 |
| Q164N-H | 114.91 | 8.037 |
| G165N-H | 107.187 | 7.97 |
| M166N-H | 119.363 | 7.367 |
| D167N-H | 121.165 | 8.689 |
| T168N-H | 112.343 | 7.831 |
| S169N-H | 118.35 | 8.692 |
| N170N-H | 117.377 | 7.308 |
| M171N-H | 123.079 | 9.158 |
| D172N-H | 120.855 | 8.217 |
| V173N-H | 121.993 | 7.884 |
| F174N-H | 121.667 | 8.201 |
| V175N-H | 116.844 | 8.048 |
| Q176N-H | 116.834 | 7.382 |
| Q177N-H | 119.529 | 8.725 |
| Y178N-H | 126.228 | 8.94 |
| A179N-H | 119.731 | 8.21 |
| D180N-H | 118.623 | 8.848 |
| T181N-N | 120.097 | 7.808 |
| V182N-H | 121.4 | 7.787 |
| K183N-H | 121.704 | 7.555 |
| Y184N-H | 119.578 | 8.067 |
| L185N-H | 117.85 | 8.129 |
| S186N-H | 115.046 | 8.459 |
| E187N-H | 119.247 | 7.044 |
| K188N-H | 122.838 | 7.195 |
| K189N-H | 128.963 | 7.947 |

8.8 VcDsbA-fragment binding shifts

8.8.1 $^1\text{H}^{15}\text{N}$ HSQC chemical shifts of VcDsbA_Apo and solution complexes of fragments M2 & M3 and M5. (Chapter 4)

| Residue | VcDsbA_Apo | | VcDsbA- M2 | | VcDsbA- M3 | | VcDsbA- M3 | |
|---------|------------|-------|------------|-------|------------|-------|------------|-------|
| | N | NH | N | NH | NH | H | NH | H |
| Q2N-H | 118.321 | 7.661 | 118.278 | 7.669 | 118.275 | 7.666 | 118.273 | 7.67 |
| F3N-H | 119.644 | 7.206 | 119.733 | 7.249 | 119.724 | 7.241 | 119.713 | 7.241 |
| E5N-H | 129.49 | 9.047 | 129.574 | 9.034 | 129.533 | 9.029 | 129.612 | 9.038 |
| G6N-H | 116.818 | 8.901 | 116.773 | 8.903 | 116.782 | 8.885 | 116.773 | 8.899 |
| E7N-H | 120.836 | 7.48 | 120.882 | 7.48 | 120.833 | 7.484 | 120.779 | 7.49 |
| H8N-H | 114.645 | 8.527 | 114.677 | 8.522 | 114.725 | 8.526 | 114.784 | 8.531 |
| Y9N-H | 110.516 | 7.232 | 110.576 | 7.25 | 110.54 | 7.245 | 110.457 | 7.243 |
| Q10N-H | 119.455 | 9.191 | 119.485 | 9.245 | 119.48 | 9.199 | 119.455 | 9.212 |
| V11N-H | 125.949 | 9.43 | 125.961 | 9.426 | 125.927 | 9.405 | 126.059 | 9.463 |
| L12N-H | 129.244 | 9.35 | 129.207 | 9.324 | 129.246 | 9.338 | 129.191 | 9.339 |
| K13N-H | 117.442 | 8.456 | 117.4 | 8.445 | 117.506 | 8.437 | 117.505 | 8.441 |
| T14N-H | 112.143 | 7.424 | 112.081 | 7.428 | 112.13 | 7.43 | 112.044 | 7.424 |
| A16N-H | 122.262 | 8.499 | 122.3 | 8.472 | 122.224 | 8.472 | 122.215 | 8.467 |
| S17N-H | 115.15 | 7.741 | 115.162 | 7.755 | 115.172 | 7.747 | 115.17 | 7.744 |
| S19N-H | 115.993 | 7.6 | 116.04 | 7.605 | 116.016 | 7.605 | 116.01 | 7.606 |
| V21N-H | 125.465 | 9.113 | 125.436 | 9.102 | 125.418 | 9.106 | 125.35 | 9.102 |
| V22N-H | 126.631 | 8.89 | 126.757 | 8.896 | 126.678 | 8.891 | 126.666 | 8.881 |
| S23N-H | 120.841 | 9.324 | 120.863 | 9.32 | 120.997 | 9.329 | 121.008 | 9.342 |
| E24N-H | 122.262 | 8.499 | 122.3 | 8.472 | 122.224 | 8.472 | 122.215 | 8.467 |
| F25N-H | 128.259 | 9.723 | 128.366 | 9.784 | 128.364 | 9.787 | 128.38 | 9.762 |
| F26N-H | 121.923 | 9.002 | 121.995 | 9.008 | 121.945 | 9.005 | 121.965 | 9.006 |
| S27N-H | 110.397 | 6.675 | 110.445 | 6.695 | 110.525 | 6.697 | 110.472 | 6.7 |
| Y29N-H | 115.567 | 8.499 | 115.724 | 8.511 | 115.708 | 8.508 | 115.756 | 8.5 |
| C30N-H | 122.088 | 7.493 | 121.981 | 7.475 | 122.044 | 7.474 | 121.867 | 7.474 |
| C33N-H | 115.838 | 8.712 | 116.124 | 8.738 | 116.086 | 8.73 | 115.56 | 8.785 |
| N34N-H | 119.644 | 7.206 | 119.733 | 7.249 | 119.724 | 7.241 | 119.713 | 7.241 |

Appendices

| | | | | | | | | |
|--------|---------|-------|---------|-------|---------|-------|---------|-------|
| T35N-H | 113.476 | 7.755 | 113.528 | 7.764 | 113.528 | 7.764 | 113.622 | 7.76 |
| F36N-H | 118.834 | 7.385 | 118.891 | 7.398 | 118.933 | 7.41 | 118.906 | 7.393 |
| E37N-H | 119.324 | 7.124 | 119.365 | 7.118 | 119.443 | 7.121 | 119.367 | 7.127 |
| I39N-H | 118.457 | 7.059 | 118.374 | 6.985 | 118.56 | 7.019 | 118.637 | 7.036 |
| I40N-H | 121.463 | 7.437 | 121.007 | 7.433 | 121.432 | 7.439 | 121.442 | 7.435 |
| A41N-H | 120.752 | 8.067 | 120.975 | 8.057 | 120.804 | 8.035 | 120.977 | 8.033 |
| Q42N-H | 116.805 | 7.181 | 116.878 | 7.2 | 116.837 | 7.191 | 116.769 | 7.184 |
| L43N-H | 122.335 | 8.785 | 122.527 | 8.741 | 122.406 | 8.744 | 122.486 | 8.824 |
| K44N-H | 116.454 | 8.173 | 116.435 | 8.14 | 116.423 | 8.164 | 116.573 | 8.181 |
| Q45N-H | 115.394 | 7.118 | 115.265 | 7.115 | 115.344 | 7.093 | 115.333 | 7.089 |
| Q46N-H | 114.084 | 7.357 | 114.265 | 7.35 | 114.267 | 7.359 | 114.025 | 7.35 |
| L47N-H | 121.316 | 6.943 | 121.224 | 6.918 | 121.321 | 6.935 | 121.311 | 6.925 |
| E49N-H | 121.924 | 8.568 | 121.986 | 8.56 | 121.972 | 8.557 | 121.924 | 8.548 |
| G50N-H | 112.715 | 8.804 | 112.477 | 8.768 | 112.604 | 8.776 | 112.538 | 8.769 |
| A51N-H | 122.125 | 7.707 | 122.151 | 7.693 | 122.159 | 7.695 | 122.153 | 7.697 |
| K52N-H | 121.557 | 7.693 | 121.767 | 7.717 | 121.65 | 7.706 | 121.688 | 7.72 |
| F53N-H | 122.335 | 8.785 | 122.527 | 8.741 | 122.406 | 8.744 | 122.486 | 8.824 |
| Q54N-H | 128.105 | 8.465 | 128.029 | 8.451 | 128.065 | 8.455 | 128.063 | 8.447 |
| K55N-H | 126.597 | 8.539 | 126.493 | 8.529 | 126.545 | 8.52 | 126.56 | 8.522 |
| V58N-H | 116.805 | 7.181 | 116.878 | 7.2 | 116.837 | 7.191 | 116.769 | 7.184 |
| S59N-H | 117.801 | 9.961 | 117.916 | 9.986 | 117.98 | 9.982 | 117.988 | 9.99 |
| F60N-H | 114.978 | 6.96 | 114.973 | 6.945 | 115.015 | 6.948 | 115.052 | 6.941 |
| M61N-H | 118.303 | 6.435 | 118.302 | 6.449 | 118.32 | 6.449 | 118.274 | 6.447 |
| G62N-H | 108.168 | 8.544 | 108.282 | 8.544 | 108.273 | 8.542 | 108.257 | 8.539 |
| G63N-H | 112.138 | 8.623 | 112.027 | 8.596 | 112.044 | 8.599 | 112.092 | 8.602 |
| M65N-H | 118.321 | 7.661 | 118.278 | 7.669 | 118.275 | 7.666 | 118.273 | 7.67 |
| G66N-H | 109.698 | 7.73 | 109.783 | 7.746 | 109.756 | 7.743 | 109.778 | 7.746 |
| Q67N-H | 120.614 | 8.462 | 120.641 | 8.448 | 120.634 | 8.447 | 120.645 | 8.445 |
| A68N-H | 123.004 | 8.168 | 123.041 | 8.137 | 123.048 | 8.138 | 123.023 | 8.134 |
| M69N-H | 118.992 | 8.915 | 118.983 | 8.91 | 119.033 | 8.91 | 119.026 | 8.913 |
| S70N-H | 118.549 | 8.506 | 118.54 | 8.504 | 118.562 | 8.506 | 118.559 | 8.508 |
| K71N-H | 122.335 | 8.785 | 122.527 | 8.741 | 122.406 | 8.744 | 122.486 | 8.824 |
| A72N-H | 126.02 | 9.473 | 126.015 | 9.459 | 126.058 | 9.463 | 125.988 | 9.464 |
| Y73N-H | 119.966 | 8.602 | 119.979 | 8.608 | 119.99 | 8.608 | 120.059 | 8.615 |
| A74N-H | 119.852 | 8.655 | 119.949 | 8.656 | 119.933 | 8.659 | 119.963 | 8.664 |

Appendices

| | | | | | | | | |
|---------|---------|-------|---------|-------|---------|-------|---------|-------|
| T75N-H | 117.407 | 8.586 | 117.426 | 8.576 | 117.42 | 8.58 | 117.453 | 8.581 |
| M76N-H | 121.478 | 8.263 | 121.435 | 8.261 | 121.452 | 8.258 | 121.48 | 8.255 |
| I77N-H | 117.32 | 7.248 | 117.429 | 7.256 | 117.388 | 7.258 | 117.386 | 7.266 |
| A78N-H | 124.374 | 8.55 | 124.357 | 8.528 | 124.354 | 8.533 | 124.346 | 8.541 |
| L79N-H | 113.846 | 8.259 | 113.825 | 8.244 | 113.842 | 8.245 | 113.845 | 8.244 |
| E80N-H | 116.634 | 7.963 | 116.664 | 7.965 | 116.682 | 7.967 | 116.633 | 7.97 |
| V81N-H | 108.079 | 8.414 | 108.087 | 8.408 | 108.08 | 8.418 | 108.068 | 8.419 |
| E82N-H | 125.194 | 8.772 | 125.319 | 8.768 | 125.283 | 8.772 | 125.33 | 8.799 |
| D83N-H | 116.308 | 8.615 | 116.367 | 8.6 | 116.359 | 8.603 | 116.43 | 8.606 |
| K84N-H | 116.939 | 7.497 | 116.924 | 7.472 | 116.923 | 7.475 | 116.936 | 7.478 |
| M85N-H | 113.082 | 8.511 | 113.118 | 8.498 | 113.07 | 8.501 | 113.136 | 8.495 |
| V86N-H | 122.335 | 8.785 | 122.527 | 8.741 | 122.406 | 8.744 | 122.486 | 8.824 |
| V88N-H | 117.112 | 6.608 | 117.092 | 6.614 | 117.096 | 6.617 | 117.113 | 6.614 |
| M89N-H | 121.478 | 8.263 | 121.435 | 8.261 | 121.452 | 8.258 | 121.48 | 8.255 |
| F90N-H | 117.061 | 8.294 | 117.085 | 8.272 | 117.06 | 8.301 | 116.973 | 8.307 |
| N91N-H | 116.906 | 8.473 | 116.947 | 8.48 | 116.961 | 8.481 | 116.963 | 8.483 |
| I93N-H | 114.084 | 7.357 | 114.265 | 7.35 | 114.267 | 7.359 | 114.025 | 7.35 |
| H94N-H | 115.197 | 8.327 | 115.159 | 8.325 | 115.161 | 8.321 | 115.122 | 8.347 |
| T95N-H | 117.277 | 8.001 | 117.317 | 8.017 | 117.268 | 8.011 | 117.258 | 8.019 |
| L96N-H | 117.881 | 8.303 | 117.927 | 8.306 | 117.963 | 8.307 | 118.007 | 8.312 |
| R97N-H | 109.376 | 6.48 | 109.436 | 6.482 | 109.5 | 6.487 | 109.513 | 6.493 |
| K98N-H | 117.269 | 8.095 | 117.32 | 8.101 | 117.337 | 8.102 | 117.345 | 8.102 |
| K101N-H | 123.39 | 9.254 | 123.361 | 9.26 | 123.402 | 9.255 | 123.378 | 9.245 |
| D102N-H | 113.729 | 7.455 | 113.75 | 7.464 | 113.758 | 7.463 | 113.706 | 7.464 |
| E103N-H | 117.206 | 8.993 | 117.251 | 8.995 | 117.214 | 8.992 | 117.262 | 8.992 |
| Q104N-H | 123.269 | 8.311 | 123.182 | 8.303 | 123.2 | 8.304 | 123.166 | 8.306 |
| E105N-H | 122.488 | 8.843 | 122.553 | 8.836 | 122.553 | 8.836 | 122.486 | 8.824 |
| L106N-H | 119.945 | 7.764 | 119.996 | 7.762 | 119.976 | 7.759 | 119.971 | 7.757 |
| R107N-H | 119.079 | 7.582 | 119.1 | 7.579 | 119.11 | 7.581 | 119.136 | 7.593 |
| Q108N-H | 119.69 | 7.81 | 119.657 | 7.822 | 119.677 | 7.811 | 119.718 | 7.825 |
| I109N-H | 117.873 | 7.502 | 117.948 | 7.511 | 117.919 | 7.508 | 117.97 | 7.51 |
| F110N-H | 115.394 | 7.118 | 115.265 | 7.115 | 115.344 | 7.093 | 115.393 | 7.115 |
| L111N-H | 121.363 | 7.771 | 121.403 | 7.772 | 121.385 | 7.778 | 121.426 | 7.774 |
| D112N-H | 120.865 | 8.986 | 120.951 | 8.993 | 120.963 | 8.99 | 120.93 | 8.992 |
| E113N-H | 116.805 | 7.181 | 116.878 | 7.2 | 116.837 | 7.191 | 116.769 | 7.184 |

Appendices

| | | | | | | | | |
|---------|---------|-------|---------|-------|---------|-------|---------|-------|
| G114N-H | 106.538 | 7.954 | 106.581 | 7.951 | 106.566 | 7.955 | 106.545 | 7.954 |
| I115N-H | 122.61 | 7.294 | 122.671 | 7.306 | 122.669 | 7.306 | 122.644 | 7.314 |
| D116N-H | 128.189 | 8.572 | 128.156 | 8.539 | 128.178 | 8.541 | 128.119 | 8.527 |
| A117N-H | 129.982 | 8.786 | 129.995 | 8.767 | 129.992 | 8.768 | 129.967 | 8.769 |
| A118N-H | 118.22 | 8.064 | 118.217 | 8.072 | 118.203 | 8.072 | 118.187 | 8.074 |
| K119N-H | 119.162 | 8.132 | 119.101 | 8.101 | 119.101 | 8.099 | 119.072 | 8.089 |
| F120N-H | 120.555 | 8.61 | 120.525 | 8.587 | 120.517 | 8.589 | 120.502 | 8.586 |
| D121N-H | 119.827 | 9.07 | 119.831 | 9.068 | 119.856 | 9.068 | 119.827 | 9.068 |
| A122N-H | 119.877 | 7.685 | 119.958 | 7.675 | 119.954 | 7.675 | 119.874 | 7.677 |
| A123N-H | 118.834 | 7.385 | 118.891 | 7.398 | 118.909 | 7.398 | 118.906 | 7.393 |
| Y124N-H | 120.39 | 9.051 | 120.371 | 9.042 | 120.38 | 9.041 | 120.389 | 9.04 |
| N125N-H | 109.444 | 6.873 | 109.488 | 6.868 | 109.48 | 6.869 | 109.506 | 6.865 |
| G126N-H | 107.906 | 7.522 | 107.85 | 7.515 | 107.89 | 7.526 | 107.887 | 7.525 |
| F127N-H | 119.099 | 8.378 | 119.093 | 8.366 | 119.136 | 8.363 | 119.12 | 8.37 |
| A128N-H | 124.967 | 8.464 | 124.927 | 8.439 | 124.942 | 8.441 | 124.952 | 8.442 |
| V129N-H | 121.312 | 8.66 | 121.216 | 8.627 | 121.205 | 8.625 | 121.207 | 8.622 |
| D130N-H | 119.324 | 7.124 | 119.365 | 7.118 | 119.443 | 7.121 | 119.367 | 7.127 |
| S131N-H | 111.986 | 8.244 | 112.074 | 8.23 | 112.048 | 8.233 | 112.089 | 8.23 |
| M132N-H | 122.741 | 7.883 | 122.712 | 7.89 | 122.729 | 7.885 | 122.75 | 7.887 |
| V133N-H | 117.144 | 7.811 | 117.122 | 7.803 | 117.101 | 7.807 | 117.106 | 7.803 |
| R134N-H | 117.004 | 7.593 | 117.048 | 7.585 | 117.08 | 7.59 | 117.075 | 7.591 |
| R135N-H | 120.278 | 7.629 | 120.263 | 7.627 | 120.272 | 7.626 | 120.27 | 7.628 |
| F136N-H | 120.718 | 8.717 | 120.67 | 8.719 | 120.718 | 8.717 | 120.665 | 8.717 |
| D137N-H | 116.291 | 8.046 | 116.231 | 8.014 | 116.271 | 8.021 | 116.274 | 8.018 |
| K138N-H | 122.723 | 8.489 | 122.763 | 8.5 | 122.774 | 8.499 | 122.758 | 8.497 |
| Q139N-H | 117.345 | 8.849 | 117.31 | 8.842 | 117.336 | 8.839 | 117.362 | 8.839 |
| D140N-H | 120.428 | 8.007 | 120.43 | 7.987 | 120.477 | 7.989 | 120.554 | 7.994 |
| Q141N-H | 120.555 | 8.61 | 120.525 | 8.587 | 120.517 | 8.589 | 120.502 | 8.586 |
| D142N-H | 120.841 | 9.324 | 120.863 | 9.32 | 120.997 | 9.329 | 121.008 | 9.342 |
| S143N-H | 112.268 | 7.355 | 112.292 | 7.378 | 112.282 | 7.374 | 112.279 | 7.378 |
| G144N-H | 107.631 | 7.352 | 107.652 | 7.356 | 107.667 | 7.366 | 107.647 | 7.368 |
| L145N-H | 119.333 | 6.892 | 119.206 | 6.875 | 119.217 | 6.859 | 119.284 | 6.887 |
| T146N-H | 106.393 | 8.425 | 106.359 | 8.4 | 106.411 | 8.398 | 106.446 | 8.407 |
| G147N-H | 108.704 | 7.472 | 108.691 | 7.481 | 108.713 | 7.476 | 108.844 | 7.494 |
| V148N-H | 108.38 | 8.319 | 108.273 | 8.283 | 108.309 | 8.271 | 108.105 | 8.238 |

| | | | | | | | | |
|---------|---------|--------|---------|--------|---------|--------|---------|--------|
| A150N-H | 122.551 | 8.399 | 122.545 | 8.404 | 122.628 | 8.402 | 122.625 | 8.415 |
| V151N-H | 125.521 | 9.467 | 125.441 | 9.456 | 125.62 | 9.48 | 125.85 | 9.551 |
| V152N-H | 126.904 | 9.326 | 127.007 | 9.36 | 126.997 | 9.333 | 126.992 | 9.328 |
| V153N-H | 130.328 | 10.055 | 130.334 | 10.054 | 130.344 | 10.052 | 130.448 | 10.057 |
| N154N-H | 128.346 | 10.353 | 128.366 | 10.357 | 128.352 | 10.367 | 128.347 | 10.362 |
| N155N-H | 113.109 | 9.313 | 113.257 | 9.316 | 113.167 | 9.314 | 113.21 | 9.313 |
| R156N-H | 117.678 | 7.452 | 117.632 | 7.47 | 117.721 | 7.457 | 117.68 | 7.459 |
| Y157N-H | 117.881 | 8.303 | 117.927 | 8.306 | 117.963 | 8.307 | 118.007 | 8.312 |
| L158N-H | 127.971 | 9.468 | 128.019 | 9.457 | 127.991 | 9.458 | 127.953 | 9.448 |
| V159N-H | 129.049 | 8.944 | 129.073 | 8.949 | 129.037 | 8.935 | 129.032 | 8.946 |
| Q160N-H | 127.983 | 8.327 | 128.004 | 8.403 | 128.012 | 8.331 | 127.977 | 8.358 |
| G161N-H | 113.016 | 8.533 | 113.118 | 8.498 | 113.07 | 8.501 | 113.565 | 8.562 |
| Q162N-H | 119.241 | 9.076 | 118.868 | 9.065 | 118.838 | 9.036 | 119.126 | 9.071 |
| S163N-H | 115.138 | 8.079 | 115.064 | 8.076 | 115.148 | 8.067 | 115.024 | 8.055 |
| A164N-H | 125.536 | 7.627 | 125.75 | 7.612 | 125.519 | 7.63 | 125.672 | 7.592 |
| K165N-H | 121.283 | 9.18 | 121.319 | 9.174 | 121.748 | 9.207 | 120.498 | 9.28 |
| S166N-H | 110.955 | 7.654 | 110.883 | 7.639 | 110.938 | 7.653 | 111.086 | 7.681 |
| L167N-H | 123.612 | 8.904 | 123.382 | 8.969 | 123.588 | 8.868 | 123.627 | 8.906 |
| D168N-H | 115.426 | 7.961 | 115.438 | 7.979 | 115.537 | 7.948 | 115.152 | 7.842 |
| E169N-H | 119.379 | 7.679 | 119.47 | 7.724 | 119.427 | 7.685 | 119.342 | 7.669 |
| Y170N-H | 121.257 | 7.808 | 121.025 | 7.845 | 121.419 | 7.818 | 121.371 | 7.815 |
| F171N-H | 117.061 | 8.294 | 117.085 | 8.272 | 117.06 | 8.301 | 116.973 | 8.307 |
| D172N-H | 120.008 | 8.301 | 120.023 | 8.228 | 120.099 | 8.255 | 119.45 | 8.284 |
| L173N-H | 124.588 | 8.214 | 124.743 | 8.286 | 124.61 | 8.243 | 124.719 | 8.346 |
| V174N-H | 119.69 | 7.81 | 119.657 | 7.822 | 119.677 | 7.811 | 119.718 | 7.825 |
| N175N-H | 116.325 | 8.332 | 116.389 | 8.326 | 116.352 | 8.312 | 116.327 | 8.345 |
| Y176N-H | 121.777 | 8.325 | 121.809 | 8.353 | 121.868 | 8.341 | 121.825 | 8.345 |
| L177N-H | 121.992 | 8.21 | 121.952 | 8.159 | 122.012 | 8.205 | 122.027 | 8.187 |
| L178N-H | 117.102 | 7.713 | 117.171 | 7.735 | 117.131 | 7.709 | 117.123 | 7.705 |
| T179N-H | 107.906 | 7.522 | 107.85 | 7.515 | 107.89 | 7.526 | 107.887 | 7.525 |
| L180N-H | 124.384 | 7.119 | 124.498 | 7.132 | 124.433 | 7.139 | 124.495 | 7.151 |
| K181N-H | 127.475 | 7.963 | 127.496 | 7.933 | 127.343 | 7.923 | 127.265 | 7.917 |

8.8.2 ^1H , ^{15}N , $^{13}\text{C}_\alpha$ resonance assignments of VcDsbA in complex with the benzimidazole fragment N2 at a concentration of 1 mM recorded from CACB(CO)NH experiment. (Chapter 4)

| | VcDsbA_Apo | | | VcDsbA - N2 | | |
|-----------------|------------|-------------|--------|-------------|-------------|------------|
| Residue | N | C $_\alpha$ | NH | N | C $_\alpha$ | NH |
| K4N-F3CA-K4H | 119.65 | 53.409 | 7.222 | 119.65 | 53.409 | 7.222 |
| E5N-K4CA-E5H | 127.416 | 53.128 | 11.552 | 127.416 | 53.052 | 11.55 2 |
| G6N-E5CA-G6H | 129.486 | 54.99 | 9.008 | 129.509 | 54.99 | 9.014 |
| E7N-G6CA-E7H | 116.759 | 42.961 | 8.88 | 116.759 | 42.961 | 8.88 |
| H8N-E7CA-H8H | 120.771 | 54.977 | 7.472 | 120.771 | 54.977 | 7.472 |
| Y9N-H8CA-Y9H | 114.685 | 56.586 | 8.511 | 114.685 | 56.586 | 8.511 |
| Q10N-Y9CA-Q10H | 110.538 | 53.278 | 7.228 | 110.538 | 53.278 | 7.228 |
| V11N-Q10CA-V11H | 119.488 | 51.03 | 9.19 | 119.488 | 51.03 | 9.19 |
| L12N-V11CA-L12H | 125.886 | 60.349 | 9.39 | 125.886 | 60.349 | 9.39 |
| K13N-L12CA-K13H | 129.221 | 51.338 | 9.326 | 129.221 | 51.338 | 9.326 |
| T14N-K13CA-T14H | 117.414 | 53.751 | 8.421 | 117.414 | 53.751 | 8.421 |
| S17N-A16CA-S17H | 122.283 | 49.169 | 8.462 | 122.246 | 49.169 | 8.466 |
| S23N-V22CA-S23H | 126.667 | 59.147 | 8.882 | 126.674 | 59.147 | 8.891 |
| E24N-S23CA-E24H | 121.2 | 53.153 | 9.315 | 121.2 | 53.163 | 9.315 |
| F25N-E24CA-F25H | 121.995 | 50.313 | 8.468 | 121.978 | 50.289 | 8.442 |
| F26N-F25CA-F26H | 128.162 | 54.76 | 9.723 | 128.236 | 54.729 | 9.687 |
| S27N-F26CA-S27H | 121.95 | 52.338 | 8.996 | 121.908 | 52.322 | 8.972 |
| F28N-F27CA-F28H | 110.408 | 61.727 | 6.647 | 110.44 | 61.727 | 6.653 |
| Y29N-F28CA-Y29H | 132.501 | 60.055 | 11.432 | 132.501 | 59.915 | 11.43 2 |
| C30N-Y29CA-C30H | 115.752 | 57.866 | 8.499 | 115.676 | 57.946 | 8.502 |
| N34N-C33CA-N34H | 116.146 | 61.578 | 8.721 | 116.135 | 61.604 | 8.694 |
| T35N-N34CA-T35H | 120.043 | 53.401 | 7.231 | 120.156 | 53.331 | 7.224 |
| F36N-T35CA-F36H | 112.736 | 62.367 | 7.683 | 112.693 | 62.386 | 7.644 |
| E37N-F36CA-E37H | 119.137 | 54.866 | 7.414 | 119.11 | 54.866 | 7.387 |
| I40N-I39CA-I40H | 118.59 | 61.785 | 7.003 | 118.552 | 61.895 | 6.972 |
| A41N-I40CA-A41H | 121.228 | 58.433 | 7.306 | 121.133 | 58.377 | 7.32 |

Appendices

| | | | | | | |
|-----------------|---------|--------|-------|---------|--------|-------|
| Q42N-A41CA-Q42H | 120.742 | 52.823 | 8.041 | 120.773 | 52.823 | 8.018 |
| L43N-Q42CA-L43H | 116.64 | 56.084 | 7.128 | 116.64 | 56.163 | 7.128 |
| K44N-L43CA-K44H | 122.616 | 55.624 | 8.709 | 122.578 | 55.624 | 8.668 |
| Q45N-K44CA-Q45H | 116.303 | 57.277 | 8.146 | 116.303 | 57.233 | 8.146 |
| Q46N-Q45CA-Q46H | 115.288 | 54.751 | 7.075 | 115.23 | 54.801 | 7.072 |
| L47N-Q46CA-L47H | 114.297 | 51.583 | 7.361 | 114.297 | 51.583 | 7.362 |
| G50N-E49CA-G50H | 121.91 | 55.787 | 8.545 | 121.922 | 55.787 | 8.547 |
| A51N-G50CA-A51H | 112.573 | 42.516 | 8.768 | 112.573 | 42.48 | 8.768 |
| K52N-A51CA-K52H | 122.119 | 48.806 | 7.682 | 122.119 | 48.806 | 7.686 |
| F53N-K52CA-F53H | 121.621 | 52.56 | 7.698 | 121.621 | 52.46 | 7.698 |
| Q54N-F53CA-Q54H | 122.355 | 53.886 | 8.738 | 122.414 | 53.886 | 8.735 |
| K55N-K54CA-K55H | 128.019 | 51.707 | 8.433 | 128.019 | 51.641 | 8.433 |
| S59N-V58CA-S59H | 116.885 | 56.405 | 7.149 | 116.819 | 56.523 | 7.165 |
| F60N-S59CA-F60H | 117.914 | 57.218 | 9.971 | 117.918 | 57.241 | 9.97 |
| M61N-F60CA-M61H | 114.982 | 56.767 | 6.935 | 114.97 | 56.785 | 6.931 |
| G62N-M61CA-G62H | 118.296 | 52.584 | 6.432 | 118.269 | 52.584 | 6.428 |
| G63N-G62CA-G63H | 108.207 | 41.011 | 8.525 | 108.207 | 41.044 | 8.525 |
| G66N-M65CA-G66H | 118.263 | 51.258 | 7.658 | 118.265 | 51.305 | 7.653 |
| Q67N-G66CA-Q67H | 109.802 | 47.431 | 7.732 | 109.802 | 47.348 | 7.73 |
| A68N-Q67CA-A68H | 120.564 | 56.778 | 8.436 | 120.564 | 56.778 | 8.436 |
| M69N-A68CA-M69H | 122.983 | 52.58 | 8.116 | 122.983 | 52.539 | 8.12 |
| S70N-M69CA-S70H | 118.97 | 53.59 | 8.9 | 118.969 | 53.59 | 8.893 |
| K71N-S70CA-K71H | 118.541 | 59.956 | 8.49 | 118.541 | 60.348 | 8.49 |
| A72N-K71CA-A72H | 122.976 | 57.707 | 8.686 | 122.903 | 57.707 | 8.685 |
| T75N-A74CA-T75H | 119.921 | 52.711 | 8.652 | 119.898 | 52.701 | 8.649 |
| M76N-T75CA-M76H | 117.398 | 63.549 | 8.568 | 117.382 | 63.529 | 8.567 |
| I77N-M76CA-I77H | 121.366 | 57.374 | 8.254 | 121.364 | 57.417 | 8.243 |
| A78N-I77CA-A78H | 117.345 | 62.104 | 7.248 | 117.362 | 62.2 | 7.247 |
| L79N-A78CA-L79H | 124.318 | 52.763 | 8.527 | 124.309 | 52.724 | 8.524 |
| E80N-L79CA-E80H | 113.806 | 51.962 | 8.239 | 113.797 | 51.974 | 8.231 |
| V81N-E80CA-V81H | 116.668 | 57.942 | 7.957 | 116.656 | 57.894 | 7.953 |
| E82N-V81CA-E82H | 124.976 | 56.853 | 8.44 | 124.923 | 56.853 | 8.43 |
| D83N-E82CA-D83H | 125.231 | 59.071 | 8.757 | 125.231 | 59.071 | 8.757 |
| K84N-D83CA-K84H | 116.308 | 54.494 | 8.589 | 116.308 | 54.494 | 8.589 |
| M85N-K84CA-M85H | 116.872 | 54.87 | 7.464 | 116.872 | 54.87 | 7.464 |

Appendices

| | | | | | | |
|--------------------|---------|--------|-------|---------|--------|-------|
| V86N-M85CA-V86H | 113.072 | 52.161 | 8.484 | 113.072 | 52.161 | 8.484 |
| V88N-P87CA-V88H | 122.149 | 63.26 | 8.769 | 122.149 | 63.26 | 8.769 |
| M89N-V88CA-M89H | 117.048 | 65.014 | 6.6 | 117.048 | 65.103 | 6.6 |
| F90N-M89CA-F90H | 121.508 | 54.239 | 8.217 | 121.477 | 54.389 | 8.22 |
| N91N-F90CA-N91H | 117.001 | 60.736 | 8.299 | 117.026 | 60.736 | 8.291 |
| R92N-N91CA-R92H | 116.931 | 54.363 | 8.468 | 116.931 | 54.357 | 8.468 |
| I93N-R92CA-I93H | 121.83 | 55.654 | 8.566 | 121.839 | 55.654 | 8.556 |
| H94N-I93CA-H94H | 114.467 | 62.882 | 7.326 | 114.467 | 62.736 | 7.326 |
| T95N-H94CA-T95H | 115.146 | 56.416 | 8.303 | 115.134 | 56.435 | 8.302 |
| L96N-T95CA-L96H | 117.231 | 62.782 | 8.003 | 117.214 | 62.782 | 7.996 |
| R97N-L96CA-R97H | 118.094 | 53.149 | 8.301 | 118.094 | 53.134 | 8.301 |
| K98N-R97CA-K98H | 109.403 | 53.825 | 6.477 | 109.399 | 53.836 | 6.472 |
| D102N-K101CA-D102H | 123.457 | 54.98 | 9.242 | 123.457 | 54.98 | 9.242 |
| E103N-D102CA-E103H | 113.683 | 50.36 | 7.453 | 113.683 | 50.36 | 7.453 |
| Q104N-E103CA-Q104H | 117.17 | 58.21 | 8.974 | 117.17 | 58.21 | 8.974 |
| E105N-Q104CA-E105H | 123.143 | 57.1 | 8.291 | 123.143 | 57.1 | 8.291 |
| L106N-E105CA-L106H | 122.545 | 57.01 | 8.823 | 122.545 | 57.06 | 8.823 |
| R107N-L106CA-R107H | 119.91 | 55.79 | 7.75 | 119.891 | 55.79 | 7.742 |
| Q108N-R107CA-Q108H | 119.089 | 55.66 | 7.57 | 119.089 | 55.75 | 7.57 |
| I109N-Q108CA-I109H | 119.58 | 55.51 | 7.814 | 119.58 | 55.51 | 7.803 |
| F110N-I109CA-F110H | 117.931 | 63.12 | 7.499 | 117.931 | 63.1 | 7.499 |
| L111N-F110CA-L111H | 115.151 | 58.16 | 7.105 | 115.173 | 58.16 | 7.105 |
| D112N-L111CA-D112H | 121.34 | 55.72 | 7.763 | 121.34 | 55.72 | 7.763 |
| E113N-D112CA-E113H | 120.897 | 54.14 | 8.976 | 120.897 | 54.16 | 8.976 |
| G114N-E113CA-G114H | 116.885 | 52.34 | 7.19 | 116.819 | 52.34 | 7.182 |
| I115N-G114CA-I115H | 106.522 | 42.89 | 7.944 | 106.522 | 42.89 | 7.944 |
| D116N-I115CA-D116H | 122.669 | 57.84 | 7.297 | 122.669 | 57.84 | 7.297 |
| A117N-D116CA-A117H | 128.121 | 52.47 | 8.53 | 128.121 | 52.47 | 8.53 |
| A118N-A117CA-A118H | 129.963 | 53.3 | 8.754 | 129.963 | 53.3 | 8.754 |
| K119N-A118CA-K119H | 118.144 | 52.54 | 8.057 | 118.144 | 52.54 | 8.057 |
| F120N-K119CA-F120H | 119.058 | 56.57 | 8.087 | 119.058 | 56.57 | 8.087 |
| D121N-F120CA-D121H | 120.485 | 60.65 | 8.584 | 120.48 | 60.65 | 8.577 |
| A122N-D121CA-A122H | 119.798 | 55.39 | 9.056 | 119.798 | 55.39 | 9.056 |
| A123N-A122CA-A123H | 119.937 | 51.79 | 7.662 | 119.937 | 51.79 | 7.662 |
| Y124N-A123CA-Y124H | 118.839 | 51.66 | 7.381 | 118.839 | 51.66 | 7.381 |

Appendices

| | | | | | | |
|------------------------|---------|-------|--------|---------|-------|------------|
| N125N-Y124CA-N125H | 120.332 | 59.24 | 9.029 | 120.332 | 59.24 | 9.029 |
| G126N-N125CA-G126H | 109.444 | 50.38 | 6.855 | 109.444 | 50.35 | 6.855 |
| F127N-G126CA-F127H | 107.758 | 42.45 | 7.512 | 107.702 | 42.45 | 7.509 |
| A128N-F127CA-A128H | 119.087 | 54.99 | 8.357 | 119.087 | 54.98 | 8.357 |
| V129N-A128CA-V129H | 108.023 | 53.3 | 8.403 | 108.033 | 53.3 | 8.398 |
| D130N-v129CA-D130H | 121.134 | 64.36 | 8.614 | 121.134 | 64.36 | 8.614 |
| S131N-D130CA-S131H | 119.33 | 55.55 | 7.114 | 119.33 | 55.52 | 7.114 |
| M132N-S131CA- M132H | 111.994 | 59.34 | 8.221 | 111.977 | 59.42 | 8.219 |
| V133N-M132CA- V133H | 122.692 | 57.13 | 7.875 | 122.671 | 57.13 | 7.873 |
| R134N-V133CA-R134H | 117.103 | 64.03 | 7.799 | 117.083 | 63.89 | 7.795 |
| R135N-R134CA-R135H | 117.062 | 56.87 | 7.576 | 117.06 | 56.88 | 7.578 |
| F136N-R135CA-F136H | 120.241 | 57.19 | 7.612 | 120.277 | 57.18 | 7.615 |
| D137N-F136CA-D137H | 120.661 | 55.33 | 8.707 | 120.653 | 55.29 | 8.706 |
| K138N-D137CA-K138H | 116.193 | 55.27 | 8.006 | 116.231 | 55.24 | 8.011 |
| Q139N-K138CA-Q139H | 122.725 | 56.88 | 8.483 | 122.723 | 56.88 | 8.481 |
| F140N-Q139CA-F140H | 117.253 | 56.87 | 8.826 | 117.252 | 56.88 | 8.821 |
| Q141N-F140CA-Q141H | 120.447 | 59.26 | 7.978 | 120.388 | 59.31 | 7.98 |
| D142N-Q141CA-D142H | 120.424 | 56.45 | 8.556 | 120.448 | 56.45 | 8.564 |
| S143N-D142CA-S143H | 120.864 | 54.31 | 9.314 | 120.864 | 54.31 | 9.314 |
| G144N-S143CA-G144H | 112.149 | 58.21 | 7.347 | 112.149 | 58.23 | 7.347 |
| L145N-G144CA-L145H | 107.633 | 43.26 | 7.349 | 107.605 | 43.26 | 7.338 |
| T146N-L145CA-T146H | 119.215 | 52.1 | 6.847 | 119.215 | 52.16 | 6.847 |
| G147N-T146CA-G147H | 106.323 | 58.71 | 8.385 | 106.365 | 58.71 | 8.396 |
| V148N-G147CA-V148H | 108.662 | 42.6 | 7.457 | 108.662 | 42.55 | 7.457 |
| A150N-P149CA-A150H | 108.27 | 59.47 | 8.253 | 108.235 | 59.56 | 8.243 |
| V151N-A150CA-V151H | 122.67 | 49.38 | 8.372 | 122.526 | 49.34 | 8.339 |
| V152N-V151CA-V152H | 125.607 | 59.49 | 9.432 | 125.591 | 59.43 | 9.512 |
| V153N-V152CA-V153H | 126.934 | 57.5 | 9.324 | 126.934 | 57.53 | 9.314 |
| N154N-V153CA-N154H | 130.278 | 59.15 | 10.049 | 130.252 | 59.15 | 10.05 9 |
| N155N-N154CA-N155H | 128.297 | 51.86 | 10.348 | 128.297 | 51.86 | 10.34 8 |
| R156N-R155CA-R156H | 113.1 | 52.92 | 9.301 | 113.1 | 52.92 | 9.301 |

Appendices

| | | | | | | |
|--------------------|---------|-------|-------|---------|-------|-------|
| Y157N-R156CA-Y157H | 117.655 | 53.82 | 7.443 | 117.655 | 53.82 | 7.443 |
| L158N-Y157CA-L158H | 117.848 | 53.79 | 8.286 | 117.848 | 53.79 | 8.286 |
| V159N-L158CA-V159H | 127.921 | 51.88 | 9.45 | 127.921 | 51.88 | 9.45 |
| Q160N-V159CA-Q160H | 129.024 | 62.11 | 8.922 | 129.024 | 62.16 | 8.922 |
| G161N-Q160CA-G161H | 127.903 | 51.99 | 8.298 | 127.943 | 52.23 | 8.256 |
| Q162N-G161CA-Q162H | 112.75 | 45.31 | 8.512 | 112.75 | 45.13 | 8.534 |
| S163N-Q162CA-S163H | 119.384 | 54.74 | 9.069 | 119.041 | 54.57 | 8.929 |
| A164N-S163CA-A164H | 115.212 | 56.97 | 8.071 | 115.111 | 56.88 | 8.073 |
| K165N-A164CA-K165H | 125.362 | 49.09 | 7.624 | 125.255 | 49.09 | 7.634 |
| S166N-K165CA-S166H | 122.044 | 54.5 | 9.155 | 121.484 | 54.3 | 9.032 |
| L167N-S166CA-L167H | 111.11 | 53.97 | 7.644 | 111.476 | 53.97 | 7.685 |
| D168N-L167CA-D168H | 123.719 | 55.1 | 8.832 | 123.719 | 55.21 | 8.832 |
| E169N-D168CA-E169H | 115.52 | 54.85 | 7.935 | 115.336 | 54.84 | 7.944 |
| Y170N-E169CA-Y170H | 119.344 | 56.83 | 7.654 | 119.34 | 56.84 | 7.628 |
| F171N-Y170CA-F171H | 121.541 | 59.93 | 7.823 | 121.021 | 60.06 | 7.784 |
| D172N-F171CA-D172H | 116.993 | 55.84 | 8.275 | 117.316 | 55.77 | 8.276 |
| L173N-D172CA-L173H | 120.03 | 54.98 | 8.249 | 120.204 | 55.03 | 8.247 |
| V174N-L173CA-V174H | 124.518 | 55.51 | 8.221 | 124.358 | 55.51 | 8.18 |
| N175N-V174CA-N175H | 119.651 | 65.46 | 7.805 | 119.651 | 65.53 | 7.805 |
| Y176N-N175CA-Y176H | 116.268 | 53.78 | 8.306 | 116.356 | 53.8 | 8.28 |
| L177N-Y176CA-L177H | 121.761 | 57.9 | 8.316 | 121.666 | 57.82 | 8.277 |
| L178N-L177CA-L178H | 121.953 | 54.68 | 8.207 | 122 | 54.7 | 8.277 |
| T179N-L178CA-T179H | 117.067 | 54.63 | 7.71 | 117.019 | 54.58 | 7.742 |
| L180N-T179CA-L180H | 107.949 | 59.42 | 7.508 | 107.989 | 59.41 | 7.517 |
| K181N-L180CA-K181H | 124.418 | 52.76 | 7.114 | 124.424 | 52.76 | 7.122 |

References

1. (a) Walsh, C., Where will new antibiotics come from? *Nat Rev Microbiol* **2003**, *1* (1), 65-70; (b) Heras, B.; Shouldice, S. R.; Totsika, M.; Scanlon, M. J.; Schembri, M. A.; Martin, J. L., DSB proteins and bacterial pathogenicity. *Nat Rev Microbiol* **2009**, *7* (3), 215-25.
2. (a) Levy, S. B.; Marshall, B., Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med* **2004**, *10* (12 Suppl), S122-9; (b) Normark, B. H.; Normark, S., Evolution and spread of antibiotic resistance. *J Intern Med* **2002**, *252* (2), 91-106.
3. (a) Hawser, S. P.; Bouchillon, S. K.; Hoban, D. J.; Badal, R. E.; Hsueh, P. R.; Paterson, D. L., Emergence of high levels of extended-spectrum-beta-lactamase-producing gram-negative bacilli in the Asia-Pacific region: data from the Study for Monitoring Antimicrobial Resistance Trends (SMART) program, 2007. *Antimicrob Agents Chemother* **2009**, *53* (8), 3280-4; (b) Nicasio, A. M.; Kuti, J. L.; Nicolau, D. P., The current state of multidrug-resistant gram-negative bacilli in North America. *Pharmacotherapy* **2008**, *28* (2), 235-49.
4. Vincent, J. L.; Rello, J.; Marshall, J.; Silva, E.; Anzueto, A.; Martin, C. D.; Moreno, R.; Lipman, J.; Gomersall, C.; Sakr, Y.; Reinhart, K., International study of the prevalence and outcomes of infection in intensive care units. *J Am Med Assoc* **2009**, *302* (21), 2323-9.
5. Opal, S. M.; Calandra, T., Antibiotic Usage and Resistance: Gaining or Losing Ground on Infections in Critically Ill Patients? *J Am Med Assoc* **2009**, *302* (21), 2367-2368.
6. Theuretzbacher, U., Future antibiotics scenarios: is the tide starting to turn? *Int J Antimicrob Ag* **2009**, *34* (1), 15-20.
7. The antibacterial lead discovery challenge. *Nat Rev Drug Discov* **2010**, *9* (10), 751-752.
8. (a) Livermore, D. M.; Hope, R.; Brick, G.; Lillie, M.; Reynolds, R., Non-susceptibility trends among Enterobacteriaceae from bacteraemias in the UK and Ireland, 2001-06. *J Antimicrob Chemother* **2008**, *62 Suppl 2*, ii41-54; (b) Mesko Meglic, K.; Koren, S.; Palepou, M. F.; Karisik, E.; Livermore, D. M.; Pike, R.; Andlovic, A.; Jeverica, S.; Krizan-Hergouth, V.; Muller-Premru, M.; Seme, K.; Woodford, N., Nationwide survey of CTX-M-type extended-spectrum beta-lactamases among *Klebsiella pneumoniae* isolates in Slovenian hospitals. *Antimicrob Agents Chemother* **2009**, *53* (1), 287-91.
9. Peek, J. A.; Taylor, R. K., Characterization of a periplasmic thiol:disulfide interchange protein required for the functional maturation of secreted virulence factors of *Vibrio cholerae*. *Proc Natl Acad Sci U S A* **1992**, *89* (13), 6210-4.
10. Stenson, T. H.; Weiss, A. A., DsbA and DsbC are required for secretion of pertussis toxin by *Bordetella pertussis*. *Infect Immun* **2002**, *70* (5), 2297-303.

11. Wulfing, C.; Rappuoli, R., Efficient production of heat-labile enterotoxin mutant proteins by overexpression of *dsbA* in a *degP*-deficient *Escherichia coli* strain. *Arch Microbiol* **1997**, *167* (5), 280-3.
12. Finlay, B. B.; Falkow, S., Common themes in microbial pathogenicity revisited. *Microbiol Mol Biol Rev* **1997**, *61* (2), 136-69.
13. (a) Braun, P.; Ockhuijsen, C.; Eppens, E.; Koster, M.; Bitter, W.; Tommassen, J., Maturation of *Pseudomonas aeruginosa* elastase. Formation of the disulfide bonds. *J Biol Chem* **2001**, *276* (28), 26030-5; (b) Urban, A.; Leipelt, M.; Eggert, T.; Jaeger, K. E., DsbA and DsbC affect extracellular enzyme formation in *Pseudomonas aeruginosa*. *J Bacteriol* **2001**, *183* (2), 587-96.
14. (a) Bardwell, J. C., Building bridges: disulphide bond formation in the cell. *Mol Microbiol* **1994**, *14* (2), 199-205; (b) Raina, S.; Missiakas, D., Making and breaking disulfide bonds. *Annu Rev Microbiol* **1997**, *51*, 179-202.
15. (a) Bardwell, J. C.; McGovern, K.; Beckwith, J., Identification of a protein required for disulfide bond formation in vivo. *Cell* **1991**, *67* (3), 581-9; (b) Missiakas, D.; Georgopoulos, C.; Raina, S., Identification and characterization of the *Escherichia coli* gene *dsbB*, whose product is involved in the formation of disulfide bonds in vivo. *Proc Natl Acad Sci U S A* **1993**, *90* (15), 7084-8.
16. (a) Ha, U. H.; Wang, Y.; Jin, S., DsbA of *Pseudomonas aeruginosa* is essential for multiple virulence factors. *Infect Immun* **2003**, *71* (3), 1590-5; (b) Dacheux, D.; Epaulard, O.; de Groot, A.; Guery, B.; Leberre, R.; Attree, I.; Polack, B.; Toussaint, B., Activation of the *Pseudomonas aeruginosa* type III secretion system requires an intact pyruvate dehydrogenase *aceAB* operon. *Infect Immun* **2002**, *70* (7), 3973-7; (c) Watarai, M.; Tobe, T.; Yoshikawa, M.; Sasakawa, C., Disulfide oxidoreductase activity of *Shigella flexneri* is required for release of Ipa proteins and invasion of epithelial cells. *Proc Natl Acad Sci U S A* **1995**, *92* (11), 4927-4931; (d) Miki, T.; Okada, N.; Danbara, H., Two Periplasmic Disulfide Oxidoreductases, DsbA and SrgA, Target Outer Membrane Protein SpiA, a Component of the *Salmonella* Pathogenicity Island 2 Type III Secretion System. *J Biol Chem* **2004**, *279* (33), 34631-34642.
17. Inaba, K.; Ito, K., Structure and mechanisms of the DsbB-DsbA disulfide bond generation machine. *Biochim Biophys Acta* **2008**, *1783* (4), 520-9.
18. Kadokura, H.; Katzen, F.; Beckwith, J., Protein disulfide bond formation in prokaryotes. *Annu Rev Biochem* **2003**, *72*, 111-35.
19. (a) Dailey, F. E.; Berg, H. C., Mutants in disulfide bond formation that disrupt flagellar assembly in *Escherichia coli*. *Proc Natl Acad Sci U S A* **1993**, *90* (3), 1043-7; (b) Dumoulin, A.; Grauschopf, U.; Bischoff, M.; Thony-Meyer, L.; Berger-Bachi, B., *Staphylococcus aureus* DsbA is a membrane-bound lipoprotein with thiol-disulfide oxidoreductase activity. *Arch Microbiol* **2005**, *184* (2), 117-28; (c) Paxman, J. J.; Borg, N. A.; Horne, J.; Thompson, P. E.; Chin, Y.; Sharma, P.; Simpson, J. S.; Wielens, J.; Piek, S.; Kahler, C. M.; Sakellaris, H.; Pearce, M.; Bottomley, S. P.; Rossjohn, J.; Scanlon, M. J., The structure of the bacterial oxidoreductase enzyme DsbA in complex with a peptide reveals a basis for substrate specificity in the catalytic cycle of DsbA enzymes. *J Biol Chem* **2009**, *284* (26), 17835-45.
20. (a) Martin, Crystal structure of the DsbA protein required for disulphide bond formation. *Nature* **1993**, *365* (6445), 464; (b) Martin, J. L., Thioredoxin--a fold for all reasons. *Structure* **1995**, *3* (3), 245-50.

21. (a) Marra, A., Targeting virulence for antibacterial chemotherapy: identifying and characterising virulence factors for lead discovery. *Drugs R D* **2006**, *7* (1), 1-16; (b) Rasko, D. A.; Sperandio, V., Anti-virulence strategies to combat bacteria-mediated disease. *Nat Rev Drug Discov* **2010**, *9* (2), 117-28.
22. Lee, Y. M.; Almqvist, F.; Hultgren, S. J., Targeting virulence for antimicrobial chemotherapy. *Curr Opin Pharmacol* **2003**, *3* (5), 513-9.
23. (a) Wunderlich, M.; Glockshuber, R., Redox properties of protein disulfide isomerase (DsbA) from *Escherichia coli*. *Protein Sci* **1993**, *2* (5), 717-26; (b) Grauschopf, U.; Winther, J. R.; Korber, P.; Zander, T.; Dallinger, P.; Bardwell, J. C., Why is DsbA such an oxidizing disulfide catalyst? *Cell* **1995**, *83* (6), 947-55.
24. Guddat, L. W.; Bardwell, J. C.; Glockshuber, R.; Huber-Wunderlich, M.; Zander, T.; Martin, J. L., Structural analysis of three His32 mutants of DsbA: support for an electrostatic role of His32 in DsbA stability. *Protein Sci* **1997**, *6* (9), 1893-900.
25. Zapun, A.; Cooper, L.; Creighton, T. E., Replacement of the active-site cysteine residues of DsbA, a protein required for disulfide bond formation in vivo. *Biochemistry* **1994**, *33* (7), 1907-14.
26. Hu, S. H.; Peek, J. A.; Rattigan, E.; Taylor, R. K.; Martin, J. L., Structure of TcpG, the DsbA protein folding catalyst from *Vibrio cholerae*. *J Mol Biol* **1997**, *268* (1), 137-46.
27. Yu, J.; Kroll, J. S., DsbA: a protein-folding catalyst contributing to bacterial virulence. *Microbes Infect* **1999**, *1* (14), 1221-8.
28. Nelson, J. W.; Creighton, T. E., Reactivity and ionization of the active site cysteine residues of DsbA, a protein required for disulfide bond formation in vivo. *Biochemistry* **1994**, *33* (19), 5974-83.
29. (a) Couprie, J.; Vinci, F.; Dugave, C.; Quemeneur, E.; Moutiez, M., Investigation of the DsbA mechanism through the synthesis and analysis of an irreversible enzyme-ligand complex. *Biochemistry* **2000**, *39* (22), 6732-42; (b) Darby, N. J.; Creighton, T. E., Catalytic mechanism of DsbA and its comparison with that of protein disulfide isomerase. *Biochemistry* **1995**, *34* (11), 3576-87.
30. Dartigalongue, C.; Nikaido, H.; Raina, S., Protein folding in the periplasm in the absence of primary oxidant DsbA: modulation of redox potential in periplasmic space via OmpL porin. *EMBO J* **2000**, *19* (22), 5980-8.
31. (a) Fabianek, R. A.; Hennecke, H.; Thony-Meyer, L., Periplasmic protein thiol:disulfide oxidoreductases of *Escherichia coli*. *FEMS Microbiol Rev* **2000**, *24* (3), 303-16; (b) Messens, J.; Collet, J. F., Pathways of disulfide bond formation in *Escherichia coli*. *Int J Biochem Cell Biol* **2006**, *38* (7), 1050-62.
32. (a) Hiniker, A.; Collet, J. F.; Bardwell, J. C., Copper stress causes an in vivo requirement for the *Escherichia coli* disulfide isomerase DsbC. *J Biol Chem* **2005**, *280* (40), 33785-91; (b) Rietsch, A.; Belin, D.; Martin, N.; Beckwith, J., An in vivo pathway for disulfide bond isomerization in *Escherichia coli*. *Proc Natl Acad Sci U S A* **1996**, *93* (23), 13048-53.
33. Hu, S.-H.; Peek, J. A.; Rattigan, E.; Taylor, R. K.; Martin, J. L., Structure of TcpG, the DsbA protein folding catalyst from *Vibrio cholerae*. *J Mol Biol* **1997**, *268*, 137-146.

34. Guddat, L. W.; Bardwell, J. C.; Zander, T.; Martin, J. L., The uncharged surface features surrounding the active site of Escherichia coli DsbA are conserved and are implicated in peptide binding. *Protein Sci* **1997**, *6* (6), 1148-56.
35. Shouldice, S. R.; Heras, B.; Jarrott, R.; Sharma, P.; Scanlon, M. J.; Martin, J. L., Characterization of the DsbA oxidative folding catalyst from Pseudomonas aeruginosa reveals a highly oxidizing protein that binds small molecules. *Antioxid Redox Signal* **2010**, *12* (8), 921-31.
36. Heras, B.; Totsika, M.; Jarrott, R.; Shouldice, S. R.; Guncar, G.; Achard, M. E.; Wells, T. J.; Argente, M. P.; McEwan, A. G.; Schembri, M. A., Structural and functional characterization of three DsbA paralogues from Salmonella enterica serovar typhimurium. *J Biol Chem* **2010**, *285* (24), 18423-32.
37. Vivian, J. P.; Scoullar, J.; Rimmer, K.; Bushell, S. R.; Beddoe, T.; Wilce, M. C.; Byres, E.; Boyle, T. P.; Doak, B.; Simpson, J. S.; Graham, B.; Heras, B.; Kahler, C. M.; Rossjohn, J.; Scanlon, M. J., Structure and function of the oxidoreductase DsbA1 from Neisseria meningitidis. *J Mol Biol* **2009**, *394* (5), 931-43.
38. Palani, K.; Burley, S.K.; Swaminathan, S., Crystal structure of a thiol:disulfide interchange protein dsbA from Bordetella parapertussis. 2009-05-19 ed.; 2009.
39. Kurz, M.; Iturbe-Ormaetxe, I.; Jarrott, R.; Shouldice, S. R.; Wouters, M. A.; Frei, P.; Glockshuber, R.; O'Neill, S. L.; Heras, B.; Martin, J. L., Structural and functional characterization of the oxidoreductase alpha-DsbA1 from Wolbachia pipientis. *Antioxid Redox Signal* **2009**, *11* (7), 1485-500.
40. Heras, B.; Kurz, M.; Jarrott, R.; Shouldice, S. R.; Frei, P.; Robin, G.; Cemazar, M.; Thony-Meyer, L.; Glockshuber, R.; Martin, J. L., Staphylococcus aureus DsbA does not have a destabilizing disulfide. A new paradigm for bacterial oxidative folding. *J Biol Chem* **2008**, *283* (7), 4261-71.
41. Crow, A.; Lewin, A.; Hecht, O.; Carlsson Moller, M.; Moore, G. R.; Hederstedt, L.; Le Brun, N. E., Crystal structure and biophysical properties of Bacillus subtilis BdbD. An oxidizing thiol:disulfide oxidoreductase containing a novel metal site. *J Biol Chem* **2009**, *284* (35), 23719-33.
42. (a) Zhou, Y.; Cierpicki, T.; Jimenez, R. H.; Lukasik, S. M.; Ellena, J. F.; Cafiso, D. S.; Kadokura, H.; Beckwith, J.; Bushweller, J. H., NMR solution structure of the integral membrane enzyme DsbB: functional insights into DsbB-catalyzed disulfide bond formation. *Mol Cell* **2008**, *31* (6), 896-908; (b) Inaba, K.; Murakami, S.; Suzuki, M.; Nakagawa, A.; Yamashita, E.; Okada, K.; Ito, K., Crystal structure of the DsbB-DsbA complex reveals a mechanism of disulfide bond generation. *Cell* **2006**, *127* (4), 789-801.
43. Jander, G.; Martin, N. L.; Beckwith, J., Two cysteines in each periplasmic domain of the membrane protein DsbB are required for its function in protein disulfide bond formation. *EMBO J* **1994**, *13* (21), 5121-7.
44. (a) Bardwell, J. C. A.; Lee, J. O.; Jander, G.; Martin, N.; Belin, D.; Beckwith, J., A pathway for disulfide bond formation *in vivo*. *Proc Natl Acad Sci U S A* **1993**, *90*, 1038-1042; (b) Regeimbal, J.; Bardwell, J. C., DsbB catalyzes disulfide bond formation de novo. *J Biol Chem* **2002**, *277* (36), 32706-13.
45. Dutton, R. J.; Boyd, D.; Berkmen, M.; Beckwith, J., Bacterial species exhibit diversity in their mechanisms and capacity for protein disulfide bond formation. *Proc Natl Acad Sci U S A* **2008**, *105* (33), 11933-8.

46. Rinaldi, F. C.; Meza, A. N.; Guimaraes, B. G., Structural and biochemical characterization of *Xylella fastidiosa* DsbA family members: new insights into the enzyme-substrate interaction. *Biochemistry* **2009**, *48* (15), 3508-18.
47. Heras, B.; Kurz, M.; Shouldice, S. R.; Martin, J. L., The name's bond.....disulfide bond. *Curr Opin Struct Biol* **2007**, *17* (6), 691-8.
48. Quan, S.; Schneider, I.; Pan, J.; Von Hacht, A.; Bardwell, J. C., The CXXC motif is more than a redox rheostat. *J Biol Chem* **2007**, *282* (39), 28823-33.
49. Shouldice, S. R.; Heras, B.; Walden, P. M.; Totsika, M.; Schembri, M. A.; Martin, J. L., Structure and Function of DsbA, a Key Bacterial Oxidative Folding Catalyst. *Antioxid Redox Signal* **2011**.
50. (a) Holmgren, A., Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. *J Biol Chem* **1979**, *254* (19), 9627-9632; (b) Smith, A. M.; Chan, J.; Oksenberg, D.; Urfer, R.; Wexler, D. S.; Ow, A.; Gao, L. P.; McAlorum, A.; Huang, S.-G., A high-throughput turbidometric assay for screening inhibitors of protein disulfide isomerase activity *The Society of Biomolecular Screening* **2004**, *9* (7), 614-620.
51. Ruddock, L. W.; Hirst, T. R.; Freedman, R. B., pH-dependence of the dithiol-oxidizing activity of DsbA (a periplasmic protein dithiol: disulphide oxidoreductase) and protein disulphide-isomerase: studies with a novel simple peptide substrate. *Biochemical Journal* **1996**, *315*, 1001-1005.
52. (a) Westphal, V.; Spetzler, J. C.; Meldal, M.; Christensen, U.; Winther, J. R., Kinetic analysis of the mechanism and specificity of protein-disulfide isomerase using fluorescence-quenched peptides. *J Biol Chem* **1998**, *273* (39), 24992-9; (b) Spetzler, J. C.; Westphal, V.; Winther, J. R.; Meldal, M., Preparation of fluorescence quenched libraries containing interchain disulphide bonds for studies of protein disulphide isomerases. *J Pept Sci* **1998**, *4* (2), 128-37.
53. Christiansen, C.; St Hilaire, P. M.; Winther, J. R., Fluorometric polyethyleneglycol-peptide hybrid substrates for quantitative assay of protein disulfide isomerase. *Anal Biochem* **2004**, *333* (1), 148-55.
54. (a) Stathopoulos, C.; Hendrixson, D. R.; Thanassi, D. G.; Hultgren, S. J.; St Geme, J. W., 3rd; Curtiss, R., 3rd, Secretion of virulence determinants by the general secretory pathway in gram-negative pathogens: an evolving story. *Microbes Infect* **2000**, *2* (9), 1061-72; (b) Thanassi, D. G.; Hultgren, S. J., Multiple pathways allow protein secretion across the bacterial outer membrane. *Curr Opin Cell Biol* **2000**, *12* (4), 420-30; (c) Thanassi, D. G.; Hultgren, S. J., Assembly of complex organelles: pilus biogenesis in gram-negative bacteria as a model system. *Methods* **2000**, *20* (1), 111-26.
55. (a) Sinha, S.; Langford, P. R.; Kroll, J. S., Functional diversity of three different DsbA proteins from *Neisseria meningitidis*. *Microbiology* **2004**, *150*, 2993-3000; (b) Jonda, S.; Huber-Wunderlich, M.; Glockshuber, R.; Mossner, E., Complementation of DsbA deficiency with secreted thioredoxin variants reveals the crucial role of an efficient dithiol oxidant for catalyzed protein folding in the bacterial periplasm. *EMBO J* **1999**, *18* (12), 3271-81.
56. Quan, S.; Schneider, I.; Pan, J.; Hacht, A. V.; Bardwell, J. C. A., The CXXC motif is more than a redox rheostat. *J Bio Chem* **2007**, *282*, 28823-28833.
57. Vallee, B. L.; Ulmer, D. D., Biochemical effects of mercury, cadmium, and lead. *Annu Rev Biochem* **1972**, *41* (10), 91-128.

58. Figueiredo-Pereira, M. E.; Yakushin, S.; Cohen, G., Disruption of the intracellular sulfhydryl homeostasis by cadmium-induced oxidative stress leads to protein thiolation and ubiquitination in neuronal cells. *J Biol Chem* **1998**, *273* (21), 12703-9.
59. Debarbieux, L.; Beckwith, J., The reductive enzyme thioredoxin 1 acts as an oxidant when it is exported to the Escherichia coli periplasm. *Proc Natl Acad Sci U S A* **1998**, *95* (18), 10751-6.
60. (a) Li, C.; Reddy, T. R.; Fischer, P. M.; Dekker, L. V., A Cy5-labeled S100A10 tracer used to identify inhibitors of the protein interaction with annexin A2. *Assay Drug Dev Technol* **2010**, *8* (1), 85-95; (b) Berg, T., Small-molecule inhibitors of protein-protein interactions. *Curr Opin Drug Discov Devel* **2008**, *11* (5), 666-74; (c) Weber, L., Patented inhibitors of p53-Mdm2 interaction (2006 - 2008). *Expert Opin Ther Pat* **2010**, *20* (2), 179-91.
61. (a) Cramer, J. W.; Mattioni, B. E.; Savin, K. A., Strategies for conducting ADME studies during lead generation in the drug discovery process. *IDrugs* **2010**, *13* (12), 857-61; (b) Gao, Q.; Yang, L.; Zhu, Y., Pharmacophore based drug design approach as a practical process in drug discovery. *Curr Comput Aided Drug Des* **2010**, *6* (1), 37-49.
62. Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; Schacht, A. L., How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* **2010**, *9* (3), 203-14.
63. Leeson, P. D.; Springthorpe, B., The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov* **2007**, *6* (11), 881-90.
64. (a) von Nussbaum, F.; Brands, M.; Hinzen, B.; Weigand, S.; Habich, D., Antibacterial natural products in medicinal chemistry--exodus or revival? *Angew Chem Int Ed Engl* **2006**, *45* (31), 5072-129; (b) Feher, M.; Schmidt, J. M., Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inf Comput Sci* **2003**, *43* (1), 218-27.
65. Newman, D. J.; Cragg, G. M., Natural products as sources of new drugs over the last 25 years. *J Nat Prod* **2007**, *70* (3), 461-77.
66. Service, R. F., Combinatorial chemistry hits the drug market. *Science* **1996**, *272* (5266), 1266-8.
67. Xiang, X. D.; Sun, X.; Briceno, G.; Lou, Y.; Wang, K. A.; Chang, H.; Wallace-Freedman, W. G.; Chen, S. W.; Schultz, P. G., A combinatorial approach to materials discovery. *Science* **1995**, *268* (5218), 1738-40.
69. (a) Moos, W. H.; Hurt, C. R.; Morales, G. A., Combinatorial chemistry: oh what a decade or two can do. *Mol Divers* **2009**, *13* (2), 241-5; (b) Breinbauer, R.; Mentel, M., Combinatorial chemistry and the synthesis of compound libraries. *Methods Mol Biol* **2009**, *572*, 73-80.
70. Hann, M. M.; Oprea, T. I., Pursuing the leadlikeness concept in pharmaceutical research. *Curr Opin Chem Biol* **2004**, *8* (3), 255-63.
71. Schonbrun, E.; Abate, A. R.; Steinvurzel, P. E.; Weitz, D. A.; Crozier, K. B., High-throughput fluorescence detection using an integrated zone-plate array. *Lab Chip* **2010**, *10* (7), 852-6.
72. Astle, J. M.; Simpson, L. S.; Huang, Y.; Reddy, M. M.; Wilson, R.; Connell, S.; Wilson, J.; Kodadek, T., Seamless bead to microarray screening: rapid identification of the highest affinity protein ligands from large combinatorial libraries. *Chem Biol* **2010**, *17* (1), 38-45.

73. Rishton, G. M., Nonleadlikeness and leadlikeness in biochemical screening. *Drug Discov Today* **2003**, 8 (2), 86-96.
74. Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J., Recent Developments in Fragment-Based Drug Discovery. *J Med Chem* **2008**, 51 (13), 3661-3680.
75. Hann, M. M.; Leach, A. R.; Harper, G., Molecular complexity and its impact on the probability of finding leads for drug discovery. *J Chem Inf Comput Sci* **2001**, 41 (3), 856-64.
76. (a) Valler, M. J.; Green, D., Diversity screening versus focussed screening in drug discovery. *Drug Discov Today* **2000**, 5 (7), 286-293; (b) New horizons in chemical space. *Nat Rev Drug Discov* **2004**, 3 (5), 375.
77. (a) Bohacek, R. S.; McMartin, C.; Guida, W. C., The art and practice of structure-based drug design: a molecular modeling perspective. *Med Res Rev* **1996**, 16 (1), 3-50; (b) Hajduk, P. J.; Greer, J., A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov* **2007**, 6 (3), 211-9.
78. Erlanson, D. A.; Jahnke, W., The Concept of Fragment-based Drug Discovery. In *Fragment-based Approaches in Drug Discovery*, Jahnke, W.; Erlanson, D. A., Eds. 2006; pp 1-10.
79. Huwe, C. M., Synthetic library design. *Drug Discov Today* **2006**, 11 (15-16), 763-7.
80. (a) Hajduk, P. J.; Greer, J., A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov* **2007**, 6 (3), 211-219; (b) Murray, C. W.; Rees, D. C., The rise of fragment-based drug discovery. *Nature Chem.* **2009**, 1, 187-192; (c) Bembenek, S. D.; Tounge, B. A.; Reynolds, C. H., Ligand efficiency and fragment-based drug discovery. *Drug discovery today* **2009**, 14 (5-6), 278-283; (d) Chessari, G.; Woodhead, A. J., From fragment to clinical candidate-a historical perspective. *Drug discovery today* **2009**, 14 (13-14), 668-675.
81. Jencks, W. P., On the attribution and additivity of binding energies. *Proc Natl Acad Sci U S A* **1981**, 78 (7), 4046-50.
82. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W., Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, 274 (5292), 1531-4.
83. Erlanson, D. A.; McDowell, R. S.; O'Brien, T., Fragment-based drug discovery. *J Med Chem* **2004**, 47 (14), 3463-82.
84. Huth, J. R.; Sun, C.; Sauer, D. R.; Hajduk, P. J., Utilization of NMR-derived fragment leads in drug design. *Methods Enzymol* **2005**, 394, 549-71.
85. Fischer, M.; Hubbard, R. E., Fragment-based ligand discovery. *Mol Interv* **2009**, 9 (1), 22-30.
86. (a) Fejzo, J.; Lepre, C. A.; Peng, J. W.; Bemis, G. W.; Ajay; Murcko, M. A.; Moore, J. M., The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. *Chem Biol* **1999**, 6 (10), 755-69; (b) Hartshorn, M. J.; Murray, C. W.; Cleasby, A.; Frederickson, M.; Tickle, I. J.; Jhoti, H., Fragment-based lead discovery using X-ray crystallography. *J Med Chem* **2005**, 48 (2), 403-13.
87. (a) Baurin, N.; Baker, R.; Richardson, C.; Chen, I.; Foloppe, N.; Potter, A.; Jordan, A.; Roughley, S.; Parratt, M.; Greaney, P.; Morley, D.; Hubbard, R. E., Drug-like annotation and duplicate analysis of a 23-supplier chemical database totalling 2.7 million compounds. *J Chem Inf Comput Sci* **2004**, 44 (2), 643-51; (b) Artis, D. R.; Lin, J. J.; Zhang, C.; Wang, W.; Mehra, U.; Perreault, M.; Erbe, D.; Krupka, H. I.; England, B. P.;

- Arnold, J.; Plotnikov, A. N.; Marimuthu, A.; Nguyen, H.; Will, S.; Signaevsky, M.; Kral, J.; Cantwell, J.; Settachatgull, C.; Yan, D. S.; Fong, D.; Oh, A.; Shi, S.; Womack, P.; Powell, B.; Habets, G.; West, B. L.; Zhang, K. Y.; Milburn, M. V.; Vlasuk, G. P.; Hirth, K. P.; Nolop, K.; Bollag, G.; Ibrahim, P. N.; Tobin, J. F., Scaffold-based discovery of indeglitazar, a PPAR pan-active anti-diabetic agent. *Proc Natl Acad Sci U S A* **2009**, *106* (1), 262-7.
88. Handbook, M. R., Maybridge Collection 45th Anniv. ed 2007.
89. Hajduk, P. J., Fragment-based drug design: how big is too big? *J Med Chem* **2006**, *49* (24), 6972-6.
90. Lipinski, C. A., Drug-like properties and the causes of poor solubility and poor permeability. *J Pharmacol Toxicol Methods* **2000**, *44* (1), 235-49.
91. Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D., Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* **2002**, *45* (12), 2615-23.
92. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H., A 'rule of three' for fragment-based lead discovery? *Drug Discov Today* **2003**, *8* (19), 876-7.
93. Hopkins, A. L.; Groom, C. R.; Alex, A., Ligand efficiency: a useful metric for lead selection. *Drug discovery today* **2004**, *9* (430-431), 430.
94. (a) Reynolds, C. H.; Bembenek, S. D.; Tounge, B. A., The role of molecular size in ligand efficiency. *Bioorg Med Chem Lett* **2007**, *17* (15), 4258-61; (b) Leach, A. R.; Hann, M. M.; Burrows, J. N.; Griffen, E. J., Fragment screening: an introduction. *Mol Biosyst* **2006**, *2* (9), 430-46.
95. (a) Hajduk, P. J.; Huth, J. R.; Tse, C., Predicting protein druggability. *Drug Discov Today* **2005**, *10* (23-24), 1675-82; (b) Hajduk, P. J.; Huth, J. R.; Fesik, S. W., Druggability indices for protein targets derived from NMR-based screening data. *J Med Chem* **2005**, *48* (7), 2518-25.
96. (a) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R., Fragment-based lead discovery. *Nat Rev Drug Discov* **2004**, *3* (8), 660-72; (b) Carr, R. A.; Congreve, M.; Murray, C. W.; Rees, D. C., Fragment-based lead discovery: leads by design. *Drug Discov Today* **2005**, *10* (14), 987-92; (c) Congreve, M.; Murray, C. W.; Blundell, T. L., Structural biology and drug discovery. *Drug Discov Today* **2005**, *10* (13), 895-907.
97. Hajduk, P. J., Fragment-Based Drug Design: How Big Is Too Big? *J Med Chem* **2006**, *49* (24), 6972-6976.
98. Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J., Recent developments in fragment-based drug discovery. *J Med Chem* **2008**, *51* (13), 3661-80.
99. Feyfant, E.; Cross, J. B.; Paris, K.; Tsao, D. H., Fragment-based drug design. *Methods Mol Biol* **2011**, *685*, 241-52.
100. Alex, A. A.; Flocco, M. M., Fragment-based drug discovery: what has it achieved so far? *Curr Top Med Chem* **2007**, *7* (16), 1544-67.
101. Howard, S.; Berdini, V.; Boulstridge, J. A.; Carr, M. G.; Cross, D. M.; Curry, J.; Devine, L. A.; Early, T. R.; Fazal, L.; Gill, A. L.; Heathcote, M.; Maman, S.; Matthews, J. E.; McMenamin, R. L.; Navarro, E. F.; O'Brien, M. A.; O'Reilly, M.; Rees, D. C.; Reule, M.; Tisi, D.; Williams, G.; Vinkovic, M.; Wyatt, P. G., Fragment-based discovery of the pyrazol-4-yl urea (AT9283), a multitargeted kinase inhibitor with potent aurora kinase activity. *J Med Chem* **2009**, *52* (2), 379-88.

102. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W., Discovering high-affinity ligands for proteins: SAR by NMR. (structure-activity relationships; nuclear magnetic resonance). *Science* **1996**, v274 (n5292), p1531(4).
103. Hajduk, P. J., SAR by NMR: putting the pieces together. *Mol Interv* **2006**, 6 (5), 266-72.
104. Abe, A.; Kenny, B.; Stein, M.; Finlay, B. B., Characterization of two virulence proteins secreted by rabbit enteropathogenic *Escherichia coli*, EspA and EspB, whose maximal expression is sensitive to host body temperature. *Infect Immun* **1997**, 65 (9), 3547-55.
105. Erlanson, D. A.; Wells, J. A.; Braisted, A. C., Tethering: fragment-based drug discovery. *Annu Rev Biophys Biomol Struct* **2004**, 33, 199-223.
106. (a) Yang, W.; Fucini, R. V.; Fahr, B. T.; Randal, M.; Lind, K. E.; Lam, M. B.; Lu, W.; Lu, Y.; Cary, D. R.; Romanowski, M. J.; Colussi, D.; Pietrak, B.; Allison, T. J.; Munshi, S. K.; Penny, D. M.; Pham, P.; Sun, J.; Thomas, A. E.; Wilkinson, J. M.; Jacobs, J. W.; McDowell, R. S.; Ballinger, M. D., Fragment-based discovery of nonpeptidic BACE-1 inhibitors using tethering. *Biochemistry* **2009**, 48 (21), 4488-96; (b) Braisted, A. C.; Oslob, J. D.; Delano, W. L.; Hyde, J.; McDowell, R. S.; Waal, N.; Yu, C.; Arkin, M. R.; Raimundo, B. C., Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. *J Am Chem Soc* **2003**, 125 (13), 3714-5; (c) Arkin, M. R.; Randal, M.; DeLano, W. L.; Hyde, J.; Luong, T. N.; Oslob, J. D.; Raphael, D. R.; Taylor, L.; Wang, J.; McDowell, R. S.; Wells, J. A.; Braisted, A. C., Binding of small molecules to an adaptive protein-protein interface. *Proc Natl Acad Sci U S A* **2003**, 100 (4), 1603-8.
107. Jan Kahmann, D. W. a. T. H., The impact of physical based methods screening and their delivery of better quality hits. *Drug Discov. Today Technol.* **2008**, 5 (1), e15-e22.
108. (a) Moy, F. J.; Haraki, K.; Mobilio, D.; Walker, G.; Powers, R.; Tabei, K.; Tong, H.; Siegel, M. M., MS/NMR: a structure-based approach for discovering protein ligands and for drug design by coupling size exclusion chromatography, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Anal Chem* **2001**, 73 (3), 571-81; (b) Ciulli, A.; Williams, G.; Smith, A. G.; Blundell, T. L.; Abell, C., Probing hot spots at protein-ligand binding sites: a fragment-based approach using biophysical methods. *J Med Chem* **2006**, 49 (16), 4992-5000.
109. Dyson, H. J.; Wright, P. E., Equilibrium NMR studies of unfolded and partially folded proteins. *Nat Struct Biol* **1998**, 5 Suppl, 499-503.
110. Zartler, E. R.; Mo, H., Practical aspects of NMR-based fragment discovery. *Curr Top Med Chem* **2007**, 7 (16), 1592-9.
111. (a) Jhoti, H., Fragment-based drug discovery using rational design. *Ernst Schering Found Symp Proc* **2007**, (3), 169-85; (b) Jhoti, H.; Cleasby, A.; Verdonk, M.; Williams, G., Fragment-based screening using X-ray crystallography and NMR spectroscopy. *Curr Opin Chem Biol* **2007**, 11 (5), 485-93.
112. Huth, J. R.; Mendoza, R.; Olejniczak, E. T.; Johnson, R. W.; Cothron, D. A.; Liu, Y.; Lerner, C. G.; Chen, J.; Hajduk, P. J., ALARM NMR: a rapid and robust experimental method to detect reactive false positives in biochemical screens. *J Am Chem Soc* **2005**, 127 (1), 217-24.
113. Meyer, B.; Peters, T., NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Ed Engl* **2003**, 42 (8), 864-90.

114. Gronquist, M.; Meinwald, J.; Eisner, T.; Schroeder, F. C., Exploring uncharted terrain in nature's structure space using capillary NMR spectroscopy: 13 steroids from 50 fireflies. *J Am Chem Soc* **2005**, *127* (31), 10810-1.
115. (a) Jahnke, W.; Rudisser, S.; Zurini, M., Spin label enhanced NMR screening. *J Am Chem Soc* **2001**, *123* (13), 3149-50; (b) McCoy, M. A.; Senior, M. M.; Gesell, J. J.; Ramanathan, L.; Wyss, D. F., Solution structure and dynamics of the single-chain hepatitis C virus NS3 protease NS4A cofactor complex. *J Mol Biol* **2001**, *305* (5), 1099-110.
116. Leach, A. R.; Hann, M. M.; Burrows, J. N.; Griffen, E. J., Fragment screening: an introduction. *Molecular Biosystem* **2006**, *2*, 429-446.
117. (a) Pellecchia, M., Solution nuclear magnetic resonance spectroscopy techniques for probing intermolecular interactions. *Chem Biol* **2005**, *12* (9), 961-71; (b) Pellecchia, M.; Sem, D. S.; Wuthrich, K., NMR in drug discovery. *Nat Rev Drug Discov* **2002**, *1* (3), 211-9.
118. Lepre, C. A.; Moore, J. M.; Peng, J. W., Theory and applications of NMR-based screening in pharmaceutical research. *Chem Rev* **2004**, *104* (8), 3641-76.
119. Klages, J.; Coles, M.; Kessler, H., NMR-based screening: a powerful tool in fragment-based drug discovery. *Analyst* **2007**, *132* (7), 693-705.
120. (a) Murray, C. W.; Blundell, T. L., Structural biology in fragment-based drug design. *Curr Opin Struct Biol* **2010**, *20* (4), 497-507; (b) Sun, C.; Hajduk, P. J., Nuclear magnetic resonance in target profiling and compound file enhancement. *Curr Opin Drug Discov Devel* **2006**, *9* (4), 463-70.
121. Dalvit, C.; Fasolini, M.; Flocco, M.; Knapp, S.; Pevarello, P.; Veronesi, M., NMR-Based screening with competition water-ligand observed via gradient spectroscopy experiments: detection of high-affinity ligands. *J Med Chem* **2002**, *45* (12), 2610-4.
122. Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B., WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J Biomol NMR* **2001**, *21* (4), 349-59.
123. Ayed, A.; Mulder, F. A.; Yi, G. S.; Lu, Y.; Kay, L. E.; Arrowsmith, C. H., Latent and active p53 are identical in conformation. *Nat Struct Biol* **2001**, *8* (9), 756-60.
124. (a) Derrick, T. S.; McCord, E. F.; Larive, C. K., Analysis of protein/ligand interactions with NMR diffusion measurements: the importance of eliminating the protein background. *J Magn Reson* **2002**, *155* (2), 217-25; (b) Wishart, D., NMR spectroscopy and protein structure determination: applications to drug discovery and development. *Curr Pharm Biotechnol* **2005**, *6* (2), 105-20.
125. (a) Pellecchia, M.; Meininger, D.; Dong, Q.; Chang, E.; Jack, R.; Sem, D. S., NMR-based structural characterization of large protein-ligand interactions. *J Biomol NMR* **2002**, *22* (2), 165-73; (b) Sem, D. S.; Pellecchia, M., NMR in the acceleration of drug discovery. *Curr Opin Drug Discov Devel* **2001**, *4* (4), 479-92.
126. (a) Mattos, C.; Ringe, D., Locating and characterizing binding sites on proteins. *Nat Biotechnol* **1996**, *14* (5), 595-9; (b) English, A. C.; Groom, C. R.; Hubbard, R. E., Experimental and computational mapping of the binding surface of a crystalline protein. *Protein Eng* **2001**, *14* (1), 47-59.
127. Blundell, T. L.; Jhoti, H.; Abell, C., High-throughput crystallography for lead discovery in drug design. *Nat Rev Drug Discov* **2002**, *1* (1), 45-54.

128. Ochi, T.; Bolanos-Garcia, V. M.; Stojanoff, V.; Moreno, A., Perspectives on protein crystallisation. *Prog Biophys Mol Biol* **2009**, *101* (1-3), 56-63.
129. Blundell, T. L.; Patel, S., High-throughput X-ray crystallography for drug discovery. *Curr Opin Pharmacol* **2004**, *4* (5), 490-6.
130. Zartler, E. R. a. S., M.J., , *Fragment-Based Drug Discovery*. John Wiley & Sons, Ltd.: 2008.
131. Jahnke, W. a. E., D.A., , *Fragment-Based Approaches in Drug Discovery*. Wiley-VCH: **2006**.
132. Kahmann, J., The impact of physical based method s screening and their delivery of better quality hits. *Drug Discov. Today Technol.* **2008**, 15-22.
133. Giannetti, A. M.; Koch, B. D.; Browner, M. F., Surface plasmon resonance based assay for the detection and characterization of promiscuous inhibitors. *J Med Chem* **2008**, *51* (3), 574-80.
134. Neumann, T.; Junker, H. D.; Schmidt, K.; Sekul, R., SPR-based fragment screening: advantages and applications. *Curr Top Med Chem* **2007**, *7* (16), 1630-42.
135. Abad, M. C.; Askari, H.; O'Neill, J.; Klinger, A. L.; Milligan, C.; Lewandowski, F.; Springer, B.; Spurlino, J.; Rentzeperis, D., Structural determination of estrogen-related receptor gamma in the presence of phenol derivative compounds. *J Steroid Biochem Mol Biol* **2008**, *108* (1-2), 44-54.
136. Lo, M. C.; Aulabaugh, A.; Jin, G.; Cowling, R.; Bard, J.; Malamas, M.; Ellestad, G., Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* **2004**, *332* (1), 153-9.
137. Matulis, D.; Kranz, J. K.; Salemm, F. R.; Todd, M. J., Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. *Biochemistry* **2005**, *44* (13), 5258-66.
138. Freire, E., A thermodynamic approach to the affinity optimization of drug candidates. *Chem Biol Drug Des* **2009**, *74* (5), 468-72.
139. Falconer, R. J.; Collins, B. M., Survey of the year 2009: applications of isothermal titration calorimetry. *J Mol Recognit* **2011**, *24* (1), 1-16.
140. Nielsen, A. L.; Jorgensen, F. S.; Olsen, L.; Christensen, S. F.; Benie, A. J.; Bjornholm, T.; St Hilaire, P. M., A diversity optimized combinatorial library for the identification of Fc-fragment binding ligands. *Biopolymers* **2010**, *94* (2), 192-205.
141. Chung, C. T.; Miller, R. H., A rapid and convenient method for the preparation and storage of competent bacterial cells. *Nucleic Acids Res* **1988**, *16* (8), 3580.
142. Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227* (5259), 680-5.
143. (a) Shevchenko, A.; Wilm, M.; Vorm, O.; Jensen, O. N.; Podtelejnikov, A. V.; Neubauer, G.; Mortensen, P.; Mann, M., A strategy for identifying gel-separated proteins in sequence databases by MS alone. *Biochem Soc Trans* **1996**, *24* (3), 893-6; (b) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M., Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* **1996**, *68* (5), 850-8.
144. Studier, F. W., Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **2005**, *41* (1), 207-34.
145. (a) Cerny, G.; Teuber, M., Differential release of periplasmic versus cytoplasmic enzymes from Escherichia coli B by polymixin B. *Arch Mikrobiol* **1971**, *78* (2), 166-79; (b) Dixon, R. A.; Chopra, I., Leakage of periplasmic proteins from Escherichia coli

- mediated by polymyxin B nonapeptide. *Antimicrob Agents Chemother* **1986**, 29 (5), 781-8.
146. Marley, J.; Lu, M.; Bracken, C., A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* **2001**, 20 (1), 71-5.
147. Santos, M. J.; Teixeira, J. A.; Rodrigues, L. R., Fractionation and recovery of whey proteins by hydrophobic interaction chromatography. *J Chromatogr B Analyt Technol Biomed Life Sci* **2011**.
148. Senczuk, A. M.; Klinke, R.; Arakawa, T.; Vedantham, G.; Yigzaw, Y., Hydrophobic interaction chromatography in dual salt system increases protein binding capacity. *Biotechnol Bioeng* **2009**, 103 (5), 930-5.
149. Pan, D.; Hill, A. P.; Kashou, A.; Wilson, K. A.; Tan-Wilson, A., Electrophoretic transfer protein zymography. *Anal Biochem* **2011**.
150. Yigzaw, Y.; Hinckley, P.; Hewig, A.; Vedantham, G., Ion exchange chromatography of proteins and clearance of aggregates. *Curr Pharm Biotechnol* **2009**, 10 (4), 421-6.
151. Nettleship, J. E.; Brown, J.; Groves, M. R.; Geerlof, A., Methods for protein characterization by mass spectrometry, thermal shift (ThermoFluor) assay, and multiangle or static light scattering. *Methods Mol Biol* **2008**, 426, 299-318.
152. Brickman, E.; Beckwith, J., Analysis of the regulation of Escherichia coli alkaline phosphatase synthesis using deletions and phi80 transducing phages. *J Mol Biol* **1975**, 96 (2), 307-16.
153. Smith, A. M.; Chan, J.; Oksenberg, D.; Urfer, R.; Wexler, D. S.; Ow, A.; Gao, L.; McAlorum, A.; Huang, S. G., A high-throughput turbidometric assay for screening inhibitors of protein disulfide isomerase activity. *J Biomol Screen* **2004**, 9 (7), 614-20.
154. Sinha, S.; Langford, P. R.; Kroll, J. S., Functional diversity of three different DsbA proteins from Neisseria meningitidis. *Microbiology* **2004**, 150 (Pt 9), 2993-3000.
155. Motulsky, H. J. a. A. C. *Fitting models to biological data using linear and nonlinear regression. A practical guide to curve fitting.*, 5.0 GraphPad Software Inc., : 2007.
156. Grzesiek, S.; Bax, A., Amino acid type determination in the sequential assignment procedure of uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched proteins. *J Biomol NMR* **1993**, 3 (2), 185-204.
157. Craik, D. J., Horne, J. and Scanlon, M. J, *Multidimensional Nuclear Magnetic Resonance of Biomolecules*. **2008**.
158. Grzesiek, S., Bax, A., Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson* **1992**, 96, 432-440.
159. Wittekind, M., Mueller, L. , HNCACB, a high sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha and beta carbon resonances in proteins. *J Magn Reson* **1993**, 101, 201-205.
160. Clubb, R. T.; Thanabal, V.; Wagner, G., A Constant-Time Three-Dimensional Triple-Resonance Pulse Scheme to Correlate Intraresidue $^1\text{H}^{\text{N}}$, ^{15}N , and ^{13}C Chemical Shifts in ^{15}N - ^{13}C -Labeled Proteins. *J Magn. Reson.* **1992**, 97, 213-217.
161. Sun, Z. Y.; Hyberts, S. G.; Rovnyak, D.; Park, S.; Stern, A. S.; Hoch, J. C.; Wagner, G., High-resolution aliphatic side-chain assignments in 3D HCcoNH experiments with joint H-C evolution and non-uniform sampling. *J Biomol NMR* **2005**, 32 (1), 55-60.

162. Kover, K. E.; Groves, P.; Jimenez-Barbero, J.; Batta, G., Molecular recognition and screening using a ^{15}N group selective STD NMR method. *J Am Chem Soc* **2007**, *129* (37), 11579-82.
163. Groves, P.; Kover, K. E.; Andre, S.; Bendorowicz-Pikula, J.; Batta, G.; Bruix, M.; Buchet, R.; Canales, A.; Canada, F. J.; Gabius, H. J.; Laurents, D. V.; Naranjo, J. R.; Palczewska, M.; Pikula, S.; Rial, E.; Strzelecka-Kiliszek, A.; Jimenez-Barbero, J., Temperature dependence of ligand-protein complex formation as reflected by saturation transfer difference NMR experiments. *Magn Reson Chem* **2007**, *45* (9), 745-8.
164. Piotto, M.; Saudek, V.; Sklenar, V., Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR* **1992**, *2* (6), 661-5.
165. (a) Grzesiek, S.; Bax, A., Spin-locked multiple quantum coherence for signal enhancement in heteronuclear multidimensional NMR experiments. *J Biomol NMR* **1995**, *6* (3), 335-9; (b) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A., NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **1995**, *6* (3), 277-93.
166. Goddard TD, K. D. *SPARKY 3* 2006.
167. Hohwy, M.; Spadola, L.; Lundquist, B.; Hawtin, P.; Dahmen, J.; Groth-Clausen, I.; Nilsson, E.; Persdotter, S.; Von Wachenfeldt, K.; Folmer, R. H. A.; Edman, K., Novel prostaglandin D synthase inhibitors generated by fragment-based drug design. *J Med Chem* **2008**, *51* (7), 2178-2186.
168. DeLano, W. L. *The PyMOL Molecular Graphics System*, Version 1.2r3pre; Schrödinger, LLC.: 2002.
169. (a) Huang, D.; Caflisch, A., Library screening by fragment-based docking. *J Mol Recognit* **2010**, *23* (2), 183-93; (b) Abad-Zapatero, C.; Metz, J. T., Ligand efficiency indices as guideposts for drug discovery. *Drug Discov Today* **2005**, *10* (7), 464-9; (c) Barelier, S.; Pons, J.; Gehring, K.; Lancelin, J.-M.; Krimm, I., Ligand Specificity in Fragment-Based Drug Design. *J Med Chem* **2010**, *53* (14), 5256-5266.
170. Clore, G. M.; Bax, A.; Driscoll, P. C.; Wingfield, P. T.; Gronenborn, A. M., Assignment of the side-chain ^1H and ^{13}C resonances of interleukin-1 beta using double- and triple-resonance heteronuclear three-dimensional NMR spectroscopy. *Biochemistry* **1990**, *29* (35), 8172-84.
171. Horne, J.; Scanlon, M. J., Backbone and side chain ^1H , ^{15}N and ^{13}C assignments for the reduced form of the oxidoreductase protein DsbA from *Vibrio cholerae*. *Biomol NMR Assign* **2007**, *1* (1), 75-6.
172. Olejniczak, E. T.; Xu, R. X.; Fesik, S. W., A 4D HCCH-TOCSY experiment for assigning the side chain ^1H and ^{13}C resonances of proteins. *J Biomol NMR* **1992**, *2* (6), 655-9.
173. Carvalho, A. L.; Trincão, J.; Romão, M. J., X-ray crystallography in drug discovery. *Methods Mol Biol* **2009**, *572*, 31-56.
174. Jain, D.; Lamour, V., Computational tools in protein crystallography. *Methods Mol Biol* **2010**, *673*, 129-56.
175. (a) Potterton, L.; McNicholas, S.; Krissinel, E.; Gruber, J.; Cowtan, K.; Emsley, P.; Murshudov, G. N.; Cohen, S.; Perrakis, A.; Noble, M., Developments in the CCP4 molecular-graphics project. *Acta Crystallogr D Biol Crystallogr* **2004**, *60* (Pt 12 Pt 1), 2288-94; (b) Krissinel, E. B.; Winn, M. D.; Ballard, C. C.; Ashton, A. W.; Patel, P.; Potterton, E. A.; McNicholas, S. J.; Cowtan, K. D.; Emsley, P., The new CCP4

- Coordinate Library as a toolkit for the design of coordinate-related applications in protein crystallography. *Acta Crystallogr D Biol Crystallogr* **2004**, 60 (Pt 12 Pt 1), 2250-5.
176. Minor, W.; Tomchick, D.; Otwinowski, Z., Strategies for macromolecular synchrotron crystallography. *Structure* **2000**, 8 (5), R105-10.
177. (a) Kudlicki, A.; Rowicka, M.; Otwinowski, Z., The crystallographic fast Fourier transform. Recursive symmetry reduction. *Acta Crystallogr A* **2007**, 63 (Pt 6), 465-80; (b) W., O. Z. a. M., *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, C.W. Carter, Jr. & R. M. Sweet, Eds.,: **1997**; Vol. 276.
178. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C., PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **2002**, 58 (Pt 11), 1948-54.
179. (a) Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004**, 60 (Pt 12 Pt 1), 2126-32; (b) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallographica Section D* **2010**, 66 (4), 486-501.
180. Rowicka, M.; Kudlicki, A.; Zelinka, J.; Otwinowski, Z., Coordinate transformations in modern crystallographic computing. *Acta Crystallogr A* **2004**, 60 (Pt 6), 542-9.
181. Kirchmair, J.; Markt, P.; Distinto, S.; Wolber, G.; Langer, T., Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection--what can we learn from earlier mistakes? *J Comput Aided Mol Des* **2008**, 22 (3-4), 213-28.
182. (a) *Protein and Ligand Preparation Wizard*, version 2.2; Schrödinger, LLC: New York, NY., 2008; (b) Huuskonen, J.; Livingstone, D. J.; Manallack, D. T., Prediction of drug solubility from molecular structure using a drug-like training set. *SAR QSAR Environ Res* **2008**, 19 (3-4), 191-212.
183. *SiteMap*, version 2.2; **2008**.Schrödinger, LLC: New York, NY.,
184. (a) Kawatkar, S.; Wang, H.; Czerminski, R.; Joseph-McCarthy, D., Virtual fragment screening: an exploration of various docking and scoring protocols for fragments using Glide. *J Comput Aided Mol Des* **2009**; (b) Repasky, M. P.; Shelley, M.; Friesner, R. A., Flexible ligand docking with Glide. *Curr Protoc Bioinformatics* **2007**, Chapter 8, Unit 8 12; (c) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* **2004**, 47 (7), 1750-9.
185. (a) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem* **2006**, 49 (21), 6177-96; (b) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* **2004**, 47 (7), 1739-49.
186. Pearlman, D. A.; Charifson, P. S., Improved scoring of ligand-protein interactions using OWFEG free energy grids. *J Med Chem* **2001**, 44 (4), 502-11.

187. Stark, J.; Powers, R., Rapid protein-ligand costructures using chemical shift perturbations. *J Am Chem Soc* **2008**, *130* (2), 535-45.
188. Sandor, M.; Kiss, R.; Keseru, G. M., Virtual fragment docking by Glide: a validation study on 190 protein-fragment complexes. *J Chem Inf Model* **2010**, *50* (6), 1165-72.
189. Goodford, P. J., A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem* **1985**, *28* (7), 849-57.
190. Halgren, T., New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des* **2007**, *69* (2), 146-8.
191. Schrödinger *Maestro*, version 8.5, , LLC, : New York, NY, 2009.
192. 'Instant JChem was used for structure database management, s. a. p. *Instant JChem*, 5.4.1; ChemAxon Budapest, 1037 Hungary, 2009
193. Cummings, M. D.; DesJarlais, R. L.; Gibbs, A. C.; Mohan, V.; Jaeger, E. P., Comparison of automated docking programs as virtual screening tools. *J Med Chem* **2005**, *48* (4), 962-76.
194. Ehresmann, B.; Martin, B.; Horn, A. H.; Clark, T., Local molecular properties and their use in predicting reactivity. *J Mol Model* **2003**, *9* (5), 342-7.
195. Manallack, D. T., The use of local surface properties for molecular superimposition. *J Mol Model* **2008**, *14* (9), 797-805.
196. Yu, J.; Webb, H.; Hirst, T. R., A homologue of the Escherichia coli DsbA protein involved in disulphide bond formation is required for enterotoxin biogenesis in Vibrio cholerae. *Mol Microbiol* **1992**, *6* (14), 1949-58.
197. Yu, J.; McLaughlin, S.; Freedman, R. B.; Hirst, T. R., Cloning and active site mutagenesis of Vibrio cholerae DsbA, a periplasmic enzyme that catalyzes disulfide bond formation. *J Biol Chem* **1993**, *268* (6), 4326-30.
198. Vajda, S.; Guarnieri, F., Characterization of protein-ligand interaction sites using experimental and computational methods. *Curr Opin Drug Discov Devel* **2006**, *9* (3), 354-62.
199. Moriaud, F.; Doppelt-Azeroual, O.; Martin, L.; Oguievetskaia, K.; Koch, K.; Vorotyntsev, A.; Adcock, S. A.; Delfaud, F., Computational Fragment-Based Approach at PDB Scale by Protein Local Similarity. *J Chem Inf Model* **2009**.
200. (a) Schneider, G.; Fechner, U., Computer-based de novo design of drug-like molecules. *Nat Rev Drug Discov* **2005**, *4* (8), 649-63; (b) Bohm, H. J., A novel computational tool for automated structure-based drug design. *J Mol Recognit* **1993**, *6* (3), 131-7.
201. Mohan, V.; Gibbs, A. C.; Cummings, M. D.; Jaeger, E. P.; DesJarlais, R. L., Docking: successes and challenges. *Curr Pharm Des* **2005**, *11* (3), 323-33.
202. (a) Jambon, M.; Imberty, A.; Deleage, G.; Geourjon, C., A new bioinformatic approach to detect common 3D sites in protein structures. *Proteins* **2003**, *52* (2), 137-45; (b) Ramensky, V.; Sobol, A.; Zaitseva, N.; Rubinov, A.; Zosimov, V., A novel approach to local similarity of protein binding sites substantially improves computational drug design results. *Proteins* **2007**, *69* (2), 349-57.
203. (a) Landon, M. R.; Lieberman, R. L.; Hoang, Q. Q.; Ju, S.; Caaveiro, J. M.; Orwig, S. D.; Kozakov, D.; Brenke, R.; Chuang, G. Y.; Beglov, D.; Vajda, S.; Petsko, G. A.; Ringe, D., Detection of ligand binding hot spots on protein surfaces via fragment-

- based methods: application to DJ-1 and glucocerebrosidase. *J Comput Aided Mol Des* **2009**; (b) Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S., Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics* **2009**, *25* (5), 621-7; (c) Landon, M. R.; Lancia, D. R., Jr.; Yu, J.; Thiel, S. C.; Vajda, S., Identification of hot spots within druggable binding regions by computational solvent mapping of proteins. *J Med Chem* **2007**, *50* (6), 1231-40.
204. (a) Miranker, A.; Karplus, M., Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins* **1991**, *11* (1), 29-34; (b) Stultz, C. M.; Karplus, M., MCSS functionality maps for a flexible protein. *Proteins* **1999**, *37* (4), 512-29.
205. Clark, D. E.; Pickett, S. D., Computational methods for the prediction of 'drug-likeness'. *Drug Discov Today* **2000**, *5* (2), 49-58.
206. Stultz, C. M. a. K., M., *Fragment Docking to Proteins with the Multi-copy Simultaneous Search Methodology*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, : **2006**; p 125-146.
207. Martin, J. L.; Bardwell, J. C.; Kuriyan, J., Crystal structure of the DsbA protein required for disulphide bond formation in vivo. *Nature* **1993**, *365* (6445), 464-8.
208. Horne, J.; d'Auvergne, E. J.; Coles, M.; Velkov, T.; Chin, Y.; Charman, W. N.; Pranker, R.; Gooley, P. R.; Scanlon, M. J., Probing the flexibility of the DsbA oxidoreductase from *Vibrio cholerae*--a 15N - 1H heteronuclear NMR relaxation analysis of oxidized and reduced forms of DsbA. *J Mol Biol* **2007**, *371* (3), 703-16.
209. Guddat, L. W.; Bardwell, J. C.; Martin, J. L., Crystal structures of reduced and oxidized DsbA: investigation of domain motion and thiolate stabilization. *Structure* **1998**, *6* (6), 757-67.
210. Chen, J.; Zhang, Z.; Stebbins, J. L.; Zhang, X.; Hoffman, R.; Moore, A.; Pellecchia, M., A fragment-based approach for the discovery of isoform-specific p38alpha inhibitors. *ACS Chem Biol* **2007**, *2* (5), 329-36.
211. (a) Kortvelyesi, T.; Dennis, S.; Silberstein, M.; Brown, L., 3rd; Vajda, S., Algorithms for computational solvent mapping of proteins. *Proteins* **2003**, *51* (3), 340-51; (b) O'Boyle, N. M.; Brewerton, S. C.; Taylor, R., Using buriedness to improve discrimination between actives and inactives in docking. *J Chem Inf Model* **2008**, *48* (6), 1269-78.
212. Silberstein, M.; Dennis, S.; Brown, L.; Kortvelyesi, T.; Clodfelter, K.; Vajda, S., Identification of substrate binding sites in enzymes by computational solvent mapping. *J Mol Biol* **2003**, *332* (5), 1095-113.
213. (a) Liebeschuetz, J. W.; Jones, S. D.; Morgan, P. J.; Murray, C. W.; Rimmer, A. D.; Roscoe, J. M.; Waszkowycz, B.; Welsh, P. M.; Wylie, W. A.; Young, S. C.; Martin, H.; Mahler, J.; Brady, L.; Wilkinson, K., PRO_SELECT: combining structure-based drug design and array-based chemistry for rapid lead discovery. 2. The development of a series of highly potent and selective factor Xa inhibitors. *J Med Chem* **2002**, *45* (6), 1221-32; (b) Kick, E. K.; Roe, D. C.; Skillman, A. G.; Liu, G.; Ewing, T. J.; Sun, Y.; Kuntz, I. D.; Ellman, J. A., Structure-based design and combinatorial chemistry yield low nanomolar inhibitors of cathepsin D. *Chem Biol* **1997**, *4* (4), 297-307.
214. Constantine, K. L.; Davis, M. E.; Metzler, W. J.; Mueller, L.; Claus, B. L., Protein-ligand NOE matching: a high-throughput method for binding pose evaluation that

- does not require protein NMR resonance assignments. *J Am Chem Soc* **2006**, *128* (22), 7252-63.
215. (a) Bertini, I.; Fragai, M.; Giachetti, A.; Luchinat, C.; Maletta, M.; Parigi, G.; Yeo, K. J., Combining in silico tools and NMR data to validate protein-ligand structural models: application to matrix metalloproteinases. *J Med Chem* **2005**, *48* (24), 7544-59; (b) Constantine, K. L.; Mueller, L.; Metzler, W. J.; McDonnell, P. A.; Todderud, G.; Goldfarb, V.; Fan, Y.; Newitt, J. A.; Kiefer, S. E.; Gao, M.; Tortolani, D.; Vaccaro, W.; Tokarski, J., Multiple and single binding modes of fragment-like kinase inhibitors revealed by molecular modeling, residue type-selective protonation, and nuclear overhauser effects. *J Med Chem* **2008**, *51* (19), 6225-9.
216. (a) Rognan, D.; Mukhija, S.; Folkers, G.; Zerbe, O., NMR-restrained docking of a peptidic inhibitor to the N-terminal domain of the phosphoenolpyruvate:sugar phosphotransferase enzyme I. *J Comput Aided Mol Des* **2001**, *15* (2), 103-15; (b) Polshakov, V. I.; Morgan, W. D.; Birdsall, B.; Feeney, J., Validation of a new restraint docking method for solution structure determinations of protein-ligand complexes. *J Biomol NMR* **1999**, *14* (2), 115-22.
217. (a) Ivetac, A.; McCammon, J. A., Mapping the druggable allosteric space of G-protein coupled receptors: a fragment-based molecular dynamics approach. *Chem Biol Drug Des* **2010**, *76* (3), 201-17; (b) Lichtarge, O.; Bourne, H. R.; Cohen, F. E., An evolutionary trace method defines binding surfaces common to protein families. *J Mol Biol* **1996**, *257* (2), 342-58.
218. Martinez, R. M.; Jude, B. A.; Kirn, T. J.; Skorupski, K.; Taylor, R. K., Role of FlgT in anchoring the flagellum of *Vibrio cholerae*. *J Bacteriol* **2010**, *192* (8), 2085-92.
219. Sun, D.; Lafferty, M. J.; Peek, J. A.; Taylor, R. K., Domains within the *Vibrio cholerae* toxin coregulated pilin subunit that mediate bacterial colonization. *Gene* **1997**, *192* (1), 79-85.
220. Hopkins, A. L.; Groom, C. R.; Alex, A., Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* **2004**, *9* (10), 430-1.
221. Edwards, P. D.; Albert, J. S.; Sylvester, M.; Aharony, D.; Andisik, D.; Callaghan, O.; Campbell, J. B.; Carr, R. A.; Chessari, G.; Congreve, M.; Frederickson, M.; Folmer, R. H.; Geschwindner, S.; Koether, G.; Kolmodin, K.; Krumrine, J.; Mauger, R. C.; Murray, C. W.; Olsson, L. L.; Patel, S.; Spear, N.; Tian, G., Application of fragment-based lead generation to the discovery of novel, cyclic amidine beta-secretase inhibitors with nanomolar potency, cellular activity, and high ligand efficiency. *J Med Chem* **2007**, *50* (24), 5912-25.
222. Mpamhanga, C. P.; Spinks, D.; Tulloch, L. B.; Shanks, E. J.; Robinson, D. A.; Collie, I. T.; Fairlamb, A. H.; Wyatt, P. G.; Frearson, J. A.; Hunter, W. N.; Gilbert, I. H.; Brenk, R., One scaffold, three binding modes: novel and selective pteridine reductase 1 inhibitors derived from fragment hits discovered by virtual screening. *J Med Chem* **2009**, *52* (14), 4454-65.
223. de Kloe, G. E.; Bailey, D.; Leurs, R.; de Esch, I. J., Transforming fragments into candidates: small becomes big in medicinal chemistry. *Drug Discov Today* **2009**, *14* (13-14), 630-46.
224. Waldrop, G. L., Smaller is better for antibiotic discovery. *ACS Chem Biol* **2009**, *4* (6), 397-9.

225. Martin, J. L.; Waksman, G.; Bardwell, J. C.; Beckwith, J.; Kuriyan, J., Crystallization of DsbA, an Escherichia coli protein required for disulphide bond formation in vivo. *J Mol Biol* **1993**, *230* (3), 1097-100.
226. Inaba, K.; Takahashi, Y. H.; Ito, K.; Hayashi, S., Critical role of a thiolate-quinone charge transfer complex and its adduct form in de novo disulfide bond generation by DsbB. *Proc Natl Acad Sci U S A* **2006**, *103* (2), 287-92.
227. Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K., Structure-based discovery of beta2-adrenergic receptor ligands. *Proc Natl Acad Sci U S A* **2009**, *106* (16), 6843-8.
228. Guha, R.; Howard, M. T.; Hutchison, G. R.; Murray-Rust, P.; Rzepa, H.; Steinbeck, C.; Wegner, J.; Willighagen, E. L., The Blue Obelisk-interopability in chemical informatics. *J Chem Inf Model* **2006**, *46* (3), 991-8.
229. Baldi, P.; Benz, R. W., BLASTing small molecules--statistics and extreme statistics of chemical similarity scores. *Bioinformatics* **2008**, *24* (13), i357-65.
230. (a) Wielens, J.; Headey, S. J.; Deadman, J. J.; Rhodes, D. I.; Parker, M. W.; Chalmers, D. K.; Scanlon, M. J., Fragment-based design of ligands targeting a novel site on the integrase enzyme of human immunodeficiency virus 1. *Chemmedchem* **2011**, *6* (2), 258-61; (b) Wielens, J.; Headey, S. J.; Deadman, J. J.; Rhodes, D. I.; Parker, M. W.; Chalmers, D. K.; Scanlon, M. J., Fragment-Based Design of Ligands Targeting a Novel Site on the Integrase Enzyme of Human Immunodeficiency Virus 1. *Chemmedchem* **2010**.
231. Zoete, V.; Grosdidier, A.; Michielin, O., Docking, virtual high throughput screening and in silico fragment-based drug design. *J Cell Mol Med* **2009**, *13* (2), 238-48.
232. van Montfort, R. L.; Collins, I., Fragment-based methods in drug discovery: it's the small things that matter. *Curr Top Med Chem* **2009**, *9* (18), 1676-7.
233. Campos-Olivas, R., NMR screening and hit validation in fragment based drug discovery. *Curr Top Med Chem* **2011**, *11* (1), 43-67.
234. Berne, B. J.; Weeks, J. D.; Zhou, R., Dewetting and hydrophobic interaction in physical and biological systems. *Annu Rev Phys Chem* **2009**, *60*, 85-103.
235. Minichino, A.; Habash, J.; Raftery, J.; Helliwell, J. R., The properties of (2Fo - Fc) and (Fo - Fc) electron-density maps at medium-to-high resolutions. *Acta Crystallogr D Biol Crystallogr* **2003**, *59* (Pt 5), 843-9.
236. O'Neill, M. A.; Gaisford, S., Application and use of isothermal calorimetry in pharmaceutical development. *Int J Pharm* **2011**.
237. Borsi, V.; Calderone, V.; Fragai, M.; Luchinat, C.; Sarti, N., Entropic contribution to the linking coefficient in fragment based drug design: a case study. *J Med Chem* **2010**, *53* (10), 4285-9.
238. Crisman, T. J.; Bender, A.; Milik, M.; Jenkins, J. L.; Scheiber, J.; Sukuru, S. C.; Fejzo, J.; Hommel, U.; Davies, J. W.; Glick, M., "Virtual fragment linking": an approach to identify potent binders from low affinity fragment hits. *J Med Chem* **2008**, *51* (8), 2481-91.