STRUCTURAL AND FUNCTIONAL ANALYSES OF COPPER-SENSING OPERON REGULATOR PROTEIN (CsoRGz) OF Geobacillus zalihae STRAIN T1

ASHWAANI MANGAVELU

FBSB 2018 20
STRUCTURAL AND FUNCTIONAL ANALYSES OF COPPER-SENSING OPERON REGULATOR PROTEIN (CsoR<sub>Gz</sub>) OF *Geobacillus zalihae* STRAIN T1

By

ASHWAANI MANGAVELU

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirements for the Degree of Master of Science

March 2018
COPYRIGHT

All material contained within the thesis, including without limitation text, logos, icons, photographs, and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia
Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

STRUCTURAL AND FUNCTIONAL ANALYSES OF COPPER-SENSING OPERON REGULATOR PROTEIN (CsoRGz) OF Geobacillus zalihae STRAIN T1

By

ASHWAANI MANGAVELU

March 2018

Chairman : Normi Binti Mohd Yahaya, PhD
Faculty : Biotechnology and Biomolecular Sciences

Copper sensor regulator protein (CsoR) is widespread in most Gram positive bacteria. It is categorized under metalloregulatory protein that tightly regulates the passage of copper(I) ions to maintain cell viability, metabolism and preventing cell toxicity. Previously, gene annotation of a locally isolated Geobacillus zalihae strain revealed the presence of a CsoR-like hypothetical protein (CsoRGz) that shared only 30-38% sequence identity to structurally characterized CsoR proteins of various bacterial strains. This highlights the possibility of structural novelty of CsoRGz. As the actual 3D structure of CsoRGz is not available, information on possible novelty of its structure as well as its 3D copper(I) binding site and motif is unknown. This study aimed to elucidate the structure of CsoRGz via X-ray crystallography, identify the functional metal-binding residues and residues involved in dimerization of CsoRGz protein and compare its structure with other structurally characterized CsoR proteins. For this study, Escherichia coli BL21 Star® (DE3) inserted with pET28b-CsoRGz vector was used to express the recombinant CsoRGz protein. CsoRGz protein was optimally expressed in LB medium at 37°C with IPTG induction at 0.1 mM and purified to homogeneity by affinity chromatography followed by gel filtration. The protein was stable when dialysed at pH 6.5 in the buffer containing 10 mM MES, 100 mM NaCl, 0.2 mM EDTA, 0.2 mM DTT and 5% glycerol. The purified CsoRGz protein was crystallized via sitting drop vapour diffusion method at 15 °C. Trigonal crystals corresponding to space group P3121 were grown in 5% (v/v) tacsimate TM (pH 7.0), 0.1 M HEPES (pH 7.0), 10% (w/v) polyethylene glycol monomethyl ether 5,000. The crystal was successfully diffracted at 1.83Å using an in-house X-ray beam with completeness of 100% with unit-cell parameters as follows; a=44.73, b=44.73, c=82.37 Å, where α=ß=90° and γ=120°. The Matthew’s coefficient analysis revealed that the crystal structure had one molecule per asymmetric unit with the solvent content of 32.82%. Its conserved copper-binding residues, Cys46-His71-Cys75, were
found to be located at α2 helix onwards of the crystal structure similar to other CsoR proteins. The dimeric structure of CsoRGz was indicated using Protein, Interface and Structure and Assemblies (PISA) whereby the monomers were held together by disulphide bonds, hydrogen bonds and salt bridges and the residues involved in these respective interactions were duly mapped. CsoRGz showed largely similar global topology to other known Cu(I) bound CsoRs with minimal differences in the characteristics of the Cu(I) binding pocket whereby CsoRGz had a more hydrophobic environment than other reported CsoRs.
Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai
memenuhi keperluan untuk ijazah Master Sains

**ANALISA STRUKTUR DAN FUNGSI PROTEIN PENGAWAL ATUR OPERON KUPRUM (CsoR<sub>GZ</sub>) DARIPADA Geobacillus zalihae STRAIN T1**

Oleh

ASHWAANI MANGAVELU

Mac 2018

**Pengerusi** : Normi Binti Mohd Yahaya, PhD
**Faculti** : Bioteknologi dan Sains Biomolekul

Protein pengawal atur kuprum (CsoR) didapati dengan meluas dalam kebanyakan bakteria Gram positif. Ia dikategorikan di bawah protein pengawal atur logam yang mengawal atur keluar dan masuk kuprum(I) untuk mengekalkan daya tahan dan metabolisme sel dan mencegah ketoksikan sel. Anotasi genom strain Geobacillus zalihae yang dipencilkan tempatan mendedahkan kehadiran protein hipotetikal seakan CsoR (CsoR<sub>GZ</sub>) yang berkongsi hanya 30-38% identiti jujukan dengan protein CsoR yang telah dicirikan strukturnya daripada pelbagai strain bakteria. Ini menggariskan kemungkinan struktur CsoR<sub>GZ</sub> yang novel. Memandangkan struktur 3D asal CsoR<sub>GZ</sub> masih tiada, maklumat berkaitan kemungkinan struktur novel dan juga tapak dan motif pengikat kuprum(I) 3Dnya masih tidak diketahui. Jadi, kajian ini bertujuan untuk merumagka struktur CsoR<sub>GZ</sub> melalui kristolografi sinar-X, memetakan residu fungsian pengikat logam dan yang terlibat di dalam pendimeran protein CsoR<sub>GZ</sub> dan membandingkannya strukturnya dengan protein CsoR lain yang telah dicirikan strukturnya. Untuk tujuan ini, protein rekombinan CsoR<sub>GZ</sub> telah dihasilkan di dalam Escherichia coli BL21 Star<sup>®</sup> (DE3) yang telah dimasukkan dengan pET28b-CsoR<sub>GZ</sub>. Pengekspresan protein CsoR<sub>GZ</sub> telah dihasilkan secara optimum dalam media LB pada 37°C dengan induksi IPTG pada 0.1 mM dan dipencilkan sehingga homogeneti dengan menggunakan kromatografi affiniti diikuti dengan penapisan gel. Protein tersebut adalah stabil apabila didialisis pada pH 6.5 di dalam penimbal yang mengandungi 10 mM MES, 100 mM NaCl, 0.2 mM EDTA, 0.2 mM DTT dan 5% gliserol. Protein CsoR<sub>GZ</sub> yang dipencilkan telah dihablurkan melalui kaedah resapan wap *sitting drop* pada 15 °C. Hablur trigonal sepadan dengan kumpulan ruang P3<sub>1</sub>21 telah ditumbuhkan dengan 5% (v/v) tacsimate TM pH 7.0, 0.1 M HEPES pH 7.0, 10% (w/v) polietilena glikol monometill eter 5,000. Hablur tersebut berjaya dibelakukan kepada resolusi 1.83 Å menggunakan sunber sinar X *in house* dengan kelengkapan 100% dengan parameter sel unit seperti berikut; a = 44.73, b = 44.73, c = 82.37 Å, di
mana α=β=90° dan γ=120°. Analisis nilai pekali Matthew menunjukkan bahawa hablur ini mempunyai satu molekul per unit asimetri dengan kandungan pelarut sebanyak 32.82%. Residu terpelihara yang mengikat kuprum, iaitu Cys46-His71-Cys75, telah dijumpai terletak pada heliks α2 dan seterusnya serupa dengan struktur protein CsoR yang lain. Struktur dimer CsoRGz telah dianalisis dengan menggunakan program Protein, Antara Muka dan Struktur dan Perhimpunan (PISA) di mana dua monomer dipegang bersama oleh ikatan disulfida, ikatan hidrogen dan penyambung garam. Struktur CsoRGz memperlihatkan topologi global yang sama seperti protein CsoR yang lain dengan perbezaan yang sedikit dalam ciri-ciri poket mengikat Cu(I)nya di mana CsoRGz mempunyai persekitaran hidrofobik yang lebih tinggi berbanding dengan CsoR lain yang telah dilaporkan.
ACKNOWLEDGEMENTS

Special appreciation to my supervisor, Dr. Normi binti Mohd Yahaya for giving me the opportunity to work on this project, for her valuable guidance, constant encouragement and support throughout this Masters study. Thank you for being such an inspiration and role model besides being a second mother to me. I would also like to thank my supervisory committee members, Professor Dr. Raja Noor Zaliha Binti Raja Abd. Rahman, Professor Dr. Abu Bakar Bin Salleh, Associate Professor Dr. Adam Leow Thean Chor and Associate Professor Dr. Mohd Shukuri bin Mohamad Ali and other EMTech Research group members for their guidance, encouragement, insightful comments and helpful discussions.

My sincere thanks to Dr. Anuar Bin Jonet from Malaysian Genome Institute (MGI) for his help and cooperation in X-ray diffraction and data collection. Special thanks for all the postgraduate students of EMTech and Protein Engineering laboratory members for their continuous support, knowledge and friendship. Special thanks also extended to the staff of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia (UPM) for their help and cooperation. I am also thankful to Putra Research Grant UPM (Grant No.: 9502200) and Graduate Research Fellowship (GRF) for the financial support for my research studies.

Last but not least, I am eternally grateful to my parents, Mangavelu Manickam and Tami Selvi Urun and my siblings for their utmost support. Thank you so much Jivita Thestnamoorthi for your continuous encouragement for me in pursuing my Masters studies and believing in my capabilities. Lastly, to those who indirectly contributed in this research, your kindness means a lot to me. Thank you.
I certify that a Thesis Examination Committee has met on 29 March 2018 to conduct
the final examination of Ashwaani a/p Mangavelu on her thesis entitled "Structural
and Functional Analyses of Copper-Sensing Operon Regulator Protein (CsoRGe) of
*Geobacillus zalihae* Strain T1" in accordance with the Universities and University
Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106]
15 March 1998. The Committee recommends that the student be awarded the Master
of Science.

Members of the Thesis Examination Committee were as follows:

**Noor Azmi Shaharuddin, PhD**
Senior Lecturer
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Chairman)

**Tan Wen Siang, PhD**
Professor
Faculty of Biotechnology and Biomolecular Sciences
Universiti Putra Malaysia
(Internal Examiner)

**Ng Chyan Leong, PhD**
Senior Lecturer
Universiti Kebangsaan Malaysia
Malaysia
(External Examiner)

\[Signature\]

**NOR AINI AB. SHUKOR, PhD**
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 24 May 2018
This thesis was submitted to the Senate of the Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

**Normi Binti Mohd Yahaya, PhD**  
Associate Professor  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Chairman)

**Raja Noor Zaliha Raja Abd. Rahman, PhD**  
Professor  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Member)

**Abu Bakar Bin Salleh, PhD**  
Professor  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Member)

**Adam Leow Thean Chor, PhD**  
Associate Professor  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Member)

**Mohd Shukuri Bin Mohammad Ali, PhD**  
Associate Professor  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Member)

---

**ROBIAH BINTI YUNUS, PhD**  
Professor and Dean  
School of Graduate Studies  
Universiti Putra Malaysia

Date:

vii
Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software

Signature: __________________________ Date: __________________________

Name and Matric No.: Ashwaani Mangavelu, GS42467
Declaration by Members of Supervisory Committee

This is to confirm that:
- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) were adhered to.

Signature: 
Name of Chairman of Supervisory Committee: Associate Professor Dr. Normi Binti Mohd Yahaya

Signature: 
Name of Member of Supervisory Committee: Professor Dr. Raja Noor Zaliha Raja Abd. Rahman

Signature: 
Name of Member of Supervisory Committee: Professor Dr. Abu Bakar Bin Salleh

Signature: 
Name of Member of Supervisory Committee: Associate Professor Dr. Adam Leow Thean Chor

Signature: 
Name of Member of Supervisory Committee: Associate Professor Dr. Mohd Shukuri Bin Mohammad Ali
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRAK</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>APPROVAL</td>
<td>vi</td>
</tr>
<tr>
<td>DECLARATION</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td>xvii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xviii</td>
</tr>
</tbody>
</table>

## CHAPTER

1 **INTRODUCTION**

2 **LITERATURE REVIEW**

   2.1 Metal Homeostasis and Metalloregulatory Protein 3
   2.2 Copper homeostasis and bacterial pathogenicity 5
   2.3 Copper Responsive Gene Regulation in Bacteria 6
      2.3.1 Copper Sensing Operon Regulator (CsoR) Protein 7
         2.3.1.1 CsoR from *M. tuberculosis* 8
         2.3.1.2 CsoR in *B. subtilis* 9
         2.3.1.3 CsoR in *Thermus thermophiles* 10
   2.4 Copper-Ligand Coordination Chemistry 11
   2.5 Crystal structures of CsoR proteins 13
      2.5.1 Dimeric CsoR proteins 13
      2.5.2 Tetrameric CsoR Proteins 14
   2.6 CsoR Protein from *Geobacillus zalihae* (*CsoR*<sub>Gz</sub>) 17

3 **MATERIALS AND METHODS**

   3.1 Materials 19
   3.2 Bacterial strains and plasmids 19
   3.3 Growth Media Preparation 19
      3.3.1 Luria Bertani (LB) Broth 19
      3.3.2 Luria Bertani (LB) Agar 19
   3.4 Bacterial Glycerol Stock Preparation 20
   3.5 Optimization of recombinant CsoR<sub>Gz</sub> protein production in Luria Bertani (LB) medium: Effect of cultivation temperature and isopropyl-β-D-1-thiogalactopyranoside (IPTG) concentration 20
   3.6 Purification of CsoR<sub>Gz</sub> Recombinant Protein 20
      3.6.1 Affinity Chromatography 20
3.6.2 Gel Filtration Chromatography

3.7 Optimization of Dialysis, Protein Cut-off Size and Storage Buffer Conditions

3.8 Crystallization Screening of CsoRGz Protein

3.9 X-ray Diffraction and Structure Elucidation of CsoRGz

3.9.1 Cryoprotectant and Crystal Preparation

3.9.2 X-ray diffraction Data Collection

3.9.3 Phase Determination

3.9.3.1 Space Group Assignment

3.9.3.2 Molecular Replacement

3.9.3.3 Model Building and Refinement

3.10 Structure Validation of CsoRGz

3.11 Determination of Biological Assembly of CsoRGz

3.12 Structure Analysis of CsoRGz

3.13 Docking of Crystal Structure of CsoRGz with Cu(I)

4 RESULTS AND DISCUSSION

4.1 Primary and Secondary Analysis of CsoRGz Protein

4.1.1 Primary Structure Analysis

4.1.2 Secondary Structure Analysis

4.2 Optimization of Recombinant CsoRGz Protein Production in LB Medium

4.2.1 Optimization of Induction Temperature

4.2.2 Optimization of IPTG Concentration

4.3 Purification of CsoRGz Recombinant Protein

4.3.1 Affinity Chromatography

4.3.2 Gel Filtration Chromatography

4.4 Buffer Optimization for CsoRGz Protein

4.4.1 pH of Storage Buffer

4.4.2 Type of Storage Buffer

4.4.3 Protein Dialysis

4.5 Crystallization Screening for CsoRGz Protein

4.6 X-ray Diffraction Analysis

4.6.1 X-ray Diffraction and Data Collection of CsoRGz

4.7 Phase Determination

4.7.1 Template Search

4.7.2 Space Group Assignment

4.7.3 Template Preparation using MrBUMP

4.7.4 Molecular Replacement (MR)

4.8 Evaluation of CsoRGz Crystal Structure

4.8.1 Ramachandran Plot Analysis on Crystal Structure of CsoRGz Protein

4.8.2 ERRAT Analysis on Crystal Structure CsoRGz Protein

4.8.3 PROSESS (Protein Structure Evaluation Suite & Server)

4.9 Structural Analysis on Crystal structure CsoRGz Protein

4.9.1 Overall Tertiary Structure of CsoRGz Protein
4.10 Incomplete Electron Density Map Due to the Presence of Intrinsically Disordered Protein Regions 53
4.11 Average B-factors of All Atoms in the Protein Structure 56
4.12 Distinguishing Non-crystallographic Symmetry (NCS) from Crystallographic Symmetry 57
4.13 Biological Assembly of CsoR\(_Gz\) Protein 59
4.14 Bonds Between the Monomers 61
4.15 Docking of Crystal Structure of CsoR\(_Gz\) with Cu(I) 62

5 CONCLUSION AND RECOMMENDATION FOR FUTURE RESEARCH 67
5.1 Conclusion 67
5.2 Recommendation for future research 67

REFERENCES 69
APPENDICES 79
BIODATA OF STUDENT 97
LIST OF PUBLICATIONS 98
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Metalloregulatory Family and its Regulatory Proteins</td>
<td>4</td>
</tr>
<tr>
<td>2.2</td>
<td>9 Classes of copper sensing regulators and copper-responsive gene regulation in bacteria</td>
<td>7</td>
</tr>
<tr>
<td>2.3</td>
<td>Experimentally verified copper-responsive regulators, phylogenetic group and their target genes</td>
<td>8</td>
</tr>
<tr>
<td>3.1</td>
<td>Coordination and the dimensions of the simulation cells for docking</td>
<td>26</td>
</tr>
<tr>
<td>4.1</td>
<td>Amino acid composition of CsoRGz protein</td>
<td>28</td>
</tr>
<tr>
<td>4.2</td>
<td>Protein storage buffers and its constituents used for CsoR proteins from various species and their corresponding field of studies</td>
<td>36</td>
</tr>
<tr>
<td>4.3</td>
<td>Comparison on the retainability and concentration of CsoRGz-like protein after dialysis and centrifugal concentration using 3.5 and 10.0 kDa MWCO</td>
<td>37</td>
</tr>
<tr>
<td>4.4</td>
<td>Crystallization condition and time taken for the protein crystal formation obtained during crystal screening</td>
<td>39</td>
</tr>
<tr>
<td>4.5</td>
<td>X-ray data statistics for collected diffraction and processed data for CsoRGz protein crystal</td>
<td>44</td>
</tr>
<tr>
<td>4.6</td>
<td>Structural identity search was done for CsoRGz with CsoR structures deposited in PDB by using Phyre²</td>
<td>45</td>
</tr>
<tr>
<td>4.7</td>
<td>Final statistics of the refined CsoRGz protein crystal structure</td>
<td>48</td>
</tr>
<tr>
<td>4.8</td>
<td>The assembly summary for the dimeric structure of CsoRGz protein generated by PISA</td>
<td>60</td>
</tr>
<tr>
<td>4.9</td>
<td>Bonds formed by the two CsoRGz monomers and the corresponding distances between the bonds formed</td>
<td>61</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Schematic representation of the <em>csoR</em> - <em>copZA</em> locus. Palindromic sequence consists of (GTAGCCCACCCC-N4-GGGGTGGGATAC) overlapping the -35/-10 promoter region</td>
</tr>
<tr>
<td>2.1</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>Mechanism of regulation of copper homeostasis by CsoR protein via <em>CopZA</em> operon in <em>B. subtilis</em></td>
</tr>
<tr>
<td>2.2</td>
<td>10</td>
</tr>
<tr>
<td>2.3</td>
<td>Chemical structure of histidine, cysteine and methionine with side chains that are associated with copper binding</td>
</tr>
<tr>
<td>2.3</td>
<td>11</td>
</tr>
<tr>
<td>2.4</td>
<td>The active sites of copper chaperone and sensor that form different coordination geometry with Cu(II)</td>
</tr>
<tr>
<td>2.4</td>
<td>12</td>
</tr>
<tr>
<td>2.5</td>
<td>Crystal structure of <em>Mtb</em> CsoR homodimer</td>
</tr>
<tr>
<td>2.5</td>
<td>13</td>
</tr>
<tr>
<td>2.6</td>
<td>Crystal structure of CsoR protein from <em>T. thermophilus</em> HB8</td>
</tr>
<tr>
<td>2.6</td>
<td>15</td>
</tr>
<tr>
<td>2.7</td>
<td>Crystal structure of CsoR protein of <em>S. lividans</em></td>
</tr>
<tr>
<td>2.7</td>
<td>16</td>
</tr>
<tr>
<td>2.8</td>
<td>Crystal structure of CsoR protein from <em>Geobacillus thermodenitrificans</em></td>
</tr>
<tr>
<td>2.8</td>
<td>17</td>
</tr>
<tr>
<td>3.1</td>
<td>Crystallization phase diagram showing the metastable, labile/nucleation and precipitation zone</td>
</tr>
<tr>
<td>3.1</td>
<td>22</td>
</tr>
<tr>
<td>3.2</td>
<td>Crystallization of protein via sitting drop method. In this method, the protein-liquor drop sits on a shelf or a post within the reservoir and it is sealed with tape.</td>
</tr>
<tr>
<td>3.2</td>
<td>23</td>
</tr>
<tr>
<td>3.3</td>
<td>Dimeric CsoR&lt;sub&gt;Gz&lt;/sub&gt; protein model with grid box at site 1 (A) and site 2 (B) viewed using Discovery studio</td>
</tr>
<tr>
<td>3.3</td>
<td>27</td>
</tr>
<tr>
<td>4.1</td>
<td>Secondary structure of CsoR&lt;sub&gt;Gz&lt;/sub&gt; protein is composed of alpha helices and random coils</td>
</tr>
<tr>
<td>4.1</td>
<td>29</td>
</tr>
<tr>
<td>4.2</td>
<td>Expression of recombinant CsoR&lt;sub&gt;Gz&lt;/sub&gt; protein at different temperatures by <em>E. coli</em> BL21 Star&lt;sup&gt;TM&lt;/sup&gt; (DE3) induced with 0.1 mM IPTG in LB medium</td>
</tr>
<tr>
<td>4.2</td>
<td>30</td>
</tr>
<tr>
<td>4.3</td>
<td>Expression of recombinant CsoR&lt;sub&gt;Gz&lt;/sub&gt; protein at 37°C by <em>E. coli</em> BL21 Star&lt;sup&gt;TM&lt;/sup&gt; (DE3) induced with different concentrations of IPTG in LB medium</td>
</tr>
<tr>
<td>4.3</td>
<td>31</td>
</tr>
</tbody>
</table>
4.4 Purification chromatogram of recombinant CsoR\textsubscript{Gz} protein using affinity chromatography with linear elution method

4.5 Purification of recombinant CsoR\textsubscript{Gz} protein using gel filtration chromatography with step-wise elution method

4.6 SDS-PAGE analysis of CsoR\textsubscript{Gz} protein retained after dialysis and centrifugal concentration using various MWCOs

4.7 Crystals of CsoR\textsubscript{Gz} viewed under polarised microscope and its corresponding diffraction images

4.8 Single crystals obtained using 10 mg/ml CsoR\textsubscript{Gz} protein with formulation containing 5% (v/v) tacsimate TM pH 7.0, 0.1 M HEPES pH 7.0, 10% (w/v) polyethylene glycol monomethyl ether 5,000

4.9 Electron density map of poorly fitted template due to incorrect space group assignment (space group P31) viewed using COOT at 1 sigma level

4.10 The criteria for the solution-quality assessment for MrBUMP

4.11 Chain C of CsoR from \textit{Thermus thermophiles} HB8

4.12 Ramachandran plot of crystal structure of CsoR\textsubscript{Gz} protein where all 60 residues of the protein are situated at the most favoured and additional allowed regions.

4.13 ERRAT plot with overall quality factor of 100.0% for crystal structure of CsoR\textsubscript{Gz} protein

4.14 Results of PROSESS software shows the crystal structure of CsoR\textsubscript{Gz} protein has an average overall quality

4.15 Overall crystal structure of CsoR\textsubscript{Gz} solved using Molecular Replacement method containing one molecule in asymmetric unit, viewed using YASARA

4.16 Missing electron density map at the N- and C-terminal ends of the structure, viewed using COOT

4.17 Analysis on the CsoR\textsubscript{Gz} amino acid sequence using DisEMBL

4.18 Graph generated by IUPred predicting the regions on the amino acid sequence of CsoR\textsubscript{Gz}
4.19 B-factor profile of CsoR$_{Gz}$ crystal structure with the average B-factor value of 34.

4.20 (A) Electron density map with ‘extra density’ found at the residue Cys46 and Cys75 (B) Fold axis was located at in the density map where the points related by proper symmetry have similar density.

4.21 The local symmetry related residues located at both of the active sites, Cys46 and Cys75 where Cys46 residue interacts with the symmetric residue Cys75’ and Cys75 residue interacts with the symmetric residue Cys46’

4.22 The front view (A) and side view (B) of dimeric structure of CsoR$_{Gz}$ protein

4.23 Close up view of the two Cu(I) binding sites in the structure of CsoR$_{Gz}$ protein where site 1 (C46’-H71-C75) and site 2 (C46-H71’-C75’) are located at both the ends of the dimeric structure

4.24 Close-up view of metal-binding site 1 of CsoR$_{Gz}$ dimer after docking of Cu(I) atom (green sphere) for docking mode 18 with RMSD value of 4.786 Å

4.25 Close-up view of metal-binding site 1 after docking of Cu(I) atom (green sphere) for mode 16 with RMSD value of 8.455 Å and 2.935 Å (B).

4.26 Close-up view of amino acid residues involved in providing a hydrophobic environment in the Cu(I) binding site of CsoR$_{Gz}$ within 5 Å radius from the metal binding site.

4.27 Amino acid residues involved in the stabilization of the dimeric structure of CsoR$_{Gz}$ protein by forming hydrogen bonds, disulphide bonds, salt bridges and hydrophobic interactions connecting the monomers
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>List of Chemicals, Reagents, Softwares and Equipment</td>
</tr>
<tr>
<td>B</td>
<td>Composition of SDS-PAGE gel</td>
</tr>
<tr>
<td>C</td>
<td>Standard Curve for Bradford Assay</td>
</tr>
<tr>
<td>D</td>
<td>Protein and gene sequence of CsoR$_{Gz}$</td>
</tr>
<tr>
<td>E</td>
<td>Crystal Screening Formulations</td>
</tr>
<tr>
<td>i.</td>
<td>Hampton Research Crystal Screening Kit I</td>
</tr>
<tr>
<td>ii.</td>
<td>Hampton Research Crystal Screening Kit II</td>
</tr>
<tr>
<td>iii.</td>
<td>Hampton Research Index I</td>
</tr>
<tr>
<td>iv.</td>
<td>Hampton Research Index II</td>
</tr>
<tr>
<td>v.</td>
<td>Molecular Dimension JCSG Kit I</td>
</tr>
<tr>
<td>vi.</td>
<td>Molecular Dimension JCSG Kit II</td>
</tr>
<tr>
<td>F</td>
<td>Plasmid map of pET28b(+)</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

α  Alpha
Å  Angstrom
ATP  Adenosine Tri-phosphate
APS  Ammonium persulfate
Au  Aurum
bp  Base pair
β  Beta
Cd  Cadmium
CaCl₂  Calcium chloride
cAMP  Cyclic adenosine monophosphate
cDNA  Complementary deoxyribonucleic acid
Crp  cAMP receptor Protein
csc  Chromosomally encoded sucrose catabolism genes
Co  Cobalt
Cu(I)  Copper (I) ion
CsoR  Copper Sensitive Operon Repressor
°C  Degree Celcius
DNA  Deoxyribonucleic acid
DUF  Domain of Unknown Function
EDTA  Ethylene-diamine-tetraacetic acid
Fe  Ferrum
g  Gram
HTH  Helix Turn Helix
His  Histidine
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPTG</td>
<td>Isopropyl-Beta-D-Thiogalactoside</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>kDA</td>
<td>KiloDalton</td>
</tr>
<tr>
<td>L</td>
<td>Litre</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>Mn</td>
<td>Manganese</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>µL</td>
<td>Microliter</td>
</tr>
<tr>
<td>µm</td>
<td>Micrometer</td>
</tr>
<tr>
<td>mM</td>
<td>Millimolar</td>
</tr>
<tr>
<td>Ni</td>
<td>Nickel</td>
</tr>
<tr>
<td>A&lt;sub&gt;600nm&lt;/sub&gt;</td>
<td>Optical density at wavelength 600 nanometer</td>
</tr>
<tr>
<td>%</td>
<td>Percentage</td>
</tr>
<tr>
<td>NaPO&lt;sub&gt;4&lt;/sub&gt;</td>
<td>Sodium phosphate</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium deodecyl sulphate Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>3D</td>
<td>Three Dimensional</td>
</tr>
<tr>
<td>TCS</td>
<td>Two Components Regulatory System</td>
</tr>
<tr>
<td>Zn</td>
<td>Zinc</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Copper is required as a micronutrient for all living cells, as it is known as an essential component for large numbers of metalloenzymes. However, elevated level of copper ions in micro and macro organisms can be detrimental (Liu et al., 2006). Therefore, tightly controlling the levels of cellular metal ion via homeostasis is required to avoid cell toxicity as well as for cell viability and metabolism.

To cope with unfavourable copper concentration, bacteria utilize copper-induced defence mechanisms. An example of such copper homeostasis mechanism involves CsoR protein. It is a metalloregulatory protein that that functions to regulate the copper ion concentration in bacterial cells. It is responsible in controlling gene expression which allows organisms to adapt to critical level of essential metal ions, particularly copper, thus enabling cells tolerate heavy metal pollutants present in their environment (Giedroc & Arunkumar, 2007). It functions as a copper-sensing repressor, regulates expression of copper-homeostasis genes. It was first discovered in Mycobacterium tuberculosis (Liu et al., 2007). Besides M. tuberculosis, CsoR-like proteins are also widespread in other prokaryotes such as in Bacillus subtilis (Smaldone & Helmann, 2007), Staphylococcus aureus (Baker et al., 2011), Streptomyces lividans (Chaplin et al., 2015) and Listeria monocytogenes (Corbett et al., 2011).

In general, csoR gene is located upstream of copZA operon that encodes a copper-ATPase and a copper chaperone. Copper chaperone function to transport the excess copper to copper ATPase, while copper ATPase function to remove excess copper out of the cell (Hirooka et al., 2012). When copper is scarce, CsoR represses the transcription by binding to the operator of the copZA operon which overlaps copZ promoter. When copper concentration exceeds, CsoR binds to the excessive copper ions and detaches itself from the operator, thus unblocking the promoter region for transcription of copZA to take place (Rademacher & Masepohl, 2012). Previously, a scan on the genome sequence of G. zalihae revealed the presence of a CsoR-like hypothetical protein gene located at the upstream of the copZA operon. Bioinformatics analysis done by Musa (2016) on the protein sequence of CsoR-like protein of G. zalihae showed that it contained a CsoR-DUF domain similar to other well-characterized CsoR proteins. However, CsoR_{Gz} only shared up to 30-38% sequence identity to structurally characterized CsoR proteins of various bacterial origins (Musa, 2016). This suggests the possibility of its structure being novel compared to other structurally characterized CsoR proteins.

It is important to note that the 3D structure of CsoR_{Gz} have not been determined via X-ray crystallography as well as metal-binding ligands interaction using docking method. Whilst several tools such as isothermal titration calorimetry is able to shed light on the ability of the protein to interact with metals, it is not able to identify which
amino acids the metals interact with. Combination of X-ray crystallography and docking is able to give forth such information upon the interaction of proteins with metals. Hence, this present study aims to:

1. elucidate the structure of CsoR<sub>GC</sub> via X-ray crystallography
2. identify the functional metal-binding residue and dimer interface of CsoR<sub>GC</sub>
3. compare the similarities and differences of CsoR<sub>GC</sub> with other structurally characterized CsoR proteins.
REFERENCES


72


