COMPUTATIONAL ANALYSIS ON PROTEIN-LIGAND INTERACTION OF XYLITOL-PHOSPHATE DEHYDROGENASE ENZYMES FOR XYLITOL PRODUCTION

SITI AISYAH BINTI RAZALI

UNIVERSITI TEKNOLOGI MALAYSIA
COMPUTATIONAL ANALYSIS ON PROTEIN-LIGAND INTERACTION OF XYLITOL-PHOSPHATE DEHYDROGENASE ENZYMES FOR XYLITOL PRODUCTION

SITI AISYAH BINTI RAZALI

A thesis submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy

Faculty of Science
Universiti Teknologi Malaysia

JULY 2018
“To my wonderful family for their endless support and motivation. 
Ummi and Abah, thank you for your love and patience”
ACKNOWLEDGEMENT

Alhamdulillah, all praises be to Allah, to Whom I am grateful for guidance in this journey to seek His knowledge. Thank you, Allah, for Your endless blessing, love and giving me the strength to do this research. I wish to express my deepest appreciation to my supervisor, Prof. Dr Shahir Shamsir for his guidance, patience and moral support during this journey. I thank Allah for giving me an opportunity to meet and work with a great supervisor like him. May Allah bless him with good health, success and happiness. I would also like to convey an appreciation to my co-supervisor, Prof. Dr Rosli Md Illias for his critical comments and suggestions.

I would like to express my gratitude to those who have encouraged and guided me to complete this thesis. My heartfelt appreciation goes to my teammates, Amy, Sarah, Ann, Shaiful, Chew, Kak Syakila, Hafifi and Farah for being greatly tolerant, supportive perpetually and for all the fun we had during this journey. Also, I thank my fellow Bioinformatics lab mates for the stimulating discussions and knowledge sharing sessions. Special thanks to beautiful Kak Leha, Kak Zuraidah, Kak Linda and En Awang for the technical support provided. I would like to thank the Malaysian Ministry of Education and Malaysian Genome Institute for the scholarship and research financial aid. I am also indebted to Kak Dilin, Dr Kheng Onn and all members of Genetic Engineering Laboratory for their assistance in providing significant information.

Lastly, my highest gratitude is to my parents, Razali Mohd Ali and Noraini Abd Rashid, my siblings, Fatin, Umyra, Syukrie and my love, Farhan for their support, encouragement and prayer. It would be impossible to finish this research without many people who supported and believed in me.
ABSTRACT

Xylitol is a high-value low-calorie sweetener used as sugar substitute in food and pharmaceutical industry. Xylitol phosphate dehydrogenase (XPDH) catalyses the conversion of D-xylulose 5-phosphate (XU5P) and D-ribulose 5-phosphate (RU5P) to xylitol and ribitol respectively in the presence of nicotinamide adenine dinucleotide hydride (NADH). Although these enzymes have been shown to produce xylitol, however there is a limited understanding of the mechanism of the catalytic events of these reactions and the detailed mechanism has yet to be elucidated. Understanding of the catalytic activity of these enzymes would provide novel information for protein engineering to improve xylitol production. The main goal of this work is to analyse the conformational changes of XPDH-bound ligands such as Zn$^{2+}$ NADH, XU5P, and RU5P to elucidate the key amino acids involved in the substrate binding. In silico modelling, comparative molecular dynamic simulations, interaction analysis and conformational study were carried out on three XPDH enzymes of the Medium-chain dehydrogenase (MDR) family; XPDH from Lactobacillus rhamnosus (LrXPDH) and Clostridium difficile (CdXPDH, Cd1XPDH) in order to elucidate the atomistic details of conformational transition, especially on the open and closed state of XPDH. The critical residues involved in substrate binding and conformational changes were mutated using in silico site-directed mutagenesis. The result showed that residues Cys37, His58, Glu59, and Glu142 form an active site pocket within the catalytic domain. In the coenzyme domain, NADH is shown to bind to highly conserved glycine-rich motif; GXGXXG (residues 166-171). The results also revealed that XPDH consists of a dual mechanism that can catalyse hydride transfer to dissimilar substrates (XU5P and RU5P), which His58 and Ser39 would act as the proton donor for reduction of XU5P and RU5P respectively. The structural comparison and MD simulations displayed a significant difference in the conformational dynamics of the catalytic and coenzyme loops between Apo and XPDH-complexes and highlight the contribution of newly found triad residues (W48, I259, and W285). The study also identified the effect of S39A and W285A mutations on substrate binding and conformational changes. The study successfully elucidated the mechanistic aspect of catalysis mechanism and dynamical event of XPDH enzymes at molecular level. The results from this study would assist future mutagenesis study and enzyme modification work to increase the catalysis efficiency of xylitol production in the industry.
ABSTRAK

Xylitol adalah pemanis rendah kalori yang bernilai tinggi dan digunakan sebagai pengganti gula dalam industri makanan dan industri farmaseutikal. Xylitol fosfat dehidrogenase (XPDH) menjadi pemangkin penukaran xilulosa 5-fosfat (XU5P) kepada xylitol dan D-ribulosa 5-fosfat (RU5P) kepada ribitol dengan menggunakan nikotinamida adenina dinukleotida hidrida (NADH). Walaupun enzim ini telah terbukti menghasilkan xylitol, tetapi pemahaman terhadap mekanisme tindak balas ini adalah terhad dan belum dijelaskan secara terperinci. Pemahaman terhadap pemangkinan ini akan memberikan maklumat baru dalam kejuruteraan protein untuk meningkatkan pengeluaran xylitol. Matlamat utama kajian ini adalah untuk menganalisis perubahan bentuk enzim XPDH dan ligan seperti Zn2+ NADH, XU5P, dan RU5P serta menjelaskan jujuk asid amino yang terlibat dalam pengikat substrat. Pemodelan dalam siliko, perbandingan simulasi dinamik molekul, analisis interaksi dan kajian sama bentuk telah dijalankan pada tiga enzim XPDH daripada keluarga dehidrogenase Medium (MDR); iaitu XPDH dari Lactobacillus rhamnosus (LrXPDH) dan Clostridium difficile (CdXPDH, Cd1XPDH) untuk menjelaskan peralihan bentuk secara butiran atom, terutamanya dalam keadaan terbuka dan tertutup XPDH. Asid amino yang terlibat dalam pengikat substrat dan perubahan bentuk telah dimutasi menggunakan tapak siliko mutagenesis berarah. Hasil kajian menunjukkan bahawa jujuk asid amino Cys37, His58, Glu59, Glu142 membentuk poket tapak aktif dalam domain pemangkin. Dalam domain koenzim, NADH terikat dengan motif terabadi, GXGXXG (jujuk amino 166-171) yang kaya dengan glisina. Kajian ini juga mendedahkan XPDH mempunyai dwi mekanisme yang boleh memangkinkan pemindahan hidrida ke substrat yang berbeza (XU5P dan RU5P), iaitu His58 dan Ser39 akan bertindak sebagai penderma proton untuk pengurangan XU5P dan RU5P. Perbandingan struktur dan simulasi MD mendedahkan perbezaan yang signifikan dalam bentuk dinamik dari gelung mangkinan dan koenzim antara apo dan kompleks XPDH serta menonjolkan sumbangan jujuk amino triad yang baru dijumpai (W48, I259, dan W285). Kajian ini juga mengenal pasti kesan mutasi S39A dan W285A pada perubahan pengikat substrat dan analisis perubahan bentuk. Kajian ini berjaya menjelaskan aspek mekanisma pemangkinan mekanistik dan peristiwa dinamik enzim XPDH di peringkat molekul. Hasil dari kajian ini akan membantu kajian mutagenesis di masa depan dan kerja pengubahsuaian enzim untuk meningkatkan kecekapan pemangkinan pengeluaran xylitol dalam industri.
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td></td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>v</td>
</tr>
<tr>
<td>ABSTRAK</td>
<td></td>
<td>vi</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td></td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td></td>
<td>xiv</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td></td>
<td>xx</td>
</tr>
<tr>
<td>LIST OF APPENDICES</td>
<td></td>
<td>xxi</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>Overview</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Problem Statement</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Research Objectives</td>
<td>4</td>
</tr>
<tr>
<td>1.4</td>
<td>Scope of Study</td>
<td>4</td>
</tr>
<tr>
<td>1.5</td>
<td>Significance of Study</td>
<td>5</td>
</tr>
<tr>
<td>1.6</td>
<td>Thesis Organization</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>LITERATURE REVIEW</td>
<td>7</td>
</tr>
<tr>
<td>2.1</td>
<td>Xylitol</td>
<td>7</td>
</tr>
</tbody>
</table>
2.1.1 Natural Occurrence of Xylitol 8
2.1.2 Physical and Chemical Properties of Xylitol 9
2.1.3 Xylitol Market Value 12
2.1.4 Application of Xylitol 15
2.1.5 Production of Xylitol 17

2.2 Xylitol Phosphate Dehydrogenase (XPDH) 27
2.2.1 Substrate Specificity of XPDH Enzymes 27
2.2.2 Metabolic Pathways of XPDH 28

2.3 The Computational Studies of Polyol Dehydrogenase (PDH) 30
2.3.1 Sequence Analysis 34
2.3.2 Structure Analysis 34
2.3.3 Protein-ligand interaction 35
2.3.4 Protein Engineering 37

3 RESEARCH METHODOLOGY 41
3.1 Operational Framework of the Research 41
3.2 Phase 1: Sequence-based Analysis 43
3.2.1 Physicochemical Characterization 43
3.2.2 Secondary Structure Prediction 44
3.2.3 Sequence Alignment 44
3.2.4 Phylogenetic Study 44
3.2.5 Molecular Function (MF) Evaluation 45
3.2.6 In silico Mutation Screening 45
3.3 Phase 2: Structure-based Analysis 45
3.3.1 Model Development 47
3.3.2 Virtual Mutation 47
3.3.3 Structure Refinement 48
3.3.4 Evaluation of the Model 48

3.4 Phase 3: Protein Substrate Interaction 49

3.4.1 Binding Site Prediction 51
3.4.2 Molecular Docking 51

3.5 Phase 4: Protein Stability and Dynamic 52

3.5.1 Preparation Stage 55
3.5.2 Setup Stage 55
3.5.3 Simulation Stage 59
3.5.4 Analysis Stage 61

3.6 Summary of Software and Database 63

4 PROTEIN SUBSTRATE INTERACTION OF WILD-TYPE XPDH ENZYMES 70

4.1 Phase 1: Sequence-based Analysis 70

4.1.1 Physicochemical Characterization 70
4.1.2 Secondary Structure Comparison 75
4.1.3 Sequence Alignment 77
4.1.4 Phylogenetic Study 79
4.1.5 Molecular Function (MF) Evaluation 80

4.2 Phase 2: Structure-based analysis 85

4.2.1 Template Identification 85
4.2.2 Model Development 89
4.2.3 Structural Analysis 93
4.2.4 Structure Refinement 98
4.2.5 Evaluation of the model 99

4.3 Protein substrate interaction 105

4.3.1 Zinc and NADH Binding Site Prediction 105
4.3.2 Substrate Binding Site Prediction 109
4.3.3 Molecular Docking 113

4.4 Phase 4: Protein Stability and Dynamic 136
4.4.1 Protein Stability 136
4.4.2 Protein Interaction 140
4.4.3 Conformational Changes and Overall Dynamic Behavior 142

5 PROTEIN ENGINEERING OF CdXPDH COMPLEX 156

5.1 Phase 1: Sequence-based Analysis 156
5.1.1 Selection of Residues Using *In silico* Mutation Screening 156
5.1.2 Physicochemical Characterization 159
5.1.3 Sequence Alignment 162

5.2 Phase 2: Structure-based Analysis 163
5.2.1 Model Development 163
5.2.2 Structural Analysis 167
5.2.3 Structure Refinement 171
5.2.4 Evaluation of the model 172

5.3 Phase 3: Protein Substrate Interaction 178
5.3.1 Binding Site Prediction 178
5.3.2 Molecular Docking 181

5.4 Phase 4: Protein Stability and Dynamic 197
5.4.1 Protein Stability 197
5.4.2 Protein Interaction 201
5.4.3 Conformational Changes and Overall Dynamic Behavior 203
6 CONCLUSION

6.1 Research conclusion 217

6.2 Recommendation for Future Work 219

REFERENCES 220
APPENDICES 241
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Physiochemical properties of xylitol (Mussatto, 2012)</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>The summary of xylitol production from D-glucose using microbiological and enzymatic methods.</td>
<td>23</td>
</tr>
<tr>
<td>2.3</td>
<td>Substrate specificity of LrXPDH and CdXPDH</td>
<td>27</td>
</tr>
<tr>
<td>2.4</td>
<td>Production of xylitol and ribitol by XDH, XPDH and APDH</td>
<td>28</td>
</tr>
<tr>
<td>2.5</td>
<td>The computational studies of ArDH and XDH</td>
<td>31</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary of the trajectories subjected to the molecular dynamic simulations</td>
<td>54</td>
</tr>
<tr>
<td>3.2</td>
<td>List of software and databases used for <em>in silico</em> analysis of XPDH proteins - <em>Phase 1</em></td>
<td>64</td>
</tr>
<tr>
<td>3.3</td>
<td>List of software and databases used for <em>in silico</em> analysis of XPDH proteins – <em>Phase 2</em></td>
<td>66</td>
</tr>
<tr>
<td>3.4</td>
<td>List of software and databases used for <em>in silico</em> analysis of XPDH proteins – <em>Phase 3</em></td>
<td>68</td>
</tr>
<tr>
<td>3.5</td>
<td>List of software and databases used for <em>in silico</em> analysis of XPDH proteins – <em>Phase 4</em></td>
<td>69</td>
</tr>
<tr>
<td>4.1</td>
<td>Physicochemical characteristic of XPDH enzymes</td>
<td>72</td>
</tr>
<tr>
<td>4.2</td>
<td>The predicted Gene Ontology (GO) terms of LrXPDH for molecular function evaluation.</td>
<td>82</td>
</tr>
<tr>
<td>4.3</td>
<td>The predicted Gene Ontology (GO) terms of CdXPDH for molecular function evaluation.</td>
<td>83</td>
</tr>
</tbody>
</table>
4.4 The predicted Gene Ontology (GO) terms of Cd1XPDH for molecular function evaluation. 84
4.5 The predicted Gene Ontology (GO) summary of XPDH enzymes for molecular function evaluation. 85
4.6 XPDH enzymes’ top three proposed templates by different servers 86
4.7 Summary of successfully produced models of XPDH using MODELLER program 91
4.8 Structural alignment evaluation of the best XPDH models with their template GPDH 92
4.9 Summary of XPDH secondary structure elements topology 97
4.10 Summary of model validation using different tools. 104
4.11 The predicted tunnels of XPDH enzymes at the catalytic site 112
4.12 Summary of the trajectories subjected to the molecular dynamics simulations and the average RMSD values. 137
5.1 *In silico* mutation screening by using multiple tools 157
5.2 Physicochemical characteristic of WT CdXPDH and its mutants 161
5.3 Summary of successfully produced models of CdXPDH mutants using MODELLER program 165
5.4 Structural alignment evaluation of the best CdXPDH models with their template GPDH. 166
5.5 Summary of CdXPDH models validation using different tools. 177
5.6 The predicted tunnels of XPDH enzymes at the catalytic site 180
5.7 Summary of the trajectories subjected to the molecular dynamics simulations and the average RMSD values. 198
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>a) 2D and b) 3D representation of xylitol chemical structure</td>
<td>8</td>
</tr>
<tr>
<td>2.2</td>
<td>Natural occurrence of xylitol in fruits and vegetables</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>Asia pacific xylitol market overview (2009-2020) in metric tons</td>
<td>13</td>
</tr>
<tr>
<td>2.4</td>
<td>U.S Xylitol Market size, by application, 2013-2023 (Kilo Tons)</td>
<td>14</td>
</tr>
<tr>
<td>2.5</td>
<td>Global xylitol chewing gum (2009-2020) market by geographic region in metric ton</td>
<td>14</td>
</tr>
<tr>
<td>2.6</td>
<td>Xylitol production method</td>
<td>18</td>
</tr>
<tr>
<td>2.7</td>
<td>The chemical process for manufacturing xylitol</td>
<td>20</td>
</tr>
<tr>
<td>2.8</td>
<td>Three-step fermentation proces</td>
<td>24</td>
</tr>
<tr>
<td>2.9</td>
<td>Two-step fermentation proces</td>
<td>25</td>
</tr>
<tr>
<td>2.10</td>
<td>The metabolic pathways of the bioconversion of D-glucose into five carbons sugars.</td>
<td>29</td>
</tr>
<tr>
<td>2.11</td>
<td>The 3D structure of ArDH (PDB ID: 3M6I).</td>
<td>35</td>
</tr>
<tr>
<td>2.12</td>
<td>Sequence alignment of the structural zinc binding motif from different XDH</td>
<td>36</td>
</tr>
<tr>
<td>2.13</td>
<td>Comparative-modeling-based 3D structure of XDH.</td>
<td>38</td>
</tr>
<tr>
<td>3.1</td>
<td>Operational framework of the research</td>
<td>42</td>
</tr>
<tr>
<td>3.2</td>
<td>Process of structure-based analysis</td>
<td>46</td>
</tr>
<tr>
<td>3.3</td>
<td>Process of molecular docking analysis</td>
<td>50</td>
</tr>
<tr>
<td>3.4</td>
<td>Process of molecular dynamic simulation</td>
<td>53</td>
</tr>
</tbody>
</table>
3.5 The XPDH model was placed at the center of the cubic box.

3.6 The XPDH model in solvated cubic MD simulation box.

3.7 The XPDH model in the neutralized system.

4.1 The amino acid sequences of a) LrXPDH, b) CdXPDH and c) Cd1XPDH.

4.2 Summary of amino acids composition in XPDH enzymes

4.3 Summary of amino acids characterized groups' percentage in XPDH enzymes

4.4 XPDH secondary structure prediction by using GOR IV

4.5 Conserved domain analysis of XPDH enzymes.

4.6 Sequence alignment for XPDH enzymes with the closest structural homologues

4.7 Molecular phylogenetic tree derived from several amino acid sequences of Medium Dehydrogenase/Reductase (MDR) enzymes using MEGA 7 software.

4.8 The predicted terms within the Gene Ontology (GO) hierarchy for LrXPDH Molecular Function (MF).

4.9 The predicted terms within the Gene Ontology (GO) hierarchy for CdXPDH Molecular Function (MF).

4.10 The predicted terms within the Gene Ontology (GO) hierarchy for Cd1XPDH Molecular Function (MF).

4.11 Superimposition of XPDH enzymes

4.12 3D model and topology diagram of XPDH secondary structure elements.

4.13 Deep cleft in XPDH enzymes.

4.14 The root mean square (RMSD) of XPDH enzymes during 10ns structure refinement.

4.15 Ramachandran plot for XPDH enzymes model before and after structure refinement.

4.16 Error values for residues as predicted by ERRAT.

4.17 The VERIFY3D curve for LrXPDH, CdXPDH, and Cd1XPDH models.
The sequence alignment between XPDH enzymes and MDR consensus sequence (Cdd: cd08236). 106
The sequence alignment between XPDH enzymes and MDR consensus sequence (Cdd: cd08236). 107
The sequence alignment between XPDH enzymes and MDR consensus sequence (Cdd: cd08236). 108
The predicted tunnels for XPDH substrate binding. The tunnels were prepared by using MOLE software. 111
The catalytic Zn$^{2+}$ binding site of XPDH 114
The structural Zn$^{2+}$ binding site of XPDH 115
Top view of XPDH NADH binding 118
The interaction of LrXPDH with NADH in the coenzyme binding domain. 119
The interaction of CdXPDH with NADH in the coenzyme binding domain. 120
The interaction of Cd1XPDH with NADH in the coenzyme binding domain. 121
The binding mode of D-xylulose 5-phosphate (XU5P) in the catalytic site of LrXPDH. 124
The binding mode of D-xylulose 5-phosphate (XU5P) in the catalytic site of CdXPDH. 125
The binding mode of D-xylulose 5-phosphate (XU5P) in the catalytic site of Cd1XPDH. 126
The binding mode of D-ribulose 5-phosphate (RU5P) in the catalytic site of LrXPDH. 128
The binding mode of D-ribulose 5-phosphate (RU5P) in the catalytic site of CdXPDH. 129
The binding mode of D-ribulose 5-phosphate (RU5P) in the catalytic site of Cd1XPDH. 130
Reduction of D-xylulose 5-Phosphate (D-xylulose-5P) to D-xylitol 5-Phosphate (Xylitol-5P) by NADH in the catalytic site of XPDH. 133
4.35 Reduction of D-ribulose 5-Phosphate (D-ribulose-5P) to D-ribitol 5-Phosphate (D-ribitol-5P) by NADH in the catalytic site of XPDH.

4.36 Backbone RMSD of a) LRXPDH b) CdXPDH and c) Cd1XPDH during 20,000 ps simulations.

4.37 Atomic distance analysis of the MD trajectories of XPDH complexes.

4.38 Structural conformation of LrXPDH Apo in the substrate binding pocket.

4.39 Structural conformation of LrXPDH Complex I in the substrate binding pocket.

4.40 Structural conformation of LrXPDH Complex II in the substrate binding pocket.

4.41 Structural conformation of CdXPDH Apo in the substrate binding pocket.

4.42 Structural conformation of CdXPDH Complex I in the substrate binding pocket.

4.43 Structural conformation of CdXPDH Complex II in the substrate binding pocket.

4.44 Structural conformation of Cd1XPDH Apo in the substrate binding pocket.

4.45 Structural conformation of Cd1XPDH Complex I in the substrate binding pocket.

4.46 Structural conformation of Cd1XPDH Complex II in the substrate binding pocket.

5.1 Schematic mutation structures of Serine into an Alanine at position 39 and Tryptophan into an Alanine at position 285.

5.2 Sequence alignment for WT CdXPDH with its mutants

5.3 Superimposition of CdXPDH enzyme

5.4 Model development of WT CdXPDH and S39A CdXPDH

5.5 Model development of WT CdXPDH and W285A CdXPDH.
5.6 Comparison of solvent accessibility between WT CdXPDH and its mutants.

5.7 The root mean square (RMSD) of WT and mutant CdXPDH during 10ns structure refinement.

5.8 Ramachandran plot for WT CdXPDH and its mutants before and after structure refinement.

5.9 Error values for residues as predicted by ERRAT.

5.10 The VERIFY3D curve for WT CdXPDH and its mutants before and after structure refinement.

5.11 The predicted tunnels for CdXPDH substrate binding. The tunnels were prepared by using MOLE software.

5.12 The catalytic Zn$^{2+}$ binding site of CdXPDH

5.13 The structural Zn$^{2+}$ binding site of CdXPDH

5.14 Close-up view of the wild-type and mutants NADH binding.

5.15 The interaction of S39A CdXPDH with NADH in the coenzyme binding domain.

5.16 The interaction of W285A CdXPDH with NADH in the coenzyme binding domain.

5.17 The binding mode of D-xylulose 5-phosphate (XU5P) in the catalytic site of S39A CdXPDH.

5.18 The binding mode of D-ribulose 5-phosphate (RU5P) in the catalytic site of S39A CdXPDH.

5.19 The binding mode of D-xylulose 5-phosphate (XU5P) in the catalytic site of W285A CdXPDH.

5.20 The binding mode of D-ribulose 5-phosphate (RU5P) in the catalytic site of W285A CdXPDH.

5.21 Reduction of D-xylulose 5-Phosphate (D-xylulose-5P) to D-xylitol 5-Phosphate (Xylitol-5P) by NADH in the catalytic site of S39A CdXPDH.

5.22 Backbone RMSD of WT CdXPDH and its mutants during 20,000 ps simulations.

5.23 Atomic distance analysis of the MD trajectory of CdXPDH complexes.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.24</td>
<td>Structural conformation of WT Apo in the substrate binding pocket.</td>
<td>204</td>
</tr>
<tr>
<td>5.25</td>
<td>Structural conformation of WT Complex I in the substrate binding pocket.</td>
<td>205</td>
</tr>
<tr>
<td>5.26</td>
<td>Structural conformation of WT Complex II in the substrate binding pocket.</td>
<td>206</td>
</tr>
<tr>
<td>5.27</td>
<td>Structural conformation of S39A Apo in the substrate binding pocket.</td>
<td>209</td>
</tr>
<tr>
<td>5.28</td>
<td>Structural conformation of S39A Complex I in the substrate binding pocket.</td>
<td>210</td>
</tr>
<tr>
<td>5.29</td>
<td>Structural conformation of S39A Complex II in the substrate binding pocket.</td>
<td>211</td>
</tr>
<tr>
<td>5.30</td>
<td>Structural conformation of W285A Apo in the substrate binding pocket.</td>
<td>214</td>
</tr>
<tr>
<td>5.31</td>
<td>Structural conformation of W285A Complex I in the substrate binding pocket.</td>
<td>215</td>
</tr>
<tr>
<td>5.32</td>
<td>Structural conformation of S39A Complex II in the substrate binding pocket.</td>
<td>216</td>
</tr>
</tbody>
</table>
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LrXPDH</td>
<td><em>Lactobacillus rhamnosus</em> xylitol phosphate dehydrogenase</td>
</tr>
<tr>
<td>CdXPDH</td>
<td><em>Clostridium difficile</em> xylitol phosphate dehydrogenase</td>
</tr>
<tr>
<td>XU5P</td>
<td>D-xylulose 5-phosphate</td>
</tr>
<tr>
<td>RU5P</td>
<td>D-ribulose 5-phosphate</td>
</tr>
<tr>
<td>WT</td>
<td>Wild-type</td>
</tr>
<tr>
<td>MDR</td>
<td>Medium-chain dehydrogenase</td>
</tr>
<tr>
<td>NAD</td>
<td>Nicotinamide adenine dinucleotide</td>
</tr>
<tr>
<td>XDH</td>
<td>Xylitol dehydrogenase</td>
</tr>
<tr>
<td>ArDH</td>
<td>Arabitol dehydrogenase</td>
</tr>
<tr>
<td>GPDH</td>
<td>Galactitol-1-phosphate 5-dehydrogenase</td>
</tr>
<tr>
<td>PDH</td>
<td>Polyol dehydrogenase</td>
</tr>
<tr>
<td>EC</td>
<td>Enzyme Commission</td>
</tr>
<tr>
<td>CDD</td>
<td>Conserved domain database</td>
</tr>
<tr>
<td>GOR</td>
<td>Garnier-Osguthorpe-Robson</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>GO</td>
<td>Gene Ontology</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root mean square deviations</td>
</tr>
<tr>
<td>SPC</td>
<td>Simple point charge</td>
</tr>
<tr>
<td>PME</td>
<td>Particle Mesh Ewald</td>
</tr>
<tr>
<td>LINC</td>
<td>LINear Constraint Solver</td>
</tr>
<tr>
<td>GRAVY</td>
<td>Grand average of hydropathicity</td>
</tr>
<tr>
<td>NPS</td>
<td>Network Protein Sequence</td>
</tr>
<tr>
<td>MSA</td>
<td>Multiple sequence alignment</td>
</tr>
</tbody>
</table>
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>List of EC Numbers</td>
<td>242</td>
</tr>
<tr>
<td>B</td>
<td>Protein sequence of Cd1XPDH</td>
<td>244</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Overview

Today, an increasing number of researchers are focusing on xylitol production as an alternative sugar for healthy eating. Because of their unique properties, they have potential and desirable for food industry such as sugar-free chewing gum, cookies, desserts and soft drink (Mussatto, 2012). Xylitol can also improve the storage properties, taste, and colour of food product (Ur-Rehman et al., 2015). For the pharmaceutical industry, xylitol is the suitable low-calorie sweetener that is recommended for the diabetic patient as it can be metabolized in the absence of insulin (Storey et al., 2007). The global market for xylitol is currently estimated to be over US$750 million per year and priced at US$ 6-7 per kg (Global Market Insights, 2016). Xylitol has 12% share of total polyol market, which is the second largest after sorbitol (Albuquerque et al., 2014).

This sugar is found naturally in fruits and vegetables as well as in yeast, seaweed, and mushrooms. It can be extracted by solid-liquid extraction, but it becomes a major economic problem due to its small proportion of the raw materials (Winkelhausen and Kuzmanova, 1998). Industrially, xylitol produced by catalytic reduction of pure D-xylose, however the chemical method of xylitol manufacturing is laborious and expensive (Rafiqul and Sakinah, 2013a; X.-H. Qi et al., 2016).
Alternatively, this problem could be solved by using D-glucose as the low-cost raw material (Cheng et al., 2014a). D-glucose can be converted into xylitol by using xylitol-phosphate dehydrogenase from *Lactobacillus rhamnosus* and *Clostridium difficile* with the highest yield 22-23% (Povelainen and Miasnikov, 2007a).

The study of XPDH classification is needed in order to know the remarkable mechanism and metabolic pathway to produce xylitol. Oxidoreductases are divided into three classes which are short-chain dehydrogenase (SDR), medium chain dehydrogenase/reductase (MDR) and long-chain dehydrogenase. These enzymes are specifically acting on the CHO group of a donor molecule with NAD⁺ or NADP⁺ as the acceptor (Auld and Bergman, 2008). Xylitol-phosphate dehydrogenase from *Lactobacillus rhamnosus* ATCC 15820 (LrXPDH), XPDH from *Clostridium difficile* CD630 (CdXPDH) and XPDH from *Clostridium difficile* CD196 (Cd1XPDH) belong to the MDR family. All these three proteins consist of two domains; a catalytic domain and a nicotinamide cofactor (NADH) binding domain. The 3D structure and the active site of XPDH enzymes remained to be identified and the interaction of substrate binding has not been studied in detail at the atomic level. The present research is the first study of the sequences and structural characterization, protein-ligand interaction and protein engineering of XPDH enzymes that can produce xylitol from D-glucose.

Combination of comparative modelling, molecular docking, and molecular dynamics simulation can help to understand the action mode of substrates and the catalytic mechanism of XPDH enzymes. The computational study has been powerful tools for researchers to predict protein structure and ligand-protein interaction. *In silico*, site-directed mutagenesis will establish novel strategies to increase efficiency of XPDH enzymes activity and improve xylitol production.
1.2 Problem Statement

Xylitol-phosphate dehydrogenase from *Lactobacillus rhamnosus* ATCC 15820 (LrXPDH), XPDH from *Clostridium difficile* CD630 (CdXPDH) and XPDH from *Clostridium difficile* CD196 (Cd1XPDH), three enzymes from Medium-chain dehydrogenase family that are capable to catalyze the reduction of both D-xylulose 5-phosphate and D-ribulose 5-phosphate to xylitol (Povelainen and Miasnikov, 2007a; Abdullah, 2018). However, the three dimensional (3D) structures of all XPDH enzymes are relatively unknown and the interaction of substrate binding has not been studied in detail at the atomic level. Hence, the comparative modelling and molecular docking studies may reveal the structural active site and interaction of XPDH enzymes with their ligands.

Due to the substrate specificity of XPDH, the xylitol production was accompanied by co-production of ribitol. *In silico* site-directed mutagenesis is required for the understanding rationale of the conversion. Furthermore, the effect of the mutation on the stability of XPDH enzymes has remained unexplored. Molecular dynamic simulations are powerful tools to study the stability of the mutants. It is important to highlight that there is no computational approach for XPDH enzyme to date. *In silico* study of XPDH may provide biotechnologically interesting potential as well as improve the production of xylitol.
1.3  **Research Objectives**

The main goal of this research is to analyse the protein-ligand interaction of xylitol-phosphate dehydrogenase enzymes for xylitol production. There are several objectives need to be achieved in this research project:

1. To investigate the primary sequence characteristics and the three-dimensional structures of wild-type and mutant xylitol phosphate dehydrogenase (XPDH) enzymes.
2. To identify the key binding residues and analyse the interaction of the substrates with XPDH-complex at the catalytic and coenzyme domain.
3. To elucidate the details mechanism of xylitol phosphate dehydrogenase (XPDH) enzymes.
4. To study the effect of the mutation on the stability of XPDH enzymes based on amino acid substitution and comparative molecular dynamic simulation.
5. To elucidate the atomistic details of conformational changes on the open and closed state of XPDH enzymes.

1.4  **Scope of Study**

This study is exclusively bioinformatics and computational analysis which include model development, protein interaction, protein engineering, protein stability and dynamics. All the data were derived from the primary database and analyzed using high performance computing facilities in FBME. In this works, three Xylitol-phosphate dehydrogenase (XPDH) enzymes that can produce xylitol were selected; including XPDH from *Lactobacillus rhamnosus* and *Clostridium difficile*. The primary sequence and structural analysis of XPDH enzymes were done to investigate their functional characteristics and elucidate the potential protein engineering for xylitol production. The interaction of substrate binding protein will be studied using molecular docking. The simulations were performed using open source GROMACS.
(GROningen Machine for Chemical Simulation) version 5.1.4 software (Abraham et al., 2015) in order to investigate the dynamic signature and conformational behaviour of the protein-ligand complex.

1.5 Significance of Study

In this research, the sequence and structural analysis of XPDH enzymes provide the valuable structural information of molecular architecture of XPDH which offer novel details in PDH family and may be relevant to wider MDR superfamily. This analysis also help the fundamental biology on sequence-structure-function relationship of protein families. The study of protein-ligand interaction of XPDH provides an insight into the possible catalytic event, improve specificity of the substrate and provide information for the protein engineering to increase the xylitol production.

This study also successfully elucidate the mechanistic aspect of catalysis mechanism and dynamic event of XPDH enzymes at the molecular level, especially on the open and closed state of XPDH which has been impossible to determine by experimental technique. In silico site directed mutagenesis in this study will provide the fundamental information contribution of key residues in XPDH catalysis and molecular dynamic.

Overall, this thesis makes a significant contribution to the field of knowledge by offering information on structural, dynamic and computational study in order to design rational strategies to increase the efficiency of XPDH enzymes activity and improve xylitol production.
1.6 Thesis Organization

This thesis is comprised of six chapters. Chapter 1 describes the outline of the research which includes the background of this study and the problem statement. This chapter also emphasized the objectives, the scopes and the significance of this research.

Chapter 2 include the literature review that related to the study. This chapter is focusing on reviewing other related proteins in the same family, the production and the application of the related sugar and the basic concept of this research area.

Chapter 3 present the research methodology which includes the operational frameworks in order to achieve the research goals. All the methods and materials used in this study are described in detail.

Chapter 4 shows the structure and function prediction of Xylitol phosphate dehydrogenase (XPDH). The interaction of protein-ligand binding and molecular dynamic simulation are also discussed in detail. The significant results from this chapter were used to identify the potential protein engineering (Chapter 5) for xylitol production.

Chapter 5 highlights the information of in silico site mutagenesis of CdXPDH–complex proteins. The result of the conducted experiments and discussion related to the objectives are included in this chapter.

Chapter 6 gives a conclusion of the thesis by a general discussion of the result obtained. In addition, this chapter discusses the directions for future work in order to improve the production of xylitol.
REFERENCES


225


