Folding of Bovine Pancreatic Trypsin Inhibitor (BPTI) is Faster using Aromatic Thiols and their Corresponding Disulfides

Ram Prasad Marahatta
Florida International University, rmara011@fiu.edu

Follow this and additional works at: http://digitalcommons.fiu.edu/etd

Part of the Analytical Chemistry Commons, Biological and Chemical Physics Commons, Medicinal-Pharmaceutical Chemistry Commons, and the Organic Chemistry Commons

Recommended Citation
Marahatta, Ram Prasad, "Folding of Bovine Pancreatic Trypsin Inhibitor (BPTI) is Faster using Aromatic Thiols and their Corresponding Disulfides" (2017). FIU Electronic Theses and Dissertations. 3530.
http://digitalcommons.fiu.edu/etd/3530

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.
FOLDING OF BOVINE PANCREATIC TRypsIn INHIBITOR (BPTI) IS FASTER USING AROMATIC THIOLS AND THEIR CORRESPONDING DISULFIDES

A dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY

by

Ram Prasad Marahatta

2017
To: Dean Michael R. Heithaus  
College of Arts, Sciences and Education  

This dissertation, written by Ram Prasad Marahatta, and entitled Folding of Bovine Pancreatic Trypsin Inhibitor (BPTI) is Faster using Aromatic Thiols and their Corresponding Disulfides, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____________________________________
Kevin O'Shea

_____________________________________
Prem P. Chapagain

_____________________________________
Francisco Alberto Fernandez Lima

_____________________________________
Xiaotang Wang

_____________________________________
Watson J. Lees, Major Professor

Date of Defense: November 17, 2017

The dissertation of Ram Prasad Marahatta is approved.

_____________________________________
Dean Michael R. Heithaus  
College of Arts, Sciences and Education

_____________________________________
Andrés G. Gil  
Vice President for Research and Economic Development  
and Dean of the University Graduate School

Florida International University, 2017
© Copyright 2017 by Ram Prasad Marahatta

All rights reserved.
DEDICATION

I would like to dedicate this humble work to my loving wife, cute daughter, loving mother, my siblings, and whole Marahatta family for their encouragement and patience during my doctoral studies. I would like to remember my late father Tara Nidhi Marahatta, who had an aim to teach me to this higher level of education and without his blessing from heaven, I could not do anything. I would also like to thank my relatives and friends who always supported me during my graduate study.
ACKNOWLEDGMENTS

I would like to deeply appreciate to my committee chair Dr. Watson J. Lees for providing me an opportunity to work with him. I would not be able to complete my dissertation without his guidance and persistent help. I would also like to thank my committee member Dr. Prem P. Chapagain for encouraging and supporting me to complete simulation works. I am also very thankful to my other committee members Dr. Kevin O’Shea, Dr. Francisco Fernandez-Lima, and Dr. Xiaotang Wang for their suggestions which were very helpful to me to complete my dissertation work.

I would like to thank all members of Dr. Lees’ group whom I met during my research period. I am also very thankful to my friend Rudra Mani Pokhrel for his help during my works for Chapter 5. Likewise, I acknowledge faculty members of Department of Chemistry and Biochemistry at Florida International University for their help and encouragement. Finally, I appreciate to Department of Chemistry and Biochemistry and lab coordinators as well for letting me work as a teaching assistant for five years of my study, and many thanks to graduate secretary Magali Autie for her help regarding other aspects of graduate life.
ABSTRACT OF THE DISSERTATION

FOLDING OF BOVINE PANCREATIC TRIPSYN INHIBITOR (BPTI) IS FASTER USING AROMATIC THIOLS AND THEIR CORRESPONDING DISULFIDES

by

Ram Prasad Marahatta

Florida International University, 2017

Miami, Florida

Professor Watson J. Lees, Major Professor

Improvement in the in vitro oxidative folding of disulfide containing proteins, such as extracellular and pharmaceutically important proteins, is required. Traditional folding methods using small molecule aliphatic thiol and disulfide, such as glutathione (GSH) and glutathione disulfide (GSSG) are slow and low yielding. Small molecule aromatic thiols and disulfides show great potentiality because aromatic thiols have low pKa values, close to the thiol pKa of protein disulfide isomerase (PDI), higher nucleophilicity and good leaving group ability. Our studies showed that thiols with positively charged group, quaternary ammonium salts (QAS), are better than thiols with negatively charged groups such as phosphonic acid and sulfonic acid for the folding of bovine pancreatic trypsin inhibitor (BPTI). An enhanced folding rate of BPTI was observed when the protein was folded with a redox buffer composed of a QAS thiol and its corresponding disulfide.

Quaternary ammonium salt (QAS) thiols and their corresponding disulfides with longer alkyl side chains were synthesized. These QAS thiols and their
corresponding disulfides are promising small molecule thiols and disulfides to fold reduced BPTI efficiently because these thiols are more hydrophobic and can enter the core of the protein.

Conformational changes of disulfide containing proteins during oxidative folding influence the folding pathway greatly. We performed the folding of BPTI using targeted molecular dynamics (TMD) simulation and investigated conformational changes along with the folding pathway. Applying a bias force to all atoms versus to only alpha carbons and the sulfur of cysteines showed different folding pathways. The formation of kinetic traps $N'$ and $N^*$ were not observed during our simulation applying bias force to all atoms of the starting structure. The final native conformation was obtained once the correct antiparallel $\beta$-sheets and subsequent Cys14-Cys38 distance was decreased to a bond distance level. When bias force was applied to only alpha carbons and the sulfur of cysteines, the distance between Cys14-Cys38 increased and decreased multiple times, structure similar to the conformation of $N^*$, $N^{\text{SH}}$ were formed and native protein was ultimately obtained. We concluded that there could be multiple pathways of conformational folding which influence oxidative folding.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Structure and folding of protein</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1 Protein structure</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2 <em>In vitro</em> protein folding</td>
<td>2</td>
</tr>
<tr>
<td>1.1.2.1 Process of <em>in vitro</em> protein production</td>
<td>2</td>
</tr>
<tr>
<td>1.1.2.2 Importance of protein folding</td>
<td>6</td>
</tr>
<tr>
<td>1.1.3 Folding of disulfide containing proteins</td>
<td>8</td>
</tr>
<tr>
<td>1.1.4 Folding with Protein Disulfide Isomerase (PDI)</td>
<td>8</td>
</tr>
<tr>
<td>1.1.4.1 General information</td>
<td>8</td>
</tr>
<tr>
<td>1.1.4.2 Oxidative protein folding with PDI</td>
<td>10</td>
</tr>
<tr>
<td>1.1.4.2.1 Oxidative <em>in vivo</em> protein folding with PDI</td>
<td>10</td>
</tr>
<tr>
<td>1.1.4.3 Oxidative <em>in vitro</em> protein folding with PDI</td>
<td>11</td>
</tr>
<tr>
<td>1.1.4.4 Folding with small molecules</td>
<td>12</td>
</tr>
<tr>
<td>1.2 Folding of bovine pancreatic trypsin inhibitor (BPTI)</td>
<td>16</td>
</tr>
<tr>
<td>1.2.1 Introduction of BPTI</td>
<td>16</td>
</tr>
<tr>
<td>1.2.2 Traditional <em>in vitro</em> oxidative folding of BPTI</td>
<td>20</td>
</tr>
<tr>
<td>1.2.3 Updated folding pathway of BPTI</td>
<td>24</td>
</tr>
<tr>
<td>1.3 Folding of BPTI with aromatic thiols and their corresponding disulfides</td>
<td>26</td>
</tr>
<tr>
<td>1.3.1 Introduction</td>
<td>26</td>
</tr>
<tr>
<td>1.3.2 Construction of the structure of small molecule aromatic thiols</td>
<td>27</td>
</tr>
<tr>
<td>1.4 Molecular dynamics (MD) simulation study of BPTI</td>
<td>29</td>
</tr>
<tr>
<td>1.4.1 Importance of MD simulation</td>
<td>29</td>
</tr>
<tr>
<td>1.4.2 Protein folding and MD simulation</td>
<td>29</td>
</tr>
<tr>
<td>1.4.3 Molecular dynamics (MD) simulation of BPTI</td>
<td>30</td>
</tr>
<tr>
<td>2 Objectives</td>
<td>32</td>
</tr>
<tr>
<td>2.1 Synthesis of different aromatic thiols and their corresponding disulfides for the study of folding of reduced BPTI</td>
<td>32</td>
</tr>
<tr>
<td>2.2 Folding of reduced BPTI faster using different aromatic thiols and their corresponding disulfides</td>
<td>32</td>
</tr>
<tr>
<td>2.3 Molecular dynamics (MD) simulation study of BPTI intermediates to understand the folding of BPTI at a molecular level</td>
<td>32</td>
</tr>
<tr>
<td>3 Synthesis of aromatic thiols and corresponding disulfides</td>
<td>33</td>
</tr>
<tr>
<td>3.1 Abstract</td>
<td>33</td>
</tr>
<tr>
<td>3.2 Introduction</td>
<td>34</td>
</tr>
<tr>
<td>3.3 Results and discussion</td>
<td>37</td>
</tr>
<tr>
<td>3.3.1 Synthesis of quaternary ammonium salt thiol 1</td>
<td>37</td>
</tr>
<tr>
<td>3.3.2 Synthesis of quaternary ammonium salt thiol 2</td>
<td>37</td>
</tr>
<tr>
<td>3.3.3 Synthesis of quaternary ammonium salt thiol 3</td>
<td>40</td>
</tr>
<tr>
<td>3.3.4 Synthesis of quaternary ammonium salt thiol 4</td>
<td>40</td>
</tr>
</tbody>
</table>
Corresponding disulfide 5

4 Dramatic increase in the folding rate of Bovine Pancreatic Trypsin Inhibitor (BPTI) with the buffer composed of positively charged aromatic thiol and its corresponding disulfide .......................................................... 52
   4.1 Abstract .......................................................................................... 52
   4.2 Introduction ..................................................................................... 53
   4.3 Results ............................................................................................. 57
      4.3.1 Folding with negatively charged aromatic thiols and their corresponding disulfides ............................................................................. 59
      4.3.2 Folding with positively charged aromatic thiol 1 and its corresponding disulfide 5 ................................................................. 61
5 Targeted Molecular Dynamics (TMD) simulation study of conformational folding from [5-55] like conformation to native conformation of BPTI

5.1 Abstract .................................................................................. 81
5.2 Introduction .............................................................................. 82
5.3 Methods .................................................................................. 85
  5.3.1 Molecular dynamics (MD) setup for equilibrium and TMD .......... 85
5.4 Results and discussion .............................................................. 86
  5.4.1 TMD using all atoms ............................................................. 87
    5.4.1.1 Change in structural configuration of BPTI during TMD ....... 87
    5.4.1.2 Distance between CYS residues of the native disulfide bonds ... 89
    5.4.1.3 Root mean square deviation during TMD ......................... 90
    5.4.1.4 Analysis of intermediates with the change in radius of gyration .. 91
    5.4.1.5 Analysis of folding trajectories ...................................... 93
  5.4.2 TMD using Cα and sulfur of cysteines (SG) only ..................... 94
    5.4.2.1 Change in structural conformation during simulation .......... 94
    5.4.2.2 Root mean square deviation analysis ............................... 95
    5.4.2.3 Distance analysis of native disulfide bonds ....................... 96
    5.4.2.4 Analysis of intermediates with the change in radius of gyration .. 98
    5.4.2.5 Analysis of trajectories ................................................. 99
5.5 Conclusion .............................................................................. 101

6 Conclusion .............................................................................. 102

7 Future works ........................................................................... 104

References .................................................................................. 105

VITA ............................................................................................ 116
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Four levels of protein structure</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Thiol-disulfide interchange reaction during renaturation of protein</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Mechanism of aggregation of amyloid fibrils</td>
<td>7</td>
</tr>
<tr>
<td>1.4 Structure of human PDI</td>
<td>10</td>
</tr>
<tr>
<td>1.5 Open and closed CXXC motif of PDI</td>
<td>10</td>
</tr>
<tr>
<td>1.6 Proposed mechanism of thiol-disulfide exchange between Ero1p and a thioredoxin-like domain of PDI</td>
<td>11</td>
</tr>
<tr>
<td>1.7 Oxidative protein folding mechanism in vivo</td>
<td>12</td>
</tr>
<tr>
<td>1.8 Some traditional small molecule thiols and disulfides</td>
<td>14</td>
</tr>
<tr>
<td>1.9 Generic structures of aromatic thiol and disulfide</td>
<td>16</td>
</tr>
<tr>
<td>1.10 Structure of BPTI</td>
<td>18</td>
</tr>
<tr>
<td>1.11 Thermal denaturation of BPTI at pH 2.1</td>
<td>18</td>
</tr>
<tr>
<td>1.12 The BPTI folding pathway proposed by Creighton</td>
<td>22</td>
</tr>
<tr>
<td>1.13 The folding pathway of BPTI proposed by Kim</td>
<td>23</td>
</tr>
<tr>
<td>1.14 Updated oxidation-rearrangement BPTI folding pathway</td>
<td>26</td>
</tr>
<tr>
<td>3.1 Quaternary ammonium salt (QAS) thiols 1-4 and their corresponding disulfides 5-8</td>
<td>36</td>
</tr>
<tr>
<td>3.2 Sulfonic acid thiol 9 and its disulfide 10</td>
<td>36</td>
</tr>
<tr>
<td>3.3 Side products of the first step in the preparation of compound 16</td>
<td>39</td>
</tr>
<tr>
<td>3.4 Tosylation of 16</td>
<td>40</td>
</tr>
<tr>
<td>4.1 Mechanism of PDI</td>
<td>55</td>
</tr>
<tr>
<td>4.2 Different aromatic thiols and their corresponding disulfides</td>
<td>58</td>
</tr>
</tbody>
</table>
4.19 Chromatogram showing native protein and intermediates with 0.25 mM disulfide 5 and 5 mM thiol 1 ................................................................. 70

4.20 Folding of reduced BPTI with 10 mM thiol 1 and different concentrations of disulfide 5 ......................................................................................... 71

4.21 Comparison of folding of reduced BPTI with our optimal condition (10 mM thiol 1 and 0.25 mM disulfide 5) with optimal condition of traditional buffer (5 mM GSH and 5 mM GSSG) ......................................................... 71

4.22 Effect of negatively and positively charged thiols on net charge of BPTI .............................................................................................................. 74

5.1 Native structure of BPTI showing α-helices (in red), β-strands (in green), disulfide bonds (in yellow), and loops and turns (in blue) ......................... 87

5.2 The folding secondary structure of BPTI with its three disulfide bonds .... 88

5.3 Snapshots of conformations formed at different stages of conformational changes during conformational folding of BPTI using TMD simulation. Sulfur atoms of cysteines are shown in yellow balls .................................................... 88

5.4 Evolution of secondary structure from TMD simulations of BPTI [(5-55) to N] ........................................................................................................ 89

5.5 Distance between sulfur atoms in native disulfide bonds .......................... 90

5.6 Change in RMSD during TMD of BPTI ......................................................... 91

5.7 Folding trajectories demonstrating the BPTI folding pathway from [5-55] conformation to folded native state ......................................................... 93

5.8 Snapshots of different conformations formed during TMD simulation of [5-55] like conformation of BPTI to native like BPTI ........................................... 94

5.9 Snapshots of conformations formed at different stages of conformational changes during conformational folding of BPTI using TMD simulation .... 95

5.10 Change in RMSD during TMD ..................................................................... 96

5.11 Distance between sulfur atoms in native disulfide bonds .......................... 97

5.12 Formation of native and non-native disulfide bonds ............................... 98
5.13 Folding trajectories demonstrating the BPTI folding pathway from [5-55] conformation to folded native state .......................................................... 100

5.14 Snapshots of different conformations formed during TMD simulation of [5-55] like conformation of BPTI to native like BPTI ............................................. 101
## LIST OF SCHEMES

<table>
<thead>
<tr>
<th>SCHEME</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Nucleophilic substitution in thiol-disulfide interchange reaction</td>
<td>28</td>
</tr>
<tr>
<td>1.2 Rearrangement of protein disulfide with aromatic thiolate ion</td>
<td>28</td>
</tr>
<tr>
<td>1.3 Formation of protein disulfide with aromatic disulfide</td>
<td>29</td>
</tr>
<tr>
<td>3.1 Synthesis of quaternary ammonium thiol 1</td>
<td>38</td>
</tr>
<tr>
<td>3.2 Synthesis of elongated quaternary ammonium thiol 2</td>
<td>39</td>
</tr>
<tr>
<td>3.3 Synthesis of hexyl QAS thiol 3</td>
<td>40</td>
</tr>
<tr>
<td>3.4 Synthesis of octyl QAS thiol 4</td>
<td>41</td>
</tr>
<tr>
<td>3.5 Synthesis of sulfonic acid thiol 9</td>
<td>41</td>
</tr>
<tr>
<td>4.1 Intra- and intermolecular formation of disulfide bonds</td>
<td>54</td>
</tr>
<tr>
<td>4.2 Reactions of aromatic thiol as a nucleophile and/or central thiol in a redox buffer</td>
<td>73</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BMC</td>
<td>(±) Trans-1,2-bis(mercaptoacetamido)cyclohexane</td>
</tr>
<tr>
<td>BME</td>
<td>Beta mercaptoethanol</td>
</tr>
<tr>
<td>BPTI</td>
<td>Bovine pancreatic trypsin inhibitor</td>
</tr>
<tr>
<td>CHARMM</td>
<td>Chemistry at Harvard Macromolecular Mechanics</td>
</tr>
<tr>
<td>CTD</td>
<td>C-terminal domain</td>
</tr>
<tr>
<td>Cu²⁺</td>
<td>Copper ion</td>
</tr>
<tr>
<td>Cₓ</td>
<td>Alpha carbon</td>
</tr>
<tr>
<td>DIPEA</td>
<td>Diisopropylethylamine</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethylfumaramide</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>DTTox</td>
<td>Oxidized dithiothreitol</td>
</tr>
<tr>
<td>DTTred</td>
<td>Reduced dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>Ero1p</td>
<td>Endoplasmic reticulum oxidoreductase protein 1</td>
</tr>
<tr>
<td>FAD</td>
<td>Flavin adenine dinucleotide</td>
</tr>
<tr>
<td>GSH</td>
<td>Glutathione</td>
</tr>
<tr>
<td>GSSG</td>
<td>Glutathione disulfide</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical user interface</td>
</tr>
<tr>
<td>Gdn-HCl</td>
<td>Guanidinium chloride</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
</tbody>
</table>
HPLC  High performance liquid chromatography
kDa  Kilodalton
KCl  Potassium chloride
MD  Molecular dynamics
NAMD  Nanoscale molecular dynamics
NMA  N-methylmercaptoacetamide
NMR  Nuclear magnetic resonance
PA  Phosphonic acid
PCDs  Protein conformational disorders
PDB  Protein data bank
PDI  Protein disulfide isomerase
QAS  Quaternary ammonium salt
R_g  Radius of gyration
RMSD  Root mean square deviation
RP  Reverse phase
SA  Sulfonic acid
SG  Sulfur of cysteine
S_N2  Substitution nucleophilic bimolecular reaction
TFA  Trifluoroacetic acid
TMD  Targeted molecular dynamics
UV-vis  Ultraviolet-visible
VMD  Visual molecular dynamics
CHAPTER 1

Introduction

1.1 Structure and folding of protein

1.1.1 Protein structure

Gerhardus Johannes Mulder, a Dutch chemist, first described and coined the name ‘protein’ in 1838 meaning ‘the first quality.’ Proteins are polymers of amino acids and are composed of nitrogen, carbon, hydrogen, oxygen and sulfur. In humans, protein consists of 15.1 percent of a person’s dry weight. Proteins are enzymes, hormones, antibodies, and major components of tissues such as hair and muscles. Proteins play a primary role in most biological processes such as building, repairing and replacing tissues.

Protein structure is described in four levels - primary, secondary, tertiary, and quaternary (Figure 1.1). The primary structure of a protein is the actual sequence of amino acids in that protein which is set by DNA. The primary bonds of any protein are the peptide bonds between amino acids. Biosynthesis of a protein’s primary structure occurs on the ribosome with information provided by messenger RNA in a process called translation. The secondary structure of a protein is a local structural conformation, which is determined by H-bonds within the backbone. There are mainly two types of secondary structure in proteins: the \( \alpha \)-helix and the \( \beta \)-sheet. The tertiary structure of a protein is the overall three-dimensional structure of a single protein molecule which is described by distant interactions between groups. These interactions include H-bonding, Vander Waals interactions, hydrophobic packing, and disulfide bonding. The quaternary
structure of a protein is formed by interactions between individual protein subunits. These four levels of a protein structure are crucial to the stability of the folded native conformation of a protein.

Figure 1.1. Four levels of protein structure (From reference 4).

1.1.2 *In vitro* protein folding

1.1.2.1 Process of *in vitro* protein production

The advancement in recombinant DNA technology has allowed the production of therapeutic proteins in microbial hosts; *Escherichia coli* is the most
common organism used. In many cases, the protein of interest misfolds and forms insoluble aggregates known as inclusion bodies.\(^6\) The major desirable factor associated with the production of inclusion bodies either in the cytoplasm or in the periplasm of *E. coli* is the rapid formation of the desired protein.\(^7\) Additional advantages associated with the overexpression include issues related to toxicity of desired protein to the host cell are eliminated by the production of that particular protein via inclusion bodies formation,\(^8\) the large quantity of inclusion bodies expressed is highly enriched with a particular protein, and inclusion bodies are protected from proteolytic degradation.\(^8\) But, there are still some disadvantages of the expression of the protein as inclusion bodies such as; the of inactivity and insolubility of the protein within the inclusion bodies, poising a difficulty in solubility for further steps of protein production.\(^6\) Inclusion body formation is particularly important in the case of disulfide containing proteins, as the disulfide bond formation requires an oxidative environment which is not found in the cytosol of bacteria.\(^9\) The result is the aggregation of protein. However, the challenge remains with the regeneration of active and soluble proteins from these misfolded aggregates.\(^5,9\) The recent advancement in the restoration process has made it possible to produce large quantities of therapeutic proteins from highly dense inclusion bodies which reach to 90% at optimal conditions.\(^5\)

Formation of active protein from inclusion bodies involves several steps. The first step is the isolation and purification of the inclusion bodies. Isolation is done using one of two procedures. The most common method is cell lysis followed by centrifugation of the resulting suspension at modest rotor speed.\(^10\)
The other method of isolation is the separation of inclusion bodies from soluble proteins by filtration. The isolated inclusion bodies are contaminated mainly with membrane associated proteins. The membrane associated proteins and other contaminants are washed away with a mixture containing EDTA and a low concentration of denaturant or detergent such as deoxycholate, octylglucoside, or Triton X-100. The next step is the solubilization of the purified inclusion bodies which is commonly accomplished by using strong denaturants such as urea or guanidinium chloride (Gdn-HCl). During the solubilization of the inclusion bodies the protein is either completely unfolded or partially unfolded as the intramolecular and intermolecular interactions are disrupted. In the case of disulfide-containing proteins, proteins within the inclusion bodies contain both non-native inter- and intra-molecular disulfide bonds, which diminishes the solubility. Therefore, the addition of thiol containing reagents such as β-mercaptoethanol, dithiothreitol (DTT), dithioerythritol, cystamine, glutathione (GSH), or cysteine along with the denaturants mentioned above at alkaline pH is required. These reagents are necessary to break disulfide bonds via thiol-disulfide interchange reactions forming reduced protein which is more soluble than protein with mismatched or intermolecular disulfide bonds. The last step in the in vitro protein production via the formation of inclusion bodies is the renaturation of the solubilized inclusion proteins. The excess of denaturants and thiol reagents are first removed using different methods such as dilution, dialysis, gel-filtration chromatography, immobilization on a solid support, etc. followed by the exposure of reduced protein to oxidizing conditions. The refolding process is
highly dependent on pH, temperature, and ionic strength of the protein; therefore, these factors must be carefully considered during the optimization of the renaturation step.\textsuperscript{5}

The renaturation of disulfide containing proteins must be performed in a redox buffer containing an oxidizing agent and in the case of monothiols a reducing agent.\textsuperscript{5,9} For example, glutathione (GSH), the reduced form, and glutathione disulfide (GSSG), the oxidized form, mixed in the proper molar ratio can be utilized to increase the rate and yield of the formation of an active protein by facilitating thiol-disulfide interchange reactions. Other redox systems which can be used are pairs like DTT/GSSG, cysteine/cystine, cysteamine/cystamine.\textsuperscript{8} A small amount of EDTA is added during the preparation of the redox buffer to prevent the oxidation of the protein thiols and redox buffer by air in the presence of metal ions such as Cu$^{2+}$.

![Thiol-disulfide interchange reaction](image)

Figure 1.2. Thiol-disulfide interchange reaction during renaturation of protein where RS$^{-}$ is the deprotonated reduced form and RSSR is the oxidized disulfide form.

The renaturation process of disulfide containing proteins can be improved by the addition of low molecular weight additives which prevent the aggregation of wrongly folded species by suppressing intermolecular interactions.\textsuperscript{8} Different denaturants such as urea, GdnHCl, and detergents such as CHAPS, Triton X-100, SDS, CTAB can be added as small molecule additives.
1.1.2.2 Importance of protein folding

The protein folding process can potentially start in several different places: in the cell co-translationally during protein synthesis on the ribosomes, in the cytoplasm after the protein has been synthesized and released from the ribosome, or in a particular compartment of the cell such as the endoplasmic reticulum (ER) or mitochondria after the protein has been translocated through a membrane. The ER contains the specialized machinery used to fold intra-organellar, secretary, and transmembrane proteins. The highly crowded milieu (300-400 gL\(^{-1}\) occupied mainly by proteins, polysaccharides, and lipids) of the cell increases the possibility of protein aggregation. To overcome the aggregation problem, molecular chaperones are present in the cell. These chaperones prevent aggregation by not allowing the nascent protein molecules to interact with other molecules in the crowded environment of the cell. The chaperones guide protein folding either in the initial stage or later stages of the process, increasing the efficiency of the folding process.

In the case of incorrect folding, the aggregation of protein results in the formation of highly ordered amyloid fibrils (Figure.1.3), which are very stable and are responsible for neurodegenerative diseases. They have been grouped under protein conformational disorders (PCDs). In some cases, the aggregated proteins deposit in tissues, including heart and brain, resulting in diseases such as Alzheimer's disease, Parkinson's disease, spongiform encephalopathies, and type II diabetes. The misfolded proteins, as a result of a change in conformation lose their biological activity and acquire toxicity to the cell.
The misfolded proteins associated with the protein misfolding diseases have an extra stability. Therefore, during *in vitro* protein folding, it is very important to fold proteins into their correct three-dimensional shape, so as to obtain biological function, and not toxic activity.

The *in vitro* oxidative protein folding of disulfide containing proteins is done in the presence of a redox buffer containing both oxidizing and reducing equivalents. The progress of the folding process is studied by analyzing the intermediates formed, which can be isolated after quenching the reaction mixture after a certain amount of time. Oxidative protein folding method is applicable solely for the folding of disulfide containing proteins. Oxidative protein folding method has some advantages over the conformational protein folding process,
which is used to fold all types of proteins. These benefits include the use of a protein with clearly defined disulfide bonds, and the ease of trapping, isolation, and characterization of folding intermediates, as well as evaluation of the kinetics of the folding pathway, and control of kinetics by changing the redox reagents.¹⁹

### 1.1.3 Folding of disulfide containing proteins

Disulfide bond formation in secretary proteins is an essential process for their stability and biological activity as it decreases the entropy of the reduced form and thus increases the relative stability of the native form. However, in comparison with other proteins, the process of folding disulfide containing proteins is very slow and requires a redox environment.²⁰ Oxidative protein folding *in vivo* is an assisted process which takes place in the ER of the cell. The ER contains protein disulfide isomerase (PDI), a protein involved in the rearrangement of mismatched disulfide bonds and oxidation-reduction of disulfide bonds. The ER has an oxidative environment compared to the cytosol as a result of the presence of different redox enzymes and small molecule oxidants, and also because of the secretion and uptake of thiols and disulfides, respectively.²¹

### 1.1.4 Folding with Protein Disulfide Isomerase (PDI)

#### 1.1.4.1 General information

The enzyme PDI belongs to the thiol-disulfide oxidoreductase family and is located in the endoplasmic reticulum (ER) of the cell²². It was first isolated and characterized by Anfinsen and co-workers.²³ Protein disulfide isomerase (PDI) is a 55 kDa, multifunctional protein found abundantly in the lumen of ER and
primarily takes part in the formation of correct disulfide bonds via isomerization of disulfide bonds.\textsuperscript{24} The structure of PDI contains two active thioredoxin domains (\(a\) and \(a'\)), two inactive thioredoxin domains (\(b\) and \(b'\)) and an acidic C-terminal tail resulting in a “U” shaped structure (Figure 1.4).\textsuperscript{24} The active domains (\(a\) and \(a'\)) each contain one catalytically active C-G-H-C motif located near the N terminus in the \(a\) domain and near the C terminus in the \(a'\) domain. The two C-G-H-C motifs face each other and are separated by 20 Å (distance between sulfur of C61-C406).\textsuperscript{24,25} The cysteines in each active site can be either in the disulfide form (oxidized PDI), because of an intramolecular bond formation, or in the dithiol form (reduced PDI). In reduced PDI, one thiol in each active site is exposed to the solvent and the other thiol is buried in the hydrophobic core. The solvent exposed thiol is responsible for the reactivity of PDI (pKa = 6.7) with other disulfide bonds. Oxidized PDI contains one disulfide bond in a cyclic 14-membered ring form (Figure 1.5) and is unstable as compared to other disulfides as a result of a lower redox potential (\(E'_0 = -180 \text{ mV}\)).\textsuperscript{26} Thus, thiol-disulfide interchange reactions take place rapidly with oxidized and reduced PDI, making PDI a model for designing small molecule thiols and disulfides.
Figure 1.4 Structure of human PDI (PDB ID 4EKZ).\textsuperscript{22}

Figure 1.5 Open and closed CXXC motif of PDI.

1.1.4.2 Oxidative protein folding with PDI

1.1.4.2.1 Oxidative \textit{in vivo} protein folding with PDI

Once the polypeptides are formed, these nascent structures containing cysteines in the reduced thiol form are translocated to the ER with the help of signal recognition particles as well as receptors.\textsuperscript{27} In the ER, the conditions are much more oxidizing than in the cytosol ([GSH]/[GSSG] = 1:1 to 1:3 in ER, 100:1 in the cytosol), which is thermodynamically favorable for disulfide bond formation in proteins and hence PDI is present mostly in the oxidized form.\textsuperscript{25,28,29} The protein disulfide bond is formed due to the transfer of oxidizing equivalents from Ero1p to PDI\textsuperscript{29} (Figure 1.6).
The PDI itself acts as a chaperone thereby preventing proteins from forming an aggregation. The reduced proteins form disulfide bonds with the help of oxidized PDI as oxidized PDI act as an electron acceptor whereas mismatched disulfide bonds within proteins are isomerized to native disulfide bonds with the help of reduced PDI which acts as an electron donor. The isomerization of the mismatched disulfide bond is started by the attack of the solvent-exposed thiolate of PDI on disulfide bond of the substrate protein forming a mixed disulfide bond between the protein and PDI (Figure 1.7). Then, the resulting thiol of the substrate protein can attack another disulfide bond. A series of thiol-disulfide reaction results in the formation of native disulfide bonds in the substrate protein. The regaining of oxidized PDI will take place via the series of reactions involving Ero1p and its cofactor Flavin Adenine Dinucleotide (FAD), and molecular oxygen.

1.1.4.3 Oxidative in vitro protein folding with PDI

In vitro oxidative protein folding can be aided by the addition of PDI, which also catalyzes in vivo protein folding in the ER of the cell via a series of thiol-disulfide interchange reactions. For in vitro protein folding, PDI is mixed with a redox buffer which improves protein folding by preventing the unfolded proteins
from aggregating and increasing the rate of thiol-disulfide interchange reaction.\textsuperscript{32,33} The use of PDI for \textit{in vitro} protein folding is not efficient as a consequence of its high cost of production, low catalytic activity, and instability.\textsuperscript{32,34} However, \textit{in vivo} in the ER, the concentration of PDI is high.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{oxidative_folding.pdf}
\caption{Oxidative protein folding mechanism \textit{in vivo}.\textsuperscript{35}}
\end{figure}

\textbf{1.1.4.4 Folding with small molecules}

Oxidative \textit{in vitro} protein folding of disulfide containing proteins is conventionally done using a redox buffer prepared with a small molecule aliphatic thiol and/or the corresponding disulfide.\textsuperscript{36,37} As these small molecules contain either a thiol or a disulfide group in their structure; they can react with protein thiols/disulfides in multiple steps until the native protein is formed. The native protein is a kinetically stable protein structure having only native type disulfide
bonds. Some of the traditional small molecule thiols and disulfides include glutathione (GSH) and glutathione disulfide (GSSG), reduced dithiothreitol (DTT\textsuperscript{red}) and oxidized dithiothreitol (DTT\textsuperscript{ox}), reduced and oxidized \(\beta\)-mercaptoethanol (BME), reduced and oxidized cysteine, reduced and oxidized (±) trans-1,2-bis(mercaptoacetamido)cyclohexane (BMC), reduced and oxidized cystamine\textsuperscript{38,39} (Figure 1.8). In general, the reduced and oxidized forms of the small molecules are mixed to prepare the redox buffer.\textsuperscript{40,41} Out of the above-mentioned redox systems, only two of them have been studied extensively for their effect on protein folding: DTT\textsuperscript{ox}/DTT\textsuperscript{red} and GSSG/GSH.\textsuperscript{42}

The GSSG/GSH system is found \textit{in vivo}, where folding of disulfide containing proteins take place. Therefore, the GSSG/GSH redox system is the one most commonly used for oxidative protein folding experiments especially for proteins like BPTI, RNase A, and lysozyme.\textsuperscript{43} The optimal conditions to fold these proteins are: 0.2 mM GSSG and 1 mM GSH for RNase A,\textsuperscript{44} 2 mM GSSG and 7 mM GSH for lysozyme,\textsuperscript{45,46} and 2 mM GSSG and 2 mM GSH or 5 mM GSSG and 5 mM GSH for BPTI.\textsuperscript{42}

Dithiothreitol (DTT) has a reduction potential (\(E^\circ\)) of -0.327 V; therefore, it is a weak oxidizing agent.\textsuperscript{47} Even in the case of high molar concentration of oxidized DTT, the refolding rate of RNase A is very slow as compared to GSSG at much lower concentrations.\textsuperscript{96} The advantage of using DTT is that DTT catalyzes protein folding without the formation of stable mixed disulfide intermediates; hence, the characterization of the folding pathway becomes easier.\textsuperscript{36}
Figure 1.8 Some traditional small molecule thiols and disulfides.\textsuperscript{43}
For the large-scale production of proteins, less expensive oxidative folding agents are desired. Therefore, cysteine, cystamine, and β-mercaptoethanol are applicable as these compounds are cheaper than glutathione. Cystamine has a lower thiol pKa than glutathione and has a net positive charge at pH 7; hence, it is more advantageous for basic proteins like BPTI.\textsuperscript{43}

An aliphatic dithiol (±)-trans-1,2-bis(2-mercaptoacetamido)cyclohexane (BMC) has two thiol groups in close proximity and upon oxidation can form a comparatively less stable cyclic disulfide. The yield of protein folding can be increased significantly with BMC\textsuperscript{31} as demonstrated by Raines \textit{et al.} during the folding of RNase A. When 1 mM BMC was added to a folding mixture containing 1 mM GSH and 0.5 mM GSSG, the yield of protein was two-fold as compared to folding yield with \textit{N}-methylmercaptoacetamide (NMA),\textsuperscript{36} which is the monothiol analog of BMC. The additional thiol group present in BMC helps increase the folding yield of proteins in comparison to NMA. Therefore, it is desirable to add an even low concentration of BMC to the folding mixture whether \textit{in vivo} or \textit{in vitro} to increase the yield.

Aromatic thiols and their corresponding disulfides (Figure 1.9) can be utilized more efficiently as a redox buffer to fold disulfide containing proteins because of their enhanced reactivity at physiological pH and lower thiol pKa values (4-7) as compared to their aliphatic counterparts. In addition, the thiol pKa is similar to that of the solvent exposed thiol of protein disulfide isomerase (PDI).\textsuperscript{48} The enhanced reactivity is primarily dependent upon the better leaving group ability of the thiolate ion due to the higher stability of the thiolate anion as a
result of the lower pKa values of the corresponding thiols. Hence, small molecule aromatic thiols and disulfides enhance the thiol disulfide interchange reactions occurring during folding experiments.

![Aromatic thiol and disulfide](image)

Figure 1.9 Generic structures of aromatic thiol and disulfide.

The folding rates of reduced RNase A and lysozyme were increased multiple folds with aromatic thiols in comparison to GSH. For example, folding of reduced RNase A was increased by the factor of 23 at pH 6, by the factor of 12 at pH 7 and by a factor of 8 at pH 7.7 relative to GSH.\(^{38,49}\) Similarly, folding experiments on reduced lysozyme showed that the folding rate was increased by the factor of 11 and yield by 40% with aromatic thiols and the corresponding disulfides relative to GSH and GSSG at pH 7.\(^{50}\) Aromatic thiols with thiol pKa values 1-2 units lower than the pH of the folding mixture were found to the best in terms of the reactivity. Therefore, we sought to synthesize aromatic thiols with expected thiol pKa values 5.5-6.5 as we are running folding experiments at pH 7.3 to mimic physiological conditions. I expect to dramatically increase in the folding rate of reduced protein with aromatic thiols and corresponding disulfides.

1.2 Folding of bovine pancreatic trypsin inhibitor (BPTI)

1.2.1 Introduction of BPTI

Bovine pancreatic trypsin inhibitor (BPTI) with 58 amino acid residues is one of the smallest globular proteins known, and its sole function is to bind to and inhibit serine proteases such as trypsin. The sequence of amino acids in BPTI is
With 10 positively charged lysine (K) and arginine (R) residues and 4 negatively charged aspartate (D) and glutamates (E) residues, the protein is strongly basic at neutral pH and is sometimes called basic pancreatic trypsin inhibitor. However, as the usual source for BPTI is bovine pancreas, the protein is primarily referred to as bovine pancreatic trypsin inhibitor. Residue 15 of BPTI, which is lysine, has a long basic side chain on an exposed loop of the structure. Lysine is responsible for binding the specificity pocket in the active site of trypsin hence inhibiting its enzymatic activity. Trypsin inhibitors usually have conserved cysteine residues that participate in forming disulfide bonds. BPTI has $\alpha$-helical and $\beta$-sheet regions as well as three disulfide bonds between Cys5-Cys55, Cys14-Cys38, and Cys30-Cys51, which stabilize the protein's tertiary structure\(^5\). Three disulfide bonds in 58 residues make BPTI one of the most stable proteins known. BPTI was first isolated as a trypsin inhibitor from bovine pancreas in 1936.\(^5\) The crystal structure was solved by Robert Huber in 1970.\(^5\) The first NMR structure was determined by Kurt Wuthrich at ETH in Zurich in 1980.\(^5\) BPTI was the first protein to be studied computationally using molecular dynamics.\(^5\) BPTI is also the most-studies disulfide containing protein in terms of folding pathway.

BPTI is quite inert to denaturants like urea at below 100°C but denatures in very acidic solutions; the midpoint for reversible denaturation is 81°C at pH 2.1 (Figure 1.11).\(^5\) If only one of the disulfide bonds, that formed between cysteine residues 14 and 38, is reduced and then carboxymethylated, the midpoint
decreases to 59°C, hence denaturation becomes more facile than without carboxymethylation (Figure 1.11).

Figure 1.10. Structure of BPTI.58

Figure 1.11. Thermal denaturation of BPTI at pH 2.1 (From reference 59).59

When all three disulfide bonds of BPTI are reduced at room temperature, the protein unfolds. Upon oxidation under suitable conditions, native BPTI with its unique set of three disulfide bonds is formed. If BPTI were reduced to yield six cysteine residues and randomly re-oxidized, it would produce about 7% native
protein (1/15), as there are 15 possible combinations of BPTI with three disulfide bonds (Table 1.1). The first SH group to pick a partner will have five choices, the second SH group three partners, and the last SH group one partner. Therefore there are $5 \times 3 \times 1$, or 15, possible combinations.$^{60,61}$

Table 1.1: 15 possible combinations of six cysteine residues in BPTI

<table>
<thead>
<tr>
<th>Combinations</th>
<th>1st Disulfide bond</th>
<th>2nd Disulfide bond</th>
<th>3rd Disulfide bond</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5-14</td>
<td>30-38</td>
<td>51-55</td>
</tr>
<tr>
<td>2</td>
<td>5-14</td>
<td>30-51</td>
<td>38-55</td>
</tr>
<tr>
<td>3</td>
<td>5-14</td>
<td>30-55</td>
<td>38-51</td>
</tr>
<tr>
<td>4</td>
<td>5-30</td>
<td>14-38</td>
<td>51-55</td>
</tr>
<tr>
<td>5</td>
<td>5-30</td>
<td>14-51</td>
<td>38-55</td>
</tr>
<tr>
<td>6</td>
<td>5-30</td>
<td>14-55</td>
<td>38-51</td>
</tr>
<tr>
<td>7</td>
<td>5-38</td>
<td>14-30</td>
<td>51-55</td>
</tr>
<tr>
<td>8</td>
<td>5-38</td>
<td>14-51</td>
<td>30-55</td>
</tr>
<tr>
<td>9</td>
<td>5-38</td>
<td>14-55</td>
<td>30-51</td>
</tr>
<tr>
<td>10</td>
<td>5-51</td>
<td>14-30</td>
<td>38-55</td>
</tr>
<tr>
<td>11</td>
<td>5-51</td>
<td>14-38</td>
<td>30-55</td>
</tr>
<tr>
<td>12</td>
<td>5-51</td>
<td>14-55</td>
<td>30-38</td>
</tr>
<tr>
<td>13</td>
<td>5-55</td>
<td>14-30</td>
<td>38-51</td>
</tr>
<tr>
<td>14*</td>
<td>5-55</td>
<td>14-38</td>
<td>30-51</td>
</tr>
<tr>
<td>15</td>
<td>5-55</td>
<td>14-51</td>
<td>30-38</td>
</tr>
</tbody>
</table>

*native BPTI
Previous studies of BPTI and other proteins containing disulfide bonds indicated that correct pairing of cysteines is achieved if appropriate conditions are selected and sufficient time is allowed. The meaning of the finding is that the folding path of the protein places the SH groups in position for correct pairing. The corollary of this statement is that the S-S bridges are not essential for proper refolding. But, the formation of S-S bond contributes to the extra stability of the protein structure once the protein is folded. A protein containing S-S bridges has a smaller number of conformations available in the unfolded form than does a comparable protein without the bridges. Consequently, it shows a lower entropy gain on unfolding and is therefore stabilized.

1.2.2 Traditional in vitro oxidative folding of BPTI

The in vitro oxidative folding of fully reduced BPTI was traditionally carried out in the presence of DTT\textsuperscript{ox} and GSSG. The folding pathways were determined from the study of trapped intermediates. The folding study of BPTI was originally examined by Creighton who refolded reduced BPTI (R) at pH 8.7 in the presence of DTT\textsuperscript{ox} followed by the trapping of folding intermediates chemically using iodoacetamide to alkylate free thiols. These intermediates were separated via ion-exchange chromatography followed by analysis using two-dimensional paper electrophoresis. The study identified two main 1SS (contains one disulfide bond) intermediates, five main 2SS (contains two disulfide bonds), intermediates and only one 3SS (contains three disulfide bonds) protein structures (Figure 1.12). Out of the two 1SS intermediates only one is native like (contains [30-51] disulfide bond), out of the five 2SS intermediates, three are native like, N' [30-51;
14-38], N* [5-55; 14-38], and N^{SH} [30-51; 5-55], and the one 3SS containing structure is the native protein N [30-51; 5-55; and 14-38]. The disulfide bond found in the most abundant one disulfide intermediate, [30-51], is present in all two-disulfide intermediates identified except in N*. Two of the 2SS intermediates containing non-native disulfide bonds, [30-51; 5-14] and [30-51; 5-38], were highly populated in Creighton’s study.

From Creighton's work, it was revealed that the two well populated 2SS non-native intermediates have an important role in guiding the folding pathway to the native structure. These two 2SS intermediates which contain (30-51) as one of the two disulfide bonds have similar kinetic behavior and did not form substantially during the folding process but rearranged to another disulfide intermediate N^{SH} [30-51; 5-55] which further rearranges rapidly to native BPTI N [30-51; 5-55; and 14-38]. The native like 2SS intermediate N* [5-55; 14-38] was found to be very stable kinetically and only very slowly rearranged to N^{SH} [30-51; 5-55], so the intermediate was called non-productive. The folding pathway of BPTI goes through two well populated non-native 2SS intermediates, [30-51; 5-14] and [30-51; 5-38] and the route is favored kinetically.

Later, the oxidative folding study of reduced BPTI was studied extensively by Kim and his team using a modified separation technique. The folding reaction was quenched at certain time points using an acid to stop further thiol-disulfide interchange reactions. Acid quenching has the advantage over the conventional method in that the process is reversible and the trapped and purified intermediates can be used for further folding experiments so that the folding
process can be studied in detail. The intermediates formed were separated and analyzed by reverse phase HPLC (RP-HPLC). In his study, Kim found that only six different species accumulated during folding and that all contained native disulfide bonds: [30-51], [5-55], N' [30-51; 14-38], N* [5-55; 14-38], N^{SH} [30-51; 5-55], and N [30-51; 5-55; 14-38] (Figure 1.13).

Figure 1.12 The BPTI folding pathway proposed by Creighton. R is reduced form and N is native form of BPTI. Numbers with dash sign inside bracket represent the disulfide bonds present in respective structures.

The fully reduced BPTI (R) was oxidized to all possible single disulfide intermediates which rapidly rearrange to either [30-51] or [5-55]. These single disulfide intermediates are converted quickly to two disulfide intermediates N', and N* as the thiols of Cys14 and Cys38 are solvent exposed and form a
disulfide bond easily. As $N^*$ has buried Cys30 and Cys51 in its interior, the thiols of Cys30 and Cys51 are inaccessible to oxidizing agents and hence $N^*$ cannot undergo either oxidation or rearrangement on the experimental time scale; therefore, the route that forms $N^*$ is regarded as nonproductive. On the other hand, the pathway via $N'$, which also has two buried thiols, Cys5 and Cys55, is productive because this intermediate can rearrange to the more stable intermediates $N^*$ and $N^{SH}$. The transition to $N^{SH}$ does not depend upon the concentration of oxidizing agents, which means that the transition is via an intramolecular thiol-disulfide interchange reaction. Conversion of $N'$ to $N^{SH}$ is the rate determining step in protein folding. In the final step leading to the formation of native BPTI, oxidation of Cys14 and Cys38 takes place forming the (14-38) bond.

Figure 1.13 The folding pathway of BPTI proposed by Kim. $R$ is reduced form and $N$ is the native form of BPTI. The numbers inside the box separated by a dash are disulfide bonds. $N'$, $N^*$, and $N^{SH}$ are native like 2SS intermediates. Very fast means the process can be finished within milliseconds and very slow means it could take months.

Subsequent studies by Kim et al. of the oxidative folding of reduced BPTI using GSSG concluded that the slow rate of direct oxidation of $N'$ to $N$ is because of the native-like structure of $N'$ which slows down the intermolecular reaction
with GSSG to form a mixed disulfide as well as the subsequent intramolecular reaction leading to $N$.\textsuperscript{68} The formation of $N$ by direct oxidation requires high concentration of GSSG which also leads to the formation of a dead-end doubly mixed disulfide intermediate; therefore, folding via direct oxidation of $N'$ does not form $N$ in considerable amounts. During oxidative folding, two reactions were seen in competition. One is the formation of $N$ from the singly mixed disulfide $N'(SG)$, and the other is the formation of $N'(SG)_2$ from $N'(SG)$. The formation of singly mixed disulfide $N'(SG)$ is from the oxidation of one of the free thiols of Cys5 or Cys55. At modest GSSG concentration, the formation of singly mixed disulfide $N'(SG)$ is slower than the rearrangement of $N'$ to $N^{S\text{H}}$. As the rate of formation of $N$ from $N^{S\text{H}}$ is 150 times faster than the rate of formation of singly mixed disulfide $N'(SG)$ from $N'$, the formation of $N$ takes place via intramolecular rearrangement. At high GSSG concentrations (> 500 $\mu$M), formation of doubly mixed disulfide $N'(SG)_2$ from singly mixed disulfide $N'(SG)$ is 30 times faster than the formation of $N'(SG)$ from $N'$; therefore, the process does not allow the formation of a considerable amount of $N$. The conclusion is that complete unfolding of native like intermediates $N'$ and $N^*$ as seen in BPTI folding, is common to all disulfide containing proteins.\textsuperscript{68}

1.2.3 Updated folding pathway of BPTI

Recent investigation by Yingsong and Lees on the folding of reduced BPTI via a growth type pathway and updated the BPTI folding pathway (Figure 1.14).\textsuperscript{69} A growth type pathway is a protein folding model where the rate determining step(s) in the folding reaction is the nucleation of native-like conformation
followed by the smooth formation of conformation with native disulfide bonds.\textsuperscript{70} The folding condition optimized previously by Kibria and Lees with GSH and GSSG were 5 mM GSH and 5 mM GSSG to fold reduced BPTI to 93% native protein in 48 h.\textsuperscript{42} Conformational kinetic trap N* and oxidative kinetic trap N'(SG)\textsubscript{2} which is also called doubly mixed disulfide, were observed in the reaction and were balanced under optimal conditions. The formation of protein doubly mixed disulfide with 5 mM GSH and 5 mM GSSG suggested that growth type pathways were needed to efficiently fold reduced BPTI. Wang \textit{et al.} determined the rate constants for many of the steps that occur during the folding of reduced BPTI. On the basis of these rate constants, the best folding conditions for BPTI were predicted assuming only two changes to the redox buffer of GSH/GSSG were made during the folding process. The conversion took place via the formation of N*(SG). The folding of reduced BPTI was initiated with 2 mM GSSG and 5 mM GSH to convert all reduced BPTI to disulfide intermediates, N' and N* were formed as primary intermediates occupying 50% of the total concentration. Then, 30 mM GSH was added after 15 min, total GSH concentration was 35 mM, to reduce the rest of the intermediates like N', N'(SG), N'(SG)\textsubscript{2}. The reduced proteins were converted to the thermodynamically more stable intermediate N* as these intermediates react with GSSG. After 1 h, 50 mM GSSG was added and the reaction continued for 12 h to produce 93% N. N* was previously regarded as a dead-end kinetic trap. The finding was helpful to increase the yield of native protein because folding via the productive route: N' to N\textsuperscript{SH} to N or N' to N'(SG) to N was limited by the formation of doubly mixed disulfide N'(SG)\textsubscript{2}. 
Figure 1.14 Updated oxidation-rearrangement BPTI folding pathway.\textsuperscript{65,68,69}

1.3 Folding of BPTI with aromatic thiols and their corresponding disulfides

1.3.1 Introduction

The enhanced reactivity of aromatic thiols and their corresponding disulfides towards thiol-disulfide interchange reactions and the feasibility of their synthesis in lab are very important aspects for their application in protein folding. The pKa values of aromatic thiols can be altered by changing the side chain attached to the aromatic ring. Therefore, the synthesis of small molecule aromatic thiols with nearly the same thiol pKa as the free thiol of PDI, and that are soluble in water in both the thiol and disulfide forms, should enhance the reactivity of these small molecules aromatic thiols at pH 7 relative to the traditionally used aliphatic thiols such as glutathione. The enhancement of the folding rate of proteins with aromatic thiols in comparison to glutathione was demonstrated previously in our group. As described, aromatic thiols speed up the folding rate of lysozyme by 11 and 7 times and increased its yield by 40% and
25% at pH 7 and 8 respectively. Additionally, the rate of folding for RNase A was also found to be enhanced up to 23 times at pH 6, up to 12 times at pH 7, and up to 8 times at pH 7.7 with aromatic thiols as compared to glutathione. Therefore, replacing glutathione with aromatic thiols in the redox system can expedite the rate of folding of disulfide containing proteins at neutral pH and a further increase in relative rate can be achieved by decreasing the pH of the redox buffer.

1.3.2 Construction of the structure of small molecule aromatic thiols

The structure of aromatic thiols should have the following characteristics: these small molecule thiols should have charged side chain so that the thiols are soluble in buffer, the thiol group should have a pKa value very close to neutral pH, and the molecules should be easily separated from protein after the folding assay is completed so that it will not hamper the analysis. The charged groups are either negatively charged like sulfonic acid and phosphoric acid, or positively charged like quaternary ammonium salt. These groups are installed in the structure during their synthesis. The thiol pKa values can be calculated from the \( \sigma_p \) a value of the substituent. The Hammett plot has a slope of -1.6, \( \rho = -1.6 \pm 0.1 \).

The formation of disulfide bonds takes place via a nucleophilic substitution mechanism. A thiolate ion (R\(_1\)S\(^-\)) attacks the disulfide bond (R\(_2\)SSR\(_3\)) and substitutes (R\(_3\)S\(^-\)), as a result, a new disulfide bond (R\(_1\)SSR\(_2\)) is formed as shown in scheme 1.1. In the reaction, R\(_1\)S\(^-\) acts as a nucleophile, R\(_3\)S\(^-\) serves as a leaving group, and R\(_2\)S\(^-\) serves as the center of attack. The rate of the thiol-
disulfide interchange reaction is dependent upon several factors including the stability of the leaving group and the nucleophilicity of the attacking group.

Scheme 1.1 Nucleophilic substitution in thiol-disulfide interchange reaction

Therefore, the basis for the design of small molecule aromatic thiols for protein folding is to enhance the rate of thiol-disulfide interchange reactions which take place between the small molecules and proteins during oxidative \textit{in vitro} protein folding. The rate of rearrangement of protein disulfide bonds found in the folding intermediates can be a slow process, as I have observed in the folding of reduced BPTI; hence, aromatic thiolate ions will expedite the process (Scheme 1.2). Native disulfide bonds formation with aromatic thiols involve reduction of non-native disulfides to thiols and rearrangement of non-native disulfide bonds to native disulfide bonds. Aromatic disulfides also help in increasing the rate of formation of protein disulfide bonds (Scheme 1.3) because aromatic thiolates are stable leaving groups. Our aim to fold BPTI faster is highly dependent upon the rate of formation and rearrangement of protein disulfides with the aid of aromatic thiols and their corresponding disulfides.

Scheme 1.2 Rearrangement of protein disulfide with aromatic thiolate ion
1.4 Molecular dynamics (MD) simulation study of BPTI

1.4.1 Importance of MD simulation

Molecular dynamics (MD) simulation is a computational method which is regarded as a primary method for the theoretical study of the biomolecules at their atomistic levels to insight into their structure, function, and dynamics. Atoms and molecules are interacted for the fixed time providing a view of the details of the interactions of the system. The method of MD simulation has its application in wide range of science including physics, chemistry, biochemistry, biophysics, and structural biology. Growing advancement of methodology and computer power is helping to study the larger systems and variety of conformational changes utilizing the longer time of simulation to find the details of a process going on inside the body which is not possible from experiments. The three-dimensional structure of proteins is predicted via the MD simulation of folding of a random coil of that particular proteins.

1.4.2 Protein folding and MD simulation

The knowledge of a folding process of a newly synthesized protein helps to understand the basic principle of life at its atomistic levels. The field of medicine can take the advantages of the molecular dynamics (MD) simulation of protein folding process as the study can facilitate the development of the
treatment methods of diseases associated with protein misfolding. The study of protein folding helps the field of nanotechnology because the investigation can reveal the principle behind the folding. Both the experimental and computational methods have been advancing to elucidate the process of conformational changes of proteins. MD simulation has been improvised and coupled with the ultrafast, inexpensive computers to study the unfolding and folding of proteins in their atomistic levels.

1.4.3 Molecular dynamics (MD) simulation of BPTI

Bovine pancreatic trypsin inhibitor (BPTI) is the first protein which was tested with molecular dynamics (MD) simulations in vacuo. After eleven years, MD simulation of BPTI done in water at the same time was reported stating that the simulation was run for higher time scales and improved the results giving better RMSD value and reducing the number of incorrect hydrogen atoms as opposed to the simulation run in vacuo. The presence of water molecules in simulation provides more realistic motion of protein and structure will attain very close to the crystal structure. Later on, simulation on BPTI has been performed in different laboratories.

Bovine pancreatic trypsin inhibitor (BPTI), as it contains three disulfide bonds, has an extra stability in its native form. During simulation, if the disulfide bonds are included as constraints, can reduce the conformational space. Saito and coworkers were the first to study the protein folding/unfolding of disulfide containing proteins assuming that folding starts with the formation of secondary structure (α-helices and β-sheets) followed by the attaining the tertiary
The simulation on finding the role of disulfide bonds in protein folding revealed that folding of disulfide containing proteins undergoes with different pathways to the native form via the formation of native-like intermediates only, but the process is mediated by the formation of non-native intermediates. The results were in agreement with both the conflicting experimental results by Creighton, and Weissman and Kim. Another study on disulfide bonds and protein folding by Scheraga and coworkers concluded that the conformational folding and the disulfide bond formation are extremely cooperative process as well as the development and breaking of disulfide bonds in the disulfide containing proteins takes place when certain conditions such as closeness of cysteine residues, orientation, and easiness of thiols to oxidizing agents are fulfilled. Lately, MD simulation works on BPTI revealed that the initially formed specifically collapsed structure of protein similar to native form guides the formation of disulfide bonds. Herein, the importance of conformations of intermediates during folding in the formation of disulfide bonds is described, taking BPTI as a model protein, and using targeted molecular dynamics (TMD) simulation. Targeted MD is a great computational tool to study the conformational changes of starting structure to an assigned target structure with the application of time-dependent geometrical constraint, usually run at ordinary temperature. As the process of conformational folding and disulfide bond formation are coupled to each other, the TMD simulation study on BPTI would give an opportunity to understand the mechanism of conformational changes and disulfide bond formation visually.
CHAPTER 2

Objectives

The aims of my dissertation research are to synthesize aromatic thiols and their corresponding disulfides, and to investigate the folding of reduced BPTI using these compounds.

2.1 Synthesis of different aromatic thiols and their corresponding disulfides for the study of folding of reduced BPTI

Four different QAS thiols and their corresponding disulfides as well as one sulfonic acid thiol and its disulfide will be prepared with the expectation that they will improve the efficiency and folding rate of reduced BPTI.

2.2 Folding of reduced BPTI faster using different aromatic thiols and their corresponding disulfides

Three different aromatic thiols (PA, SA, and QAS) are used to fold reduced BPTI at physiological pH selecting different concentrations of the thiols and the corresponding disulfides to prepare the redox buffer. The best thiol in our folding study as well as the best condition will be determined and the results compared with the traditional folding buffer GSSG/GSH.

2.3 Molecular dynamics (MD) simulation study of BPTI intermediates to understand the folding of BPTI at a molecular level

The molecular mechanism of the conformational changes at an atomistic level will be modeled. The software NAMD will be used for the simulation study and the software VMD will be used for visualization of the simulation.
CHAPTER 3

Synthesis of aromatic thiols and corresponding disulfides

3.1 Abstract

The use of redox buffers containing an aromatic thiol and its corresponding disulfide were found to increase the in vitro folding rate of disulfide containing proteins compared to buffers composed of an aliphatic thiol and its corresponding disulfide. The increased folding rate with aromatic thiols at physiological pH, 7.3, is attributed to their lower thiol pKa values (4-7), greater nucleophilicity at neutral pH, and enhanced leaving group ability. Oxidative protein folding in the endoplasmic reticulum (ER) of a cell is aided by protein disulfide isomerase (PDI). The enzyme PDI contains two active sites each with a CXXC motif where one of the cysteine is buried and the other is solvent exposed and has a thiol pKa of 6.7. I synthesized five different aromatic thiols and their corresponding disulfides with expected pKa values close to the pKa of the solvent exposed thiol of PDI: Four positively charged quaternary ammonium salt thiols and their corresponding disulfides, and one negatively charged sulfonic acid thiol and its corresponding disulfide. The folding of reduced bovine pancreatic trypsin inhibitor (BPTI) was performed with a redox buffer composed of one positively charged thiol and its disulfide, and one negatively charged thiol and its disulfide. The results showed that the positively charged thiol folded reduced BPTI faster as compared to the traditional redox buffer composed of aliphatic thiol, glutathione (GSH), and its disulfide glutathione disulfide (GSSG). Although the negatively charged aromatic thiol and its disulfide were found to fold
reduced BPTI rapidly, the protein also precipitated. It is expected that the other positively charged thiols and their corresponding disulfides will also fold reduced BPTI more efficiently.

3.2 Introduction

The number, importance, and use of protein-based drugs are expected to increase dramatically over the next five years. The total number of protein-based drugs approved worldwide is around 650. Out of these 650, about 400 of them are produced using recombinant DNA technology. An additional 1300 drugs are being developed. Almost all of these drugs contain disulfide bonds and a third of them are produced in E. coli. Overexpression of proteins in E. coli is efficient as large amounts of protein are rapidly produced. However, overexpressed disulfide-containing proteins tend to misfold and aggregate as inclusion bodies inside E. coli. The aggregated protein is then resolubilized and folded oxidatively in vitro to obtain active protein. Oxidative protein folding in vitro as well as in vivo involves the conformational folding of proteins combined with the oxidation of protein thiols to disulfides, and the reduction and rearrangement of protein disulfides. During rearrangement non-native disulfide bonds are rearranged to native disulfide bonds.

Oxidative protein folding has traditionally been performed with a redox buffer of glutathione (GSH) and glutathione disulfide (GSSG), as these aliphatic small molecules are found in the endoplasmic reticulum (ER) of eukaryotes where in vivo oxidative protein folding takes place. The slow step(s) in oxidative protein folding in vitro involve(s) thiol-disulfide interchange reactions. The thiol-
disulfide interchange reactions during protein folding occur between protein thiol/disulfide and small molecule thiol/disulfide, where thiolate ion, deprotected thiol, attacks a disulfide. As the reaction goes through the nucleophilic substitution mechanism, the rate of the reaction depends on the concentration of the thiolate ion in the solution, which ultimately depends on the thiol pKa and the pH of a solution. In eukaryotic cells, thiol-disulfide interchange reactions are catalyzed by protein disulfide isomerase (PDI). One of the properties of PDI that is believed to be important for catalysis is a nucleophilic low pKa thiol in the active site. The PDI is effective but is expensive for in vitro oxidative protein folding since it needs to be added in almost stoichiometric amounts.

To improve in vitro protein folding, I sought to prepare small molecules with nucleophilic low pKa thiols similar to PDI. Aromatic thiols have thiol pKa values (pKa = 4-7) comparable to PDI and lower than those of aliphatic thiols, such as GSH (pKa = 8.7). In addition, aromatic thiols are more nucleophilic than aliphatic thiols with similar pKa values. Gough et al. demonstrated that a redox buffer composed of aromatic thiols and disulfides increased the folding rate of ribonuclease A 10 times in comparison to a redox buffer composed of GSH and GSSG. These redox buffers composed of small molecule disulfides and corresponding thiols, in variable concentrations, oxidize protein thiols to disulfides as well as rearrange mismatched non-native disulfide bonds within protein to native disulfide bonds.

Herein, four different quaternary ammonium salt (QAS) thiols 1-4 and their corresponding disulfides 5-8 (Figure 3.1), and the sulfonic acid (SA) thiol 9 and
its disulfide 10 (Figure 3.2) were synthesized. Compounds 1, 5, 9, and 10 were synthesized using previously reported methods \(^{50,97-99}\) and utilized for the folding of BPTI. The redox buffer, composed of QAS thiol 1 and its disulfide 5, folded reduced BPTI to about 90% completion within one hour without any indication of protein precipitation. The redox buffer, composed of SA thiol 9 and its disulfide 10, folded reduced BPTI faster, but was found to be less suitable to be used as a redox buffer since it resulted in protein precipitation. Compounds 2-4 and 6-8 were synthesized for the first time.

Figure 3.1. Quaternary ammonium salt (QAS) thiols 1-4 and their corresponding disulfides 5-8.

Figure 3.2. Sulfonic acid thiol 9 and its disulfide 10.
3.3 Results and discussion

Four different quaternary ammonium thiols (1-4) and their corresponding disulfides (5-8), and one sulfonic acid thiol (9) and its disulfide (10) were synthesized. Disulfides were synthesized by dissolving the corresponding thiols in water and stirring in the presence of air until the disulfides were formed. The completion of disulfide formation was confirmed using Ellman’s reagent. Quaternary ammonium salt thiols with increased number of carbons in the side chain were selected, as thiol 1 had shown significant improvement in protein folding rates and yields.\textsuperscript{34,46} Increasing the hydrophobic side chain may result in thiols more likely to enter the hydrophobic core of proteins.

3.3.1 Synthesis of quaternary ammonium salt thiol 1

The QAS thiol 1 was synthesized as shown in scheme 3.1. To protect the SH group $p$-toluene thiol (11) and benzoyl chloride (12) were reacted in the presence of trimethylamine to form compound 13. The protected compound 13 was then subjected to a radical bromination reaction with $N$-bromosuccinamide in the presence of light to provide bromide 14. Compound 14 reacted with trimethylamine to form QAS salt 15. The benzoyl protecting group was removed from 15 by refluxing 15 with 8% HBr.

3.3.2 Synthesis of quaternary ammonium salt thiol 2

The QAS thiol 2 was prepared following scheme 3.2. Benzyl bromide 14 was reacted with $\beta$-mercaptoethanol in the presence of diisopropylethylamine (DIPEA) to form alcohol 16. Compound 16 was then reacted with carbon tetrachloride and triphenylphosphine to obtain bromide 17. Compound 17 was
reacted with trimethylamine and formed QAS salt 18. The benzoyl protecting group was removed from 18 by refluxing 18 with 8% HBr.

Scheme 3.1. Synthesis of quaternary ammonium thiol 1

There were lots of challenges during the synthesis of 2. The first step was the formation of an alcohol. Initially, it was challenging to find the correct base that would initiate the reaction. Bases tried were potassium carbonate, potassium hydroxide, sodium carbonate, cesium carbonate, silver carbonate, sodium hydroxide, and DIPEA. Only potassium carbonate and DIPEA gave positive results, however DIPEA was selected as it gave the highest yield and was easy to dissolve. As the reaction should go through a $S_N2$ mechanism, the polar aprotic solvent DMF was selected. There was also a regioselective issue. The thiolate from $\beta$-mercaptoethanol could attack either the benzylic carbon or the
carbonyl carbon forming two different products. The best ratio obtained was 1:1.

Two different side products along with 42% of product were observed in the 
$^1$HNMR spectra of the crude mixture. The proposed the following side products  
(Figure 3.3) based on the NMR.

**Scheme 3.2. Synthesis of elongated quaternary ammonium thiol 2**

![Scheme 3.2. Synthesis of elongated quaternary ammonium thiol 2](image)

Figure 3.3. Side products of the first step in the preparation of compound 16.
The second step of formation of bromide 17 was attempted via the formation of the tosylate. The product was the chloride rather than the expected tosylate. The chlorinated compound did not react with trimethylamine so the reaction was proceeded with the Appel reaction.

\[
\begin{align*}
\text{S} \quad \text{OH} \\
\text{S} \quad \text{Ph} \\
\text{O} \\
\text{Cl} \\
\text{CH}_3 \\
\text{N} \\
\text{S} \quad \text{Cl} \\
\text{S} \quad \text{Ph} \quad \text{O} \\
\text{Chlorinated product}
\end{align*}
\]

Figure 3.4. Tosylation of 16.

**3.3.3 Synthesis of quaternary ammonium salt thiol 3**

Benzyl bromide 14 was reacted with \(N,N\)-dimethylhexylamine (19) to form QAS salt 20. Compound 20 was refluxed for 35 h with 8\% HBr to deprotect the benzoyl group and form QAS thiol 3, Scheme 3.3.

\[
\begin{align*}
\text{S} \quad \text{Br} \\
\text{O} \\
\text{Ph} \\
+ \\
\text{N} \\
\text{S} \quad \text{N} \\
\text{O} \\
\text{Ph} \quad \text{Br} \quad \text{SH} \\
\text{ethanol} \\
\text{stir, 6 h, 71\%} \\
\text{8\% HBr in H}_2\text{O} \\
\text{reflux, 35 h, 97\%}
\end{align*}
\]

Scheme 3.3. Synthesis of hexyl QAS thiol 3

**3.3.4 Synthesis of quaternary ammonium salt thiol 4**

The formation of QAS thiol 4 is shown in scheme 3.4. Benzyl bromide 14 was reacted with \(N,N\)-dimethyloctylamine (21) to form QAS salt 22. Compound 22 on reflux with 48\% HBr diluted with water for 35 h formed QAS thiol 4.
3.3.5 Synthesis of sulfonylic acid thiol 9

Preparation of sulfonylic acid thiol 9 is shown in scheme 3.5. The reaction was started by treating p-aminobenzensulfonylic acid (23) with diluted sodium carbonate. The resulting sodium salt of p-aminobenzensulfonylic acid (24) was treated with nitrous acid at low temperature forming diazonium salt (25). Compound 25 was then reacted with disodium disulfide which resulted in sodium-4,4′-dithio-bis(benzenesulfonate) (26). Compound 26 was run through an H⁺ ion exchange column to form the disulfide of sulfonylic acid thiol (10). Compound 10 on treatment with triphenylphosphine in the presence of water provided sulfonylic acid thiol 9.

The following intermediates were not isolated but were formed during the synthetic pathway. Intermediate 10 was formed in between 26 and 9.
3.4 Experimental section

3.4.1 Synthesis of 1-(4-(mercaptophenyl))-N,N,N-trimethylmethanaminium bromide (1)

3.4.1.1 Synthesis of S-p-tolyl benzothiophenol (13)\(^{97}\)

In the 500-mL round bottom flask, 11.5 mL of benzoyl chloride (13.91 g, 102.6 mmol), 2.39 g of p-toluene thiol (99.7 mmol), and 100 mL of ether were mixed. The mixture was stirred and cooled in an ice-water bath. Then 18 mL of triethylamine (13.07 g, 128.2 mmol) was added to an equal volume of ether (18 mL) and the resulting solution was added dropwise to the above mixture via a dropping funnel over 30 min. The reaction mixture was subsequently left in the ice-water bath for 1.5 h and then allowed to warm to room temperature. The mixture was filtered and the solid washed with ether. The ether fractions were combined and dried with MgSO\(_4\). After rotary evaporation of the solvent, 19.00 g (83\%) of the benzoyl derivative was obtained. \(^1\)H NMR (CDCl\(_3\), 400 MHz) \(\delta\) 8.03 (dd, \(J = 8.4, 1.2\) Hz, 2H), 7.62 (tt, \(J = 7.6, 1.2\) Hz, 1H), 7.50 – 7.46 (m, 2H), 7.42 – 7.39 (m, 2H), 7.27 (d \(J = 8.0\) Hz, 2H), 2.41 (s, 3H).
3.4.1.2 Synthesis of S-4-(bromomethyl)phenyl benzothiophenol (14)

S-p-tolylbenzothiophenol (13) (10.16 g, 44.5 mmol) and N-bromosuccinimide (7.97 g, 44.7 mmol) were mixed with deoxygenated benzene (63.5 mL). The reaction mixture was irradiated with a 250W GE heat lamp, which provided sufficient light energy to start reflux. After 30 min at reflux, the mixture was cooled to 0°C, filtered and concentrated under reduced pressure. The residue was partitioned between 300 mL of CH₂Cl₂ and 150 mL of water. The aqueous layer was washed with 125 mL of CH₂Cl₂. The combined organic layers were dried with MgSO₄, filtered and concentrated. The solid was then recrystallized from hexanes to obtain 8.36 g (61%) of product. The product also contained 5% of S-4-(dibromomethyl)phenyl benzothiophenol. ¹H NMR (CDCl₃, 400 MHz) δ 8.01 (dd, J = 8.4, 1.2 Hz, 2H), 7.66 – 7.60 (t, 1H), 7.54 – 7.46 (m, 6H), 4.52 (s, 2H)

3.4.1.3 Synthesis of 1-(4-(benzoylthio)phenyl)-N,N,N-trimethylmethanaminium bromide (15)

Benzyl bromide (14) (25.0 g, 81.4 mmol) was added in a round-bottomed flask and cooled to 0°C. Trimethylamine (38.30 ml of a 33% solution in ethanol, 213.8 mmol) was then quickly added to the round-bottomed flask and the resulting mixture was stirred for 10 min. The mixture was then concentrated under reduced pressure, and the residue was recrystallized from EtOAc/EtOH (95:5) to provide 17.1 g of quaternary ammonium salt (15), 57% yield. ¹H NMR (CDCl₃, 400 MHz) δ 7.97 (d, J = 7.6 Hz, 2H), 7.81 (d, J = 8.0 Hz, 2H), 7.64 – 7.57 (m, 3H), 7.48 (t, J = 8.0 Hz, 2H), 5.19 (s, 2H), 3.41 (s, 9H).
3.4.1.4 Synthesis of 1-(4-(mercaptophenyl))-N,N,N-trimethylmethanaminium bromide (1)

Quaternary ammonium salt (15) (1.86 g, 5.08 mmol) was dissolved in 6 mL of HBr (48% w/v in water) and 30 mL of water. The mixture was then refluxed for 5 h. After cooling to 0°C, the mixture was filtered. The aqueous fraction was lyophilized several times by redissolving the residue in 10 mL of water each time. The quaternary ammonium salt thiol (1) was obtained quantitatively. \(^{1}H\) NMR (CD\(_3\)OD, 400 MHz) \(\delta\) 7.43 (s, 4H), 4.50 (s, 2H), 3.10 (s, 9H).

3.4.2 Synthesis of 2-S-(4-mercapto)benzyl)-N,N,N-trimethylethanaminium bromide [extended QAS thiol] (2)

3.4.2.1 Synthesis of 2-S-(4-(benzoylthio)benzyl)ethanol (16)

Benzyl bromide (14) (0.77 g, 2.5 mmol), \(N,N\)-diisopropylethylamine (0.32 g, 2.5 mmol), 2-mercaptoethanol (0.198 g, 2.5 mmol), and 10 mL of DMF were mixed in round-bottomed flask and stirred for 15 min. Next, 1 N HCl (100 mL) was added and the mixture was extracted with ethyl acetate (3 x 20 mL). The organic layers were combined, dried with MgSO\(_4\) and concentrated in vacuo. The residue was purified twice via silica gel chromatography using CH\(_2\)Cl\(_2\): EtOAc (10:1) and then hexane: EtOAc: CH\(_2\)Cl\(_2\) (3:1:1) as the eluent to provide alcohol (16) as a white solid, yield 0.282 g, 36%. \(^{1}H\) NMR (CDCl\(_3\), 400 MHz) \(\delta\) 8.02 (dd, \(J = 8.0, 0.8\) Hz, 2H), 7.61 (t, \(J = 7.6\) Hz, 1H), 7.51 – 7.46 (m, 4H), 7.42 (d, \(J = 8.0\) Hz, 2H), 3.77 (s, 2H), 3.70 (t, \(J = 6.0\) Hz, 2H), 2.66 (t, \(J = 6.0\) Hz, 2H), 2.05 (brs, 1H); \(^{13}C\) NMR (CDCl\(_3\), 100 MHz) \(\delta\) 190.3, 139.9, 136.7, 135.4, 133.8, 129.9, 128.9, 127.6, 126.2, 60.4, 35.5, 34.4.
3.4.2.2 Synthesis of 2-S-(4-(benzoylthio)benzyl)-1-bromoethane (17)\textsuperscript{100}

A mixture of an alcohol (16) (0.304 g, 1.0 mmol) and CBr\textsubscript{4} (0.365 g, 1.1 mmol) in CH\textsubscript{2}Cl\textsubscript{2} (2 mL) was prepared in a 25-mL flask and cooled to 0°C. Triphenyl phosphine (0.289 g, 1.1 mmol) was added via a powder funnel in portions with vigorous stirring. Then, the mixture was stirred for 15 h at room temperature. The mixture was then concentrated in vacuo and the residue purified via silica gel chromatography with a mobile phase of hexane: EtOAc (3:1) to provide 0.28 g of bromide (17) as a white solid, yield 76%. \textsuperscript{1}H NMR (CD\textsubscript{3}COCD\textsubscript{3}, 400 MHz) δ 8.03 (dd, J = 8.4, 1.2 Hz, 2H), 7.73 (t, J = 7.6 Hz, 1H), 7.60 (t, J = 8.0 Hz, 2H), 7.54 - 7.49 (m, 4H), 3.95 (s, 2H), 3.57 (t, J = 8.0 Hz, 2H), 2.92 (t, J = 8.0 Hz, 2H); \textsuperscript{13}C NMR (CD\textsubscript{3}COCD\textsubscript{3}, 100 MHz) δ 189.8, 141.4, 137.5, 136.1, 134.9, 130.7, 130.0, 128.1, 126.8, 35.8, 34.1, 31.6.

3.4.2.3 Synthesis of 2-S-(4-(benzoylthio)benzyl)-N,N,N-trimethylethanaminium bromide (18)

Bromide 17 (0.28 g, 0.76 mmol) and 33% trimethyl amine in ethanol (0.41 mL, 2.29 mmol) were mixed in a 25-mL flask. After 2 d at room temperature, the mixture was concentrated in vacuo and the residue was recrystallized from EtOAc: EtOH (95:1) to provide 0.246 g of crystals of elongated QAS salt (18), yield 75%. \textsuperscript{1}H NMR (CDCl\textsubscript{3}, 400 MHz) δ 7.98 (dd, J = 8.0, 0.8 Hz, 2H), 7.64 - 7.59 (m, 3H), 7.51 - 7.45 (m, 4H), 3.97 (s, 2H), 3.53 - 3.48 (m, 2H), 3.34 (s, 9H), 2.89 - 2.85 (m, 2H); \textsuperscript{13}C NMR (CDCl\textsubscript{3}, 100 MHz) δ 189.9, 138.7, 135.2, 135.2, 134.8, 133.0, 129.4, 127.9, 126.5, 125.5, 64.6, 52.4, 34.9, 22.8.
3.4.2.4 Synthesis of 2-S-(4-mercapto)benzyl-\(N,N,N\)-trimethylethanaminium bromide (2)

The quaternary ammonium salt (18) (1.08 g, 2.54 mmol) was dissolved in 3 mL of 48% HBr and 15 mL of water. The mixture was refluxed for 4 h and then cooled to 0°C and filtered. The filtrate was then extraction with EtOAc (2 x 30 mL). The aqueous fraction was lyophilized several times by redissolving the residue in 10 mL of water each time. The pure extended QAS thiol (2) was obtained quantitatively. \(^1\)H NMR (\(D_2O\), 400 MHz) \(\delta\) 7.34 (dd, \(J = 15.2, 8.0\) Hz, 4H), 3.78 (s, 2H), 3.32-3.28 (m, 2H), 3.01 (s, 9H), 2.80-2.76 (m, 2H); \(^{13}\)C NMR (\(D_2O\), 100 MHz) \(\delta\) 130.0, 129.9, 129.2, 128.5, 65.2. 52.8, 34.7, 22.7.

3.4.2.5 Synthesis of 1,1'-(4,4'-disulfanediylbis(1,4-phenylene)bis(2-S-\(N,N,N\)-trimethylethanaminium) bromide [Disulfide of extended QAS thiol] (6)

The elongated QAS thiol (2) (0.0053 g, 0.016 mmol) was dissolved in 40 mL of water and stirred rapidly in the presence of air. After a week, the solution was lyophilized to obtain the disulfide of elongated QAS thiol (6) quantitatively. \(^1\)H NMR (\(D_2O\), 400 MHz) \(\delta\) 7.57 (d, \(J = 8.0\) Hz, 2H), 7.36 (d, \(J = 8.0\) Hz, 2H), 3.80 (s, 2H), 3.26-3.22 (m, 2H), 2.94 (s, 9H), 2.79-2.75 (m, 2H); \(^{13}\)C NMR (\(D_2O\), 100 MHz) \(\delta\) 137.4, 135.5, 13.0, 126.3, 65.2. 52.7, 34.9, 23.2.

3.4.3 Synthesis of hexyl QAS thiol

3.4.3.1 Synthesis of 1-(4-(benzoylthio)phenyl)-\(N,N\)-dimethyl-\(N\)-hexylmethanaminium bromide (20)

Benzyl bromide 14 (2.72 g, 8.87 mmol) and \(N,N\)-dimethylhexylamine (19) (1.54 mL, 8.84 mmol) were dissolved in 9 mL of ethanol and stirred for 6 h at
room temperature. The solvent was removed under reduced pressure and the residue was washed with EtOAc (2 × 20 mL) to provide 2.75 g (71%) of product. 

$^1$H NMR (CDCl$_3$, 400 MHz) δ 8.01 (dd, $J = 8.4$, 1.2 Hz, 2H), 7.79 (d, $J = 8.0$ Hz, 2H), 7.66-7.60 (m, 3H), 7.50 (t, $J = 8.0$ Hz, 2H), 5.20 (s, 2H), 3.57-3.53 (m, 2H), 3.31 (s, 6H), 1.82 (brs, 2H), 1.34-1.28 (m, 6H), 0.89 (t, $J = 7.2$ Hz, 3H); $^{13}$C NMR (CDCl$_3$, 100 MHz) δ 189.2, 136.2, 135.5, 134.0, 133.9, 131.2, 128.8, 128.4, 127.5, 66.7, 64.1, 49.7, 31.3, 25.9, 22.9, 22.4, 13.8.

3.4.3.2 Synthesis of 1-(4-mercaptophenyl)-N,N-dimethyl-N-hexylmethanaminium bromide (3)

The quaternary ammonium salt 20 (2.75 g, 6.3 mmol) was mixed with 6 mL of 48% HBr and 30 mL of water. The mixture was then refluxed for 35 h. The contents were cooled to room temperature. The mixture was then extracted with EtOAc (2 × 30 mL). The aqueous layer was then lyophilized. The resulting residue was then repeatedly dissolved in water (10 mL) and lyophilized to provide 2.03 g of QAS thiol 3 as a white solid, 97% yield. $^1$H NMR (CD$_3$OD, 400 MHz) δ 7.42 (s, 4H), 4.49 (s, 2H), 3.33-3.29 (m, 2H), 3.01 (s, 6H), 1.87 (brs, 2H), 1.39-1.37 (m, 6H), 0.96-0.92 (m, 3H); $^{13}$C NMR (CD$_3$OD, 100 MHz) δ 138.2, 134.6, 129.9, 125.3, 68.4, 65.8, 50.3, 32.4, 27.1, 23.6, 23.5, 14.3.

3.4.4 Synthesis of octyl QAS thiol

3.4.4.1 Synthesis of 1-(4-(benzoylthio)phenyl)-N,N-dimethyl-N-octylmethanaminium bromide (22)

Benzyl bromide 14 (2.73 g, 8.88 mmol) and N,N-dimethyloctylamine (21) (1.83 mL, 8.88 mmol) were dissolved in 9 mL of ethanol and stirred for 17 h at
room temperature. The solvent was removed in vacuo. The residue was washed with EtOAc (2 × 10 mL) to remove unreacted starting materials and provide 3.91 g (95%) of product. $^1$H NMR (CDCl$_3$, 400 MHz) $\delta$ 7.99 (d, $J = 7.6$ Hz, 2H), 7.79 (d, $J = 8.4$ Hz, 2H), 7.65-7.57 (m, 3H), 7.49 (t, $J = 8.0$ Hz, 2H), 5.23 (s, 2H), 3.57-3.53 (m, 2H), 3.31 (s, 6H), 1.81 (brs, 2H), 1.35-1.24 (m, 10H), 0.87 (t, $J = 6.8$ Hz, 3H); $^{13}$C NMR (CDCl$_3$, 100 MHz) $\delta$ 189.1, 136.2, 135.5, 134.0, 133.9, 131.2, 128.8, 128.4, 127.5, 66.7, 64.1, 49.7, 31.6, 29.2, 29.0, 26.3, 22.9, 22.5, 14.0.

3.4.4.2 Synthesis of 1-(4-mercaptophenyl)-N,N-dimethyl-N-octylmethanaminium bromide (4)

The quaternary ammonium salt 22 (3.24 g, 6.97 mmol) was mixed with 7.2 mL of 48% HBr and 37.5 mL of water. The mixture was then refluxed for 35 h. The contents were cooled to room temperature. The mixture was then transferred to a separatory funnel and extracted successively with EtOAc (3 × 40 mL). The EtOAc layers were then back extracted with 20 mL of H$_2$O. The combined aqueous layers were then lyophilized. Lyophilization was repeated several times by dissolving the residue in 10 mL of water each time. The QAS thiol (4) was obtained as a white solid producing 1.86 g, 74% yield. $^1$H NMR (CD$_3$OD, 400 MHz) $\delta$ 7.42 (s, 4H), 4.49 (s, 2H), 3.33-3.28 (m, 2H), 3.01 (s, 6H), 1.87 (brs, 2H), 1.40-1.32 (m, 10H), 0.93-0.89 (m, 3H); $^{13}$C NMR (CD$_3$OD, 100 MHz) $\delta$ 138.2, 134.6, 129.9, 125.3, 68.4, 65.8, 50.3, 32.9, 30.22, 30.21, 27.4, 23.7, 23.6, 14.4.
3.4.5 Synthesis of 4-mercaptobenzene sulfonic acid (9)

3.4.5.1 Preparation of sodium-4,4’-dithiobis(benzenesulfonate) (26)

\[ p\text{-Aminobenzenesulfonic~acid~(23)~(47.5~g,~0.250~mol)~and~anhydrous~Na}_3\text{CO}_3~(13.25~g,~0.130~mol)~were~dissolved~in~water~(500~mL)~by~warming.~The~solution~was~then~cooled~to~15^\circ C.~A~solution~of~sodium~nitrite~(5~M)~was~prepared~by~adding~NaNO}_2~(18.5~g,~0.250~mol)~in~water~(50~mL)~and~added~to~the~solution~above.~The~mixture~was~then~slowly~added~to~a~mixture~of~conc~HCl~(52.5~mL,~0.640~mol)~and~300~g~of~crushed~ice.~A~suspension~of~diazo~compound~25~was~formed~and~then~stirred~for~15~min~in~an~ice~bath.~

Sodium~sulfide~nonahydrate~(65.2~g,~0.270~mol)~and~powdered~sulfur~(8.50~g,~0.270~mol)~were~dissolved~in~water~(75~mL)~by~heating~on~a~hot~plate~at~100^\circ C.~Then~a~10%~NaOH~solution~(10.0~g~NaOH~in~100~mL~H}_2\text{O})~was~added~and~the~mixture~was~cooled~to~0^\circ C~in~an~ice~bath.~Sodium~disulfide~was~formed.~The~diazo~solution~was~added~to~the~disulfide~solution~over~a~period~of~30~min,~along~with~50~g~of~ice~to~maintain~the~temperature~below~5^\circ C.~After~the~addition~was~completed,~the~vessel~was~removed~from~the~ice~bath~and~allowed~to~come~to~room~temperature.~Once~the~evolution~of~nitrogen~gas~ceased,~which~took~2~h,~the~reaction~mixture~was~acidified~to~pH~2~with~conc~HCl.~The~precipitated~sulfur~was~removed~by~filtration~and~the~filtrate~was~heated~to~remove~hydrogen~sulfide.~The~heating~was~continued~until~the~volume~reached~ca.~500~mL.~The~solution~was~then~cooled~to~room~temperature~and~neutralized~with~10%~NaOH.~The~volume~was~concentrated~to~400~mL.~The~precipitate~was~filtered~off~and~the~filtrate~was~kept~overnight~to~provide~crystallized~sodium-4,4’-
dithiobis(benzenesulfonate) (26). The crystallized product was separated by filtration and recrystallized from 80% ethanol (650 mL) to provide 21.73 g of 26, yield 42%. $^1$H NMR (D$_2$O, 400 MHz) $\delta$ 7.73 (d, $J = 8.0$ Hz, 4H), 7.64 (d, $J = 8.4$ Hz, 4H)

3.4.5.2 Synthesis of 4-mercaptobenzene sulfonic acid (9)

Sodium-4,4′-dithiobis(benzenesulfonate) (26) (4.22 g, 10 mmol) was dissolved in water and subjected to ion-exchange column chromatography (20 mL of Dowex 50WX2-200 mesh, H-form) with water as eluent. The eluted acidic fractions (150 mL) were evaporated to 35 mL and freeze dried to provide sulfonic acid disulfide (10). The residue was dissolved in methanol (100 mL), and then triphenylphosphine (6.1 g, 23.3 mmol) and 2 mL of water were added. The mixture was stirred overnight at room temperature. The solution was then evaporated under reduced pressure. The residue was dissolved in CH$_2$Cl$_2$ (60 mL) and extracted twice with water (100 mL then 50 mL). The aqueous layers were combined and washed with CH$_2$Cl$_2$ (4 × 50 mL). The aqueous layer was then freeze dried and the residue was recrystallized from benzene to give 1.56 g of 4-mercaptobenzenesulfonic acid (9) as pale yellow crystals, yield 83%. $^1$H NMR (D$_2$O, 400 MHz) $\delta$ 7.62 (d, $J = 8.0$ Hz, 2H), 7.39 (d, $J = 8.4$ Hz, 2H)

3.5 Conclusion

We have developed and implemented efficient methods to successfully synthesize thiols 2-4 and their corresponding disulfides 6-8. The difficult step was the reaction of β-mercaptoethanol with benzyl bromide 14. During the preparation
of thiol 2, three products were formed which needed to be separated carefully using silica gel column chromatography. In addition, the thiols needed to be kept in the cold to prevent air oxidation, as this can occur easily if the thiols are dissolved in water and left at room temperature. The successful preparation of these thiols and disulfides will provide new opportunities for folding disulfide containing proteins.
CHAPTER 4

Dramatic increase in the folding rate of Bovine Pancreatic Trypsin Inhibitor (BPTI) with the buffer composed of positively charged aromatic thiol and its corresponding disulfide.

4.1 Abstract

Many of the most important pharmaceuticals on the market today are proteins. A significant fraction of these proteins are produced in E. coli and need to be oxidatively folded \textit{in vitro} to obtain active protein, as they contain disulfide bonds. The drawbacks of traditional oxidative protein folding \textit{in vitro} are that it is slow and low yielding, due to protein precipitation. \textit{In vitro} oxidative protein folding traditionally involves the use of a small molecule aliphatic thiol and disulfide, such as glutathione (GSH) and glutathione disulfide (GSSG), as a redox buffer to produce the active form of the protein. To improve oxidative protein folding, small molecule thiols and disulfides were proposed that mimic the nucleophilic low pKa thiol of protein disulfide isomerase (PDI) to fold proteins. PDI catalyzes the folding of disulfide containing proteins in the endoplasmic reticulum (ER) of eukaryotic cells. I demonstrated that aromatic thiols with positively charged groups improved the folding rate of a reduced protein relative to GSH/GSSG. The folding of BPTI was followed by reverse phase HPLC of samples obtained by quenching folding reactions at specific times with formic acid. More than 90% of native protein was obtained in an hour instead of the two days taken with GSH and GSSG.
4.2 Introduction

Disulfide bonds play a vital role in stabilizing the biologically active form of numerous extracellular proteins and almost all pharmaceutically relevant proteins. Many of these proteins can be produced efficiently in bacteria; however, the proteins often aggregate as inactive inclusion bodies inside the cells. The advantage of expression of a protein in inclusion bodies is that they can be produced in large amounts, are easy to separate and are protected from proteolytic degradation. Formation of inclusion bodies may also be the best method of protein production if the protein of interest is toxic to the bacteria or host cell. To produce active protein, the protein within the inclusion bodies is resolubilized and then oxidized \textit{in vitro}. The disadvantage of the formation of inclusion bodies is the need to refold the aggregated protein \textit{in vitro}. Oxidative \textit{in vitro} protein folding involves conformational folding combined with oxidation of protein thiols to native disulfide bonds. Formation of native disulfide bonds involves reduction of protein disulfides to thiols and rearrangement of non-native disulfide bonds to native disulfide bonds (Scheme 4.1). The reaction is highly important at it is also responsible during the enzymatic formation and breaking of disulfide bonds in proteins.

Bovine pancreatic trypsin inhibitor (BPTI) with 58 amino acid residues is one of the smallest globular proteins known, and its sole function is to bind to and inhibit serine proteases such as trypsin. For example, BPTI can inhibit trypsin-like West Nile virus (WNV) NS2B-NS3 protease. Trypsin inhibitors usually have conserved cysteine residues that participate in forming disulfide bonds (Figure
Bovine pancreatic trypsin inhibitor (BPTI) has \( \alpha \)-helical and \( \beta \)-sheet regions as well as three disulfide bonds between Cys5 - Cys55, Cys14 - Cys38 and Cys30 - Cys51, which stabilize the protein’s tertiary structure.\(^{52}\) Three disulfide bonds in 58 residues make BPTI one of the most stable proteins known. When all three disulfide bonds of BPTI are reduced at room temperature, the protein unfolds. Upon oxidation under suitable conditions, native BPTI with its unique set of three disulfide bonds is formed.

**Scheme 4.1. Intra- and intermolecular formation of disulfide bonds\(^{31}\)**

The folding pathway for BPTI is shown in the sketch (Figure 1.14), which is based on previous studies of BPTI folding using GSH/ GSSG. \( N' \) and \( N^* \) are two disulfide-containing stable intermediates. \( N'(SG) \) and \( N'(SG) \) are singly mixed disulfide-containing intermediates formed when \( N' \) and \( N^* \) react with glutathione disulfide, respectively. The intermediate \( N'(SG)_2 \) is a doubly mixed disulfide formed when the free thiol of \( N'(SG) \) reacts with GSSG. \( N^{SH} \) is a rearrangement product of \( N' \) and \( N^* \) which undergoes oxidation to native protein.

*In vitro* protein folding can be aided by the addition of protein disulfide isomerase (PDI), which is found in the endoplasmic reticulum (ER) and catalyzes *in vivo* protein folding via a series of thiol-disulfide interchange reactions.\(^{31}\) The
use of PDI for *in vitro* protein folding is not efficient due its high cost of production, low catalytic activity and instability.\textsuperscript{34} However, *in vivo* in the ER, the concentration of PDI is high. The design of small molecule thiols and disulfides that improve protein folding is based on PDI. PDI is composed of two active sites each consists of a CXXC motif required for its oxidoreductase activity where C is a cysteine residue and X is any other amino acid.\textsuperscript{31,49,103} In each active site, one cysteine thiol is exposed to solvent and other is buried in the hydrophobic core of PDI. The solvent exposed thiol in PDI can attack a substrate disulfide bond resulting in the formation of a mixed disulfide bond. The mixed disulfide bond can be attacked by another thiol to form a new disulfide bond with the release of PDI (Figure 4.1). The solvent exposed thiol has a low pKa value (6.7) and at neutral pH is very reactive with disulfides as compared to small molecule aliphatic thiols.\textsuperscript{48} Therefore, the synthesis of small molecule thiols with thiol pKa values similar to that of the solvent exposed thiol in PDI is likely to increase the rate of *in vitro* folding of disulfide-containing proteins.

![Figure 4.1. Mechanism of PDI.](image-url)
A redox buffer is composed of a small molecule disulfide and its corresponding thiol in variable concentrations. During protein folding, the small molecule disulfide oxidizes protein thiols to protein disulfides and the small molecule thiol rearranges non-native protein disulfide bonds to native disulfide bonds. \(^{49}\) Both processes are completed via thiol-disulfide interchange reactions between the protein and the small molecules. There are several different factors that make aromatic disulfides/thiols better redox buffer than aliphatic disulfides/thiols. Different types of aromatic thiols with corresponding disulfides are utilized to prepare a redox buffer. The thiol pKa values of aromatic thiols range from 4-7 which is close to pKa value of the solvent exposed thiol of PDI. Aromatic thiols with either electron releasing or electron withdrawing substituents at the \textit{para} position are utilized. In addition, aromatic thiols show enhanced nucleophilicity at neutral pH and are better leaving group than aliphatic thiols. If the \textit{para} substituent is charged, then the aromatic thiol should have improved water solubility.

Historically, aliphatic thiols/disulfides such as GSH, GSSG and oxidized dithiothreitol (DTT\textsuperscript{ox}) were used for the oxidative folding of proteins. \textit{In vivo}, GSH and GSSG reduce protein disulfides and oxidize cysteine thiols to disulfides respectively. \(^{93}\) Aliphatic thiols such as GSH, DDT\textsuperscript{red}, \(\beta\)-mercaptoethanol, (±) trans-1, and 2-bis(mercaptoacetamido)cyclohexane (BMC) have thiol pKa values higher than 7. \(^{104}\) In comparison, aromatic thiols have a range of pKa values from 4 to 7 and are more reactive towards disulfides than aliphatic thiols with comparable pKa values. Therefore, the thiol pKa value and enhanced reactivity
of aromatic thiols mimic the properties of PDI. Therefore, selective synthesis of small molecule aromatic thiols with nearly the same thiol pKa as that of the free thiol of PDI, which are soluble in water in both the thiol and disulfide forms, should enhance the reactivity of these small molecule aromatic thiols at pH 7 relative to the traditionally used aliphatic thiol such as glutathione.

Herein, I describe the dramatic increase in folding rate of reduced BPTI with a redox buffer composed of an aromatic thiol and its corresponding disulfide relative to the traditional folding buffer GSH/GSSG. Out of three aromatic thiols investigated, quaternary ammonium salt thiol and its corresponding disulfide was the best for folding reduced BPTI in comparison with GSH/GSSG where BPTI is folded to 90% native protein in 48 h with 5 mM GSH and 5 mM GSSG. The folding intermediates that accumulated during folding with aromatic thiols and their corresponding disulfides were not the same as those observed with GSH/GSSG (N' and N*).

4.3 Results

Three different p-substituted aromatic thiols 27, 9, 1 and their corresponding disulfides 28, 10, 5 were chosen for the preparation of redox buffers to fold reduced BPTI (Figure 4.2). These thiols and disulfides were selected as the pKa values of the thiols are charged at neutral pH to similar to the pKa value of solvent exposed thiol of PDI (6.7). These compounds enhance water solubility. Both positively and negatively charged thiols were selected to examine the effect of charge on protein folding at physiological pH.
The protein folding experiments with aromatic thiols and their corresponding disulfides were performed at pH 7.3 using 30 µM of reduced BPTI. Variable concentrations of disulfides and thiols were used for the preparation of the redox buffer along with 0.20 M KCl, 0.10 M bis-tris-propane, and 1 mM EDTA. The folding experiment was run at 25°C under argon. The reactions were quenched with formic acid at six different time points by removing 300 µL aliquots from the reaction mixture at each time point. Each aliquot was analyzed by reverse phase HPLC. A pH of 7.3 was selected to mimic physiological conditions of protein folding and to compare the results with previously reported work. The curve of percentage native protein formed versus refolding time was plotted for every experiment performed and compared to other conditions. The highest percentage of native protein formed in the shortest time without observed protein precipitation was considered the best condition. Finally, the result was compared to the optimal conditions of the traditional folding buffer GSSG/GSH (5 mM/5 mM).

![Figure 4.2. Different aromatic thiols and their corresponding disulfides. 27: phosphonic acid thiol (PA); 28: phosphonic acid disulfide; 9: sulfonic acid thiol](image-url)
(SA); 10: sulfonic acid disulfide; 1: quaternary ammonium salt thiol (QAS); 5: quaternary ammonium salt diusulfide.50,97–99,105

4.3.1 Folding with negatively charged aromatic thiols and their corresponding disulfides

Two different negatively charged aromatic thiols, phosphonic acid thiol (PA) 27 and sulfonic acid thiol (SA) 9, were selected to study the folding of reduced BPTI. Thiol 27 has two negative charges and thiol 9 has one negative charge. Initially, reaction mixtures were prepared by adding selected concentrations of thiol 27 and its disulfide 28 to the folding mixture containing reduced BPTI and bis-tris-propane buffer at pH 7.3. Two different concentrations of disulfide 28 were chosen, 0.09 and 0.25 mM. The folding experiment was first done using 0.09 mM disulfide 28 and 1 mM thiol 27. Protein precipitation was observed almost immediately after protein folding started, as the solution became cloudy. The precipitation was worse with higher concentrations of thiols.

Next, the folding of reduced BPTI was investigated using aromatic thiol 9 and its disulfide 10. For 0.09 mM disulfide 10 concentration, 1, 2, 5, and 10 mM thiol 9 were investigated and for 0.25 mM concentration of disulfide 10, 1, 2, 5, and 10 mM thiol 9 were investigated. A reaction mixture containing 0.25 mM of disulfide 10 in combination with 1 mM of the thiol 9 folded BPTI to 82% while a reaction mixture with 2 mM thiol achieved 91% in 6 h. With 5 mM and 10 mM thiol 9, more than 90% native protein was obtained within an hour. With thiol 9 and 0.25 mM disulfide 10, all combinations of redox buffer precipitated protein. The concentration of disulfide 10 was reduced and used 0.125 mM to see if
protein precipitation was reduced. No precipitation was with 2 mM thiol 9 and folding was almost complete in 8 h forming 95% native BPTI. While with 5 mM thiol 9, 92% of native BPTI was formed in 4 h; however, protein precipitation was observed at about 6 h. The 1 mM thiol was not used as folding was slower with 2 mM thiol and also the 10 mM thiol was not used as protein was precipitated with 5 mM thiol. The concentration of the disulfide 10 was further reduced to 0.09 mM and used it in combination with 1, 2, 5, and 10 mM thiol 9. With 5 and 10 mM thiol 9 (Figure 4.3). The folding yield was higher with higher concentration of thiols and folding was faster too.

![Graph](image)

<table>
<thead>
<tr>
<th>Refolding time (min)</th>
<th>1 mM</th>
<th>2 mM</th>
<th>5 mM</th>
<th>10 mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>10.7</td>
<td>16.2</td>
<td>36.0</td>
<td>47.1</td>
</tr>
<tr>
<td>60</td>
<td>24.8</td>
<td>36.8</td>
<td>53.1</td>
<td>66.5</td>
</tr>
<tr>
<td>120</td>
<td>58.0</td>
<td>72.0</td>
<td>82.0</td>
<td>92.1</td>
</tr>
<tr>
<td>240</td>
<td>77.0</td>
<td>87.7</td>
<td>93.9</td>
<td>97.4</td>
</tr>
<tr>
<td>360</td>
<td>90.0</td>
<td>93.5</td>
<td>96.2</td>
<td>98.3</td>
</tr>
<tr>
<td>480</td>
<td>94.3</td>
<td>94.7</td>
<td>97.4</td>
<td>98.9</td>
</tr>
</tbody>
</table>

**Figure 4.3.** Folding of reduced BPTI with 0.09 mM disulfide 10 and different concentrations of thiol 9.

The protein area obtained from the analysis is shown in Figure 4.4 which showed that 0.09 mM of disulfide 10 with all concentrations of thiol 9, areas of total protein is lower than the expected. However, 0.09 mM of disulfide 10 and 1 mM of thiol 9, area of protein is very close to expected value, as there is no observed protein precipitation. Therefore, 0.09/1 mM disulfide/thiol combination is the best condition for these sets of experiments. The overall conclusion from
these folding experiments using negatively charged aromatic thiols and their corresponding disulfides was that the protein precipitates. Therefore, to fold basic proteins like BPTI faster and to obtain higher yield at shorter time, these thiols should not be considered.

**Figure 4.4.** Total area of protein in different folding conditions with thiol 9 and disulfide 10.

### 4.3.2 Folding with positively charged aromatic thiol 1 and its corresponding disulfide 5

It was proposed that a positively charged small molecule thiol would increase the net positive charge in the mixed disulfide or between the positively charged protein and small molecule and thus increase solubility. With negatively charged small molecules, the net charge of the mixed disulfide between the small molecule and positively charged protein is decreased. For this purpose, quaternary ammonium salt (QAS) thiol 1 and its disulfide 5 were used. Two
different disulfide concentrations, 0.09 mM and 0.25 mM, were first selected to optimize the condition. Eight different conditions of folding buffer were prepared using 0.09 and 0.25 mM disulfide and 1, 2, 5, and 10 mM thiol 1. The results are shown in the following Figures 4.5 and 4.6. With 1 mM thiol 1, 90% folded protein was observed in 4 h. With 2 mM thiol 1, 95% folded protein was observed in 2 h. With 5 mM and 10 mM thiol 1, folding resulted in more than 90% native protein in 1 h. In all these conditions, protein precipitation was not observed. Comparing these conditions, 0.25 mM of disulfide 5 with 5 and 10 mM of thiol 1 were our best conditions for folding reduced BPTI. In these experiments, the drawbacks of folding with QAS were the decrease of native protein percentage with higher thiol concentrations (Figures 4.5 and 4.6) and the increase in total protein area (Figure 4.7) with higher time points. Adventitious oxidation of thiol 1 was observed to be a problem at longer time points (Figure 4.8). I hypothesized that this was the reason for the described drawbacks. Therefore, thiol oxidation needs to be controlled for better results.

![Graph showing refolding time versus % native protein](image)

**Figure 4.5.** Folding of reduced BPTI with 0.09 mM disulfide 5 and different concentrations of thiol 1.
Figure 4.6. Folding of reduced BPTI with 0.25 mM disulfide 5 and different concentrations of thiol 1.

<table>
<thead>
<tr>
<th>Refolding time (min)</th>
<th>1 mM</th>
<th>2 mM</th>
<th>5 mM</th>
<th>10 mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>25.4</td>
<td>36.3</td>
<td>49.3</td>
<td>50.3</td>
</tr>
<tr>
<td>15</td>
<td>42.8</td>
<td>59.7</td>
<td>72.7</td>
<td>74.6</td>
</tr>
<tr>
<td>60</td>
<td>73.4</td>
<td>87.4</td>
<td>89.7</td>
<td>86.0</td>
</tr>
<tr>
<td>120</td>
<td>88.6</td>
<td>94.6</td>
<td>91.2</td>
<td>95.7</td>
</tr>
<tr>
<td>240</td>
<td>95.6</td>
<td>95.5</td>
<td>90.9</td>
<td>85.3</td>
</tr>
<tr>
<td>480</td>
<td>95.0</td>
<td>95.1</td>
<td>90.6</td>
<td>85.9</td>
</tr>
</tbody>
</table>

Figure 4.7. Total area of protein in different folding conditions with thiol 1 and disulfide 5.

Two intermediates were seen during the analysis with retention times of 38 min and 42 min, the amount of these intermediates decreased with the refolding time (Figures 4.9 and 4.10). These intermediates are promising for the determination of rate constants in the future.
Figure 4.8. Oxidation of thiol 1 in different folding conditions with thiol 1 and disulfide 5.

Figure 4.9. Prominent intermediate during folding of reduced BPTI with 0.25 mM disulfide 5 and thiol 1 at 38 min.

Different experiments were performed to minimize the oxidation of thiol 1 and to achieve a consistent protein area over different time points. When the reaction without protein was run, the oxidation of thiol 1 was observed, indicating poor quenching. The reaction with 5 mM thiol 1 only was also run and quenched with different concentrations of different acids and observed that 80 µL HCOOH was the best to quench the 300 µL reaction mixture (Figure 4.11).
The experiments were performed by mixing protein, thiol, disulfide, water and buffer in different ways and found that mixing all components of folding mixture in an argon environment reduced thiol oxidation dramatically. The QAS reaction (0.25 mM disulfide 5 and different concentrations of thiol 1) was repeated again with 4 h as the last time point. All solutions were bubbled with argon before mixing and mixing itself was done with argon gas bubbling. The result is summarized in graphical form (Figure 4.12).
Figure 4.12. Folding of reduced BPTI with 0.25 mM disulfide 5 and different concentrations of thiol 1.

Although the thiol oxidation was diminished, it was still observed that the total protein area was going up with higher time points. I proposed that the waiting time before the injection in HPLC was causing the increase in the total area of protein. Therefore, reactions were quenched at 5 min utilizing three different methods: analyzed by immediate injection, injection after few hours storing in ice bath, and inject next day storing at -20°C. It was found that total protein area in all samples had a close range of total area (Figure 4.13). In the previous experiments, where samples waited in the HPLC auto sampler for a long time before injection, the total protein area increased. However, the storage time for these samples were too long due to longer analysis time of HPLC which took 110 min for one sample. To minimize the storage time, the HPLC analysis time was decreased to 40 min by changing the HPLC gradient to 90% solvent A in 0 min and 60% solvent A in 40 min. Aliquots after quenching were stored at 4°C before being injected into the HPLC. With the new program, the maximum waiting time was 2 h.
Figure 4.13. Analysis of total protein area with different storage methods.

Then, folding reactions were run containing 0.25 mM disulfide 5 and 1, 2, 5, and 10 mM thiol 1 concentrations to see the effect. It was observed that total protein area was consistent and almost no thiol oxidation was observed (Figure 4.14). It was also observed that the native protein percentage was not going down with higher time points. It was concluded that the decrease in native protein percentage was because of the increase in total protein area with longer time points. The problem of decreasing native protein percentage was solved through the series of experiments. It was observed that reduced BPTI folds to more than 90% native BPTI within an hour (Figure 4.15). Likewise, there was no protein precipitation with any concentrations of thiol 1 and disulfide 5. The observation showed that the protein’s total area was close to the area of native protein. It was concluded is that 0.25 mM disulfide 5 and 10 mM thiol 1 give the best results.
Figure 4.14. Total protein area in new HPLC method.

Figure 4.15. Folding of reduced BPTI with 0.25 mM disulfide 5 and different concentrations of thiol 1.

The problem associated with new HPLC method was that the intermediate peaks were not well resolved. This did not cause any problem for the overall results because the focus was only on the folding yield of protein (Figures 4.16, 4.17, 4.18, and 4.19).

Then 10 mM thiol 1 was used and changed the disulfide 5 concentrations from 0.25 mM to 1 mM and 5 mM to see the effect. With 1 mM disulfide 5, the
result was very close to 0.25 mM, but with 5 mM disulfide concentration, the result was worse (Figure 4.20).

**Figure 4.16.** Chromatogram showing native protein and intermediates with 0.25 mM disulfide 5 and 10 mM thiol 1.

**Figure 4.17.** Chromatogram showing native protein and intermediates with 0.25 mM disulfide 5 and 1 mM thiol 1.
Figure 4.18. Chromatogram showing native protein and intermediates with 0.25 mM disulfide 5 and 2 mM thiol 1.

Figure 4.19. Chromatogram showing native protein and intermediates with 0.25 mM disulfide 5 and 5 mM thiol 1.

Then the folding of BPTI was compared using QAS with that of GSSG/GSH. The results indicated that QAS is better than GSSG/GSH for folding reduced BPTI in terms of time and yield (Figure 4.21).

From the folding results with reduced BPTI and positively charged aromatic thiol, QAS, and its corresponding disulfide, it was found that 0.25 mM QAS disulfide with 5 and 10 mM of its thiol as a redox buffer were the best conditions for BPTI folding. In this condition, BPTI can fold within 1 h to over 90%. Therefore, the method improved significantly over the traditional method of
BPTI folding using GSSG/GSH, which needed 48 h to fold to 90%. To conclude, with the thiols that have negatively charged para-substituents, protein precipitation was observed resulting in the loss of protein. In addition, protein folding was slower than QAS.

Figure 4.20. Folding of reduced BPTI with 10 mM thiol 1 and different concentrations of disulfide 5.

Figure 4.21. Comparison of folding of reduced BPTI with our optimal condition (10 mM thiol 1 and 0.25 mM disulfide 5) with optimal condition of traditional buffer (5 mM GSH and 5 mM GSSG).

4.4 Discussion

Our hypothesis was that a redox buffer composed of aromatic thiol and the corresponding disulfide would fold protein faster than a redox buffer
composed of aliphatic thiol and the corresponding disulfide. To demonstrate the
dramatic increase in the folding rate of reduced BPTI with a redox buffer
composed of aromatic thiols and corresponding disulfides, three different
aromatic thiols and corresponding disulfides were selected. The pKa of these
thiols, their structure, and pH of the solution were considerefor the selection.
Aromatic thiols have good nucleophilic character, and are better leaving groups
than glutathione at physiological pH.\textsuperscript{94,110} Thiolates of aromatic thiols 27, 9, and
1; and glutathione thiolate have approximately equal nucleophilic character, but
at physiological pH, large proportions of glutathione (pKa = 8.7) is in its inactive
thiol form while large proportions of aromatic thiols 27, 9, and 1 (pKa = 5.5-5.7)
are in their thiolate form.\textsuperscript{48} Previous work showed that at neutral pH, the
observed rate constant of the reaction with aromatic thiol is 6 times faster than
 glutathione, a small molecule aliphatic thiol, and its disulfide.\textsuperscript{49} Therefore, at pH
7.3, thiolates of thiols 27, 9, and 1 are expected to react faster than glutathione.
The leaving group ability of any thiol is inversely proportional to its pKa
values.\textsuperscript{48,94} Hence, thiolates of 27, 9, and 1 are better leaving groups than
 glutathione as their pKa values are lower.

During\textit{ in vitro} protein folding with a redox buffer containing small molecule
aromatic thiols and disulfides, there are eight possible mechanisms of thiol-
disulfide interchange reactions, six of them are shown in scheme 4.2 where I-IV
involves small molecules as a nucleophile and III-VI involve the small molecules
as a central thiol.\textsuperscript{38,49} The thiolate from an aromatic thiol can attack a protein
disulfide (reaction I), a small molecule disulfide (reaction IV), or a mixed disulfide
between the small molecule and protein (reaction II and III). In reactions, I-IV, these thiolates act as a nucleophile whereas in reactions, III-VI, the small molecule thiols act as a central thiol. In reaction V, the aromatic thiol acts as both a leaving group and a central thiol. Reactions I, III, V, and VI show a net change in protein disulfide bonds. Therefore, aromatic thiols are expected to involve in rate limiting steps in reactions I, III, V and VI. The rate constants for these reactions are expected to be greater than with glutathione.

**Scheme 4.2** Reactions of aromatic thiol as a nucleophile and/or central thiol in a redox buffer

Reduced BPTI was folded using redox buffers composed of one of three different aromatic thiols, QAS, PA, or SA and its corresponding disulfide. It was determined that the redox buffer composed of the positively charged aromatic
thiol, QAS thiol, and its disulfide could fold reduced BPTI faster and more efficiently than redox buffers composed of negatively charged SA thiol/disulfide or PA thiol/disulfide. The decrease in the folding yield of native BPTI from reduced BPTI with a redox buffer composed of 27 and 28 or 9 and 10 was proposed to be due to the decrease in net positive charge of unfolded BPTI as these negatively charged small molecules form mixed disulfides with reduced BPTI and decrease the net positive charge of BPTI (Figure 4.22). The result is the formation of aggregation. Therefore, to increase in the folding yield of BPTI, a decrease in aggregation is required.111

![Figure 4.22. Effect of negatively and positively charged thiols on net charge of BPTI.](image)
Positively charged quaternary ammonium thiol 1 and its disulfide 5 were used to prepare a redox buffer. The use of a positively charged thiol and disulfide will result in an increase in the net positive charge of reduced BPTI so that folding goes smoothly and rapidly. The results obtained from the folding of reduced BPTI using QAS thiol 1 and its disulfide 5 at pH 7.3 proved that a redox buffer composed of an aromatic thiol and corresponding disulfide increases the rate of folding dramatically relative to glutathione at their optimal conditions. Also, positively charged thiols and disulfides are better for BPTI folding. It is also noteworthy that thiols with lower pKa values undergo the thiol disulfide interchange reactions faster hence increasing the rate of formation of native protein at their optimal concentration. At higher thiol/disulfide concentrations, the folding rate of BPTI decreases.

To conclude, folding of reduced BPTI increased dramatically in the presence of a redox buffer composed of positively charged aromatic thiol QAS and its corresponding disulfide as compared to traditional redox buffer composed of glutathione/glutathione disulfide. The folding with QAS thiol and its disulfide took about an hour to fold reduced BPTI to 90% while glutathione/glutathione took about 48 h to fold reduced BPTI to 90%. The optimal condition of QAS thiol was 10 mM and disulfide was 0.25 mM at pH 7.3. Use of thiols/disulfides with negatively and positively charged groups allowed us to understand their effect on protein precipitation. For BPTI, use of positively charged thiols/disulfides with pKa close to that of free thiol of PDI helps to fold BPTI rapidly in high yield.
4.5 Experimental section

4.5.1 Materials

The protein used for the folding study, BPTI, was purchased from Roche Applied Science under the trade name Aprotinin and was used to prepare reduced BPTI directly. Trizma base, guanidinium chloride (GdnHCl), bis-tris propane, EDTA, potassium chloride (KCl), dithiothreitol (DTT), GSH, and GSSG were purchased from Sigma-Aldrich. Sephadex™ G-25 Fine was purchased from GE Healthcare and made wet for 24 h with 0.01 N HCl before packing the chromatographic column. Trifloroacetic acid (TFA) and acetonitrile (ACN) were HPLC grade and were purchased from Fischer Scientific. Concentrated hydrochloric acid was also purchased from Fischer scientific. Nanopure deionized water was prepared using a Branstead D3750 and was deoxygenated by bubbling argon for 30 min through it before using in each experiment. Aromatic thiols 27, 9, and 1 and their corresponding disulfides 28, 10, and 5 were prepared as described previously. The UV-Vis spectra were measured using a Cary 300 spectrophotometer. The HPLC analysis was performed on a Hitachi D-7000 system, which was connected with a L-7400 UV-Vis detector and a column oven. The pH measurements were done using a VWR symphony SB20 pH meter which was calibrated before its use. BPTI was lyophilized using a Labconco FreeZone 2.5 L Benchtop freeze dryer. The columns used for reverse phase HPLC (RP-HPLC) were an Alltech Macrosphere C18 preparative column (250 × 22 mm); a Vydac C18 semi-preparative column (250 × 10 mm); and a Vydac C18 analytical column (250 × 4.6 mm).
4.5.2 Preparation of reduced BPTI

Reduced BPTI was prepared from Aprotinin in Trizma buffer using reducing agents. The Trizma buffer was prepared by dissolving Trizma base, KCl and EDTA in deionized water. The pH of the buffer was adjusted to 8.7 using concentrated HCl. The solution was prepared on a 100 mL scale by maintaining the final concentration of Trizma as 0.01 M, KCl 0.20 M, and EDTA 1.0 mM. The reduction mixture was prepared in a 15 mL centrifuge tube using Trizma buffer in which GdnHCl and DTT were added. The volume of 10 mL was prepared by maintaining the final concentration of GdnHCl 6 M and DTT 0.05 M. The solution was equilibrated in 25°C water bath before its application. Aprotinin was added to the reduction mixture at a concentration of 6.5 mg/mL. The resulting mixture was kept at 25°C in a water bath for 1 h to reduce BPTI followed by the addition of 0.2 N HCl to adjust the pH between 2-3 and quench the reduction. The reduced BPTI was then purified by gel filtration on a sephadex G-25 ccolumn using 0.01 N HCl as the mobile phase. The fractions were analyzed by a UV-Vis spectrophotometer at 280 nm. The fractions that contain reduced BPTI were combined and then lyophilized before further purification by HPLC. The dried reduced BPTI was then dissolved in 0.01 N HCl and purified by RP-HPLC on a C18 preparative column. The mobile phase used was a mixture of solvent A (0.1% TFA in water) and solvent B (90% acetonitrile with 0.1% TFA in water). The elution gradient used was 0 min, 90% solvent A; 20 min, 65% solvent A; 100 min, 61% solvent A; 120 min, 50% solvent A with a flow rate of 5 mL/min. The fractions containing reduced BPTI were identified by absorbance at 280 nm. The
fractions were then further analyzed using a Vydac C18 analytical column by monitoring the absorbance at 229 nm. The flow rate of 1 mL/min was used with the elution gradient of 0 min, 90% solvent A; 15 min, 73% solvent A; 35 min, 71% solvent A; 50 min, 69% solvent A; 70 min, 65% solvent A. The pure reduced BPTI was then lyophilized and dissolved in 0.01 N HCl to prepare a stock solution of concentration 1 mg/mL. The stock solution was stored in -20°C before use.

4.5.3 Oxidative folding of reduced BPTI

To start a folding reaction, 1.5X folding buffer was first prepared. In 80 mL deionized water, 4.234 g of bis-tris propane, 2.236 g of potassium chloride (KCl), and 0.0558 g of EDTA were dissolved and the pH was adjusted to 7.3 with concentrated HCl. Then, deionized water was added until the volume was 100 mL in a volumetric flask. The buffer was deoxygenated by passing argon gas through it for 30 min. The thiols and their corresponding disulfides were purified using a Vydac C18 semi-preparative HPLC column heated to 50°C before using them in a folding buffer. Injection of samples was done manually, 2 mL at a time, and a flow rate of 3 mL/min was utilized. The gradient used was 0 min, 90% solvent A; 15 min, 75% solvent A; 35 min, 73% solvent A; 50 min, 72% solvent A; 100 min, 65% solvent A. The absorbance was monitored at 252 nm. Pure fractions of thiols and disulfides were collected, lyophilized, dissolved in deoxygenated buffer, and kept at -20°C before use. The final folding buffer for the reaction contains 0.10 M bis-tris propane, 0.20 M KCl, 1.0 mM EDTA, and indicated concentration of aromatic thiols and their corresponding disulfides.
Folding reactions were conducted under argon at 25°C. At each specific time, a 300 μL aliquot was removed and quenched by the addition of formic acid. Initially 20 μL formic acid was used but it was found from experiments done to improve folding efficiency that 80 μL formic acid quenched better than 20 μL, therefore, I used 80 μL in later experiments. Each aliquot was analyzed by RP-HPLC on a Vydac C18 analytical column. A flow rate of 1 mL/min was used. For the traditional HPLC analysis method, linear gradient was used: 0 min, 90% solvent A; 15 min, 75% solvent A; 35 min, 73% solvent A; 50 min, 72% solvent A; 110 min, 70% solvent A. The absorbance was monitored at 229 nm and the column temperature was maintained at 50°C. All peak areas were summed and the total area was assigned a value of 100%. An improved HPLC analysis method was developed for the faster analysis of samples. The flow rate was 1 mL/min and elution gradient used was 0 min, 90% solvent A; 40 min 60% solvent A.

4.6 Conclusion

Folding of reduced BPTI was investigated using both negatively and positively charged aromatic thiols/disulfides. With the redox buffer composed of negatively charged thiols, phosphonic acid thiol and sulfonic acid thiol, and their corresponding disulfides BPTI was precipitated during folding. Therefore, the conclusion was that the negatively charged aromatic thiols and their corresponding disulfides are potentially poor folding agents for basic proteins. With the positively charged thiol, QAS, and its disulfide as the redox buffer, reduced BPTI was folded without precipitation and reached 90% native form within an hour. The area of native protein and total protein area were consistent
with no protein precipitation. The percentage of native protein in 1 h time point was highest with 10 mM thiol, and 0.25 mM QAS disulfide.
CHAPTER 5

Targeted Molecular Dynamics (TMD) simulation study for conformational folding from the [5-55] like conformation to the native conformation of BPTI

5.1 Abstract

Oxidative folding of extracellular proteins includes conformational folding and disulfide bond formation which are coupled to each other during the folding process until the formation of the native protein structure. Therefore, conformational changes of these secreted proteins during folding play an important role in the formation of native disulfide bond and to complete the folding process. Targeted molecular dynamics (TMD) was used to study the conformational changes and the formation of disulfide bonds in bovine pancreatic trypsin inhibitor (BPTI) as a model protein containing disulfide bonds. The initial structure was constructed with Cys5 and Cys55 close enough to form a disulfide bond but all other cysteines are far apart and not close enough to form disulfide bonds. The TMD simulations were carried out in two different ways and investigated the conformational changes during the simulation. The folding process and the formation of the native protein was studied using visual analysis, evolution of secondary structure, disulfide bond formation, RMSD calculation, radius of gyration, and hydrogen-bond formation. With the initial simulation, targeting all atoms of the initial structure, it was found that conformational changes played an important role in the formation of the disulfide bonds as seen for the decrease in distance of Cys14-Cys38 which was not observed until the last stage of the TMD. The formation of the kinetic traps N' and N* were not
observed during our initial simulations. The final native conformation was obtained once the correct antiparallel $\beta$-sheets and subsequent Cys14-Cys38 distance came closer to form disulfide bond. In the second simulation, targeting only alpha carbons and sulfur of cysteines of the starting structure to the final structure, formation of native protein was achieved via the formation of $N^*$ followed by $N^{\text{SH}}$. The Cys14-Cys38 distance was decreased to a value close to S-S distance value and observed increased many times until the final native structure was formed.

5.2 Introduction

Folding studies of disulfide containing proteins are significant because of the diseases associated with misfolded proteins containing disulfide bonds. Experimentally, folding studies are performed by quenching the folding reaction at various stages and studying the formation of disulfide bonds. The conformational changes of these intermediates are difficult to determine with experimental procedures alone. Therefore, highly standardized computational methods are required to complement the experiments and understand the formation of disulfide bonds and protein folding.

Molecular Dynamics (MD) computational simulation is a valuable tool for investigating conformational changes of macromolecules because it can give atomic level details of conformational changes. Examples of TMD simulations used to study conformation changes include the conformational changes during the opening and closing of GroEL to unravel its ATP binding mechanism, transition from open form to closed form of ion channels to understand their
function as a ion transport such as KcsA potassium ion channels\textsuperscript{117,118}, and structural changes of the C-terminal domain (CTD) from $\alpha$-form to $\beta$-form of RfaH.\textsuperscript{119} Targeted molecular dynamics (TMD) is a robust simulation tool used to examine the process of transition from one conformation to another at normal temperature, reducing the distance continuously to the target confirmation by applying time dependent geometrical constraint.\textsuperscript{120} During TMD, the root mean square distance (RMSD), which is a measure of physical distance between two structures, to the target is continuously reduced; the system forces initial state to find the path to final state; and the Cartesian distances between the conformations describes the progress of reaction.\textsuperscript{121} The RMSD between the current coordinates and the target structure is computed in every time step by aligning the target to the current coordinates. Reasons of calculating RMSD are to find out the time point when conformation changes and to define the folding procedure. The following equation is used to calculate the force on each atom.

\begin{equation}
U_{TMD} = \left[ \frac{k}{2N} [\text{rmsd}(t) - \text{RMSD}(t)]^2 \right]
\end{equation}

where, $\text{rmsd}(t)$ is the instantaneous best-fit RMSD of the coordinates of simulated initial structure from the coordinates of targeted structure, and $\text{RMSD}(t)$ is the linear distance from initial RMSD at the first TMD step to the final RMSD at the last TMD step.\textsuperscript{89} The symbol $k$ is the spring constant which is scaled down by the number $N$ of the targeted atoms.

Herein, the formation of disulfide bonded like intermediates along with the conformational changes was studied using bovine pancreatic trypsin inhibitor (BPTI) as a model protein, which contains 58 residues with three disulfide bonds.
Experimental work by Creighton\textsuperscript{86} on the renaturation of reduced BPTI showed that the non-native single-disulfide like intermediates were formed closing Cys5 and Cys30, Cys30 and Cys55, and Cys5 and Cys51 in 1/4\textsuperscript{th} of the total isolable intermediates. Later, Weissman and Kim reexamined the folding of the reduced BPTI\textsuperscript{65} using improvised techniques such as acid quenching (quenching technique) and reverse phase high-performance liquid chromatography (HPLC) (separation technique) and found that only the native-like disulfide bonds were present in the isolated intermediates. Simulation study using a lattice model\textsuperscript{85} and theoretical study using proximity rule\textsuperscript{122} showed that non-native type disulfide bonds were formed transiently. However, only the intermediates with native type disulfide bonds were formed during the later stages of folding.

Targeted MD was used to study to see the effect of conformational changes in disulfide bond formation starting with the [5-55] like intermediate of BPTI, which is challenging to perform experimentally. The crystal structure of native BPTI was obtained from the protein data bank (PDB ID: 4PTI). The conformation was modified with pymol\textsuperscript{123} so that only the Cys5-Cys55 distance was close enough to form bond and all the other native Cys-Cys were far apart. Some important structural criteria (proximity and orientation of cysteine residues) of folding required by disulfide containing proteins were considered during the selection of the starting structure. Starting from the modified structure, TMD was performed, targeting the initial structure to the final native structure in explicit water. TMD was selected because it can provide important information on conformational changes and molecular motions that are hard to study in shorter
computational times.\textsuperscript{119} BPTI has been extensively used as a model system in many molecular dynamics simulation studies.\textsuperscript{80,124} However, to the best of our knowledge, our work represents the first investigation of the conformational changes in BPTI using TMD simulations. The conformational changes in BPTI during the folding was investigated and compared the results with previous MD simulations and experiments.

5.3 Methods

The x-ray structure of native BPTI was obtained from the crystal structure of BPTI in the protein data bank (PDB ID: 4PTI). The starting structure containing all cysteine residues but without the presence of actual disulfide bond in the structure was constructed. Cysteine 5 and 55 were close enough in the model to form disulfide bond. All atoms were used during simulation. The change in the conformation of [5-55] to native structure was first simulated in explicit solvent with CHARMM36 force fields.\textsuperscript{125}

5.3.1 Molecular dynamics (MD) setup for equilibration and TMD

The initial, [5-55] like structure for running TMD was prepared from the native structure obtained from the protein data bank (PDB ID: 4PTI). By manually adjusting some of the dihedral angles, we obtained an open form structure from the native structure which still retains the secondary structural elements. The resulting protein structure was solvated in the rectangular box of water molecules having minimum solute-wall distance of 10 Å with TIP3 water. The system was neutralized by adding chloride ions using Monte-Carlo ion placing method. I added 0.15 M KCl which generated 30 positive and 36 negative ions. CHARMM-
GUI\textsuperscript{126} web-server was used for solvation and preparation of the input files needed for simulations. The dimension of the simulation box was $80 \times 80 \times 80 \ \text{Å}^3$ with the total number of 34828 atoms in the system. Molecular dynamics (MD) simulations were performed with CHARMM\textsuperscript{36} force field using NAMD.\textsuperscript{127} The particle mesh Ewald method\textsuperscript{128} was used to calculate long range interactions using a 12 Å nonbonded cutoff. The energy minimization was done for 20 ps. The temperature of the system was maintained at 303.15 K using Langevin dynamics with a damping coefficient of 1 ps\textsuperscript{-1}. The system was energy minimized for 20 ps and equilibrated for 180 ps with NPT run using 2 fs integration time step. The RATTLE and SETTLE algorithms were used to restrain protein bonds, and to maintain water geometry respectively. A 1 ns production run was performed, followed by a TMD run for 15 ns. A biasing force of force constant per atom of $k_{\text{atom}} = \frac{k}{N} = 0.589 \text{ kcal/ (mol Å}^2\text{)}$ was used for the TMD simulation.

5.4 Results and discussion

The native structure of BPTI has two $\alpha$-helix, two $\beta$-strand, and three disulfide bonds as shown in Figure 5.1. The first $\alpha$-helix (H1 in the figure 5.1) has one and a half turns from Asp3 to Leu6 and is close to the N terminus. The second $\alpha$-helix (H2 in the figure) has three turns from Ala48 to Gly56 and is close to the C terminus. The first $\beta$-strand (B1 in the figure) lies from Ile18 to Asn24 and second $\beta$-strand (B2 in figure) lies from Leu29 to Tyr35. These two $\beta$-strands are anti-parallel to each other. Three disulfide bonds (5-55), (14-38), and (30-51) make the tertiary structure of BPTI stable by linking the secondary structures. The maximum dimension of native BPTI is 30 Å.\textsuperscript{59}
Figure 5.1 Native structure of BPTI showing $\alpha$-helices (in red), $\beta$-strands (in green), disulfide bonds (in yellow), and loops and turns (in blue).\textsuperscript{59,88} The crystal structure is taken from protein data bank (PDB ID: 4PTI).

The topology of BPTI looks like a piece of string folded twice by itself as shown in Figure 5.2. This simplified structure indicates that the (5-55) disulfide bond joins the two terminal $\alpha$-helices, the (14-38) disulfide bond joins the two $\beta$-strands near the loose end of $\beta$-hairpins. The third disulfide bond (30-51) connects the three turn $\alpha$-helix (H2) with the $\beta$-sheets; therefore, the compact native structure of BPTI is formed.

5.4.1 TMD using all atoms

5.4.1.1 Changes in structural configuration of BPTI during TMD

The [5-55] like initial structure is an open form BPTI which has the rmsd of 15 Å compared to the native structure. Targeted MD simulations of 5, 10, and 15 ns length were performed and each resulted in the conversion of the [5-55] like initial structure to the native structure of BPTI. The process of conversion of the
one disulfide intermediate [5-55] to N is shown in Figure 5.3, where snapshots taken at various stages of structural transformation demonstrate the protein’s conformations at respective stages.

**Figure 5.2** The secondary structure of BPTI with its three disulfide bonds.59,88

**Figure 5.3** Snapshots of conformations formed at different stages of conformational changes during conformational folding of BPTI using TMD simulation. Sulfur atoms of cysteines are shown in yellow balls.
The graph of the evolution of the secondary structures as a function of time during the TMD simulation is shown below in Figure 5.4. We found that the residues 18-24 transforms from a coil/turn/loop-to a $\beta$-strand after 14 ns of simulation, which occurred after decreasing S-S distance of Cys30 and Cys51 to the value close to native bond. Similarly, for residues 28-36, and 44-56, in which turn/loop/coli-to-$\beta$-strand transitions were observed after decreasing S-S distance Cys30 and Cys51 close enough to form a bond (Figure 5.4). All other residues do not show secondary structure changes during the simulation.

\textbf{Figure 5.4} Evolution of secondary structure from TMD simulations of BPTI [(5-55) to N].

5.4.1.2 Distance between CYS residues of the native disulfide bonds

With the use of TMD, the conversion of a single disulfide containing intermediate, [5-55] to native BPTI, [5-55, 14-38, 30-51], was achieved on a 15
Within 2.5 ns, the Cys30-Cys51 distance was decreased to 5.3 Å with from 21.5 Å in the starting structure. On the other hand, the Cys5-Cys55 distance was increased to 11.3 Å within 2 ns from 3.73 Å in the starting structure. The Cys30-Cys51 distance was seen closing and then increasing multiple times before a compact structure of BPTI was obtained. Formation of compact native conformation was observed at the late stage of simulation. The Cys14-Cys38 distance was decreased close to real S-S bond distance after the formation of \(N^{SH}\) [5-55, 30-51] like conformation, which proved the experimental finding that \(N^{SH}\) rearranges rapidly to N (Figure 5.5).

![Figure 5.5 Distance between sulfur atoms in native disulfide bonds.](image)

**Figure 5.5** Distance between sulfur atoms in native disulfide bonds.

### 5.4.1.3 Root mean square deviation during TMD

The root mean square deviation (RMSD) was computed from the atomic trajectories for targeted MD run. The plot of the RMSD of all backbone and heavy
atoms in the TMD-simulated structure relative to the corresponding target structure is shown in Figure 5.6. The result showed that all backbone atoms and heavy atoms used in our TMD simulations reached the target structure within 15 ns (with the accuracy of 1 Å). The initial RMSD of 8.7 Å at first trajectory continuously decreased to 1.12 Å at 3000th trajectory forming the native structure.

**Figure 5.6** Change in RMSD during TMD of BPTI.

5.4.1.4 Analysis of intermediates with the change in radius of gyration

The relationship between the conformational changes and the folding was studied by taking and analyzing snapshots of folding trajectories as depicted in Figure 5.7. The initial structure, with Cys5 and Cys55 distance close enough to form bond, had a radius of gyration (R_g) of 14.1 Å, was targeted to native structure as targeted MD simulation was started, and observed that R_g was reduced to 11.4 Å. The β-sheets and α-helices were already present in the initial structure. The folding was proceeded by crumpling up of the β-strands and
subsequent ordering. The folding was proceeded with the rearrangement of starting structure to a non-native conformation [51-55] followed by formation of [5-51]. During the process, Cys5-Cys55 distance was decreased and then increased time and forth which promoted to decrease the distance between Cys30 and Cys51. The distance of Cys30 and Cys51 was also seen decreasing and then increasing many times rearranging to some non-native like intermediates and compacting the structure of protein. The Cys5-Cys55 distance was decreased to form bond followed by the decrease in Cys30-Cys51 distance which lead to the formation of conformation of N$^{SH}$. After the flipping of the Cys30 and Cys51 residues, correct antiparallel $\beta$-strands formation took place followed by rapid closeness of Cys14-Cys38 distance leading to the compact structure of BPTI was observed. The most interesting observation during our simulation was the absence of N$'$ and N$^*$ which were key intermediates formed during oxidative folding of BPTI. Once the antiparallel $\beta$-strands were seen formed correctly during the late stage of the simulation (after the decrease in Cys30-Cys51 distance close to the bond real length), immediate decrease in the Cys14-Cys38 distance was observed that ultimately led to the native structure of BPTI. An important observation here is that the structure of N$^{SH}$ is very close to the native form structure, and that Cys14-Cys38 distance is the last to come close enough to form a bond during the folding process of BPTI. The observations are consistent with experimental findings.
Figure 5.7 Folding trajectories demonstrating the BPTI folding pathway from [5-55] conformation to folded native state. The curve shows the decrease in the radius of gyration ($R_g$) with respect to folding time. Some of the conformations of trajectory (a-j) were shown to show how $R_g$ decreases along with the progress of folding.

5.4.1.5 Analysis of the folding trajectories

In folding trajectory analysis, the complete unfolding of a conformation of single disulfide intermediate, [5-55], was found followed by the formation of single non-native intermediate [51-55]. Then, [51-55] was rearranged to another non-native like single disulfide intermediate [5-51]. The [5-51] intermediate was then rearranged to native single disulfide intermediate [30-51] which was found to be broken and formed several times. The decrease of Cys30-Cys51 distance for the last time flips one of the $\beta$-strand leaded the formation of native two-disulfide.
intermediate $N^{SH}$. The $N^{SH}$ was converted immediately to N due to the closeness of Cys14-Cys38 residues. Figure 5.8 demonstrates the transformation to target structure via the formation of different conformations.

![Figure 5.8](image)

**Figure 5.8** Snapshots of different conformations formed during TMD simulation of [5-55] like conformation of BPTI to native like BPTI. Sulfur atoms of cysteines are shown as yellow balls.

### 5.4.2 TMD using C$_α$ and sulfur of cysteines (SG) only

#### 5.4.2.1 Change in structural conformation during simulation

The change in conformation of BPTI during TMD simulation is as shown in Figure 5.9 where the snapshots taken at different time points show the progress of conformational changes from initial structure to the target structure. The 15 ns
simulation targeting only C\textsubscript{\alpha} and SG atoms only was performed at this time as opposed to targeting all heavy atoms in previous simulation. The transformation to the native structure was achieved at around 13 ns simulation bringing cysteines close to that distance as found in native disulfide bonds.

**Figure 5.9** Snapshots of conformations formed at different stages of conformational changes during conformational folding of BPTI using TMD simulation. Sulfur atoms of cysteines are shown in yellow balls.

**5.4.2.2 Root mean square deviation analysis**

The root mean square deviation (RMSD) was computed from the atomic trajectories for TMD run. The plot of the RMSD of alpha carbons and sulfur of cysteine in the TMD-simulated structure relative to the corresponding target structure is shown in Figure 5.10. The result showed that TMD simulations reached the target structure within 15 ns. The initial RMSD of 8.70 Å at first trajectory continuously decreased to 1.17 Å at 3000\textsuperscript{th} trajectory forming the native
structure which suggests that TMD successful to understand the conformational folding mechanism.

![Graph showing RMSD vs. time](image)

**Figure 5.10** Change in RMSD during TMD.

### 5.4.2.3 Distance analysis of native disulfide bonds

The formation of native BPTI [5-55, 14-38, 30-51] from the starting conformation [5-55] was achieved in 15 ns time scale. The Cys14-Cys38 distance was seen closing and then increasing multiple times during simulation. Initially the Cys14-Cys38 distance was set 8.4 Å which was decreased to 3.4 Å within 2 ns. At the meantime, Cys5-Cys55 distance was increased to 12.3 Å with in 1 ns from 3.4 Å in the starting structure. The Cys30-Cys51 distance was sharply decreased to 8.4 Å from 20.1 Å of starting structure within 1.5 ns, then increased to 18.6 Å in 4 ns. It was then seen decreasing gradually. It was observed that Cys5-Cys55 and Cys14-Cys38 distances decreasing and then increasing multiple times to form a compact structure of BPTI. Formation of $N^\text{SH}$
took place transiently at 13 ns followed by the formation of N which proved the experimental finding that $N^{SH}$ rearranges rapidly to N. Figure 5.11 shows the closing and increasing the Cys5-Cys55 and Cys14-Cys38 distances for bringing the Cys30 and Cys51 close to each other and for compacting the BPTI structure.

**Figure 5.11** Distance between sulfur atoms in native disulfide bonds.

Formation of few non-native intermediates were also observed in our simulation as shown in Figure 5.12. The first non-native intermediate [51-55] was supposed to form as Cys51-Cys55 distance was very small enough to form a bond. At around 8.5 ns, the Cys30 and Cys55 were very close, then were Cys5 and Cys30. Rearrangement at this stage allowed formation of $N^*$ which rearranged to $N^{SH}$ for a very short time and immediately went to N.
5.4.2.4 Analysis of intermediates with the change in radius of gyration

The figure 5.13 shows the decrease in radius of gyration with folding time. The snapshots of folding trajectories (a-j) at various steps of folding shows how the folding of BPTI is related to the formation of disulfide bonds by considering of radius of gyration. The starting structure has a radius of gyration ($R_g$) of 13.6 Å which on TMD simulation reached the final native structure having $R_g$ of 10.6 Å. As folding proceeded, the collapse and subsequent ordering of $\beta$-strands was observed. The folding was proceeded with the rearrangement of starting structure to non-native conformation [51-55] followed by [5-51]. The formation of
[14-38] was observed at about 2 ns. During the process, Cys5-Cys55 and Cys14-Cys38 distances were decreased and then increased multiple times which helped to compact the structure and proceed towards the native like conformation. The Cys5 and Cys38 were seen very close at around 8.5 ns. The [5-30] non-native intermediate rearranges and compacts the structure of protein by forming native like two-disulfide intermediate N* at about 12.5 ns which further rearranges to another native like two-disulfide intermediate N^{SH} transiently.

The very native like conformation N^{SH} has Cys5-Cys55 and Cys30-Cys51 in the close proximity. The Cys14-Cys38 distance decreased immediately leading to native structure. In the present simulation, N* was formed and rearranged to N^{SH}. Here too, Cys14-Cys38 distance was decreased after the formation of N^{SH} which is important in the folding process of BPTI. Our finding from simulation is very close to the experimental findings.

5.4.2.5 Analysis of trajectories

The folding proceeded via the complete unfolding followed by immediate formation of non-native [51-55] intermediate which rearranged back to [5-55] but with different conformation which then rearranged to [14-38] intermediate. The Cys5-Cys55 and Cys14-Cys38 distances were decreasing and then increasing multiple times while the conformation is changing more towards the native one. Another single disulfide containing non-native intermediate [5-30] was observed which rearranged to N*. The N* intermediate was then rearranged to N^{SH}. The decrease in Cys14-Cys38 distance was then observed bringing the stable
conformation of native BPTI. The pathway of transformation depicted in Figure 5.14 showing the intermediate structures.

![Diagram showing the folding pathway and intermediate structures of BPTI.](image)

Figure 5.13 Folding trajectories demonstrating the BPTI folding pathway from [5-55] conformation to folded native state. The curve shows the decrease in the radius of gyration ($R_g$) with respect to folding time. Some of the conformations of trajectory (a-j) were shown to show how $R_g$ decreases along with the progress of folding.
5.5 Conclusion

The actual disulfide bonds were made absent during the preparation of input files for simulation for targeted MD. The role of disulfide bond in very important for the folding of BPTI resulting in the much better approach for the final native structure. Targeted MD was run for single domain for 15 ns. The presented TMD simulations were focused on orientation of cysteine residue for disulfide bond formation as a result of the conformational changes during the folding process. The formation of stable native structure at the end of simulation without the involvement of actual disulfide bonds in simulation proves that our model can be used for the study of conformational folding mechanism of disulfide containing protein. The variations of the radius of gyration ($R_g$) along with the folding trajectories analyzed to investigate the pathway of BPTI folding.

Figure 5.14 Snapshots of different conformations formed during TMD simulation of [5-55] like conformation of BPTI to native like BPTI. Sulfur atoms of cysteines are shown as yellow balls.
CHAPTER 6

Conclusion

Folding of reduced BPTI was investigated oxidatively using three different aromatic thiols and their corresponding disulfides namely PA, SA, and QAS. The effect of charge(s) of the side chain group was also studied. The optimal folding condition using a redox buffer composed of QAS thiol and its disulfide was determined by plotting a graph of percentage of native protein formed versus refolding time for every condition selected and then comparing the graphs. Hence, an efficient folding method of reduced BPTI was determined and compared with the previously published results using GSH/GSSG. In best condition, the native form of BPTI was produced in over 90% yield in less than an hour while using GSH/GSSG under optimal condition, took 2 days to produce over 90% native protein.

As folding was most efficient using QAS thiol and its corresponding disulfide, it was sought to prepare different QAS thiols with more hydrophobic groups for further study. Three different QAS thiols (2, 3, and 4) and their corresponding disulfides (6, 7, and 8) were successfully prepared. The purity of these compounds was determined by nuclear magnetic spectroscopy and HPLC.

The folding of reduced BPTI was also studied by taking advantage of sophisticated computational programming methods developed to study molecular dynamics. Targeted molecular dynamics was used for our folding study. The folding process was studied in two different ways. In one method, all atoms of the initial structure were targeted to the target structure. It was seen that the correct
antiparallel β-sheets formation is required to form the (14-38) disulfide bond. The formation of N' [30-51; 14-38] and N* [5-55; 14-38] intermediates were not observed. The [5-55] intermediate rearranged to N^{SH} [5-55; 30-51] after a series of thiol-disulfide interchange process simultaneously with the conformational folding. Almost at the end of simulation, flipping of one β-sheet allowed the protein to form N^{SH} which immediately transformed to native protein as the (14-38) disulfide bond was formed. In other method, only the α-carbons and sulfur atoms of the cysteine residues (SG) were targeted to the target structure. The study showed the formation and breaking of (14-38) bonds multiple times. The formation of native intermediate N* [5-55; 14-38] was observed followed by the formation of N^{SH} [5-55; 30-51] transiently, which was transformed to native structure immediately. It was concluded from both of our studies that, conformational changes in the intermediate formed plays a crucial role to form disulfide bonds.
CHAPTER 7

Future Works

Folding of BPTI was very successful using both experimental and computational methods. To further improve the folding rate and yield, the use of redox buffer made up of more hydrophobic QAS thiols 2, 3, and 4 and their corresponding disulfides 6, 7 and 8 is proposed. As the hydrophobicity increases, the pKa value decreases. It is expected that the pKa of these QAS thiols will be close to 7 hence folding will go more smoothly forming less intermediates. It is also expected that the yield will also increase even though the folding will go at a slower rate.

The use of different computational methods is also proposed to investigate the folding process of reduced BPTI. Use of different shaped random structures of reduced BPTI and run simulations using different molecular dynamics is also suggested. One can mimic the oxidative folding conditions using computer programming so that better insight on the folding process can be found which will help experimentalists to think about setting up experiments in a different way.
References


(61) Conn, P. M. Molecular Biology of Protein Folding; Progress in Molecular Biology and Translational Science; Elsevier Science, 2009.


(89) Targeted Molecular Dynamics (TMD) (NAMD 2.10 User’s Guide).


(123) The PyMOL Molecular Graphics System, Version 1.7.4.5 Schrödinger, LLC.


VITA

RAM PRASAD MARAHATTA

Chhoprak-5, Gorkha, Nepal

1999-2004  B.SC., Chemistry
Tribhuvan University
Kathmandu, Nepal

2003-2005  Science Teacher
Namuna Machhindra Higher Secondary School
Lalitpur, Nepal

2009  M.Sc., Chemistry
Tribhuvan University
Kathmandu, Nepal

2009-2012  Lecturer of Chemistry
Notre Dame College
Bandipur, Nepal

2012-2017  Doctoral Candidate and Teaching Assistant
Florida International University
Miami, Florida

2016  MS in Chemistry
Florida International University
Miami, Florida

PUBLICATIONS AND PRESENTATIONS
