Effect of solute–solvent interactions on protein stability and ligand binding

by

Ikbae Son

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Pharmaceutical Sciences
University of Toronto

© Copyright 2017 by Ikbae Son
Abstract

Effect of solute–solvent interactions on protein stability and ligand binding

Ikbae Son
Doctor of Philosophy
Graduate Department of Pharmaceutical Sciences
University of Toronto
2017

My doctoral thesis is aimed at characterizing the effect of solute–solvent interactions on protein stability and ligand binding events. To quantify non-specific protein–solvent interactions, we measured and analyzed partial molar volume, $V^\circ$, and compressibility, $K_S^\circ$, of proteins within a framework of statistical thermodynamic formalisms. We characterized the binding of tri-$N$-acetylglucosamine [(GlcNAc)$_3$] and cAMP to lysozyme and cAMP-binding domain of EPAC1, respectively. Our volumetric measurements reveal that the protein–ligand complexes are less dynamic compared to their unbound states while complex formation is accompanied by the release of water of hydration to the bulk. To characterize protein–urea interactions, we determined and analyzed volumetric properties of four globular proteins at urea concentrations ranging from 0 to 8 M. We interpreted urea-dependent volumetric measurements in terms of the solvent exchange model in which the binding of urea to a protein proceeds with a release of two waters of hydration. Comparison of urea-dependent volumetric data for a folded protein with the similar data obtained on small molecules mimicking protein groups suggests the lack of cooperative effects involved in protein–urea interactions. Changes in volumetric properties associated with urea-induced unfolding transitions are consistent with solvent-accessible surface areas of unfolded proteins being roughly twice as large as those of their respective native states. Specifically, we reported the first experimental characterization of the thermodynamic profile of urea binding to a native protein. We measured the volumetric parameters of lysozyme as a function of urea within a temperature range of 18 to 45°C. Based on the van’t Hoff analysis of the temperature dependence of the equilibrium constants for protein–urea binding events we evaluated the full thermodynamic profile of protein–urea interactions. Finally, we combined the experimental and theoretical approaches to investigate the influence of the urea on the binding
of (GlcNAc)$_3$ to lysozyme. Analysis of urea dependence of the binding free energy reveals that the protein–ligand binding reaction is governed by a close interplay between the free energy contributions of the excluded volume effect and direct solute–solvent interactions.
Acknowledgements

I would like to thank my research supervisor, Prof. Tigran V. Chalikian, for his encouragement and guidance throughout my thesis project. I would like to extend my gratitude to my advisory committee members: Prof. Robert B. Macgregor Jr., Prof. Régis Pomès, Prof. Jeffery Henderson, and Prof. Heiko Heeklotz, for their time, valuable input, and guidance. I also thank Prof. Gregory Poon to serve as my external appraiser and Prof. Hue Sun Chan for serving as my external examiner for my thesis defense.

I would like to express gratitude to all my lab colleagues, Dr. Yuen Shek, Dr. Yang Lee, Dr. Helen Pan, Gloria Kim, Lutan Liu, Nisha Patel, Lora Stephanian, and Claire Wang for making graduate school more enjoyable and being good friends. Special thanks to Dr. David Dubins who gave me tremendous suggestions and fruitful guidance on computer simulations.

I am indebted to my beloved family, dad, mom, my little sister, and all my friends for their moral support throughout my entire graduate work.

Lastly, I would like to acknowledge Canadian Institute Health Research (CIHR) Training Program in Protein Folding and Interaction Dynamics for financial support.
Contents

1 General Introduction

1.1 Water is Needed for Protein to Function ........................................ 1
  1.1.1 Structural support ...................................................... 2
  1.1.2 Protein folding ....................................................... 3
  1.1.3 Molecular association ............................................... 6
  1.1.4 Quantifying changes in hydration ................................... 7

1.2 Cosolvents Influence Protein Stability and Function ...................... 10
  1.2.1 Direct and indirect mechanisms of urea-induced unfolding .......... 10
  1.2.2 Solubility studies and transfer model ................................ 11
  1.2.3 Preferential interactions ............................................ 14
  1.2.4 Vapor pressure osmometry ......................................... 15

1.3 Volumetric Techniques .......................................................... 17
  1.3.1 Investigation of protein–water interactions ........................ 17
  1.3.2 Investigation of protein–urea interactions ........................ 21

1.4 Research Aims ........................................................................ 23
  1.4.1 Quantification and characterization of changes in protein–solvent interactions accompanying protein–ligand association .......................... 24
    1.4.1.1 Lysozyme ............................................................ 24
## 2 Characterization of \((\text{GlcNAc})_3\) Binding to Lysozyme

### 2.1 Abstract

### 2.2 Introduction

### 2.3 Materials and Methods

#### 2.3.1 Materials

#### 2.3.2 Fluorescence

#### 2.3.3 Volumetric and densimetric measurements

#### 2.3.4 Determination of intrinsic volumes and solvent-accessible surface areas

### 2.4 Results

### 2.5 Discussion

#### 2.5.1 Evaluation of the binding-induced changes in hydration

#### 2.5.2 Evaluation of the binding-induced change in the intrinsic compressibility of lysozyme

### 2.6 Conclusion

## 3 Characterization of cAMP Binding to EPAC1

### 3.1 Abstract

### 3.2 Introduction

### 3.3 Materials and Methods

#### 3.3.1 Materials
3.3.2 Ultrasonic and densimetric measurements ........................................ 45
3.3.3 Determination of intrinsic volumes and solvent-accessible surface areas . 46
3.4 Results and Discussion ........................................................................... 46
  3.4.1 Binding affinity .................................................................................. 46
  3.4.2 A change in hydration ....................................................................... 48
  3.4.3 A change in protein dynamics ............................................................... 50
  3.4.4 A change in configurational entropy .................................................... 51
3.5 Conclusion ............................................................................................... 52

4 Interaction of Urea with Native and Unfolded Proteins ........................ 54
  4.1 Abstract .................................................................................................. 54
  4.2 Introduction ............................................................................................ 55
  4.3 Materials and Methods .......................................................................... 56
    4.3.1 Materials .......................................................................................... 56
    4.3.2 Concentration determinations ............................................................ 57
    4.3.3 Solution preparation .......................................................................... 58
    4.3.4 Optical spectroscopy .......................................................................... 58
    4.3.5 Determination of the partial molar volumes and adiabatic compressibilities of solutes ................................................................. 58
  4.4 Results ..................................................................................................... 59
    4.4.1 Lysozyme .......................................................................................... 61
    4.4.2 Apocytochrome c .............................................................................. 66
    4.4.3 Ribonuclease A .................................................................................. 67
    4.4.4 α-Chymotrypsinogen A ...................................................................... 69
    4.4.5 The problem of baselines .................................................................. 70
5 Thermodynamic Profile of Protein–Urea Interactions

5.1 Abstract ......................................................... 72

5.2 Introduction ................................................. 73

5.3 Materials and Methods .................................... 74

5.3.1 Materials ...................................................... 74

5.3.2 Optical spectroscopy ...................................... 75

5.3.3 High-precision densimetry ............................... 75

5.3.4 Ultrasonic velocimetry .................................... 75

5.4 Results .......................................................... 76

5.5 Discussion ...................................................... 78

5.5.1 Data analysis ............................................... 78

5.5.2 Thermodynamic parameters of urea–protein interactions .... 84

5.5.3 Rationalization of the thermodynamic profile of urea–protein interactions 85

5.5.4 Volumetric considerations ............................... 87

5.6 Conclusion ...................................................... 88

6 Effect of Urea on Protein–Ligand Association

6.1 Abstract ......................................................... 89

6.2 Introduction ................................................... 90

6.3 Materials and Methods .................................... 92

6.3.1 Materials ...................................................... 92

6.3.2 Fluorescence ............................................... 92

6.4 Results and Discussion .................................... 93
### 6.5 Conclusion

7 General Conclusions and Future Perspectives

#### 7.1 Conclusions

#### 7.2 Future Directions

- **7.2.1** pH-induced folding and unfolding of lysozyme in urea solution
- **7.2.2** Thermodynamic characterization of the interaction of glycine betaine with a native protein

Bibliography
List of Tables

2.1 Molecular Volumes, $V_M$, and Solvent Accessible Surface Areas, $S_A$, of Unligated Lysozyme, Lysozyme–(GlcNAc)$_3$ Complex, Lysozyme in the Complex, and Free (GlcNAc)$_3$ .......................................................... 35

3.1 Thermodynamic Characteristics of the Association of cAMP with the CBD of EPAC1 .................................................................................................................................................. 48

3.2 Molecular Volumes, $V_M$, and Solvent Accessible Surface Areas, $S_A$, of the cAMP–EPAC1 Complex, apoEPAC1, and Free cAMP ................................................................................. 48

4.1 Extinction Coefficients (M$^{-1}$ cm$^{-1}$) of the Proteins as a Function of Urea ................................................................................................................................. 58

5.1 Urea–Protein Binding Constants Determined from the Volume and Compressibility Data and Changes in Volume, $\Delta V^\circ$, and Compressibility, $\Delta K_S^\circ$, Accompanying the Binding in an Ideal Solution at Various Temperatures ....................................................................... 82

5.2 Thermodynamic Parameters of Urea–Protein Interactions at 25 °C ......................................................................................................................... 83

6.1 The (GlcNAc)$_3$–Lysozyme Binding Constants, $K_b$, and Binding Free Energy, $\Delta G_b = -RT \ln K_b$, at Different Urea Concentrations ......................................................................................... 95
List of Figures

2.1 Relative fluorescence intensity of a solution containing lysozyme plotted against the (GlcNAc)$_3$-to-lysozyme molar ratio, $r$ .................................................. 34

2.2 Change in the relative molar sound velocity increment of lysozyme plotted against the (GlcNAc)$_3$-to-lysozyme molar ratio, $r$ .................................................. 35

2.3 Change in the partial molar volume of lysozyme plotted against the (GlcNAc)$_3$-to-lysozyme molar ratio, $r$ .................................................. 36

3.1 Change in the partial molar volume of the CBD of EPAC1 plotted against the cAMP-to-protein molar ratio, $r$ .................................................. 47

3.2 Change in the relative molar sound velocity increment of the CBD of EPAC1 plotted against the cAMP-to-protein molar ratio, $r$ .................................................. 47

4.1 Near- and far-UV CD spectra of lysozyme at different urea concentrations .... 60

4.2 Far-UV CD spectra of apocytochrome c at different urea concentrations .... 61

4.3 Urea dependences of the molar ellipticities at 219 nm for ribonuclease A and α-chymotrypsinogen A ................................................................. 62

4.4 Urea dependences of the partial molar volumes of lysozyme, apocytochrome c, ribonuclease A, and α-chymotrypsinogen A ............................................................. 63

4.5 Urea dependences of the partial molar adiabatic compressibilities of lysozyme, apocytochrome c, ribonuclease A, and α-chymotrypsinogen A ............................................................. 64

5.1 Far and near-UV CD spectra of lysozyme at 18 celsius and 45 celsius in the absence and presence of 8 M urea .............................................................. 76
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2</td>
<td>Urea dependencies of the partial molar volume, the relative molar sound velocity increment, and the partial molar adiabatic compressibility of water at 18, 25, 35, and 45 °C.</td>
<td>78</td>
</tr>
<tr>
<td>5.3</td>
<td>Urea dependencies of the partial molar volume, the relative molar sound velocity increment, and the partial molar adiabatic compressibility of lysozyme at 18, 25, 35, and 45 °C.</td>
<td>79</td>
</tr>
<tr>
<td>5.4</td>
<td>Temperature dependence of the equilibrium constant, $k$.</td>
<td>83</td>
</tr>
<tr>
<td>6.1</td>
<td>The relative fluorescence intensity of lysozyme plotted against the (GlcNAc)$_3$-to-lysozyme molar ratio, $r$, at 0, 2, 4, 6, and 8 M urea.</td>
<td>94</td>
</tr>
<tr>
<td>6.2</td>
<td>The urea-dependence of the (GlcNAc)$_3$–lysozyme binding free energy.</td>
<td>96</td>
</tr>
<tr>
<td>6.3</td>
<td>The urea-dependence of excluded volume ($\Delta\Delta G_C$) and direct solute–solvent interaction ($\Delta\Delta G_I$) components of the differential (GlcNAc)$_3$–lysozyme binding free energy ($\Delta\Delta G_b$) between a concentrated urea solution and water.</td>
<td>97</td>
</tr>
</tbody>
</table>
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASA</td>
<td>Accessible surface area</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CBD</td>
<td>cAMP-binding domain</td>
</tr>
<tr>
<td>DEP</td>
<td>Dishevelled/Egl-10/pleckstrin</td>
</tr>
<tr>
<td>EPAC</td>
<td>Exchange protein directly activated by cAMP</td>
</tr>
<tr>
<td>GB</td>
<td>Glycine betaine</td>
</tr>
<tr>
<td>(GlcNAc)$_3$</td>
<td>tri-$N$-acetylglucosamine</td>
</tr>
<tr>
<td>LEM</td>
<td>Linear extrapolation model</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamics</td>
</tr>
<tr>
<td>MSP</td>
<td>Molecular surface package</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>SEM</td>
<td>Solvent exchange model</td>
</tr>
<tr>
<td>SPT</td>
<td>Scaled particle theory</td>
</tr>
</tbody>
</table>
Chapter 1

General Introduction

1.1 Water is Needed for Protein to Function

As life, as we know it, occurs in an aqueous milieu, water can be considered as the universal solvent of life. Water is a highly polar molecule that is crucial as a solvent due to its capacity to dissolve a vast range of polar and charged compounds. For example, proteins exist as heavily hydrated structures because their solvent-exposed surfaces are predominantly composed of hydrophilic groups. In a cell, water has profound biological relevance since the structure and function of a protein are modulated by the interactions with its solvent environment [1–3]. Given the determining role of hydration in protein function, water has been classified as the “21st amino acid” [3].

The importance of water has been well recognized with the hydrophobic effect being one of the main driving forces in protein folding. Hydrophobic amino acids are sequestered from water as a result of the unfavorable hydration of hydrophobic groups [4]. Furthermore, proteins generally require about $\sim 0.4 \text{ g of water/g of protein}$ to achieve normal function [5, 6]. Although some enzymes may retain functionality in organic solvents, increasing hydration typically enhances enzymatic activity [7, 8]. The improvement in enzymatic activity suggests that water plays a role in facilitating the structural flexibility necessary for protein function [1, 9]. Both experiment and molecular dynamics simulations reveal that the structural relaxation of a protein correlates with protein–water hydrogen bond dynamics [9–11]. Frauenfelder and colleagues [9, 12] have further suggested that there is high coupling between protein motion and water dynamics: the large-scale fluctuations of the protein chain can be “slaved” to the fluctuations of bulk water, and even internal protein motions can be “slaved” to the fluctuations of the hydration shell. Consequently, the structure and dynamics of water may feed back onto those aspects of the proteins themselves, so that protein function depends on a delicate interplay between the
two [1].

The biological properties of a protein depend on its interaction with other molecules, and water plays an important role in facilitating these events. The thermodynamic contribution of water molecules involved in a binding event is known to be on the order of or even exceed the net thermodynamics of the binding reaction [13–16]. Since water properties are sensitive to the details of the interaction with the adjacent surface, the ligand-induced release of individual water molecules may contribute favorably or unfavorably to the binding affinity and specificity depending on the location of the hydration site [17–19]. In some cases, not all of the water is expelled, but remaining water molecules at the interface play a central role in the binding [19,20]. Sometimes it is the sheer presence of water molecules at the binding site that makes the association possible [18,21].

The importance of water is underscored by the effect of cosolvents. The addition of destabilizing cosolvents, such as urea, disrupts native protein structure. Since native proteins are only 5–10 kcal mol\(^{-1}\) more stable than their denatured states, a change in chemical composition of the solvent can shift the structural equilibrium toward the denatured states [2,4]. Such cosolvents have been used routinely in protein studies as tools to destabilize proteins. Although the molecular forces underlying the action of cosolvents are still debated, many experimental and theoretical studies have suggested that cosolvents influence protein stability by disrupting and/or replacing protein–water interactions [22–24].

In short, water plays a crucial role in, virtually, all aspects of the conformational stability and the functional activity of proteins. In the sections that follow, the effect of protein–solvent interactions on protein stability and function will be further examined, with particular focus on characterizing and quantifying such interactions in protein–ligand binding events and conformational transitions.

1.1.1 Structural support

While water may enhance protein flexibility, it can also provide structural support when incorporated into the protein interior. X-ray crystal structure of a bacterial surface protein, internalin B, shows that the leucine-rich motif of this protein contains a series of stacked loops that are held together by water molecules bridging the peptide chains [25]. These buried water molecules play an important role in enhancing the structural stability of a protein. Buried water molecules are highly localized with the mean residence time for buried water being much longer than that for water in the first hydration shell [26,27]. For example, the residence times for water in the first hydration shell are on the order of ∼ 500 ps or less, whereas those for buried water are between 10 ns and 1 ms [26]. Park and Saven [27] have surveyed 842 high resolution protein structures and have observed that buried waters appear preferentially near
regions without secondary structures in which most of them make primary hydrogen bonds to main-chain atoms providing structural support to the main-chain. Buried water molecules are on average 11 times more likely to form hydrogen bond with main-chain atoms near residues without α-helical and β-sheet secondary structures [27].

The location of buried water molecules is often highly conserved in structurally homologous proteins [28–30]. Sreenivasan and Axelsen [29] have examined chemical environments surrounding buried waters molecules within the structurally homologous family of eukaryotic serine proteases. These authors have observed that 21 conserved buried waters are surrounded by highly conserved residues including main chain atoms. These water molecules are often observed in identical locations of proteins in both solution and crystals [29,30].

Buried water molecules often play a crucial role in stabilizing the protein structure; they can even be considered integral structural components of the protein [28, 31, 32]. Covalt and co-workers [31] have investigated the effect of mutations in interleukin-1β on buried water molecules. Such water molecules provide structural support to the β-strands of interleukin-1β. For instance, water forms hydrogen bonds between L10 and V40 in the wild-type structure; however, a T9G mutant lacks this water molecule thereby altering protein stability [31]. Kinetic unfolding studies using nuclear magnetic resonance (NMR) spectroscopy have revealed that residues linked by water molecules are among the first to become disordered during the unfolding of interleukin-1β [32]. Consequently, the destabilized structure allows solvents to enter the interior of the protein.

1.1.2 Protein folding

In a cell, most proteins spontaneously fold into their native structures to perform their biological functions. The hydrophobic effect is considered as one of the major driving forces in protein folding. It arises from unfavorable hydration of nonpolar groups which minimizes the amount of surface exposed to water [33]. The thermodynamics of the hydrophobic effect can be evaluated by studying the transfer a nonpolar solute from their pure liquid phase into liquid water [4,34,35]. Measurements of changes in solubility associated with the transfers of small hydrocarbons from their pure liquid state to water have revealed that changes in transfer chemical potential (Δ\(\mu^\circ\)) are positive, changes in heat capacity (Δ\(C_P\)) are large and positive, changes in enthalpy (Δ\(H^\circ\)) are close to zero, and changes in entropy (Δ\(S^\circ\)) are large and negative at room temperature [35]. At room temperature, the unfavorable entropic effect of hydration of hydrophobic solutes greatly exceeds the enthalpic effect leading to the unfavorable transfer free energy. Although the hydrophobic effect is entropy-driven around 25 °C, it becomes enthalpy-driven at higher temperatures given the large positive heat capacity [35].

At room temperature, it has been proposed that nonpolar solutes are surrounded by ordered
water molecules [4]. Since nonpolar compounds cannot donate or accept hydrogen bonds, water molecules surrounding a nonpolar solute prefer to hydrogen bond with each other to avoid “wasting” their hydrogen bonding capacity [4]. In this scenario, a change in enthalpy is near zero since ordered water molecules around a nonpolar solute retain their hydrogen bonds, thereby, avoiding the enthalpic penalty. A change in entropy involved in transfer of a nonpolar molecule to water is large and negative because water molecules from oil become ordered around the solute, resulting in a loss of degrees of freedom [4]. Consequently, nonpolar solutes are forced to coalesce together to reduce their area of contact with water and minimize the unfavorable change in entropy [36].

Calorimetric studies have been conducted to evaluate the thermodynamic characteristics of transferring various protein groups from gas to water [37, 38]. As with the transfer of hydrocarbons, the transfer free energy of nonpolar groups is unfavorable due to the entropic effect. For polar compounds, changes in entropy are negative, but large and negative changes in enthalpy makes the transfer free energy favorable [38]. The large negative changes in enthalpy of polar compounds suggests favorable direct solute–solvent interactions, which include hydrogen bonding and van der Waals interactions. Negative changes in entropy of polar groups suggest a decrease in translational and orientation entropy of water of hydration due the formation of solute–solvent hydrogen bonding [39]. In contrast to nonpolar compounds, changes in heat capacity associated with transfer of a polar compound from vapor to water is negative which reflects a decrease in the enthalpic contribution with increasing temperature [40]. Makhatadze and Privalov [37] have evaluated the effect of hydration on protein folding by summing the hydration effects of solvent exposed groups. These authors have revealed that hydration of polar and aromatic groups destabilize the protein structure, while that of aliphatic groups stabilize it [37]. While an exception of aliphatic groups, the hydration of all groups which become exposed upon unfolding destabilizes the folded states, whereas this state is stabilized by favorable enthalpic interactions among internal groups [37].

Due to the unfavorable hydration of nonpolar groups, many of these groups are buried inside a protein [4, 33]. On average, 86% of the surface area of hydrophobic groups are buried in the interior of a folded protein [41]. Burial of hydrophobic residues helps proteins to fold and to retain their compact structure [4, 33]. However, the burial of hydrophobic groups is unavoidably accompanied by burial of hydrophilic groups because of the irregular distribution of hydrophobic and hydrophilic groups within the primary polypeptide sequence [41].

During the initial stages of the folding process, water molecules are involved in guiding the packing of hydrophilic residues by mediating long-range interactions between polar and charged groups [42, 43]. Papoian et al. [42] have employed molecular dynamics (MD) simulations to compare the computed structures of caspase proteins with their crystal structures. These authors have demonstrated a substantial improvement in structural predictions for large α-proteins.
when water molecules are included in the simulation. It has been found that water mediates relatively long-range (6.5–9.5 Å) connections between hydrophilic domains [42]. Water-mediated long-range interactions also play a crucial role in formation of β-sheets [43]. In contrast to α-helices which are localized within the primary amino acid sequence, β-strands may be far apart in the sequence. Zong and colleagues [43] have demonstrated that adding water-mediated interactions leads to a significantly improved prediction of the structure of α/β proteins. Analysis of the simulation results suggests that topological frustration in β-sheet formation during collapse is alleviated by protein–water interactions [43]. Since the protein chain cannot cross itself, some parts of the chain will not be able to form contacts with other parts of the chain, potentially causing frustration in the folding process [2,44]. Water molecules constrain the conformational freedom of the amino acid side chains, thereby, diminishing the effect of topological frustration [2].

A protein folding energy landscape resembles a partially rough funnel with traps where the protein can transiently reside [45,46]. For fast folding proteins, the energetic roughness (energetic frustration) of the landscape is sufficiently small, and the global characteristics of transition state ensembles are most strongly determined by native contacts (residues interacting in the fully folded structure) [2,45]. For this reason, native contacts play a dominant role in the folding process of fast folding proteins, which forms the basis of coarse-grained simulation methods like Gō models [2,45,47]. Onuchic and co-workers [47] have applied Gō models to study the folding transition of the Src homology 3 domain. The initial folding process is characterized by structural collapse of a near-native structure with partially hydrated hydrophobic core [47]. These authors have observed that the structural collapse leads to native contact formation between the diverging turn and the distal loop. In this process, water molecules mediate the search for the native structure by connecting protein residues [2]. The final folding process involves long-range tertiary contacts across two β sheets followed by expulsion of water molecules from the hydrophobic core [47].

It has been proposed that the expulsion of water molecules from hydrophobic groups leads to formation of a very thin layer of vapor separating water from the surface of the hydrophobic solute [1,48,49]. Results of X-ray reflectivity measurements suggest that there is a lower liquid density around the surface of hydrophobic alkanes [48]. When two hydrophobic solutes come close together, water molecules are released from the region between two surfaces due to the constraints on the hydrogen bond network [49]. The release of water from the intervening space creates an imbalance in pressure that provides long-range attraction between hydrophobic surfaces [49].

In the course of the folding, the protein backbone and side chains adopt a tightly packed configuration, leading to a very large loss of the configurational entropy of a protein molecule [1]. Water molecules are released from the excluded volume because the tight packing of protein
groups diminishes the volume available for the solvent molecules in contact with the protein surface [47, 50]. Harano and Kinoshita [50] have suggested that the expulsion of water molecules from the protein core may translate into a gain in the translational entropy. These authors have applied a statistical-mechanical theory to demonstrate that the excluded volume effects can offset the entropic cost of protein compaction. A relatively large gain in translational entropy has been observed in the form of water molecules being released from the protein surface [1]. In fact, for a peptide of the order of 50 residues or more, these authors have estimated that the gain in translational entropy offsets the loss of configurational entropy upon folding [50].

1.1.3 Molecular association

In addition to its role in modulating protein folding, water plays a dominant role in regulating the equilibrium in molecular association events [2, 14, 18, 51]. The binding of a ligand to a protein can be viewed as removal of water from the binding surfaces with the subsequent formation of protein–ligand interactions [19]. Any redistribution of solvent molecules between the hydration and bulk phases should affect the net thermodynamics of the molecular events [2, 13, 14]. Breiten et al. [16] have performed calorimetric characterization of several carbonic anhydrase complexes. These authors have observed that the values of binding enthalpy and entropy are dictated by the structure and thermodynamic properties of water at the binding surface and the bound ligands. Filfil and colleagues [15] have applied volumetric techniques to investigate the binding of bovine pancreatic trypsin inhibitor to bovine trypsinogen. These authors have evaluated the hydralional, $\Delta S_{\text{hyd}}$, and configurational, $\Delta S_{\text{conf}}$, changes in entropy associated with the binding event. The values of $\Delta S_{\text{hyd}}$ and $\Delta S_{\text{conf}}$ are equal to $143 \pm 52$ and $-97 \pm 54$ cal mol$^{-1}$ K$^{-1}$, respectively [15]. These results suggest that a highly favorable change in hydration entropy prevails over an unfavorable change in configurational entropy, providing the thermodynamic driving force for the protein binding.

Although the binding of a ligand generally leads to release of water molecules from the binding site, some water molecules are often retained within the protein–ligand interface [19]. Such water molecules do not simply occupy cavities in the binding site; instead they form up to four hydrogen bonds with the ligand and protein groups [1]. Many proteins make use of bound water molecules as functional units to mediate interactions with other proteins in protein–protein complexes or with substrate molecules in enzymatic complexes [19]. Lu and co-workers [20] have performed a comprehensive analysis of water molecules at the protein–ligand interfaces in 392 high-resolution crystal structures. These authors have observed that 76% of interfacial water molecules form hydrogen bonds with polar atomics groups of both the ligand and the protein [20]. The number of ligand-bound water molecules correlates with the polar van der Waals surface area of the ligand [20].
Water molecules at the protein–ligand interface are, generally, more tightly bound than water molecules interacting only with the protein surface [21, 52]. Water molecules may occupy defined sites in a protein structure, thereby, facilitating specificity of substrate binding. Chung and co-workers [53] have combined mass spectrometric and thermodynamic measurements to study interactions between the Src homology 2 domain of tyrosine kinase Src and tyrosyl phosphopeptides. The protein domain binds a range of tryosyl peptides in a specific or nonspecific manner [53]. These authors have observed that a network of water molecules is conserved within the specific complex structure, but not within the nonspecific peptide complex [53]. In this sense, such water molecules represent an extension of the protein surface that mediate highly specific hydrogen bonding interactions.

In contrast to enhancing the specificity of substrate binding, water molecules can also promote promiscuous bindings [51]. The oligopeptide binding protein, OppA, binds peptides of two to five amino acid residues regardless of their sequence [21]. The lack of specificity can be achieved due to the fact that all interactions between the protein and the peptide side chains are mediated by water [21]. These water molecules can make the binding surface highly adaptable and even promiscuous.

A better understanding of the thermodynamic and structural role of water molecules in the active site of a protein can greatly increase the efficiency of rational drug design [51]. Barillari et al. [54] have assessed the use of water molecules in drug design by classifying them according to their propensity to be displaced by ligands. By studying the thermodynamics of water molecules in the binding sites of six proteins complexed with ligands, these authors have identified two classes of water molecules: those conserved and not displaced by any of the ligands and those that are displaced by ligands [54]. It has been further found that conserved water molecules are more tightly bound, making three or four hydrogen bonds with the protein and the ligand, than displaceable water molecules [54]. Identification of conserved water molecules may be useful to design a ligand that maximizes the interaction with such water molecules, thereby, optimizing its binding to the target protein site.

1.1.4 Quantifying changes in hydration

Water molecules interacting with the protein surface are thermodynamically distinct from bulk water with respect to enthalpy, entropy, heat capacity, volume, expansibility, compressibility, etc. [55, 56]. However, it is quite difficult to draw an exact line where the hydration shell ends and the bulk phase begins. Since protein–solvent interactions can be distributed among a large number of non-covalent interactions, there is a fluctuating cloud of water molecules that are thermodynamically affected by a protein [2, 5, 57]. Although most ordered water molecules are observed within the first layer of hydration, the hydration shell may involve several layers of
water molecules [58]. Ebbinghaus and colleagues [59] have applied terahertz spectroscopy to measure hydration dynamics around five helix bundle proteins. These authors have found that the distance between proteins significantly influence the terahertz absorbance of the proteins and first hydration layer. The absorbance decreases gradually as two proteins are brought closer together. The absorbance eventually starts to increase when the separation becomes below 18 Å, which signifies overlapping solvation layers between two proteins [59]. Thus, hydration shell around each protein extends to at least $\sim 9$ Å with multiple layers of water of hydration [59].

Different tools sample different populations of water of hydration. Structural techniques, such as X-ray crystallography and nuclear magnetic resonance spectroscopy, are sensitive to highly localized and/or immobilized water molecules. Yet, disordered water molecules, which are more difficult to detect, represent a large fraction of protein hydration. In a strictly thermodynamic sense, hydration refers to all water molecules whose thermodynamic properties are distinct from bulk water, regardless of their physical proximity to the protein surface [60].

To evaluate the effect of hydration on protein–ligand association, osmotic stress studies have been applied to analyzing changes in hydration accompanying protein–ligand and protein–protein association events [61]. The addition of “inert” osmolytes to a reaction buffer influences the protein–ligand binding equilibrium by increasing the osmotic pressure. If the binding is associated with the release of water molecules from an interacting surface, the addition of osmolytes will drive the equilibrium towards the less hydrated ligand-bound state [61]. The combined binding/release of osmolyte and water can be evaluated from a change in the equilibrium constant, $K$, of an association event caused by a change in osmolyte concentration and, hence, the activity of water. The equilibrium constant is measured as a function of the activity of water, $a_1$, with the differential hydration parameter, $\Delta \Gamma_{21}$, being related to resulting changes in $K$ [61–63]:

$$\Delta \Gamma_{21} = \left( \frac{\partial \ln K}{\partial \ln a_1} \right)_{T,P} = \Delta n_1 - \frac{N_1}{N_3} \Delta n_3$$  \hspace{1cm} (1.1)

where $N_1$ and $N_3$ are the numbers of moles of water and osmolyte in solution, respectively; while $\Delta n_1$ and $\Delta n_3$ refer to the excess numbers of water and osmolyte around the solute, respectively. These parameters describe the stoichiometric contribution of solvent and cosolvent molecules to protein association reactions [61].

Reid and Rand [64] have applied osmotic stress techniques to analyzing a change in hydration accompanying the binding of D-glucose to hexokinase PII. These authors have measured the binding constant as a function of the activity of water using fluorescence spectroscopy and vapor pressure osmometry. These authors have determined that $326 \pm 35$ waters of hydration become released to the bulk upon the complex formation. Sidorova and Rau [65] have used the
osmotic stress technique to investigate the binding of netropsin to DNA using circular dichroism spectroscopy and vapor pressure osmometry. These authors have revealed that the binding of netropsin to the DNA cause the release of $55 \pm 11$ water molecules to the bulk.

Harries and co-workers [66] have used microcalorimetry to measure changes in the equilibrium constant for $\beta$-cyclodextrin–adamantane host/guest complexation caused by various osmolytes. These authors have observed a wide range of water molecules being released or taken up upon complexation. Among neutral osmolytes, glycine is the most preferentially excluded with $\Delta n_1$ being equal to $-24$, while betaine is the most preferentially bound with $\Delta n_1$ being equal to $22$ [66]. For preferentially excluded osmolytes with negligible solute–osmolyte interactions, the value of $\Delta n_3$ in Equation (1.1) becomes negligible. In this case, the differential hydration parameter is taken equal to the number of released or taken up water molecules ($\Delta \Gamma_{21} = \Delta n_1$). Although the value of $\Delta \Gamma_{21}$ should be independent of the identity of preferentially excluded osmolytes, osmotic stress studies frequently yield different values of $\Delta n_1$ for the same reaction [61,66].

The discrepancy in the values of $\Delta n_1$ observed for different excluded osmolytes suggests that the differential hydration, $\Delta \Gamma_{21}$, parameter in Equation (1.1) accounts not only for the stoichiometric binding/release of water and cosolvent molecules to a protein but also for the excluded volume effect. Excluded volume is an entropic effect originating from exclusion of water and osmolyte molecules from the space occupied by solute molecules, thereby reducing the translational entropy of the solute [67]. Considering the excluded volume effect, the differential hydration parameter, $\Delta \Gamma_{21}$, for a preferentially excluded osmolyte can be expressed as [68]:

$$\Delta \Gamma_{21} = \left( \frac{\partial \ln K}{\partial \ln a_1} \right)_{T,P} = -\frac{1}{RT} \left( \frac{\partial \Delta \Delta G_C}{\partial \ln a_1} \right)_{T,P} + \Delta n_1 \tag{1.2}$$

where the free energy of cavity formation, $\Delta \Delta G_C$, reflects the excluded volume effect. Thermodynamically, the effect of excluded volume can be modeled by creating a cavity in solution which is sufficiently large to accommodate the solute molecule but is inaccessible to solvent molecules [68]. While the second term, $\Delta n_1$, in Equation (1.2) does not depend on the identity of the osmolyte, the first term reflects the osmolyte-dependent changes in the free energy of cavity formation. A computational study has revealed that the excluded volume contribution to $\Delta \Gamma_{21}$ associated with the protein–ligand binding reaction is very significant and ranges from 30 to 70% correlating with the size of the osmolyte molecules [68]. Because of the non-stoichiometric free energy of cavity formation, different preferentially excluded cosolvents may yield significantly different values of $\Delta \Gamma_{21}$ for the same reaction [66].
1.2 Cosolvents Influence Protein Stability and Function

In addition to water, a variety of water miscible organic compounds may be present in biological systems and influence the protein stability and association events. Organic cosolvents are sometimes referred to as osmolytes, as they are frequently used by cells of a variety of organisms to counteract the osmotic loss of cell water [69]. Many organisms experience different types of environmental stress, such as drought and desiccation, that reduce the amount of intracellular water. In response, cells use osmolytes to maintain the necessary level of cellular water and, hence, normal cellular processes [69].

Cosolvents are commonly used in protein folding and stability studies since interactions of cosolvents with proteins can have a profound effect on the folding equilibria of proteins in aqueous solutions [69–72]. A cosolvent that exerts a stabilizing effect on the folded structure of a protein is known as a protecting or stabilizing cosolvent, while a cosolvent that shifts the equilibrium towards the unfolded structure is known as a denaturing or destabilizing cosolvent [72,73].

Among all cosolvents of biological relevance, urea has been studied most extensively in protein folding and stability investigations [22]. However, the exact molecular mechanism of urea is still debated. This dissertation will primarily focus on the effect protein–urea interactions on protein stability and ligand binding events.

1.2.1 Direct and indirect mechanisms of urea-induced unfolding

Two mechanisms have been proposed to explain the unfolding effect of urea on proteins [22]. According to the indirect mechanism, urea is proposed to unfold a protein by altering water structure around it [22,74]. This view has emerged from solute transfer experiments showing that hydrophobic compounds are more soluble in aqueous urea than in water [74]. It has been proposed that the addition of urea disrupts the structure of water thereby weakening the hydrophobic effect [74]. However, various studies have demonstrated that urea has a negligible effect on hydrogen bond interactions between water molecules [75, 76]. Urea fits relatively easily into the network of water molecules with the orientational dynamics of water being unaffected [77,78]. A neutron scattering study shows that the radial distribution function of urea around water in a 1:4 solution resembles that of water around water [78]. Furthermore, MD simulations suggest that urea does not disrupt the structure of water even at a high urea concentration [79,80].

The direct mechanism of urea-induced protein denaturation postulates that urea diminishes the stability of a protein through favorable interactions with constituent atomic groups, rather
than through altering the structure of water [22]. The nature and strength of urea’s interactions with various protein groups including the peptide backbone and amino acid side chains have been extensively studied given the chemical heterogeneity of the protein surface [79,81,82].

Various studies have shown that urea interacts favorably with the majority of amino acid side chains [22,83,84]. MD simulations conducted by Thirumalai’s group have revealed that the main driving force of urea-induced protein unfolding is the formation of hydrogen bonds between urea and peptide carbonyl groups and charged side chains of protein [79,85]. In another simulation study, the weakening of the hydrophobic effect is the main driving force in urea-induced protein unfolding [82,86]. It has been found that urea mainly interacts with less polar side chains and the backbone, while accumulation of urea in the vicinity of nonpolar groups displaces water molecules from the solvation shell [82]. MD simulations of individual amino acids in aqueous urea have demonstrated that urea interacts favorably with almost all amino acids leading to its preferential binding over water [84]. Furthermore, Lee et al. [83] have measured partial molar volume and adiabatic compressibility contributions of the amino acid side chains and the peptide backbone at urea concentrations ranging from 0 to 8 M. These authors have found that the association constants, $k$, for solute–urea interactions range from 0.04 to 0.39 M consistent with favorable interactions of urea with a wide range of protein groups [83]. Despite the fact that many studies favor the direct mechanism, there is still a lack of consensus on the precise molecular basis of urea-induced unfolding.

1.2.2 Solubility studies and transfer model

Solubility studies have provided insights into the contribution of protein–cosolvent interactions to protein stability [4, 87]. Such studies enable one to analyze cosolvent interactions with individual protein groups. Measurements of a change in free energy accompanying transfer of amino acid side chains and the peptide backbone from pure water to various water–cosolvent mixtures can be used to identify key thermodynamic forces which are involved in the control and maintenance of protein structure and stability. A change in free energy associated with transferring a solute from water to a cosolvent solution, $\Delta \mu_{tr}$, represents the difference in the chemical potential of the solute in water, $\mu_{21}^o$, and in the binary solvent (water-cosolvent mixture), $\mu_{23}^o$. The transfer free energy, $\Delta \mu_{tr}$, can be derived from the measurements of the differential solubility of the compound in question in water, $s_{21}$, and the binary solvent, $s_{23}$ [88]:

$$\Delta \mu_{tr} = \mu_{23}^o - \mu_{21}^o = RT \ln \left( \frac{a_{21}}{a_{23}} \right) = RT \ln \left( \frac{s_{21}}{s_{23}} \right) + RT \ln \left( \frac{\gamma_{21}}{\gamma_{23}} \right)$$

(1.3)

where $a_{21} = \gamma_{21}s_{21}$ and $a_{23} = \gamma_{23}s_{23}$ are the activities of the solute in water and in the binary solvent–cosolvent mixture, respectively; $\gamma_{21}$ and $\gamma_{23}$ represent the coefficients of activities of the
Chapter 1. General Introduction

The energy balance of the thermodynamic cycle is expressed as $\Delta G_{N\rightarrow D}^{OS} - \Delta G_{N\rightarrow D}^{W} = \Delta G_{tr,D} - \Delta G_{tr,N}$. Rearranging this relationship, one obtains $\Delta G_{N\rightarrow D}^{OS} = \Delta G_{N\rightarrow D}^{W} + \Delta G_{tr,D} - \Delta G_{tr,N}$. The differential transfer free energy of a solute in the native and unfolded states, $\Delta G_{tr,D} - \Delta G_{tr,N}$, is related to the $m$-value of linear extrapolation model [89]:

$$\Delta G_{N\rightarrow D}^{OS} = \Delta G_{N\rightarrow D}^{W} + m[\text{cosolvent}]$$

where $m$ is the proportionality coefficient; and $[\text{cosolvent}]$ is the molar concentration of cosolvent. The $m$-value represents the response of protein stability to the addition of cosolvents [22]. In a 1 M cosolvent solution, the differential transfer free energy of a solute, $\Delta G_{tr,D} - \Delta G_{tr,N}$, equals the $m$-value.
It has been proposed that the differential transfer free energy of a solute is proportional to the number of groups newly exposed upon unfolding and the transfer free energies of these groups [88, 89]. In this case, the free energy of protein unfolding in 1 M cosolvent, \( \Delta G_{N \rightarrow D}^{QS} \), is related to the free energy of unfolding in water, \( \Delta G_{N \rightarrow D}^{W} \), through [89]:

\[
\Delta G_{N \rightarrow D}^{QS} = \Delta G_{N \rightarrow D}^{W} + \sum_i n_i \alpha_i \Delta \mu'_{tr,i} 
\]

where \( n_i \) is number of residues of type \( i \) in the protein; \( \alpha_i \) is the fractional change in solvent accessibility of group \( i \) upon unfolding; and \( \Delta \mu'_{tr,i} \) is the change in free energy accompanying the transfer of group \( i \) from water to 1 M cosolvent. In this model, the \( \sum_i n_i \alpha_i \Delta \mu'_{tr,i} \) term cumulatively represents the \( m \)-value within the framework of the assumption that the interaction of various protein groups with cosolvents contributes in an additive manner to the thermodynamics of unfolding [88].

Bolen’s group [81, 88, 89] has measured transfer free energies, \( \Delta \mu'_{tr} \), for amino acid side chains and the peptide backbone in an attempt to predict the \( m \)-values for various stabilizing and destabilizing cosolvents. It has been concluded that the peptide backbone contribution to the free energy of unfolding, \( \Delta G_{N \rightarrow D}^{QS} \), dominates over the sum of side chain contributions [88, 89]. Although the peptide backbone represents only a fraction of the total surface area of newly exposed protein groups, their contribution to the estimated \( m \)-values is much greater than the side chains’ contribution [88]. In 1 M urea, the collective side chain contribution to \( \mu'_{tr} \) is positive, which is opposed by the predominant negative contribution of the peptide backbone [88, 91]. Stabilizing osmolytes, such as proline and trimethylamine N-oxide, interact favorably with most of the side chains, while interacting unfavorably with the peptide backbone which prevents and reverses protein unfolding [89].

Although solubility studies have produced valuable insights into the contribution of protein–cosolvent interactions to protein stability, the summation of individual protein groups is susceptible to unaccountable errors because individual groups do not reside in a uniform microenvironment on the protein surface [4]. For example, the hydrophobicity of nonpolar groups on the protein surface depends on the chemical and geometric properties of the neighboring groups [92]. Solubility-based studies are performed, by definition, at the solubility limit of the solute which may range from mM to several M. However, protein studies are performed at much lower concentrations (\( \sim 1 \) mM or less). While some amino acids are very soluble (e.g. glycine or serine), others are only sparingly soluble (e.g. tyrosine). The extent of solute–cosolvent interactions at the solubility limit of the solute in 1 M cosolvent may vary significantly from one amino acid to another [83]. Consequently, direct comparison between the data for solutes exhibiting significantly different solubilities may be fraught with errors.
1.2.3 Preferential interactions

In a water–cosolvent mixture, the composition of the solvent in the immediate vicinity of a protein may differ from that in the bulk [93,94]. The preferential binding parameter, Γ_{23}, gives a measure of cosolvent excess in the immediate vicinity of the protein surface relative to the concentration in the bulk [94]:

$$\Gamma_{23} = \left( \frac{\partial m_3}{\partial m_2} \right)_{T,P,\mu_1,\mu_3} = - \left( \frac{\partial \mu_2}{\partial \mu_3} \right)_{T,P,\mu_1,m_2}$$  \hspace{1cm} (1.7)

where subscripts 1, 2, and 3 denote water, protein, and cosolvent, respectively; \(\mu_i\) is the chemical potential of component \(i\); and \(m_i\) is the molal concentration of component \(i\). Preferential binding parameter reflects a change in cosolvent concentration required to maintain constant chemical potentials of water, \(\mu_1\), and cosolvent, \(\mu_3\), when a protein is added to the water-cosolvent mixture [95]. Preferential binding parameter has been interpreted as the difference in the cosolvent concentrations in the immediate vicinity of a protein and bulk solution [95,96]:

$$\Gamma = \left\langle N_3^{local} - \left( \frac{N_3^{bulk}}{N_1^{bulk}} \right) N_1^{local} \right\rangle$$  \hspace{1cm} (1.8)

where \(N\) refers to the number of water (\(N_1\)) and cosolvent (\(N_3\)) molecules in the vicinity of a protein (\(N_1^{local}\)) and in the bulk (\(N_1^{bulk}\)). A positive preferential binding parameter, \(\Gamma_{23}\), signifies accumulation of cosolvent near the protein surface relative to the bulk, while a negative \(\Gamma_{23}\) indicates preferential exclusion [95].

Lee and Timasheff [97] have measured the preferential binding parameters of twelve proteins in the destabilizing cosolvent guanidine hydrochloride. These authors have observed that the proteins exhibit positive values of \(\Gamma_{23}\) from 0 to 55 [97]. Prakash et al. [98] have observed positive preferential binding parameters ranging from 2 to 96 for a set of nine unfolded proteins in 8 M urea [98]. A positive sign in preferential binding parameter suggests that the affinity of a protein for the destabilizing cosolvents is greater than that for water, which translates into preferential exclusion of water from the protein surface [99].

In contrast to destabilizing cosolvents, Arakawa and Timasheff [100] have observed that proteins in stabilizing cosolvents exhibit negative preferential binding parameters, which further decrease with an increase in cosolvent concentrations. The preferential binding parameter of lysozyme decreases from \(-6.45\) to \(-13.5\) as the concentration of glycine betaine (GB) increases from 0.7 to 2.0 M, while the preferential binding parameter of bovine serum albumin decreases from \(-26.9\) to \(-72.5\) over the same range of GB concentrations [100]. In general, destabilizing cosolvents exhibit positive \(\Gamma_{23}\), whereas stabilizing cosolvents exhibit negative \(\Gamma_{23}\) [93]. These results
indicate that stabilizing cosolvents are preferentially excluded from the vicinity of the protein, whereas destabilizing cosolvents interact preferentially with the protein [101].

Although equilibrium dialysis experiments provide valuable insights into the distribution of water and cosolvent molecules around proteins, the preferential binding parameter describes the differential interactions of a protein with water and cosolvent over the entire solvent exposed protein surface. This technique is limited to macromolecules such as proteins and nucleic acids. As a result, equilibrium dialysis experiment cannot be applied to gain insights into specific interactions between individual protein groups and cosolvents.

1.2.4 Vapor pressure osmometry

Record’s group [102] has developed a method in which vapor pressure osmometry is used to measure preferential interaction coefficient. Vapor pressure osmometry measures the osmolality, Osm, of a solution which is related to the activity of the water, Osm = −55.5 ln a1 [102]. Specifically, the excess osmolality, ∆Osm, of the solution is a measure of favorable or unfavorable interactions between solute and cosolvent relative to their interactions with water [103]:

\[
\Delta \text{Osm} = \text{Osm}(m_2, m_3) - \text{Osm}(m_2, 0) - \text{Osm}(0, m_3)
\]  \hspace{1cm} (1.9)

where Osm($m_2, m_3$) is the osmolality of a solution with the solute and cosolvent concentrations of $m_2$ and $m_3$, respectively; Osm($m_2, 0$) represents the osmolality of a solution with the solute concentration of $m_2$ in the absence of cosolvents; and Osm($0, m_3$) is the osmolality of a solution with the cosolvent concentration of $m_3$ in the absence of the solute. Experimental values of ∆Osm yield the derivative of the chemical potential of a solute with respect to the molality of cosolvent, $\mu_{23} = (\partial \mu_2 / \partial m_3)_{m_2}$, via the following relationship [103,104]:

\[
\mu_{23} \approx RT \frac{\Delta \text{Osm}}{m_2 m_3}
\]  \hspace{1cm} (1.10)

In this approach, the solution osmolality is determined as a function of osmolyte in the presence and absence of a solute to determine the preferential binding parameter via $\Gamma_{23} = -\mu_{23}/\mu_{33}$, where $\mu_{33} = (\partial \mu_3 / \partial m_3)_{m_2}$ [103,104].

Experimental data on preferential interaction parameter have been interpreted within the framework of the local-bulk partitioning model to characterize interactions between cosolvents and solvent-exposed protein groups [105,106]. In this model, the solution is separated into two domains: the local domain near the protein surface and the bulk domain sufficiently distant from the protein where the distribution of cosolvent and water molecules is not perturbed by
the protein. This model predicts a linear dependence between preferential binding coefficient \( \Gamma_{23} \) and \( m_3^{\text{bulk}} \) and solvent-accessible surface area (ASA) of the protein [105, 106]. The local-bulk partition coefficient, \( K_P = m_3^{\text{local}} / m_3^{\text{bulk}} \), is defined as the ratio of the molal concentration of cosolvent in the local domain, \( m_3^{\text{local}} \), to that in the bulk domain, \( m_3^{\text{bulk}} \). The relationship between \( K_P \) and ASA can be expressed as [105,106]:

\[
\frac{\Gamma_{23}}{m_3^{\text{bulk}} \text{ASA}} \approx \frac{(K_P - 1)b_1}{m_1}
\]  

(1.11)

where \( b_1 \) is the average number of water molecules in the local domain per 1 Å² of protein surface.

The partition coefficient describes accumulation or exclusion of cosolvent in the vicinity of a protein [102,105,106]. A \( K_P \) between 0 and 1 is indicative of the exclusion of cosolvent from the protein surface, while \( K_P \) greater than 1 indicates accumulation. By analyzing \( \Gamma_{23} \) data using the local-bulk domain model, Record’s group [107] has obtained partition coefficients that characterize the accumulation or exclusion of various cosolvents near native bovine serum albumin (BSA). In urea, the value of \( K_P \) of 1.10±0.04 suggests accumulation of urea around BSA, whereas in glycine betaine, the value of \( K_P \) of 0.14 indicates exclusion of the cosolvent [106,107]. Zhang and colleagues [102] have reported that \( \Gamma_{23} \) depends linearly on the molal concentration of cosolvents in the bulk, \( m_3^{\text{bulk}} \), consistent with the local-bulk partitioning model. These authors have measured osmolality as a function of cosolvent concentrations in the presence and absence of BSA. For glycine betaine, the value of \( \partial \Gamma_{23} / \partial m_3^{\text{bulk}} = -49\pm4 \) which is in agreement with glycine betaine being completely excluded from the hydrated surface of BSA. Whereas, for urea, the value of \( \partial \Gamma_{23} / \partial m_3^{\text{bulk}} = 6\pm1 \) suggests a moderate extent of accumulation of urea in the vicinity of BSA [102].

Courtenay et al. [107] have dissected \( K_P \) into contributions from various protein groups by measuring partition coefficients for the native, \( K_P^{\text{nat}} \), and unfolded, \( K_P^{\text{unf}} \), protein states. For urea, similar values of \( K_P^{\text{nat}} \) (1.10±0.04) and \( K_P^{\text{unf}} \) (1.12±0.01) have been observed. These authors have proposed that the equality originates from the similarity of the polar peptide backbone fraction of ASA (∼0.13) for the folded and unfolded conformations [107]. The partition coefficient for polar peptide backbone, \( K_P^{\text{ppb}} = 2.1\pm0.1 \), is greater than that for other protein surface, \( K_P^{\text{other}} = 1.0\pm0.1 \) [107]. This result suggests that urea is strongly accumulated near polar peptide backbone compared with other types of protein ASA.

Guinn et al. [104] have characterized the interaction of urea with various types of molecular surfaces by comparing the respective partition coefficients, \( K_P \). These authors have observed
that accumulation of urea near various atoms changes in the following order:

amide O \sim \text{aromatic C} > \text{carboxylate O} > \text{amide N} > \text{hydroxyl O} > \text{aliphatic C}

The individual partition coefficients range between 1.28 for amide O and 1.03 for aliphatic C [104]. The strong accumulation of urea near amide O and carboxylate O suggests formation of hydrogen bonds between urea and the two O groups [104]. Even though urea shows only a relatively weak accumulation around aliphatic C, this interaction contributes significantly to the effect of urea on protein stability because the composition of the surface exposed in unfolding is predominantly (~ 65 %) aliphatic C [22,104]. In contrast to urea, Capp and colleagues [103] have reported that glycine betaine is excluded from amide O (\(K_P = 0.24 \pm 0.27\)) and carboxylate O (\(K_P = 0.22 \pm 0.06\)). Since GB lacks the capacity to act as a hydrogen bond donor, water molecules preferentially interact with amide O and carboxylate O. As a result, GB becomes excluded from these groups [103].

While vapor pressure osmometry has provided valuable insights into the interactions of cosolvent with proteins, the contributions of the individual functional groups have been evaluated from dissecting the model compound data based on the ASA of functional groups [104,108]. However, employing group additivity should be viewed with caution because each group may not necessarily reside in a similar microenvironment [46]. The local environment around a functional group should be taken into consideration when interpreting cosolvent–protein interactions.

### 1.3 Volumetric Techniques

#### 1.3.1 Investigation of protein–water interactions

Volumetric observables, such as the partial molar volume, \(V^o\), and compressibility, \(K^o\), of a solute, are influenced by the entire spectrum of solvent molecules interacting with the solute [109–111]. In proteins, \(V^o\) and \(K^o\) reflect, in addition to solvation, the internal spatial architecture of the folded polypeptide chain and the packing of amino acid side chains. Conformational transitions affect the volumetric properties of a protein through the elimination of intramolecular voids, loosening of intraglobular packing, and exposure of previously buried atomic groups to solvent [112].

The partial molar volume of a solute is the partial derivative of the solution volume, \(V\), with respect to the number of moles, \(N_i\), of the \(i\)-th solute at a constant temperature, \(T\), and pressure, \(P_i\):
Based on scaled particle theory (SPT), the partial molar volume of a solute is the sum of the intrinsic, \( V_M \), thermal, \( V_T \), and interaction, \( V_I \), contributions, and the ideal term, \( \beta T_0 RT \) [113, 114]:

\[
V^\circ = V_M + V_T + V_I + \beta T_0 RT
\]  

(1.13)

where \( \beta T_0 \) is the coefficient of isothermal compressibility of the solvent. The ideal term reflects the volume effect arising from the translational degree of freedom available to the solute molecule [114]. At 25 °C, the ideal term, \( \beta T_0 RT \), is equal to 1.1 cm³ mol⁻¹ [115].

The intrinsic volume, \( V_M \), of a solute is the geometric volume that is impenetrable to a spherical probe with a radius of 1.4 Å [116]. The intrinsic volume of a large macromolecule such as a protein can be calculated from structural data using the geometric algorithm presented by Richards [116]. The thermal volume, \( V_T \), is the volume of the void created around the solute due to the thermally-induced mutual vibrational motions of solute and solvent molecules as well as due to steric and structural effects [117]. Theoretically, the vibration component of \( V_T \) should subside to zero at absolute zero [118]. The thermal volume is proportional to its solvent accessible surface area, \( S_A \), with the proportionality coefficient equal to the effective thickness, \( \delta \), of the void volume [113, 117, 118]:

\[
V_T = \delta S_A
\]  

(1.14)

The interaction volume, \( V_I \), represents a change in volume due to interactions between the solute and solvent molecules. For a protein transition or protein–ligand binding events, a change in the interaction volume, \( \Delta V_I \), reflects the redistribution of water molecules between the bulk and hydration phases [112]:

\[
\Delta V_I = \sum_i \Delta n_{hi}(V_{hi} - V_0)
\]  

(1.15)

where \( \Delta n_{hi} \) is the number of water molecules taken up by the \( i \)-th domain of the ligand or the protein upon the transition or the association event; and \( V_{hi} \) and \( V_0 \) are the partial molar volumes of water of hydration of the \( i \)-th solute domain and bulk water, respectively. The partial molar volume of water solvating proteins, \( V_h \), is roughly 10% smaller than that of
bulk water [119,120]. The \((V_h - V_0)\) term is, thus, estimated to be \(-1.8 \text{ cm}^3 \text{ mol}^{-1}\) [119,120]. Therefore, a change in the number of waters of hydration, \(\Delta n_h\), associated with a protein transition or protein–ligand binding event can be estimated from \(\Delta V_I\).

Compressibility may provide unique volumetric insights into changes in hydration and dynamics of a protein. The isothermal compressibility, \(K_T\), of a system can be defined as:

\[
K_T = \beta_T V = -\left(\frac{\partial V}{\partial P}\right)_T
\]

(1.16)

where \(\beta_T = -V^{-1}(\partial V/\partial P)_T\) is the coefficient of isothermal compressibility of the system. The adiabatic compressibility of a system is the negative pressure derivative of its volume at constant entropy \((S)\):

\[
K_S = \beta_S V = -\left(\frac{\partial V}{\partial P}\right)_S
\]

(1.17)

where \(\beta_S = -V^{-1}(\partial V/\partial P)_S\) is the coefficient adiabatic compressibility of the system. The partial molar compressibility, \(K^\circ\), of a solute is defined as the partial derivative of the compressibility, \(K\), of the solution with respect to the number of moles, \(N_i\), of the \(i\)-th solute at a constant temperature, and pressure:

\[
K^\circ_T = \left(\frac{\partial K_T}{\partial N_i}\right)_{T,P}
\]

(1.18)

\[
K^\circ_S = \left(\frac{\partial K_S}{\partial N_i}\right)_{T,P}
\]

(1.19)

The relationship between \(K^\circ_T\) and \(K^\circ_S\) is given by:

\[
K^\circ_T = K^\circ_S + \frac{T\alpha_0^\circ}{\rho_0 c_p^\circ} \left(\frac{2E^\circ}{\alpha_0} - \frac{C_P^\circ}{\rho_0 c_p^\circ}\right)
\]

(1.20)

where \(c_{P_0}\) is the specific heat capacity at constant pressure of the solvent; \(\rho_0\) is density of the solvent; \(\alpha_0\) is the coefficient of thermal expansion of the solvent; \(E^\circ\) is the partial molar expansibility of the solute; and \(C_P^\circ\) is the partial molar heat capacity of the solute [121]. The difference between the isothermal and adiabatic compressibilities of a solute is not large due to the large heat capacity and small expansibility of water [121].

A change in adiabatic compressibility, \(\Delta K_S\), accompanying protein transition or ligand binding
can be presented as the sum of the intrinsic, \( \Delta K_M \), and hydration, \( \Delta K_h \), contributions [111,121]:

\[
\Delta K_S = \Delta K_M + \Delta K_h = \Delta K_M + \sum_i \Delta n_{hi}(K_{hi} - K_0)
\]

(1.21)

where \( K_{hi} \) and \( K_0 \) are the partial molar adiabatic compressibilities of water of hydration of the \( i \)-th solute domain and bulk water, respectively.

The hydration-induced change in solvent compressibility reflects the interactions of solvent molecules with protein groups. For globular proteins, the intrinsic compressibility, \( K_M = \beta_M V_M \), reflects the compressibility of the void volume within a protein due to the imperfect packing of its solvent-inaccessible interior [111,122]. The average value of \( \beta_M \) of \( 25 \times 10^{-6} \) bar\(^{-1} \) has been estimated based on the regression analysis of the partial molar adiabatic compressibility of a large number of globular proteins [112,123].

In general, intrinsic compressibility provides a measure of mechanical flexibility of globular proteins. The value of \( K_M \) is sensitive to conformational dynamics of a protein given its relationship with volume fluctuations [124,125]. The mean-square fluctuations of the intrinsic volume, \( \langle \delta V \rangle^2 \), provide quantitative insights into the transition- and binding-induced changes in protein dynamics [124,125]. The value of \( \langle \delta V \rangle^2 \) of a protein is directly proportional to \( K_M \):

\[
\langle \delta V^2 \rangle = k_B T K_M
\]

(1.22)

where \( k_B \) is the Boltzmann constant.

Volumetric properties of protein–ligand interactions have been used to study binding-induced changes in protein hydration and conformation [15,115,126–128]. Kamiyama and Gekko [129] have measured the partial specific volume, \( v^o \), and the coefficient of adiabatic compressibility, \( \beta_S = (K_S / v^o) \), of dihydrofolate reductase in the presence of various ligands. Binding of these ligands result in large changes in \( v^o \) and \( \beta_S \) of 0.734–0.754 cm\(^3\) g\(^{-1} \) and (6.6–9.8) \( \times 10^{-6} \) bar\(^{-1} \), respectively [129]. Based on the X-ray crystal structures of the complexes, the values of \( v^o \) and \( \beta_S \) increase with an increase in the volume of internal cavity of proteins. Comparison of \( \beta_S \) with the cavity distribution in the crystal structures reveals that the magnitude of conformational flexibility of a protein–ligand complex is mainly determined by the total cavity volume [129]. Filfil and colleagues [130] have measured changes in volume and compressibility associated with the binding of D-glucose to yeast hexokinase PII. These authors have observed that the coefficient of adiabatic compressibility, \( \beta_S \), of the protein’s interior decreases by 1%. Significant dehydration of the protein upon ligand binding has been observed with 332±20 water molecules being released to the bulk [130]. This value agrees well with a release of 326±35 water molecules.
measured by the osmotic stress technique [64]. The coefficient of adiabatic compressibility, \( \beta_S \), of the protein’s interior decreases by 1%. Therefore, it has been inferred that the ligand-bound state of hexokinase is less hydrated, more rigid, and less dynamic compared to its unbound state.

### 1.3.2 Investigation of protein–urea interactions

Volumetric parameters may provide novel thermodynamic insights into the differential solute–cosolvent versus solute–principal solvent interactions [131, 132]. However, due to the lack of conceptual formalism, it has been difficult to separate the volumetric contribution of protein–cosolvent interactions from that of protein–water interactions.

A recently derived formalism based on the combination of statistical thermodynamics and solvent exchange model (SEM) has enabled one to use volumetric data for gaining insights into the molecular origins of the protein–cosolvent interactions [131–134]. In this formalism, an elementary reaction of cosolvent (C) binding to a binding site on the solute (S) surface is accompanied by a release of \( r \) water (W) molecules to the bulk: [131,135]:

\[
SW_r + C \leftrightarrow SC + rW \tag{1.23}
\]

A change in free energy associated with the transfer of a protein from water to a concentrated cosolvent solution, \( \Delta G_{tr} \), is the sum of the differential free energy of cavity formation, \( \Delta \Delta G_C \), and the differential interaction free energy, \( \Delta \Delta G_I \) [67]. If a protein has \( n \) water-binding sites, \( n/r \) is the maximum number of cosolvent-binding sites. Under the assumption of \( n/r \) identical and independent cosolvent binding sites, transfer free energy, \( \Delta G_{tr} \), is given by the relationship [131,136]:

\[
\Delta G_{tr} = \Delta \Delta G_C - \left( \frac{n}{r} \right) RT \ln \left[ \left( \frac{a_1}{a_{10}} \right)^r + k \left( \frac{a_3}{a_{10}} \right) \right] \tag{1.24}
\]

where \( \Delta \Delta G_C \) is the differential free energy of cavity formation; \( a_1 \) and \( a_3 \) are the activity of water and cosolvent, respectively; \( a_{10} \) is the activity of water in the absence of cosolvent; and \( k \) is the effective equilibrium constant for the reaction in which a cosolvent molecule replaces \( r \) water molecules when binding to its binding sites. In Equation (1.24), the second term represents the free energy of direct interaction of solute with solvent molecules.

By differentiating Equation (1.24) with respect to pressure, one derives the expression for a change in volume accompanying solute transfer from water to a concentrated cosolvent solution.
\[ \Delta V^\circ = \Delta V_C - \gamma_1 n \Delta V_1^\circ + \frac{\Delta V \left( \frac{a_3}{a_1} \right) k}{1 + \left( \frac{a_3}{a_1} \right) k} \]  

(1.25)

where \( \Delta V_C \) is the differential volume of the cavity enclosing a solute in a concentrated cosolvent solution and water; \( \Delta V = \Delta V_0 + \gamma_1 r \Delta V_1^\circ - \gamma_3 \Delta V_3^\circ \) is the change in volume associated with replacement of water of hydration with cosolvent in a concentrated cosolvent solution; \( \Delta V_0 \) is the solvent exchange volume in an ideal solution; \( \Delta V_1^\circ \) and \( \Delta V_3^\circ \) are the excess partial molar volumes of water and cosolvent in a concentrated solution, respectively; and \( \gamma_1 \) and \( \gamma_3 \) are the correction factors reflecting the influence of the bulk solvent on the properties of solvating water and cosolvent, respectively [131,135].

The relationship for a change in isothermal compressibility accompanying the water-to-cosolvent transfer of a solute, \( \Delta K_T^\circ \), can be obtained by differentiating Equation (1.25) with respect to pressure:

\[ \Delta K_T^\circ = \Delta K_{TC} - \gamma_1 n \Delta K_{T1}^\circ + \frac{\Delta K_T \left( \frac{a_3}{a_1} \right) k}{1 + \left( \frac{a_3}{a_1} \right) k} + \frac{\Delta V^2 \left( \frac{a_3}{a_1} \right) k}{RT \left[ 1 + \left( \frac{a_3}{a_1} \right) k \right]^2} \]  

(1.26)

where \( \Delta K_{TC} = -\left( \frac{\partial \Delta V_C}{\partial P} \right)_T \) is the differential compressibility of the cavity enclosing a solute in water and a concentrated cosolvent solution; \( \Delta K_T = \Delta K_{T0} + \gamma_1 r \Delta K_{T1}^\circ - \gamma_3 \Delta K_{T3}^\circ \) is the change in compressibility associated with replacement of water with cosolvent in a concentrated solution; \( \Delta K_{T1}^\circ \) and \( \Delta K_{T3}^\circ \) are the excess partial molar isothermal compressibilities of water and cosolvent in a concentrated cosolvent solution, respectively; and \( \Delta K_{T0} = -\left( \frac{\partial \Delta V_0}{\partial P} \right)_T \) is the change in compressibility associated with solvent replacement in an ideal solution. Given the large heat capacity and small expansibility of water-based solutions, the difference between the partial molar adiabatic and isothermal compressibilities of a solute is small and, as a first approximation, can be ignored. Consequently, partial molar adiabatic compressibilities, \( \Delta K_S^\circ \), can be used instead of partial molar isothermal compressibilities, \( \Delta K_T^\circ \), in Equation (1.26) [131,135].

In a recent study, Lee and colleagues [135] have investigated the interaction between urea and protein groups using volumetric measurements. These authors have measured the partial molar volume and adiabatic compressibility of amino acid side chains and oligoglycines in aqueous solutions of urea at concentrations ranging from 0 to 8 M. Volumetric measurements have been analyzed based on the solvent exchange model to determine the effective equilibrium constants, \( k \), for the binding of urea to protein groups with replacement of two waters of hydration [135].
The equilibrium constants ranging from 0.04 to 0.39 M suggest that urea interacts with variety of protein groups. Furthermore, the differential free energy of solute–solvent interactions, \( \Delta \Delta G_I \), derived from the effective equilibrium constants, \( k \), range from highly favorable to slightly unfavorable [135].

Shek et al. [133] have measured the partial molar volumes and adiabatic compressibilities of amino acid side chains and oligoglycines at glycine betaine concentrations ranging from 0 to 4 M. The solvent exchange model has been applied to analyzing the equilibrium constant, \( k \), for the reaction in which a GB molecule binds to protein functional groups replacing four water molecules. The equilibrium constants range from 0.02 to 0.32 M [133]. These values are comparable to the previous study using a destabilizing cosolvent, urea [135]. This observation suggests that the contribution of direct protein–GB interactions to the total interaction free energy is not negligible. These authors have calculated the free energy of the transfer of functional groups from water to concentrated GB solutions, \( \Delta G_{tr} \), as the sum of changes in free energy of cavity formation, \( \Delta \Delta G_C \), and the differential free energy of solute–solvent interactions, \( \Delta \Delta G_I \), in a concentrated GB solution and water [133]. The results have shown that \( \Delta G_{tr} \) is determined by a fine balance between the unfavorable \( \Delta \Delta G_C \) and favorable \( \Delta \Delta G_I \) contributions [133]. Consequently, the stabilizing action of GB originates not from the weakness of its interactions with protein groups but rather from the highly unfavorable free energy of cavity formation in a concentrated GB solution relative to that in water.

Shek et al. [137] have further applied volumetric techniques to studying the influence of protein–cosolvent interactions on globular proteins. They have measured the partial molar volumes and compressibilities of cytochrome c, ribonuclease A, lysozyme, and ovalbumin in aqueous solutions of GB at concentrations between 0 and 4 M. The analysis has revealed that a highly favorable change in \( \Delta \Delta G_I \) is compensated by an unfavorable \( \Delta \Delta G_C \), yielding a modestly unfavorable free energy for the transfer of a protein from water to a GB solution [137]. Furthermore, by comparing these parameters with the similar characteristics determined for low molecular weight analogs of proteins, it has been concluded that no significant cooperative effects are involved in protein–cosolvent interactions [133].

1.4 Research Aims

Despite the significance of protein–solvent interactions, the molecular mechanisms underlying these interactions remain poorly understood [131,136]. The specifics of protein solvation studies are dictated by the fact that the molecules of the principal solvent and cosolvent may interact with chemically and geometrically distinct loci on the protein surface, they may be localized or not localized, may be immobilized or not immobilized, and hence difficult to study. There are not many experimental techniques that enable one to nonselectively identify and characterize all
solvent and cosolvent molecules influenced by the solute. Therefore, this dissertation is aimed at characterizing protein–solvent interactions as well as changes in these interactions upon various protein binding and unfolding events by applying volumetric techniques and analyzing the resulting data within the framework of statistical thermodynamic models.

1.4.1 Quantification and characterization of changes in protein–solvent interactions accompanying protein–ligand association

One objective of this thesis is to quantify and characterize changes in hydration accompanying protein–ligand associations. Volumetric techniques were applied to studying the role of hydration in the binding of tri-N-acetylglucosamine [(GlcNAc)₃] to lysozyme and the binding of cyclic adenosine monophosphate (cAMP) to the cAMP-binding domain (CBD) of exchange protein directly activated by cAMP (EPAC). These are thermodynamically and structurally well-characterized systems which makes it possible to develop an interpretation algorithm for rationalizing volume and compressibility data in terms of hydration. A combination of volumetric data and calorimetric results enables one to evaluate the energetic contributions of water molecules released or taken up upon protein-ligand association events.

Being a well-characterized system, the lysozyme–(GlcNAc)₃ complex is an attractive model for deriving a detailed understanding of the effect of urea on protein–ligand binding reactions. We analyzed the urea-dependent volumetric properties of lysozyme–(GlcNAc)₃ association. It should be noted that lysozyme remains native at concentrations of urea as high as the solubility limit of the latter [138]. At 25 °C, lysozyme remains native at pH 7.0 up to 9.5 M urea, while undergoing a urea-induced unfolding at pH 2.0 [139]. This property of lysozyme enables one to employ lysozyme as a model to study the binding of urea to a native protein throughout a wide concentration range.

1.4.1.1 Lysozyme

Lysozyme damages bacterial cell wall by catalyzing hydrolysis of a 1,4-β-linkages between the C-1 of N-acetylmuramic acid and the C-4 of N-acetylglucosamine in the peptidoglycan layer of the bacterial cell wall [140]. The structures of lysozyme associated with N-acetylglucosamines have been investigated by X-ray crystallographic, powder diffraction, and NMR spectroscopic techniques [141–145]. The three sugar rings of (GlcNAc)₃ become buried within the binding cleft of hen-egg lysozyme [141]. The specificity of the binding results mainly from an extensive network of hydrogen bonds between the protein and the inhibitor as well as the five bound water molecules [141]. The major (GlcNAc)₃-induced conformational change involves the amino acid residues in the vicinity of subsite C of the active site of the enzyme [141, 142, 144]. However,
changes on a smaller scale occur in regions remote from the active site [143].

The binding of (GlcNAc)$_3$ to lysozyme is an enthalpy-driven process which proceeds with an unfavorable change in entropy [146]. Pressure perturbation calorimetric measurements have revealed that the binding of (GlcNAc)$_3$ to lysozyme is accompanied by changes in standard state free energy, $\Delta G^\circ$, enthalpy, $\Delta H^\circ$, and entropy, $\Delta S^\circ$ of $-6.9$ kcal mol$^{-1}$, $-14.0 \pm 2.4$ kcal mol$^{-1}$, and $-23.7$ cal mol$^{-1}$ K$^{-1}$, respectively [146].

### 1.4.1.2 EPAC

EPAC is a guanine nucleotide exchange factor for the small GTPases Rap1 and 2, which is directly activated by cAMP [147, 148]. Rap proteins are GTP-binding proteins that serve as molecular switches between a GDP-bound inactive and a GTP-bound active state. EPAC1 acts by accelerating the slow intrinsic GDP dissociation from Rap, which subsequently associates with GTP. Two highly homologous isoforms of EPAC, namely, EPAC1 and EPAC2, which are found in mammalian cells, consist of a single polypeptide chain with an N-terminal regulatory and a C-terminal catalytic region [149, 150]. The regulatory domain of EPAC1 contains an N-terminal dishevelled/Egl-10/pleckstrin domain, followed by the cAMP-binding domain [148]. CBD is responsible for the cAMP-mediated EPAC activation, whereas the dishevelled/Egl-10/pleckstrin domain is responsible for its membrane binding and subsequent localization.

The binding of cAMP to the CBD of EPAC causes a hinge-like rotation of a C-terminal helix ($\alpha 6$) toward the $\beta$-barrel core characteristic of CBDs [151,152]. The binding and the associated structural reorganization of the protein are accompanied by a complex interplay of the enhanced and quenched dynamics of amino acid residues that are proximally as well as remotely located from the binding site in both the $\alpha$ and the $\beta$ subdomains of the protein [152]. Importantly, NMR studies have suggested that the cAMP-mediated alteration in the dynamic profile of the protein is involved in the entropically controlled allostery of the cAMP-dependent EPAC activation [152–158].

Thermodynamic studies have characterized the association of EPAC1 with cAMP as an exothermic process with a binding affinity, $K_b$, of $\sim 3.3 \times 10^5$ M$^{-1}$ ($\Delta G_b = -7.5$ kcal mol$^{-1}$) [159,160]. Based on the isothermal titration calorimetric data, the binding enthalpy, $\Delta H_b$, is estimated to be between $-8$ and $10$ kcal mol$^{-1}$ [159]. Thus, the cAMP–EPAC1 binding is accompanied by an unfavorable change in entropy of $-5 \pm 3$ cal mol$^{-1}$ K$^{-1}$.
1.4.2 Urea interaction with folded and unfolded proteins

We analyzed urea-dependent volumetric properties of native and unfolded proteins. We had previously developed a statistical thermodynamic-based algorithm that could be used in conjunction with volumetric data to extract thermodynamic parameters of elementary solute–cosolvent interactions [131,161]. With this algorithm, we have developed volumetric libraries of interactions of urea with low-molecular weight compounds [131]. In this thesis, we expand this line of research to studying solute–solvent interactions in four globular proteins, namely, apocytochrome c, hen egg white lysozyme, ribonuclease A, and α-chymotrypsinogen A, in binary water–urea mixture.

The specific set of proteins is chosen to cover the full range of conformational states accessible to proteins extending from fully folded to fully unfolded. Apocytochrome c is unfolded in the absence of urea, while lysozyme remained fully folded at all concentrations of urea studied (0–8 M). Hence, these proteins did not undergo any urea-induced conformational transitions, thereby, enabling one to separately study urea interactions with folded (lysozyme) and unfolded (apocytochrome c) protein states. In contrast, ribonuclease A and α-chymotrypsinogen A exhibit urea-induced unfolding transitions. When analyzing the volumetric responses of these protein to addition of urea, we use the thermodynamic parameters determined for urea interactions with folded and unfolded conformational states derived from our studying lysozyme and apocytochrome c.

1.4.3 Thermodynamics of protein–urea interactions

One objective of this work is to characterize the full thermodynamic profile of the binding of urea to a native protein. For the first time, we have used the solvent exchange model in conjunction with urea-dependent volumetric measurements to characterize the interactions of urea with a native protein. Previously, the energetic parameters for solute–urea interactions were obtained on unfolded proteins or small analogues of proteins. Based on calorimetric titration data, the average values of $\Delta G^\circ$, $\Delta H^\circ$, and $\Delta S^\circ$ for the binding of urea to unfolded cytochrome c, ribonuclease A, and lysozyme have been found to be $1.6 \pm 0.2$ kcal mol$^{-1}$, $-2.1 \pm 0.5$ kcal mol$^{-1}$, and $-13 \pm 2$ cal mol$^{-1}$ K$^{-1}$, respectively [162]. Data on calorimetric heat of dissolution measurements have revealed that the binding of urea to the peptide moiety of cyclic dipeptides is characterized by the values of $\Delta G^\circ$, $\Delta H^\circ$, and $\Delta S^\circ$ of $2.2 \pm 0.3$ kcal mol$^{-1}$, $-8.1 \pm 0.7$ kcal mol$^{-1}$, and $-35 \pm 3$ cal mol$^{-1}$ K$^{-1}$, respectively [163]. These studies have demonstrated that the unfavorable change in standard state free energy results from an unfavorable change in entropy that prevails over a favorable change in enthalpy.

At neutral pH, lysozyme retains its native conformation between 0 and 8 M urea within a wide
Chapter 1. General Introduction

We have measured the volumetric parameters of lysozyme between 18 and 45 °C over the entire range of urea concentrations. In contrast to previous works, our study enables us to evaluate the full thermodynamic profile of an elementary solute–cosolvent association reaction for a protein that retains its native conformation throughout the entire experimental range of applied temperatures and urea concentrations. While our results could not be directly related in a strictly quantitative manner to any of the previously reported data, a qualitative comparison could provide insights into the nature of interactions of urea with native and unfolded proteins.

1.5 Outline of the Thesis

Chapter 2 describes application of high precision densimetric and ultrasonic velocimetric measurements to characterizing the binding of tri-N-acetylglucosamine [(GlcNAc)_3] to lysozyme. Changes in volume, \( \Delta V \), and adiabatic compressibility, \( \Delta K_S \), accompanying this protein–ligand association event were measured. The volumetric results were analyzed in conjunction with X-ray crystallographic data in terms of the binding-induced changes in hydration and dynamic properties of lysozyme. We also discussed the importance of volumetric insights into the molecular origins of protein binding events.

Chapter 3 describes a study in which a combination of densimetric and ultrasonic velocimetric techniques was employed to characterize the volumetric properties of the association of the cAMP-binding domain of EPAC1 with cAMP. We measured changes in volume, \( \Delta V \), and adiabatic compressibility, \( \Delta K_S \), associated with the binding of cAMP to the CBD of EPAC1. We used the resulting volumetric data along with structural data to estimate changes in hydration accompanying the ligand binding to EPAC1. A change in protein dynamics was determined from the change in the intrinsic coefficient of adiabatic compressibility. We also estimated the hydration, \( \Delta S_{\text{hyd}} \), and configurational, \( \Delta S_{\text{conf}} \), contributions to the binding entropy, \( \Delta S_b \). In general, we emphasize how volumetric data can be used to evaluate changes in protein hydration and dynamics upon ligand binding.

Chapter 4 describes a statistical thermodynamic approach to analyzing urea-dependent volumetric properties of proteins. We analyzed our urea-dependent data on the partial molar volume and adiabatic compressibility of lysozyme, apocytochrome c, ribonuclease A, and \( \alpha \)-chymotrypsinogen A. Lysozyme remains folded, while apocytochrome c is unfolded throughout the experimental urea concentration, whereas, ribonuclease A and \( \alpha \)-chymotrypsinogen A undergo urea-induced unfolding transitions. Thus, our data allowed us to characterize protein–urea interactions in both the native and unfolded states. We evaluated the equilibrium constant, \( k \), and changes in volume, \( \Delta V_0 \), and compressibility, \( \Delta K_{T_0} \), for a reaction in which urea binds to a protein with a concomitant release of two water molecules from the hydration shell to the
bulk. The analysis produced the thermodynamic properties of elementary protein–urea association reactions while also yielding estimates of the effective solvent-accessible surface areas of the native and unfolded protein states.

Chapter 5 reports the first experimental characterization of the thermodynamic profile of urea binding to a native protein. We measured the volumetric parameters of lysozyme as a function of urea within a temperature range of 18 to 45 °C. At neutral pH, lysozyme remains in its native conformation between 0 and 8 M urea over the entire temperature range studied. Consequently, our measured volumetric properties solely reflect the interactions of urea with native lysozyme without being contributed by urea-induced conformational transitions. We treated our data within the framework of a statistical thermodynamic analytical model in which protein–urea interactions are viewed as solvent exchange in the vicinity of the protein. The van’t Hoff analysis of the temperature dependence of the equilibrium constant, \( k \), for the protein–urea binding reaction produced changes in free energy, \( \Delta G^\circ \), enthalpy, \( \Delta H^\circ \), and entropy, \( \Delta S^\circ \), accompanying the binding. We discussed ramifications of our results for providing insights into the combined effect of urea and temperature on the conformational preferences of proteins.

Chapter 6 combines experimental and theoretical approaches to investigate the influence of urea on a protein–ligand association event. We applied fluorescence measurements to determining the affinity of the inhibitor \((\text{GlcNAc})_3\) for lysozyme at urea concentrations ranging from 0 to 8 M. Notwithstanding that, at room temperature and neutral pH, lysozyme retains its native conformation up to the solubility limit of urea, the affinity of the ligand for the protein markedly decreases as the concentration of urea increases. We analyzed the urea dependence of the binding free energy within the framework of a statistical thermodynamics-based model that accounts for the excluded volume effect and direct solute–solvent interactions. The analysis revealed that the detrimental action of urea on the lysozyme–inhibitor binding originates from the competition between the free energy contributions of the excluded volume effect and direct solute–solvent interactions.

Chapter 7 summarizes the overall research conducted in my doctoral studies. The implications and recommendations for future studies are presented.
Chapter 2

Volumetric Characterization of Tri-\(N\)-acetylglucosamine Binding to Lysozyme


Author’s Contribution: Ikbae Son performed most experiments, analyzed the data, and participated in the preparation of the manuscript.

2.1 Abstract

Volumetric characteristics of protein recognition events determine the direction of pressure-induced shifts in the recognition reaction, while also providing insights into the structural, dynamic, and hydration changes. We report changes in volume, \(\Delta V\), and adiabatic compressibility, \(\Delta K_S\), accompanying the binding of tri-\(N\)-acetylglucosamine \([\text{GlcNAc}]_3\) to lysozyme at 25°C in a pH 5.5 sodium acetate buffer. We interpret our measured changes in volume and compressibility in terms of changes in hydration and dynamic properties of the protein. On the basis of our \(\Delta V\) data, we find that 79 \(\pm\) 44 water molecules are released to the bulk from the hydration shells of the protein and the ligand. Our \(\Delta K_S\) data suggest a 4 \(\pm\) 2\% decrease in the mean-square fluctuations of the intrinsic volume of the protein, \(\langle \delta V_M^2 \rangle\) (or a 2\% decrease in \(\delta V_M\)). Thus, the trisaccharide-bound state of the enzyme is less hydrated, more rigid, and less dynamic compared to the unbound state. In general, we discuss the importance
of volumetric insights into the molecular origins of protein recognition events.

## 2.2 Introduction

A detailed understanding of the thermodynamic forces participating in the control of ligand–protein recognition events is required for characterization and, even to a larger degree, prediction of the affinity and specificity of such events [13, 14, 18, 164–166]. Among these forces, hydration/dehydration of interacting surfaces and, possibly, more distant loci of the protein occupy a special place both due to the sizable contribution of hydration to the binding affinity and specificity and the difficulties involved in quantifying changes in hydration and evaluating its net thermodynamic effects [13, 14, 18]. Changes in hydration and the related thermodynamics have been implicated as a major factor contributing to the ubiquitous phenomenon of enthalpy-entropy compensation in protein recognition [13, 167, 168].

The free energies of waters solvating protein hydration sites may be higher or lower relative to waters in the bulk phase [17, 169]. This notion is consistent with a picture in which the ligand-induced release of water molecules from the protein hydration shell to the bulk may contribute positively or negatively to the affinity of the binding reaction. In fact, it has been proposed that “druggable” domains on the protein surface are all characterized by high-energy hydration sites which drives the binding affinity [17].

Characterization of the role of hydration in protein–ligand association requires the use of physical parameters that can discriminate between waters solvating various protein and ligand sites and waters in the bulk. Volumetric observables such as volume, compressibility, and expansibility offer one way for tackling the problem of hydration in protein recognition [55, 109, 111, 112, 170]. These parameters are sensitive to solute-induced changes in the properties of waters of hydration. Significantly, they sense the entire population of thermodynamically altered water molecules and not restricted just to highly localized or immobilized waters as is the case for X-ray and NMR. In addition, the volumetric properties of water molecules interacting with charged, polar, and nonpolar groups are distinct in both their absolute magnitude and temperature dependences, a feature that, potentially, can be used for discriminating between the types of hydration [55, 112].

A survey of the literature reveals only a few volumetric studies of protein binding events [127, 130, 171, 172]. The dearth of volumetric data is unfortunate. Volumetric data are complementary to more abundant calorimetric data, and furthermore, their combination may produce additional insights into the thermodynamic forces driving protein recognition. To implement to the full extent the possibilities provided by such a combination, larger volumetric databases on protein–ligand association are required.
The complexation between tri-\(N\)-acetylglucosamine \((\text{GlcNAc})_3\) and lysozyme has been well characterized by structural and thermodynamic means. This makes the lysozyme–\((\text{GlcNAc})_3\) an attractive target for a volumetric characterization. The thermodynamics of the binding of \((\text{GlcNAc})_3\) to hen egg-white lysozyme has been studied in a number of publications [146,173].

This is an enthalpy-driven process with an unfavorable change in entropy [146].

The structures of various avian lysozymes associated with \(N\)-acetylglucosamines have been investigated by X-ray crystallographic, powder diffraction, and NMR spectroscopic techniques [141–145]. The three sugar rings of \((\text{GlcNAc})_3\) become buried within the binding cleft of hen-egg lysozyme to a different degree [141]. The specificity of the binding results mainly from an extensive network of hydrogen bonds between the protein and the inhibitor as well as the five bound water molecules [141]. The major \((\text{GlcNAc})_3\)-induced conformational change involves the amino acid residues in the vicinity of subsite C of the active site of the enzyme [141,142,144]. However, changes on a smaller scale occur in regions remote from the active site [143]. Importantly, judging by the average side-chain B values of the protein, the binding of \((\text{GlcNAc})_3\) causes a stiffening of the atoms of the amino acid residues directly interacting with the bound inhibitor [141]. For example, in the native state, the temperature factors of the five- and six-membered rings of Trp62 are 39 (or 32 \(\text{Å}^2\)), respectively, while decreasing to 22 (or 14 \(\text{Å}^2\)), respectively, in the trisaccharide-bound state [141].

In this work, we report changes in volume and adiabatic compressibility accompanying the binding of \((\text{GlcNAc})_3\) to hen egg-white lysozyme. We interpret our volumetric results in conjunction with X-ray crystallographic data on the unligated enzyme and its complex with the inhibitor to estimate the binding-induced changes in hydration and internal dynamics of the protein. We also discuss implications of our results for understanding the various components of thermodynamic forces driving protein–ligand complexation events.

### 2.3 Materials and Methods

#### 2.3.1 Materials

Lysozyme from chicken egg white was purchased from Sigma-Aldrich Canada (Oakville, ON, Canada) and was exhaustively dialyzed against buffer. Tri-\(N\)-acetylglucosamine [tri-\(N\)-acetylchitotriose] was acquired from V-laboratories (Covington, LA) and used without further purification. All measurements were performed in a pH 5.5 buffer consisting of 10 mM sodium acetate/acetic acid and 10 mM NaCl. The concentration of lysozyme was determined from the absorbance measured at 25 °C with a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, ON, Canada) using a molar extinction coefficient \(\varepsilon_{280} = 37\,900\,\text{M}^{-1}\text{cm}^{-1}\) [174]. For the fluorescence measurements, the protein concentration was \(\sim 35\,\mu\text{M}\). For the densimetric and
Chapter 2. Characterization of (GlcNAc)₃ Binding to Lysozyme

ultrasonic velocimetric experiments, lysozyme concentration was ∼ 100 µM.

2.3.2 Fluorescence

Fluorescence intensity measurements were performed at 25 °C using an Aviv model ATF 105 spectrofluorometer (Aviv Associates, Lakewood, NJ). Fluorescence titration profiles were measured by the incremental addition of aliquots of (GlcNAc)₃ to a 10 mm path length cell containing a known amount of lysozyme. The protein samples were excited at 296 nm, and the intensity of emission light was recorded through a monochromator at 345 nm. When calculating the relative fluorescence intensity of lysozyme, we have taken into account the change in the concentration of the protein upon each addition of the titrant [(GlcNAc)₃].

2.3.3 Volumetric and densimetric measurements

All densimetric and ultrasonic velocimetric investigations reported here were conducted at 25 °C. Densities were measured with a precision of ±1.5 × 10⁻⁶ g cm⁻³ using a vibrating tube densimeter (DMA-5000, Anton Paar, Graz, Austria). The partial molar volume, \( V^\circ \), of the protein was calculated from density values using the relationship:

\[
V^\circ = \frac{M}{\rho_0} - \frac{\rho - \rho_0}{\rho_0 C}
\]

(2.1)

where \( \rho \) and \( \rho_0 \) are the densities of the protein solution and the solvent, respectively, \( C \) is the molar concentration of the protein, and \( M \) is the protein’s molecular weight.

Solution sound velocity measurements were carried out at ∼ 7.2 MHz by analyzing the amplitude-frequency characteristics of an ultrasonic resonator as described previously [175–177]. The analysis of the frequency characteristics of the resonator was performed by a Hewlett-Packard model E5100A network/spectrum analyzer (Mississauga, Ontario, Canada).

The key characteristic of a solute directly derived from ultrasonic velocimetric measurements is the relative molar sound velocity increment, \([U] = (U - U_0)/(U_0 C)\), where \( U \) and \( U_0 \) are the sound velocities in the protein solution and the neat solvent, respectively.

Values of the relative molar sound velocity increment, \([U]\), were used in conjunction with the measured partial molar volume data, \( V^\circ \), to calculate the partial molar adiabatic compressibility, \( K_S^\circ \), of the protein using the relationship [111, 178, 179]:
Characterization of \((\text{GlcNAc})_3\) Binding to Lysozyme

\[
K_S^0 = \beta_{S0} \left( 2V^o - 2[U] - \frac{M}{\rho_0} \right)
\]  

(2.2)

where \(\beta_{S0}\) is the coefficient of adiabatic compressibility of the solvent. The volumetric and ultrasonic velocimetric experiments have been performed at least three times with the average values of \([U]\) and \(V^o\) being used for \(K_S^0\) determination.

Densimetric and ultrasonic titrations were performed at 25°C by adding aliquots of the ligand \([(\text{GlcNAc})_3]\) to lysozyme solution following previously described experimental protocols [170, 180].

### 2.3.4 Determination of intrinsic volumes and solvent-accessible surface areas

The atomic coordinates of unligated lysozyme and the \((\text{GlcNAc})_3–\text{lysozyme}\) complex needed for calculating intrinsic volumes and solvent-accessible surface areas were obtained from the RSCB Protein Data Bank. For unligated lysozyme, we used the 4LYZ, 1AK1, 1VDQ, and 2YVB PDB entries as well as the protein domain extracted from the 1HEW entry. For the \((\text{GlcNAc})_3–\text{lysozyme}\) complex, we used the 1HEW entry. We calculated the solvent-accessible surface area, \(S_a\), for each structure as the sum of the accessible surface areas of all atoms in the structure. The intrinsic volumes, \(V_M\), of unligated lysozyme, free ligand, and the ligand–protein complex were calculated as molecular volumes as described by Richards [181,182].

Each PDB file was stripped of water molecules and cleaned using VMD (version 1.9.1) on a Linux platform. The program MSP (Molecular Surface Package) Version 3.9.3 was obtained from Dr. Michael Connolly at www.biohedron.com and used to calculate the solvent-accessible surface area and molecular volume for each structure, using a 1.4 Å probe radius on a Linux platform.

### 2.4 Results

Figure 1 presents the dependence of the relative fluorescence intensity of lysozyme on the ligand-to-protein binding ratio, \(r = [LT]/[PT]\), where \([LT]\) is the total concentration of \((\text{GlcNAc})_3\) and \([PT]\) is the total concentration of lysozyme. The experimental points presented in Figure 2.1 were approximated by an analytical function representing a one-to-one stoichiometric binding:

\[
X = X_0 + a\Delta X
\]  

(2.3)
where $X$ is a binding-dependent observable (in this case, $X$ is the relative fluorescence intensity), $X_0$ is the initial value of $X$ in the absence of the ligand, $\Delta X$ is the maximum change in $X$ when the protein is saturated with the ligand, and $\alpha$ is the fraction of protein molecules associated with the ligand. The fraction of ligated protein, $\alpha = [PL]/[PT]$, is given by the relationship:

$$\alpha = 0.5(r + 1) + Y^{-1} - [(0.25(r - 1)^2 + (r + 1)/Y + Y^{-2})^{1/2}$$  \hspace{1cm} (2.4)$$

where $Y = 2K_b[PT]$; $K_b = [PL]/([P][L])$ is the binding constant, $[PL]$ is the concentration of the $\text{(GlcNAc)}_3$-lysozyme complex, $[P]$ is the concentration of the unligated protein, and $[L]$ is the concentration of the free ligand.

Equation (2.4) can be derived for a one-to-one stoichiometric reaction by presenting the binding constant as $K_b = [PL]/([P][L]) = [PL]/((PT) - [PL])([LT] - [PL])$. This relationship can be easily transformed into the quadratic equation $Y\alpha^2 - (Y + Yr + 2)\alpha + Yr = 0$. Solving the latter with respect to the fraction of ligated protein, $\alpha$, yields Equation (2.3).

![Figure 2.1](image)

Figure 2.1: Relative fluorescence intensity of a solution containing lysozyme plotted against the $\text{(GlcNAc)}_3$-to-lysozyme molar ratio, $r$. The excitation and emission wavelengths are 296 and 345 nm, respectively. The initial concentration of lysozyme is 35 $\mu$M. The experimental points are fitted using Equation (2.3) (solid line).

By fitting the binding profile shown in Figure 2.1 by Equation (2.3), we calculated a binding constant, $K_b$, of $(1.2 \pm 0.2) \times 10^5 \text{ M}^{-1}$ for $\text{(GlcNAc)}_3$ association with lysozyme. Our measured binding constant is in good agreement with previous estimates of $1.3 \times 10^5 \text{ M}^{-1}$ [183] and $1.2 \times 10^5 \text{ M}^{-1}$ [146]. The observed agreement between our and reported binding constants lends credence to our experimental and data fitting protocols, although our fluorescence measurements were performed at somewhat high protein concentrations ($\sim 35 \mu$M). Using $\Delta G^\circ = -RT \ln K_b$, we calculated a binding free energy, $\Delta G^\circ$, of $6.9 \pm 0.1 \text{ kcal mol}^{-1}$. 


Figure 2.2: Change in the relative molar sound velocity increment of lysozyme plotted against the (GlcNAc)$_3$-to-lysozyme molar ratio, $r$. The excitation and emission wavelengths are 296 and 345 nm, respectively. The initial concentration of lysozyme is 90 µM. The experimental points are fitted using Equation (2.3) (solid line).

<table>
<thead>
<tr>
<th>protein</th>
<th>$S_A$, Å$^2$</th>
<th>$V_M$, cm$^3$ mol$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>lysozyme</td>
<td>6588 ± 113</td>
<td>9746 ± 58</td>
</tr>
<tr>
<td>complex</td>
<td>6777</td>
<td>10231</td>
</tr>
<tr>
<td>lysozyme in the complex</td>
<td>6773</td>
<td>9774</td>
</tr>
<tr>
<td>(GlcNAc)$_3$</td>
<td>875</td>
<td>341</td>
</tr>
</tbody>
</table>

Table 2.1: Molecular Volumes, $V_M$, and Solvent Accessible Surface Areas, $S_A$, of Unligated Lysozyme, Lysozyme–(GlcNAc)$_3$ Complex, Lysozyme in the Complex, and Free (GlcNAc)$_3$

Figures 2.2 and 2.3 show respectively the changes in relative molar sound velocity increment, $\Delta[U]$, and partial molar volume, $\Delta V$, of lysozyme in the absence and presence of the inhibitor at various (GlcNAc)$_3$-to-lysozyme binding ratios, $r$. We used Equation (2.3) to fit the binding profiles presented in Figures 2.2 and 2.3 and determine changes in relative molar sound velocity increment, $\Delta[U]_b$, and volume, $\Delta V_b$, accompanying saturation of lysozyme with (GlcNAc)$_3$. It should be noted that the binding constants, $K_b$, determined from fitting sound velocity (Figure 2.2) and volume (Figure 2.3) data are in good agreement with our fluorescence intensity-based evaluation (Figure 2.1). A change in adiabatic compressibility, $\Delta K_{Sb}$, accompanying ligand–protein association can be calculated from the values of $\Delta[U]_b$ and $\Delta V_b$ by modifying Equation (2.2) to the form $\Delta K_{Sb} = 2\beta S_0(\Delta V_b - \Delta [U]_b)$. Our determined values of $\Delta[U]_b$, $\Delta V_b$, and $\Delta K_{Sb}$ are $69 \pm 4$ cm$^3$ mol$^{-1}$, $44 \pm 2$ cm$^3$ mol$^{-1}$, and $(22 \pm 4) \times 10^{-4}$ cm$^3$ mol$^{-1}$ bar$^{-1}$, respectively. The reported uncertainties reflect the scattering of measurement data obtained from repetitive experiments and the standard deviation of data fitting by Equation (2.3).

Table 2.1 lists the intrinsic volumes, $V_M$, and solvent-accessible surface areas, $S_A$, of unligated lysozyme, (GlcNAc)$_3$–lysozyme complex, free (GlcNAc)$_3$, and lysozyme in the complex. The
2.5 Discussion

2.5.1 Evaluation of the binding-induced changes in hydration

The partial molar volume, $V^\circ$, of a solute is the sum of the following contributions [112,118,128]:

$$V^\circ = V_M + V_T + V_I + \beta T_0 RT$$  \hspace{1cm} (2.5)$$

where $V_M$ is the intrinsic volume of a solute, that is, the volume inaccessible to any part of a spherical probe with a radius of 1.4 Å that is rolled over the surface of a solute; $V_T$ is the thermal volume, that is, the volume of the effective void created around the solute due to thermally induced mutual vibrational motions of solute and solvent molecules as well as steric and structural effects; $V_I$ is the interaction volume, that is, the volume reduction due to solute-
solvent interactions; $\beta T_0$ is the coefficient of isothermal compressibility of the solvent; $R$ is the universal gas constant; and $T$ is the absolute temperature. The ideal term $\beta T_0 RT$ reflects the volume effect arising from the translational degrees of freedom available to the solute molecule.

Based on Equation (2.5), the change in volume accompanying a ligand–protein association event can be viewed as the sum of changes in the intrinsic, $\Delta V_M$, thermal, $\Delta V_T$, and interaction, $\Delta V_I$, contributions [128]:

$$
\Delta V = \Delta V_M + \Delta V_T + \Delta V_I
$$

To rationalize the observed volumetric changes in terms of hydration, we need to use the structural data on solvent accessible surface area, $S_A$, and intrinsic volume, $V_M$, presented in Table 2.1. A change in the intrinsic contribution, $\Delta V_M$, is $144 \pm 58 \text{ cm}^3 \text{ mol}^{-1} (10231 - 9746 - 341)$ as can be calculated from the data on the molecular volumes of the complex, the protein, and the ligand presented in Table 2.1. A change in thermal volume, $\Delta V_T$, correlates with a change in solvent accessible surface area, $\Delta S_A$. In general, the thermal contribution, $V_T$, to the partial molar volume of a solute is proportional to its solvent accessible surface area, $S_A$, with the proportionality coefficient equal to the effective thickness, $\delta$, of the thermal volume: $V_T = \delta S_A$ [117, 118, 128]. The value of $\delta$ depends on the size of the solute molecule; it increases sigmoidally from $\sim 0.6$ Å for small molecules to $\sim 1.0$ Å for large molecules of the size corresponding to proteins [117].

A change in thermal volume, $\Delta V_T$, is, thus, given by the relationship:

$$
\Delta V_T = \delta_L \Delta S_{AL} + \delta_P \Delta S_{AP}
$$

where $\delta_L$ and $\delta_P$ are the values of $\delta$ for the ligand and the protein, respectively, and $\Delta S_{AL}$ and $\Delta S_{AP}$ are the changes in the solvent accessible surface area of the ligand and the protein, respectively (note that $\Delta S_A = \Delta S_{AL} + \Delta S_{AP}$).

In our estimates, we take $\delta_L = 0.6$ Å (as for small molecules) and $\delta_P = 1.0$ Å (as for proteins). We further assume that $\Delta S_{AL} = \Delta S_{AP} = 0.5 \Delta S_A$. With these assumptions, a change in thermal volume, $\Delta V_T$, is given by $\Delta V_T = 0.5(\delta_L + \delta_P) \Delta S_A = 0.8 \Delta S_A$. A change in solvent-accessible surface area, $\Delta S_A$, equals $-686 \pm 113$ Å$^2 (6777 - 6588 - 875)$ as can be calculated from the data on the solvent-accessible surface areas of the complex, unligated lysozyme, and the trisaccharide given in Table 2.1. Thus, a change in thermal volume, $\Delta V_T$, accompanying the binding of (GlcNAc)$_3$ to lysozyme is equal to $-549 \pm 90$ Å$^3 (-0.8 \times 718)$ or $-330 \pm 54$ cm$^3$ mol$^{-1}$. With our estimates of $\Delta V_M$ and $\Delta V_T$, a change in interaction volume, $\Delta V_I$, can be determined from Equation (2.6) to be $143 \pm 79$ cm$^3$ mol$^{-1} (-44 - 144 + 331)$. 

The change in interaction volume, $\Delta V_I$, is the only component of $\Delta V$ that reflects redistribution of water molecules between the bulk and hydration phases. The value of $\Delta V_I$ is given by the sum:

$$\Delta V_I = \sum_i \Delta n_{hi}(V_{hi} - V_0) \quad (2.8)$$

where $\Delta n_{hi}$ is the number of water molecules taken up by the $i$th domain of the ligand or the protein upon their association and $V_{hi}$ and $V_0$ are the partial molar volumes of water of hydration of the $i$th solute domain and bulk water, respectively. Under the assumption of the uniformity of the hydration shells of the ligand and the protein, Equation (2.8) simplifies to the form $\Delta V_I = \Delta n_h(V_h - V_0)$, from which $\Delta n_h = \Delta V_I/(V_h - V_0)$. The partial molar volume of water solvating proteins and simple sugars, $V_h$, is roughly 10% smaller than that of bulk water [119, 120, 184]. Assuming $(V_h - V_0) = -1.8 \text{ cm}^3\text{ mol}^{-1}$, the value of $\Delta n_h$ is $-79 \pm 44 (-143/1.8)$. Thus, $79 \pm 44$ water molecules become released to the bulk from the hydration shells of (GlcNAc)$_3$ and lysozyme following their association.

Our determined number of water molecules released to the bulk, $79 \pm 44$, is similar to 76, the number of water molecules in direct contact with the associating surfaces. The latter can be evaluated by dividing the net change in solvent-accessible surface area, $\Delta S_A$, of 686 Å$^2$ by 9 Å$^2$, the effective cross section of a water molecule. Thus, we conclude that only waters contained within the first coordination spheres of the enzyme and the inhibitor predominantly contribute to the binding.

Dehydration of the interacting surfaces has a pronounced effect on the thermodynamic and kinetic properties of a binding reaction. The release of water molecules from the hydration shell of a protein may be fast or slow depending on the free energy barrier separating a specific water of hydration from the bulk water [185]. The slow exchange kinetics is predominantly displayed by water molecules which are doubly hydrogen bonded with protein groups [185]. Although the number of such water molecules is only 5–10% of the total number of waters of hydration, they play a prominent kinetic and thermodynamic role in controlling the conformational stability and recognition events of proteins [185].

It is difficult to estimate the thermodynamic impact of the collective release of 79 water molecules. Although the thermodynamic contribution of water molecules involved in protein binding and folding events is known to be on the order of or even exceed the net thermodynamics of the binding or folding reactions, the ligand-induced release of individual water molecules may contribute favorably or unfavorably to the net binding energetics depending on the location of the hydration site [17, 37, 169, 186–189]. Waters of hydration are highly heterogeneous with respect to their structural, dynamic, and thermodynamic properties [17, 169, 185, 190].
In particular, waters near a protein surface which form hydrogen bonds with 0, 1, 2, 3, or 4 functional groups are vastly different with respect to their enthalpic, entropic, and free energy contributions, while all of them are loosely defined as waters of protein hydration [17, 191]. It should be noted that, in globular proteins, the majority of waters of hydration form one or zero hydrogen bonds with protein groups [17]. In lysozyme, there are 195 water molecules which form a single hydrogen bond with protein groups and only 38 waters forming two or more hydrogen bonds [190]. Further studies preferably combining theoretical and experimental approaches are needed to evaluate the thermodynamic role of the 79 water molecules released to the bulk. Our current results represent the first quantitative step in that direction.

### 2.5.2 Evaluation of the binding-induced change in the intrinsic compressibility of lysozyme

With $\Delta n_h$ estimated, we now proceed to evaluate a change in the intrinsic compressibility of lysozyme accompanying its association with $(\text{GlcNAc})_3$. Note that the intrinsic compressibility, $K_M = \beta_M V_M$, of a protein is a linear function of its volume fluctuations [124,125]. The mean-square fluctuations of the intrinsic volume, $\langle \delta V_M^2 \rangle$, is related to the intrinsic coefficient of isothermal compressibility, $\beta_{TM}$, of the protein molecule [124,125]:

$$\langle \delta V^2 \rangle = k_B T V_M \beta_{TM}$$

(2.9)

where $k_B$ is Boltzmann’s constant and $T$ is the absolute temperature.

The mean-square fluctuations of the intrinsic volume, $\langle \delta V_M^2 \rangle$, represents an effective measure of protein dynamics. Thus, the value of $V_M$ can be used in conjunction with $\beta_M$ for quantitative characterization of the binding-induced change in the conformational dynamics of lysozyme as reflected in $\langle \delta V_M^2 \rangle$. A change in adiabatic compressibility, $\Delta K_S$, accompanying ligand–protein binding is given by the sum of the intrinsic, $\Delta K_M$, and hydration, $\Delta K_h$, contributions:

$$\Delta K_S = \Delta K_M + \Delta K_h$$

(2.10)

Analogous to the change in interaction volume, a change in the hydration component, $\Delta K_h$, can be presented as the sum

$$\Delta K_h = \sum_i \Delta n_{hi} (K_{hi} - K_0)$$

(2.11)
where $K_{hi}$ and $K_0$ are the partial molar adiabatic compressibilities of water of hydration of the $i$th solute domain and bulk water, respectively.

By assuming uniformity of the hydration shells of the ligand and the protein, Equation (2.11) simplifies to the form $\Delta K_h = \Delta n_h (K_h - K_0)$. The partial molar adiabatic compressibility of water solvating proteins and simple sugars, $K_h$, is roughly 20% smaller than that of bulk water [119, 192]. Assuming $(K_h - K_0) = -1.3 \times 10^{-4}$ cm$^3$ mol$^{-1}$ bar$^{-1}$ and given $\Delta n_h$ of $-79 \pm 44$, the value of $\Delta K_h$ is $(103 \pm 57) \times 10^{-4}$ cm$^3$ mol$^{-1}$ bar$^{-1}$ $(79 \times 1.3 \times 10^{-4})$. A change in the intrinsic compressibility of lysozyme, $\Delta K_M$, can be estimated from Equation (2.10) to be $-(81 \pm 57) \times 10^{-4}$ cm$^3$ mol$^{-1}$ bar$^{-1}$ $(22 \times 10^{-4} - 103 \times 10^{-4})$. Since $\Delta K_M = \beta_M \Delta V_M + V_M \Delta \beta_M$, a change in the intrinsic compressibility of lysozyme accompanying its association with the trisaccharide is given by $\Delta \beta_M = (\Delta K_M - \beta_M \Delta V_M)/V_M$. With the average intrinsic compressibility, $\beta_M$, of a globular protein of $25 \times 10^{-6}$ bar$^{-1}$ [112, 121, 193] and the values of $V_M = 9700$ cm$^3$ mol$^{-1}$ and $\Delta V_M = 74$ cm$^3$ mol$^{-1}$ $(9774 - 9700)$ from Table 2.1, we calculate $\Delta \beta_M$ of $-(1.0 \pm 0.6) \times 10^{-6}$ bar$^{-1}$. Thus, the binding of lysozyme to (GlcNAc)$_3$ renders the former $4 \pm 2$% $(1.0 \times 10^{-6}/25 \times 10^{-6})$ less compressible compared to the unligated enzyme. A similar decrease of $\sim 1$% has been observed for the intrinsic compressibility of hexokinase upon its association with glucose [127].

According to Equation (2.9), the observed 4% decrease in $\beta_M$ reflects a similar decrease in the mean-square fluctuations of the intrinsic volume, $\langle \delta V_M^2 \rangle$ (or 2% decrease in $\delta V_M$). The ligand-bound state of lysozyme is, thus, more rigid and less dynamic compared to its unbound state. The observed decrease in conformational dynamics as reflected in $\langle \delta V_M^2 \rangle$ is in agreement with the results of crystallographic and MD simulation studies [141, 194]. These studies have revealed a significant decrease in the temperature B factors of the residues directly interacting with the bound inhibitor, in particular, Trp62 [141, 194]. It should be noted that the observed decrease in $\langle \delta V_M^2 \rangle$ may reflect and correlate with a decrease in the configurational entropy of the protein accompanying its association with (GlcNAc)$_3$. In this respect, recall that the binding of (GlcNAc)$_3$ to lysozyme is an enthalpy-driven process with an unfavorable change in entropy [146]. Thus, the decrease in configurational entropy of lysozyme may be one of the contributors to the overall unfavorable change in entropy accompanying the process.

### 2.6 Conclusion

We applied high precision densimetric and ultrasonic velocimetric measurements to characterize the binding of (GlcNAc)$_3$ to lysozyme at $25^\circ$C in a pH 5.5 sodium acetate buffer. This enzyme–inhibitor association event causes changes in volume, $\Delta V$, and adiabatic compressibility, $\Delta K_S$, of $-44 \pm 2$ cm$^3$ mol$^{-1}$ and $(22 \pm 4) \times 10^{-4}$ cm$^3$ mol$^{-1}$ bar$^{-1}$, respectively. We interpreted these results in conjunction with X-ray crystallographic data in terms of changes in hydration of the
ligand and the protein and the dynamic properties of the latter. On the basis of our $\Delta V$ data, we estimate that $79 \pm 44$ water molecules are released to the bulk from the hydration shells of the protein and the ligand. Our $\Delta K_S$ data suggest a $4 \pm 2 \%$ decrease in the mean-square fluctuations of the intrinsic volume of the protein, $\langle \delta V_M^2 \rangle$ (or $2 \%$ decrease in $\delta V_M$). Thus, ligand binding stiffens the enzyme and renders it less dynamic compared to the unbound state. In general, we discuss the importance of volumetric insights into the molecular origins of protein recognition events.
Chapter 3

Ultrasonic and Densimetric Characterization of the Association of Cyclic AMP with the cAMP-Binding Domain of the Exchange Protein EPAC1

"Reprinted with permission from The Journal of Physical Chemistry B, 117, Ikbae Son, Rajeevan Selvaratnam, David N. Dubins, Giuseppe Melacini, and Tigran V. Chalikian. Ultrasonic and Densimetric Characterization of the Association of Cyclic AMP with the cAMP-Binding Domain of the Exchange Protein EPAC1, 10779-10784. Copyright (2013) American Chemical Society"

Author’s Contribution: Ikbae Son performed most experiments, analyzed the data, and participated in the preparation of the manuscript.

3.1 Abstract

We employed a combination of densimetric and ultrasonic velocimetric techniques to characterize the volumetric properties of the association of the cAMP-binding domain (CBD) of EPAC1 with cAMP at 25 °C in a pH 7.6 buffer. The binding of cAMP to the CBD of EPAC1 is accompanied by changes in volume, ΔV, and adiabatic compressibility, ΔKS, of 59±4 cm³mol⁻¹ and (34±9) × 10⁻⁴ cm³mol⁻¹bar⁻¹, respectively. We use these volumetric results in conjunction
with the structural data to estimate a change in hydration, $\Delta n_h$, accompanying the binding. We calculate that approximately 103 water molecules are released to the bulk from the associating surfaces of the protein and the ligand. This number is $\sim 30\%$ larger than the number of water molecules in direct contact with the associating surfaces while also being within the error of our $\Delta n_h$ determination. Therefore, we conclude that cAMP binding to EPAC1 may involve, in addition to the waters from within the first coordination sphere, also some waters from the second coordination sphere of the protein and cAMP. Our analysis of the compressibility data reveals that the protein becomes more rigid and less dynamic upon the cAMP binding as reflected in a $4 \pm 0.5\%$ decrease in its intrinsic coefficient of adiabatic compressibility. Finally, we estimate the hydration, $\Delta S_{\text{hyd}}$, and configurational, $\Delta S_{\text{conf}}$, contributions to the binding entropy, $\Delta S_b$. We find that the binding entropy is determined by the fine balance between the $\Delta S_{\text{hyd}}$ and $\Delta S_{\text{conf}}$ terms. In general, we discuss insights that are derived from a combination of volumetric and structural properties, in particular, emphasizing how measured changes in volume and compressibility can be interpreted in terms of hydration and dynamic properties of EPAC1 in its apo- and holo-forms.

### 3.2 Introduction

EPAC1 is a guanine nucleotide exchange factor for the small GTPases Rap1 and 2, which is directly activated by cAMP [147, 148]. Rap proteins are GTP-binding proteins that serve as molecular switches between a GDP-bound inactive and a GTP-bound active state. EPAC1 acts by accelerating the slow intrinsic GDP dissociation from Rap, which subsequently associates with GTP. Two highly homologous isoforms of EPAC, namely, EPAC1 and EPAC2, which are found in mammalian cells, consist of a single polypeptide chain with an N-terminal regulatory and a C-terminal catalytic region [149, 150]. The regulatory domain of EPAC1 contains an N-terminal dishevelled/Egl-10/pleckstrin (DEP) domain, followed by the cAMP-binding domain (CBD) [148]. CBD is responsible for the cAMP-mediated EPAC activation, whereas the DEP domain is responsible for its membrane binding and subsequent localization.

The binding of cAMP to the CBD of EPAC1 or EPAC2 causes a hinge-like rotation of a C-terminal helix ($\alpha 6$) toward the $\beta$-barrel core characteristic of CBDs [151, 152]. The binding and the associated structural reorganization of the protein are accompanied by a complex interplay of the enhanced and quenched dynamics of amino acid residues that are proximally as well as remotely located from the binding site in both the $\alpha$ and the $\beta$ subdomains of the protein [152]. Importantly, NMR studies have suggested that the cAMP-mediated alteration in the dynamic profile of the protein is involved in the entropically controlled allostery of the cAMP-dependent EPAC activation [152–158].

While enabling one to map out the cAMP-dependent enhancement/quenching of the protein
dynamics at residue resolution, NMR measurements alone do not readily provide the assessment of the global change in protein dynamics as reflected in the mean-square fluctuations of the protein volume, $\langle \delta V^2_M \rangle$. The latter assessment is important, since $\langle \delta V^2_M \rangle$ correlates with the configurational entropy of the protein in its apo- and holo-forms. The mean-square fluctuations of the protein volume, $\langle \delta V^2_M \rangle$, is directly proportional to the intrinsic coefficient of isothermal compressibility, $\beta_M$, of the protein molecule [124,125].

$$\langle \delta V^2_M \rangle = k_B T V_M \beta_M$$  (3.1)

where $k_B$ is Boltzmann’s constant and $T$ is the absolute temperature.

In this work, we study the volumetric properties of the association of the cAMP-binding domain (CBD) of EPAC1 (residues 149–318) with cAMP. To this end, we employ densimetric and ultrasonic velocimetric measurements to determine changes in volume, $\Delta V$, and adiabatic compressibility, $\Delta K_S$, accompanying the binding of cAMP to EPAC1. We use these results in conjunction with the available structural data on the highly homologous EPAC2–cAMP complex to estimate changes in protein dynamics and hydration linked to the binding event.

The dynamic and hydration changes mapped here for the CBD of EPAC1 will also serve as a foundation for understanding cAMP recognition in other cAMP-dependent systems controlled by structurally homologous CBDs, such as protein kinase A and the hyperpolarization activated and cyclic-nucleotide gated ion channels [195,196]. It should be noted in this regard that all macromolecular events taking place in living organisms are accompanied by alterations in solute–solvent interactions (hydration). The hydration-related energetics provides a major contribution to the driving force in macromolecular events, including protein and nucleic acid folding and binding reactions. Despite the significance of their contribution to the energetics of macromolecular folding and binding events, hydration changes are difficult to quantify, while evaluation of the hydration-related thermodynamic profile is even a more formidable task. This situation is unfortunate, since it lowers our ability to predict the conformational preferences of biopolymers as well as the affinity and specificity of their recognition reactions. One approach to tackling this problem is to establish hydration libraries on macromolecular folding and binding and to globally correlate these libraries with the energetics of individual events. Changes in hydration for specific macromolecular reactions can be quantified, for example, by the judicious use of volumetric measurements in conjunction with structural data on the system under question [15,112,115,126,127,130,197,198]. The work presented here on EPAC1 is a step in this direction and provides insights into the relative configurational and hydration determinants of ligand-binding entropies.
3.3 Materials and Methods

3.3.1 Materials

The cAMP-binding domain of human EPAC1 including residues 149-318 was expressed and purified according to protocols that have been published [199]. cAMP was purchased from Sigma Aldrich Canada (Oakville, Ontario, Canada). All measurements were performed in a pH 7.6 buffer consisting of 20 mM Tris and 50 mM NaCl.

The concentration of EPAC1 was determined from the absorbance measured at 25°C with a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, ON, Canada) using a molar extinction coefficient $\epsilon_{280} = 12490 \text{ M}^{-1} \text{ cm}^{-1}$. The latter has been calculated according to the additive scheme proposed by Pace et al. [200]. In densimetric and ultrasonic velocimetric experiments reported in this work, the protein concentration was on the order of $\sim 1 \text{ mg/mL}$ ($\sim 60 \mu\text{M}$).

3.3.2 Ultrasonic and densimetric measurements

All densimetric and ultrasonic velocimetric investigations reported here were conducted at 25°C. Densities were measured using a vibrating tube densimeter (DMA-5000, Anton Paar, Graz, Austria). The precision of density measurements was $\pm 1.5 \times 10^{-6} \text{ g cm}^{-3}$. The partial molar volume, $V^\circ$, of the protein was calculated from the relationship

$$V^\circ = \frac{M}{\rho_0} - \frac{\rho - \rho_0}{\rho_0 C} \tag{3.2}$$

where $\rho$ and $\rho_0$ are the densities of the protein solution and the solvent, respectively, and $M$ and $C$ are the protein’s molecular weight (19.2 kDa) and molar concentration, respectively.

The partial molar adiabatic compressibility, $K_S^\circ$, was calculated from the densimetric and ultrasonic data (see below) using the expression [111,178,179]

$$K_S^\circ = \beta_{S0} \left( 2V^\circ - 2[U] - \frac{M}{\rho_0} \right) \tag{3.3}$$

where $[U]$ is the relative molar sound velocity increment, $[U] = (U - U_0)/(U_0 C)$, $U$ and $U_0$ are the sound velocities in the protein solution and the neat solvent, respectively, and $\beta_{S0}$ is the coefficient of adiabatic compressibility of the solvent.
The sound velocities in protein solutions were determined at \( \sim 7.2 \) MHz by analyzing the amplitude-frequency characteristics of an ultrasonic resonator as described previously \([175–177]\). The analysis of the frequency characteristics of the resonator was performed by a Hewlett-Packard model E5100A network/spectrum analyzer (Mississauga, Ontario, Canada).

Densimetric and ultrasonic titrations were performed at 25 °C by adding aliquots of cAMP to EPAC1 solution following previously described experimental protocols \([201]\). All experiments have been performed at least three times with the average values of \([U]\) and \(V^o\) being used for \(K^o\) determination.

### 3.3.3 Determination of intrinsic volumes and solvent-accessible surface areas

The complex of the CBD of EPAC1 (residues 149–318) with cAMP has not been structurally characterized. However, the structure of the highly homologous isoform EPAC2 associated with cAMP has been solved by X-ray crystallography \([151]\). The atomic coordinates of the cAMP–EPAC2 complex are available from the RSCB Protein Data Bank (http://www.rcsb.org) (PDB entry 3CF6). We use this structure to calculate the solvent-accessible surface areas, \(S_A\), and intrinsic volumes, \(V_M\), for the complex, apoprotein (the complex minus cAMP), and free cAMP (the complex minus the protein). We calculated the solvent-accessible surface area, \(S_A\), for each structure as the sum of the accessible surface areas of all atoms in the structure. The intrinsic volumes, \(V_M\), of apoEPAC2, free ligand, and the ligand–protein complex were calculated as molecular volumes as described by Richards \([181,182]\).

The PDB file was stripped of water molecules and cleaned using VMD (version 1.9.1) on a Linux platform. The program MSP (Molecular Surface Package) Version 3.9.3 was obtained from Dr. Michael Connolly at www.biohedron.com and used to calculate the solvent-accessible surface area and molecular volume for each structure, using a 1.4 Å probe radius on a Linux platform.

### 3.4 Results and Discussion

#### 3.4.1 Binding affinity

Figures 1 and 2 present, respectively, the changes in volume, \(\Delta V\), and relative molar sound velocity increment, \(\Delta[U]\), of EPAC1 as a function of the ligand-to-protein binding ratio, \(r\). The binding profiles in Figures 3.1 and 3.2 have been approximated by an analytical function representing a one-to-one stoichiometric binding reaction \([198]\):
Chapter 3. Characterization of cAMP Binding to EPAC1

Figure 3.1: Change in the partial molar volume of the CBD of EPAC1 plotted against the cAMP-to-protein molar ratio, $r$. The initial concentration of CBD is 60 µM. The experimental points are fitted using Equation (3.4) (solid line).

Figure 3.2: Change in the relative molar sound velocity increment of the CBD of EPAC1 plotted against the cAMP-to-protein molar ratio, $r$. The initial concentration of CBD is 60 µM. The experimental points are fitted using Equation (3.4) (solid line).

\[ X = X_0 + a \Delta X \]  \hspace{1cm} (3.4)

where $X$ is a binding-dependent observable (in this case, volume or relative molar sound velocity increment), $X_0$ is the initial value of $X$ in the absence of the ligand, $\Delta X$ is the maximum change in $X$ upon protein saturation with the ligand, and $\alpha = [\text{PL}]/[\text{PT}]$ is the fraction of the ligated protein. The latter is computed as:

\[ \alpha = 0.5(r + 1) + Y^{-1} - [(0.25(r - 1)^2 + (r + 1)/Y + Y^{-2}]^{1/2} \]  \hspace{1cm} (3.5)
Table 3.1: Thermodynamic Characteristics of the Association of cAMP with the CBD of EPAC1

<table>
<thead>
<tr>
<th></th>
<th>( \Delta V ), cm(^3)mol(^{-1})</th>
<th>( \Delta [U] ), cm(^3)mol(^{-1})</th>
<th>( \Delta K_S ), cm(^3)mol(^{-1})bar(^{-1})</th>
<th>( K_b ) (volume), 10(^5)M(^{-1})</th>
<th>( K_b ) (compressibility), 10(^5)M(^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>complex</td>
<td>-59 ± 4</td>
<td>-97 ± 5</td>
<td>34 ± 9</td>
<td>3.9 ± 3.0</td>
<td>3.7 ± 2.1</td>
</tr>
<tr>
<td>apoEPAC1</td>
<td>8155</td>
<td>11350</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cAMP</td>
<td>480</td>
<td>168</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Molecular Volumes, \( V_M \), and Solvent Accessible Surface Areas, \( S_A \), of the cAMP–EPAC1 Complex, apoEPAC1, and Free cAMP

<table>
<thead>
<tr>
<th></th>
<th>( S_A ), Å(^2)</th>
<th>( V_M ), cm(^3)mol(^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>complex</td>
<td>7957</td>
<td>11600</td>
</tr>
<tr>
<td>apoEPAC1</td>
<td>8155</td>
<td>11350</td>
</tr>
<tr>
<td>cAMP</td>
<td>480</td>
<td>168</td>
</tr>
</tbody>
</table>

where \( Y = 2K_b[PT] \); \( K_b = [PL]/([P][L]) \) is the binding constant, \([PL]\) is the concentration of the cAMP–EPAC1 complex, \([P]\) is the concentration of the apoprotein, \([PT] = [PL] + [P] \) is the total concentration of the protein, and \([L]\) is the concentration of the free ligand.

Table 3.1 lists our evaluated changes in volume, \( \Delta V \), relative molar sound velocity increment, \( \Delta [U] \), and adiabatic compressibility, \( \Delta K_S = 2\beta S_0(\Delta V - \Delta [U]) \) (see Equation (3.3)), accompanying the complexation of cAMP with EPAC1. Table 3.1 also presents the association constants, \( K_b \), evaluated from fitting the volume and sound velocity binding profiles. The two binding constants, \( K_b \), are in excellent agreement with each other, while also being in close agreement with the published binding constants determined using isothermal titration calorimetry and a competitive \([^3H]cAMP \) radiolabel assay [159, 202]. The agreement lends credence to our experimental protocols and obtained results.

### 3.4.2 A change in hydration

In this section, we use the change in volume, \( \Delta V \), in conjunction with structural information to quantify a change in hydration accompanying the cAMP–EPAC1 binding. To this end, we use the empirical protocols reported previously [198]. The structural data on the cAMP–EPAC2 binding are presented in Table 3.2. A change in volume associated with a ligand–protein binding event can be parsed into the intrinsic, \( \Delta V_M \), thermal, \( \Delta V_T \), and interaction, \( \Delta V_I \), contributions [112, 198]:

\[
\Delta V = \Delta V_M + \Delta V_T + \Delta V_I
\] (3.6)
The intrinsic volume, $V_M$, of a solute is the volume inaccessible to any part of a spherical probe with a radius of 1.4 Å that is rolled over the surface of a solute [181, 182]. The thermal volume, $V_T$, is defined as the volume of the effective void created around the solute due to thermally induced mutual vibrational motions of solute and solvent molecules as well as steric and structural effects [117, 118, 203]. The interaction volume, $V_I$, is the change in volume due to solute–solvent interactions [118]. A change in the intrinsic contribution, $\Delta V_M$, of 82 cm$^3$ mol$^{-1}$ (11600 – 11350 – 168) is the difference between the molecular volumes of the holoprotein (11 600 cm$^3$ mol$^{-1}$) and the sum of the molecular volumes of the apoprotein (11 350 cm$^3$ mol$^{-1}$) and the free ligand (168 cm$^3$ mol$^{-1}$) (presented in Table 3.2). As a first approximation, a change in thermal volume, $\Delta V_T$, is proportional to the change in solvent accessible surface area, $\Delta S_A$ [198]

$$\Delta V_T = \delta_L \Delta S_{AL} + \delta_P \Delta S_{AP}$$  \hspace{1cm} (3.7)$$

where $\delta_L$ and $\delta_P$ are the thicknesses of thermal volume of the ligand and the protein, respectively, and $\Delta S_{AL}$ and $\Delta S_{AP}$ the changes in the solvent accessible surface area of the ligand and the protein, respectively (note that $\Delta S_A = \Delta S_{AL} + \Delta S_{AP}$).

Using a $\delta_L$ of 0.6 Å (the estimate for small molecules [117, 118, 204]) and a $\delta_P$ of 1.0 Å (the estimate for proteins [112, 117, 205]) and assuming $\Delta S_{AL} = \Delta S_{AP} = 0.5 \Delta S_A$, a change in thermal volume, $\Delta V_T$, is given by $\Delta V_T = 0.5(\delta_L + \delta_P)\Delta S_A = 0.8\Delta S_A$ [198]. The change in solvent-accessible surface area, $\Delta S_A$, equals $-678$ Å$^2$ ($7957 – 8155 – 480$) as can be calculated from the solvent-accessible surface areas of the holoprotein, apoprotein, and cAMP presented in Table 3.2. The change in thermal volume, $\Delta V_T$, accompanying the binding of cAMP to EPAC1 is, thus, equal to $-542$ Å ($-0.8 \times 678$) or $-326$ cm$^3$ mol$^{-1}$. With the estimates of $\Delta V_M$ and $\Delta V_T$, a change in interaction volume, $\Delta V_I$, can be determined from Equation (3.7) to be $185 \pm 4$ cm$^3$ mol$^{-1}$ ($-59 – 82 + 326$).

A change in interaction volume, $\Delta V_I$, reflects redistribution of water molecules between the bulk and hydration phases. The value of $\Delta V_I$ is given by the sum

$$\Delta V_I = \sum_i \Delta n_{hi}(V_{hi} - V_0)$$  \hspace{1cm} (3.8)$$

where $\Delta n_{hi}$ is the number of water molecules taken up by the $i$th solvent-exposed domain of the ligand or the protein upon their association, and $V_{hi}$ and $V_0$ are the partial molar volumes of water of hydration of the $i$th solute domain and bulk water, respectively. Under the assumption of the uniformity of the hydration shells of the ligand and the protein, Equation (3.8) simplifies to the form $\Delta V_I = \Delta n_h(V_h - V_0)$, from which $\Delta n_h = \Delta V_I/(V_h – V_0)$. The partial molar
volume of water solvating proteins and nucleic acids, $V_h$, is roughly 10% smaller than that of bulk water [119, 120, 184]. Assuming $(V_h - V_0) = -1.8 \text{ cm}^3\text{mol}^{-1}$, the value of $\Delta n_h$ is $-103 \pm 2 (-185/1.8)$. Thus, we conclude that, although the crystal structure of activated EPAC suggests that at least three water molecules remain trapped at the cAMP/EPAC interface [151], approximately 103 water molecules become released to the bulk from the hydration shells of cAMP and EPAC1 following their association.

A note of caution is in order regarding the uncertainty of the determined number of released water molecules $(\pm 2)$. It solely reflects the experimental error of our measured changes in volume, $\Delta V$, and does not include any error related to the structurally derived changes in intrinsic volume, $V_M$, and solvent accessible area, $S_A$. Presently, it is difficult to estimate the error of $V_M$ and $S_A$ determination as only a single structure of the cAMP–EPAC complex is available in the RSCB Protein Data Bank. On the basis of our previous works on protein–ligand association [115,130,198], the real error of $\Delta n_h$ determination is within the 25–50 % range (i.e., $\pm 25–50$).

Our determined number of water molecules released to the bulk, 103, is $\sim 30 %$ larger than 75, the number of water molecules in direct contact with the associating surfaces. The latter can be evaluated by dividing the net change in solvent-accessible surface area, $\Delta S_A$, of 678 Å by 9 Å, the effective cross-section of a water molecule. The difference of 30 % is within the error of our $\Delta n_h$ determination of 25–50 %. Thus, we conclude that cAMP binding may involve, in addition to the waters from within the first coordination sphere, also some waters from the second coordination sphere of the protein and cAMP.

Water molecules solvating proteins are highly heterogeneous with respect to their structural, dynamic, and thermodynamic properties [17, 169, 185, 190]. Consequently, the thermodynamic impact of the 103 released water molecules is difficult to quantify in terms of the accompanying changes in free energy, enthalpy, and entropy. However, as we have previously estimated based on Makhatadze’s and Privalov’s results [38], water molecules solvating all functional protein groups exhibit, at room temperature, a similar partial molar entropy that is $1.3 \pm 0.4 \text{ cal mol}^{-1}\text{K}^{-1}$ smaller than that of bulk water [115]. By multiplying this number by $\Delta n_h$, of 103 we calculate a highly favorable hydration contribution to the binding entropy, $\Delta S_{hyd}$, of $134 \pm 41 \text{ cal mol}^{-1}\text{K}^{-1}$ ($1.3 \times 103$). As shown below, the favorable entropic contribution of hydration is offset by an unfavorable change in configurational entropy of the protein and the ligand.

### 3.4.3 A change in protein dynamics

Protein dynamics is reflected in the mean-square fluctuations of the intrinsic volume, $\langle \delta V_M^2 \rangle$, which is proportional to the intrinsic coefficient of isothermal compressibility, $\beta_M$, of the protein.
(see Equation (3.1)). To estimate a change in $\beta_M$ brought about by EPAC1 association with cAMP, we use our measured change in compressibility, $\Delta K_S$, in conjunction with the estimated number of water molecules released to the bulk, $\Delta n_h$. Note that $\Delta K_S$ can be presented as the sum of the intrinsic, $\Delta K_M$, and hydration, $\Delta K_h$, contributions:

$$\Delta K_S = \Delta K_M + \Delta K_h$$

Under the assumption of the uniformity of the hydration shells of the ligand and the protein, we obtain $\Delta K_h = \Delta n_h(K_h - K_0)$. The partial molar adiabatic compressibility of water solvating proteins and nucleic acids, $K_h$, is roughly 15–20 % smaller than that of bulk water $(8.1 \times 10^{-4} \text{ cm}^3\text{ mol}^{-1}\text{ bar}^{-1})$ [119, 192]. With $\Delta n_h$ of $-103 \pm 2$ and $(K_h - K_0) = -1.3 \times 10^{-4} \text{ cm}^3\text{ mol}^{-1}\text{ bar}^{-1}$, the value of $\Delta K_h$ is $(134 \pm 3) \times 10^{-4} \text{ cm}^3\text{ mol}^{-1}\text{ bar}^{-1} (103 \times 1.3 \times 10^{-4})$.

From Equation (3.9), we estimate a change in the intrinsic compressibility of EPAC1, $\Delta K_M$, of $-(100 \pm 10) \times 10^{-4} \text{ cm}^3\text{ mol}^{-1}\text{ bar}^{-1} (34 \times 10^{-4} - 134 \times 10^{-4})$. Given $\Delta K_M = \beta_M \Delta V_M + V_M \Delta \beta_M$ and assuming $\Delta V_M \approx 0$, we obtain $\Delta \beta_M = \Delta K_M / V_M$. With $V_M$ of 11350 cm$^3$ mol$^{-1}$, we calculate $\Delta \beta_M$ of $-(0.9 \pm 0.1) \times 10^{-6} \text{ bar}^{-1} (100 \times 10^{-4} / 11350)$. Thus, the binding of EPAC1 to cAMP renders the holoprotein 4 ± 0.5 % $(0.9 \times 10^{-6} / 25 \times 10^{-6})$ less compressible compared to the apoprotein. Similar decreases in the intrinsic compressibilities of 1, 4, and 5 % have been observed for hexokinase upon its association with glucose [127], lysozyme upon its binding to $(\text{GlcNAc})_3$ [198], and ribonuclease A caused by its complexation with 2- or 3-CMP [115], respectively.

According to Equation (3.1), the observed 4 % decrease in $\beta_M$ reflects a similar decrease in the mean-square fluctuations of the intrinsic volume, $\langle \delta V_M^2 \rangle$ (or 2 % decrease in $\delta V_M$). The ligand-bound state of EPAC1 is, thus, more rigid and less dynamic compared to its unbound state. As shown below, the change in the protein dynamics correlates with a decrease in configurational entropy.

### 3.4.4 A change in configurational entropy

Thermodynamic studies have characterized the association of EPAC1 with cAMP as an exothermic process with a binding affinity, $K_b$, of $\sim 3.3 \times 10^5 \text{ M}^{-1}$ $(\Delta G_b = -7.5 \text{ kcal mol}^{-1})$ [159, 160]. Our estimate of the binding enthalpy, $\Delta H_b$, based on the isothermal titration calorimetric binding profiles presented by Rehmann et al. [159] is between $-8$ and $-10 \text{ kcal mol}^{-1}$. Thus, the cAMP–EPAC1 binding is accompanied by an unfavorable change in entropy, $\Delta S_b = (\Delta H_b - \Delta G_b)/T$, of $-2$ to $-8 \text{ cal mol}^{-1} \text{ K}^{-1}$ $(-5 \pm 3 \text{ cal mol}^{-1} \text{ K}^{-1})$. A change in entropy, $\Delta S_b$, for a ligand–protein association event can be presented as a sum of intrinsic (configurational), hydration, and translational terms [206]
\[ \Delta S_b = \Delta S_{\text{conf}} + \Delta S_{\text{hyd}} + \Delta S_{\text{rt}} \]  

(3.10)

where \( \Delta S_{\text{conf}} \) is the change in the configurational entropy of the ligand and the protein, \( \Delta S_{\text{hyd}} \) is the contribution due to a change in the hydration of the ligand and the protein, and \( \Delta S_{\text{rt}} \) is the change in entropy due to the loss of rotational and translational degrees of freedom upon the binding. For a 1:1 stoichiometric binding, \( \Delta S_{\text{rt}} \) equals \(-8 \text{ cal K}^{-1} \text{ mol}^{-1}\) [206].

Given our estimated hydration contribution, \( \Delta S_{\text{hyd}} \), of \(134 \pm 41 \text{ cal mol}^{-1} \text{ K}^{-1}\), the estimated change in configurational entropy, \( \Delta S_{\text{conf}} \), of the ligand and the protein is \(-131\pm41 \text{ cal K}^{-1} \text{ mol}^{-1} (-5 - 134 + 8)\). Note that \( \Delta S_{\text{conf}} \) is the sum of the changes in the configurational entropies of the ligands, \( \Delta S_{\text{conf}}(\text{L}) \), and the protein, \( \Delta S_{\text{conf}}(\text{P}) \). For a small nonpeptide ligand, \( \Delta S_{\text{conf}}(\text{L}) \) is a linear function of the number of rotatable bonds \( (N_{\text{rb}}) \) and the total number of atoms \( (N_{\text{atoms}}) \) [206]:

\[ \Delta S_{\text{conf}}(\text{L}) = -1.76N_{\text{rb}} + 0.414N_{\text{atoms}} \]  

(3.11)

With this relationship, we calculate the value of \( \Delta S_{\text{conf}}(\text{L}) \) for cAMP to be \(-44 \text{ cal K}^{-1} \text{ mol}^{-1}\). Thus, we estimate the binding-induced change in the configurational entropy of EPAC1, \( \Delta S_{\text{conf}}(\text{P}) \), to be unfavorable and equal to \(-87 \pm 41 \text{ cal K}^{-1} \text{ mol}^{-1} (-131 \pm 44)\). The configurational contribution of the binding free energy, \(-T\Delta S_{\text{conf}}(\text{P})\), is \(26 \pm 12 \text{ kcal mol}^{-1}\), which is 3.5 times as high in magnitude as the binding free energy, \( \Delta G_b \), of 7.5 kcal mol\(^{-1}\).

Configurational entropy is a major determinant of protein association events [165,166,206–211]. In absolute value, \( \Delta S_{\text{conf}} \) is, generally, larger than the entropy of a protein-binding reaction, \( \Delta S_b \), with the latter being governed by a fine balance between the large and opposing hydration and configurational contributions [207]. Consequently, determination of the \( \Delta S_{\text{conf}} \) and \( \Delta S_{\text{hyd}} \) contributions is required for developing empirical and/or theoretical predictive algorithms regarding the affinity and specificity of protein recognition events. Nevertheless, experimental determination of \( \Delta S_{\text{conf}} \) is difficult. The empirical approach described in this study is a step in that direction that enables one to estimate the order of magnitudes of \( \Delta S_{\text{conf}} \) and \( \Delta S_{\text{hyd}} \) and the degree of their compensation.

### 3.5 Conclusion

EPAC1 is a guanine nucleotide exchange factor for the small GTPases Rap1 and 2, which is directly activated by cAMP. In this work, we present a volumetric characterization of the association of the cAMP-binding domain of EPAC1 with cAMP. We discuss the conceptual
basis for resolving macroscopic properties (volume and compressibility) into microscopic events (protein hydration and dynamics). Our volumetric analysis performed in conjunction with the structural data on the cAMP/EPAC complex reveals that around 103 water molecules become released to the bulk from the interacting surfaces of the protein and the ligand. We find that the holoprotein is more rigid and less dynamic compared to its apo-form as reflected in a 4% decrease in its intrinsic coefficient of adiabatic compressibility. Our results enable us to estimate the favorable hydration contribution, $\Delta S_{hyd}$, and the unfavorable configurational contribution, $\Delta S_{conf}$, to the binding entropy, $\Delta S_b$. The sign and magnitude of $\Delta S_b$ are determined by the fine balance between the $\Delta S_{hyd}$ and $\Delta S_{conf}$ terms.
Chapter 4

Interactions of Urea with Native and Unfolded Proteins: A Volumetric Study


Author’s Contribution: Ikbae Son performed most experiments, analyzed the data, and participated in the preparation of the manuscript.

4.1 Abstract

We describe a statistical thermodynamic approach to analyzing urea-dependent volumetric properties of proteins. We use this approach to analyze our urea-dependent data on the partial molar volume and adiabatic compressibility of lysozyme, apocytochrome c, ribonuclease A, and α-chymotrypsinogen A. The analysis produces the thermodynamic properties of elementary urea–protein association reactions while also yielding estimates of the effective solvent-accessible surface areas of the native and unfolded protein states. Lysozyme and apocytochrome c do not undergo urea-induced transitions. The former remains folded, while the latter is unfolded between 0 and 8 M urea. In contrast, ribonuclease A and α-chymotrypsinogen A exhibit urea-induced unfolding transitions. Thus, our data permit us to characterize urea–protein interactions in both the native and unfolded states. We interpreted the urea-dependent volumetric
properties of the proteins in terms of the equilibrium constant, \( k \), and changes in volume, \( \Delta V_0 \), and compressibility, \( \Delta K_{T_0} \), for a reaction in which urea binds to a protein with a concomitant release of two waters of hydration to the bulk. Comparison of the values of \( k \), \( \Delta V_0 \), and \( \Delta K_{T_0} \) with the similar data obtained on small molecules mimicking protein groups reveals lack of cooperative effects involved in urea–protein interactions. In general, the volumetric approach, while providing a unique characterization of cosolvent–protein interactions, offers a practical way for evaluating the effective solvent accessible surface area of biologically significant fully or partially unfolded polypeptides.

4.2 Introduction

Chemical denaturation has been a major component of biophysical research for over a century. Consequently, a great deal of effort has gone into understanding the mechanisms of modulation of the equilibrium between the native and unfolded protein species by stabilizing and destabilizing cosolvents [22, 93, 104, 162, 212]. Although urea is the most common and frequently used cosolvent in protein studies, controversies exist about the molecular nature of its denaturing action. Currently, most researchers are leaning toward the direct as opposed to the indirect mechanism of urea-induced protein denaturation [22, 213–215]. The direct mechanism implies the existence of direct van der Waals or hydrogen bonding or other electrostatic interactions between urea and protein groups [216–218], while, in the indirect mechanism, urea exerts its influence via perturbation of the structure of water and the subsequent modification of protein–water interactions [74–76].

Volumetric properties of solutes provide a wealth of thermodynamic information describing the entire spectrum of solute–solvent interactions [109–112, 122, 132, 219–224]. In particular, in a binary mixture consisting of the principal solvent and cosolvent, the volumetric properties of solutes reflect the differential solute–principal solvent and solute–cosolvent interactions [83, 131, 132]. We have previously developed a statistical thermodynamics-based algorithm that can be used in conjunction with experimental volumetric data to extract the thermodynamic parameters of elementary solute–cosolvent interactions [83, 131, 132]. With this algorithm, we have conducted a systematic characterization of interactions of urea and glycine betaine with low-molecular-weight model compounds and interactions of glycine betaine with native globular proteins [83, 132, 133, 137].

Our results collectively make an argument in favor of extensive interactions of urea with all functional groups of proteins consistent with the direct mechanism of the effect of urea on protein stability [22, 216–218, 225–227]. More recently, our theoretical investigation based on the volumetrically determined parameters for urea– and glycine betaine–protein interactions has revealed that the mode of action of a specific cosolvent is governed by an extremely subtle bal-
ance between the thermodynamic contributions of cavity formation and direct solute–cosolvent interactions [67]. In this work, we expand this line of research to studying the solute–solvent interactions of four proteins, namely, apocytochrome c, hen egg white lysozyme, ribonuclease A, and α-chymotrypsinogen A, in binary water–urea mixtures. Our goal is to understand to what extent the thermodynamic insights we gained from studying low-molecular-weight model compounds are applicable to characterizing whole proteins. In particular, this study aims at elucidating if urea interactions with folded or unfolded protein states involves any cooperativity which is absent in low-molecular-weight compounds mimicking proteins. In this respect, we have previously found that cooperative effects are absent in the interactions of glycine betaine with folded proteins [137].

The specific set of proteins was chosen to cover the full range of conformational states accessible to proteins extending from folded to fully unfolded. Apocytochrome c is unfolded, while lysozyme remains fully folded at all concentrations of urea studied here (0–8 M). Hence, these proteins do not undergo any urea-induced conformational transitions thereby enabling one to separately study urea interactions with folded (lysozyme) and unfolded (apocytochrome c) protein states. In contrast, ribonuclease A and α-chymotrypsinogen A exhibit urea-induced denaturation transitions. When analyzing the volumetric responses of these proteins to an increase in urea concentration, we used the thermodynamic parameters determined for urea interactions with folded and unfolded conformational states derived from our studying lysozyme and apocytochrome c.

### 4.3 Materials and Methods

#### 4.3.1 Materials

The proteins hen egg white lysozyme, holocytochrome c from equine heart, and ribonuclease A and α-chymotrypsinogen A both from bovine pancreas as well as urea were purchased from Sigma-Aldrich Canada, Ltd. (Oakville, ON). All the reagents were of the highest purity commercially available. To get rid of salts and/or any other low-molecular-weight compounds that may be present in the samples, we exhaustively dialyzed the proteins against distilled water with subsequent lyophilization.

Apocytochrome c was obtained and purified from the holoprotein following the previously described experimental protocol [228]. The heme was removed by acid acetone extraction after the thioether bridges connecting the heme with the polypeptide chain of the protein had been cleaved by silver sulfate in the presence of acetic acid. Following the removal of the unreacted silver sulfate and the cleaved heme by centrifugation, the apoprotein was reacted with 2-mercaptoethanol to get rid of the bound silver. The resulting apoprotein was extensively
dialyzed against water at 4°C. The content of the holocytochrome c was assessed spectrophotometrically by measuring light absorption in the Soret region at 25°C and using an extinction coefficient, $\epsilon_{408}$, of 106 500 M$^{-1}$ cm$^{-1}$ [180]. Within the limit of the accuracy of our spectrophotometric measurements (< 0.5%), we could not detect any contamination of the apocytochrome c solution by the holoprotein. We used dry weight analysis to determine the extinction coefficient of our purified apocytochrome c. Our value of $\epsilon_{277} = 10 500 \pm 300$ M$^{-1}$ cm$^{-1}$ determined for 25°C is in good agreement with the reported $\epsilon_{277}$ of 10 800 M$^{-1}$ cm$^{-1}$ [228].

The volumetric and spectroscopic studies of lysozyme and apocytochrome c were performed in a 10 mM sodium phosphate buffer adjusted to pH 7.0, while ribonuclease A and α-chymotrypsinogen A were both investigated in a 10 mM glycine at pH 3.0. In all densimetric and ultrasonic velocimetric measurements and near-UV CD measurement, the protein concentrations were on the order of \( \sim 2 \text{ mg/mL} \). In the far-UV CD measurements, the protein concentrations were kept on the order of \( \sim 0.3 \text{ mg/mL} \).

### 4.3.2 Concentration determinations

Protein concentrations were measured spectrophotometrically using extinction coefficients that had been evaluated individually for each protein at all urea concentrations studied here. The urea-dependent extinction coefficients of proteins were determined according to the following protocol. After dialysis against water and lyophilization, the moisture content in each protein sample was evaluated by comparing the weight-determined concentration of a protein with the spectrophotometrically determined concentration using published extinction coefficients. The extinction coefficients for apocytochrome c, lysozyme, ribonuclease A, and α-chymotrypsinogen A in water are $\epsilon_{277} = 10 800$ [228], $\epsilon_{280} = 37 900$ [174], $\epsilon_{278} = 9 800$ [229], and $\epsilon_{280} = 50 630$ M$^{-1}$ cm$^{-1}$ [230], respectively. The extinction coefficient of each protein at each experimental urea concentration was measured as the ratio of light absorption at the characteristic wavelength to the weight-determined molar concentration of a protein. The latter was determined by weighing \( \sim 10 \text{ mg} \) of dry protein and dissolving it in a \( \sim 10 \text{ mL} \) of urea solution. The volume of the added urea solution was calculated as the ratio of its mass to density. The densities of urea solutions were measured with a precision of \( \pm 1.5 \times 10^{-4} \% \) using a vibrating tube densimeter (DMA-5000, Anton Paar, Gratz, Austria). All masses were weighed with a precision of \( \pm 0.02 \text{ mg} \) using a Mettler Toledo model AT201 analytical balance (Mettler Toledo Canada, Mississauga, Ontario, Canada). Table 4.1 presents the protein extinction coefficients at 0–8 M urea.
<table>
<thead>
<tr>
<th>urea (M)</th>
<th>lysozyme (280 nm)</th>
<th>apocytochrome c (277nm)</th>
<th>ribonuclease A (278nm)</th>
<th>α-chymotrypsinogen A (280 nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>37900</td>
<td>10800</td>
<td>9880</td>
<td>50630</td>
</tr>
<tr>
<td>1</td>
<td>38200</td>
<td>10950</td>
<td>9850</td>
<td>50450</td>
</tr>
<tr>
<td>2</td>
<td>38630</td>
<td>11040</td>
<td>9760</td>
<td>48270</td>
</tr>
<tr>
<td>3</td>
<td>39060</td>
<td>11150</td>
<td>9220</td>
<td>46690</td>
</tr>
<tr>
<td>4</td>
<td>39630</td>
<td>11270</td>
<td>8960</td>
<td>46860</td>
</tr>
<tr>
<td>5</td>
<td>39900</td>
<td>11320</td>
<td>8860</td>
<td>46970</td>
</tr>
<tr>
<td>6</td>
<td>40200</td>
<td>11420</td>
<td>8950</td>
<td>47290</td>
</tr>
<tr>
<td>7</td>
<td>40480</td>
<td>11530</td>
<td>9000</td>
<td>48490</td>
</tr>
<tr>
<td>8</td>
<td>40770</td>
<td>11650</td>
<td>9050</td>
<td>48220</td>
</tr>
</tbody>
</table>

Table 4.1: Extinction Coefficients (M$^{-1}$ cm$^{-1}$) of the Proteins as a Function of Urea

### 4.3.3 Solution preparation

Aqueous solutions of urea with concentrations of 1, 2, 3, 4, 5, 6, 7, and 8 M were prepared by weighing 10–50 g of urea and adding pre-estimated amounts of a respective buffer to achieve the desired molalities, $m$. The molar concentration, $C$, of urea was computed from its molal value, $m$, using $C = 1/(mρ_b) + φV/1000$, where $ρ_b$ is the density of pure buffer and $φV$ is the apparent molar volume of urea. Urea solutions were used as solvents for the proteins investigated in this study. The concentrations of the proteins were determined spectrophotometrically using the extinction coefficients listed in Table 4.1.

### 4.3.4 Optical spectroscopy

Light absorbance spectra of the proteins were measured at 25 °C using a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, ON, Canada). CD spectra were recorded at 25 °C using an AVIV model 62A DS spectropolarimeter (AVIV, Lakewood, NJ). All CD spectroscopic measurements were carried out in a 1 mm path-length cuvette.

### 4.3.5 Determination of the partial molar volumes and adiabatic compressibilities of solutes

All densities were measured at 25 °C with a precision of ±1.5 × 10$^{-4}$% using a vibrating tube densimeter (DMA-5000, Anton Paar, Gratz, Austria). The apparent molar volumes, $φV$, of the proteins were calculated from the relationship $φV = M/ρ_0 - (ρ - ρ_0)/(ρ_0C)$, where $M$ is the molecular weight of the protein, $C$ is the molar concentration of the protein, and $ρ$ and $ρ_0$ are
the densities of the solution and the solvent (water or urea solution), respectively.

Solution sound velocities, $U$, were measured at 25 °C at a frequency of 7.2 MHz using the resonator method and a previously described differential technique [175–177, 231]. The analysis of the frequency characteristics of the ultrasonic resonator cells required for sound velocity measurements was performed by a Hewlett-Packard model E5100A network/spectrum analyzer (Mississauga, ON, Canada). For the type of ultrasonic resonators used in this work, the accuracies of the sound velocity measurements are about $\pm 1 \times 10^{-4}\%$ [111, 177]. The acoustic characteristic of a solute, which can be derived directly from ultrasonic measurements, is the relative molar sound velocity increment, $[U]$. It is given by $[U] = (U - U_0)/(U_0C)$, where $U$ and $U_0$ are the sound velocities in the solution and the solvent, respectively.

The values of $[U]$ were used in conjunction with the $\phi V$ values derived from densimetric measurements to calculate the apparent molar adiabatic compressibilities, $\phi K_S$, using the relationship $\phi K_S = \beta S_0 (2\phi V - 2[U] - M/\rho_0)$, where $\beta S_0 = (\rho_0 U_0^2)^{-1}$ is the coefficient of adiabatic compressibility of the solvent [111, 178, 179]. The values of $\rho_0$, $U_0$, and $\beta S_0$ were directly determined for each urea solution from our densimetric and acoustic measurements. For each evaluation of $\phi V$ or $\phi K_S$, three to five independent measurements were carried out within a concentration range of 2–3 mg/mL. Our reported values of $\phi V$ or $\phi K_S$ represent the averages of these measurements with the errors calculated as standard deviations.

Given the small protein concentrations used in our study ($\sim 2$ mg/mL), we do not discriminate between the apparent and partial molar volumetric properties. Hence, we treat below the experimentally determined apparent molar volumes, $\phi V$, and compressibilities, $\phi K_S$, of the proteins as equivalent to their partial molar volumes, $V^\circ$, and compressibilities, $K_S^\circ$, respectively.

4.4 Results

The CD spectra of the proteins were recorded to identify and characterize their conformational states at various urea concentrations. No urea-dependent CD spectral changes were observed for lysozyme. Within the entire range of urea concentrations studied (0–8 M), lysozyme exhibits the native-like near- and far-UV CD spectra (Figure 4.1a,b). This observation suggests that lysozyme retains its native conformation even at the highest urea concentration. On the other hand, apocytochrome $c$ is unfolded even in the absence of urea as reflected in the far UV CD spectra (Figure 4.2). Figure 4.2 shows the far-UV CD spectra of apocytochrome $c$ at various concentrations of urea. The far-UV CD spectrum of apocytochrome $c$ in the absence of urea coincides with that of the acid-induced fully unfolded state of holocytochrome $c$ [180]. This observation is consistent with the picture in which apocytochrome $c$ is an essentially unfolded protein with little or no secondary structural elements remaining. Inspection of Figure 4.2
reveals that any secondary structural content (in particular, α-helices) that may be present in apocytochrome c further decreases with an increase in urea concentration as reflected in an increase in the ellipticity at $\sim 222$ nm.

Figure 4.1: Near- (panel a) and far-UV (panel b) CD spectra of lysozyme at different urea concentrations.

In contrast to lysozyme and apocytochrome c, ribonuclease A and α-chymotrypsinogen A display extensive urea-induced conformational changes as reflected in the CD-detected denat-
Figure 4.2: Far-UV CD spectra of apocytochrome c at different urea concentrations.

uration profiles illustrated in panels a and b of Figure 4.3, respectively. The denaturation profiles in panels a and b of Figure 4.3 were analyzed using the linear extrapolation model (LEM)(49) in which the fraction of protein denatured is given by \( \alpha = (1 + K^{-1})^{-1} \), where \( K = \exp[-(\Delta G_0 - m[\text{urea}]/RT)] \) is the equilibrium constant between the folded and unfolded protein states and \( \Delta G_0 \) is the thermodynamic stability of the protein in the absence of urea. The fitting of the denaturation profiles in panels a and b of Figure 4.3 produced the values of \( \Delta G_0 \) of 4.9 ± 0.3 and 8.9 ± 0.6 kcal mol\(^{-1}\) and the \( m \) values of 2.0 ± 0.1 and 2.7 ± 0.3 kcal mol\(^{-1}\) M\(^{-1}\) for ribonuclease A and \( \alpha \)-chymotrypsinogen A, respectively. We used these parameters when analyzing the urea dependences of the volumetric properties of the two proteins.

Panels a, b, c, and d of Figure 4.4 show the urea dependences of the partial molar volumes, \( V^\circ \), of lysozyme, apocytochrome c, ribonuclease A, and \( \alpha \)-chymotrypsinogen A, respectively. Panels a, b, c, and d of Figure 4.5 present the urea dependences of the partial molar adiabatic compressibilities, \( K^\circ S \), of lysozyme, apocytochrome c, ribonuclease A, and \( \alpha \)-chymotrypsinogen A, respectively. Inspection of Figure 4.4 and Figure 4.5 reveals that the partial molar volumes, \( V^\circ \), and adiabatic compressibilities, \( K^\circ S \), of lysozyme and apocytochrome c monotonically increase with an increase in urea concentration. On the other hand, the urea dependences of the partial molar volumes and compressibilities of ribonuclease A and \( \alpha \)-chymotrypsinogen A exhibit transition profiles consistent with the CD data presented in panels a and b of Figure 4.3.

4.4.1 Lysozyme

At neutral pH, lysozyme retains its native conformation within the entire range of urea concentrations studied here (0–8 M). Consequently, the urea dependences of its partial molar volume, \( V^\circ \) (Figure 4.4a), and compressibility, \( K^\circ S \) (Figure 4.5a), predominantly reflect changes
in protein solvation when solute–water interactions are increasingly replaced by solute–urea interactions. To fit the experimental volumetric data in Figure 4.4a, we used an analytical expression derived for a change in volume, $\Delta V^\circ$, accompanying transfer of a solute from water to a concentrated water–cosolvent mixture [83,131,132]:

$$\Delta V^\circ = \Delta V_C - \gamma_1 n \Delta V^\circ_1 + \frac{\Delta V \left( \frac{a_3}{a_1} \right) \left( \frac{a_3}{a_1} \right) k}{1 + \left( \frac{a_3}{a_1} \right) k}$$

(4.1)

where $a_1$ and $a_3$ are the activities of water and cosolvent (urea), respectively; $n$ is the number of water molecules solvating a solute in pure water (hydration number); $r$ is the number of water molecules released to the bulk from the hydration shell of a solute following its association.

Figure 4.3: Urea dependences of the molar ellipticities at 219 nm for ribonuclease A (a) and $\alpha$-chymotrypsinogen A (b). Experimental data were fitted by Equation (4.3) (solid lines).
Figure 4.4: Urea dependences of the partial molar volumes of lysozyme (a), apocytochrome c (b), ribonuclease A (c), and \( \alpha \)-chymotrypsinogen A (d). Experimental data in panels A and B were fitted by Equation (4.1), while those in panels C and D were fitted by Equation (4.3) (solid lines). The pre- and postdenaturational baselines of the protein denaturation profiles in panels C and D were fitted by Equation (4.1) (dashed red lines).

with the cosolvent; \( \Delta V_C \) is the differential cavity volume in a concentrated urea solution and water; \( k \) is the equilibrium constant for a reaction in which a cosolvent molecule binds the solute replacing \( r \) water molecules; \( \Delta V = \Delta V_0 + \gamma_1 r \Delta V_1^o - \gamma_3 \Delta V_3^o \) is the solvent exchange volume in a concentrated cosolvent solution; \( \Delta V_0 \) is the exchange volume in an ideal solution; \( \Delta V_1^o \) and \( \Delta V_3^o \) are the excess partial molar volumes of water and cosolvent in a concentrated solution; and \( \gamma_1 \) and \( \gamma_3 \) are the correction factors reflecting the influence of the bulk solvent on the properties of solvating water and cosolvent, respectively.

The number of water molecules, \( r \), substituted by urea is 2 [131]. The values of \( \gamma_1 \) and \( \gamma_3 \) may change from 0 (the properties of the solvation shell change in parallel with those of the bulk) to 1 (the properties of the solvation shell are independent of those of the bulk). For a protein, the correction factor \( \gamma_1 \) is roughly equal to the fraction of nonpolar and charged groups in the solvent accessible surface area of a solute, \( \gamma_1 = [(S_n + S_C)/S_A] \), while \( \gamma_3 \) is effectively 0 [83, 131, 137]. The solvent accessible surface areas of nonpolar, \( S_n \), and charged, \( S_C \), groups
and the total solvent accessible surface area, $S_A$, of lysozyme as well as the other proteins studied here were taken from our previous publication [203]. For lysozyme, $\gamma_1$ is 0.62.

The excess partial molar volumes of water, $\Delta V^0_C$, as a function of urea concentration have been reported [131]. The hydration number of a protein, $n$, can be estimated as the ratio of its solvent-accessible surface area, $S_A$, of 6685 Å$^2$ to 9 Å$^2$ ($S_W$), the effective cross-sectional area of a water molecule; $n = S_A/S_W$. For lysozyme, the value of $n$ is 743.

In principle, the values of $\Delta V_C$ in Equation (4.1) can be calculated as a function of urea concentration based on scaled particle theory [203, 232]. There are a number of complexities that render the SPT-based evaluations of the changes in cavity volume, $\Delta V_C$, for proteins approximate and rather unreliable [137]. Consequently, we do not attempt to evaluate the $\Delta V_C$ term in Equation (4.1). Instead, we fit the experimental data shown in Figure 4.4 by Equation (4.1) without explicitly taking into account the $\Delta V_C$ term. In such a treatment, the cosolvent-induced change in the cavity volume, $\Delta V_C$, will appear as an added contribution to
the values of $\Delta V_0$ determined from the fit.

For analyzing the compressibility data in Figure 4.5a, we used the expression derived for a change in isothermal compressibility, $\Delta K_T^o$, accompanying transfer of a solute from water to a water–cosolvent mixture [131–133]:

$$
\Delta K_T^o = \Delta K_{TC} - \gamma_1 n \Delta K_{T1}^o + \frac{\Delta K_T \left( \frac{n}{r} \right) \left( \frac{a_3}{a_1^3} \right) k}{1 + \left( \frac{a_3}{a_1^3} k \right)} + \frac{\Delta V^2 \left( \frac{n}{r} \right) \left( \frac{a_3}{a_1^3} \right) k}{RT \left[ 1 + \left( \frac{a_3}{a_1^3} k \right) \right]^2}
$$

(4.2)

where $\Delta K_{TC} = -(\partial \Delta V_C / \partial P)_T$ is the differential compressibility of the cavity enclosing a solute in water and cosolvent solution; $\Delta K_T = \Delta K_{T0} + \gamma_1 r \Delta K_{T1}^o - \gamma_3 \Delta K_{T3}^o$ is the change in compressibility associated with replacement of water by cosolvent at the binding site of a solute in a concentrated cosolvent (urea) solution; $\Delta K_{T1}^o$ and $\Delta K_{T3}^o$ are the excess partial molar isothermal compressibilities of water and cosolvent in a concentrated cosolvent solution; and $\Delta K_{T0} = -(\partial \Delta V_0 / \partial P)_T$ is the change in compressibility associated with replacement of water with cosolvent at the binding site in an ideal solution. The excess partial molar adiabatic compressibility of water, $\Delta K_{S1}^o$, has been measured as a function of urea concentration in a previous work [131].

Although Equation (4.2) has been derived for isothermal compressibility, we use it for treating our adiabatic compressibility data. In aqueous solutions, the difference between the partial molar adiabatic and isothermal compressibilities of solutes is small given the large heat capacity and small thermal expansibility of water [233]. Fitting the data presented in Figure 4.4a and Figure 4.5a produces an equilibrium constant, $k$, as well as changes in volume, $\Delta V_0$, and compressibility $\Delta V_{T0}$, accompanying the binding of urea to the protein with a concomitant release of two water molecules to the bulk. The equilibrium constants determined from fitting the volume (Figure 4.4a) and compressibility (Figure 4.5a) data are practically identical and equal to $0.12 \pm 0.01$ and $0.11 \pm 0.01$ M, respectively. The values of $\Delta V_0$ and $\Delta K_{T0}$ are $0.087 \pm 0.002$ cm$^3$ mol$^{-1}$ and $(1.65 \pm 0.04) \times 10^{-4}$ cm$^3$ mol$^{-1}$ bar$^{-1}$, respectively.

We compared the values of $k$, $\Delta V_0$, and $\Delta K_{T0}$ with similar data obtained on small molecules mimicking protein groups such as N-acetyl amino acid amides [83]. The comparison revealed that the values obtained for lysozyme are within the range observed for low molecular weight model compounds. For example, the values of $k$, $\Delta V_0$, and $\Delta K_{T0}$ evaluated for urea interactions with a glutamine side chain are $0.09 \pm 0.01$ M, $0.12 \pm 0.01$ cm$^3$ mol$^{-1}$, and $(2.12 \pm 0.05) \times 10^{-4}$ cm$^3$ mol$^{-1}$ bar$^{-1}$, respectively. This is an important observation that suggests lack of cooperative effects involved in urea–protein interactions. A similar inference has been drawn earlier for the interactions of native proteins with the stabilizing osmolyte glycine betaine [137]. By extension, one may hypothesize that protein interactions with cosolvents, in general, do not
involve cooperativity. Further volumetric studies involving a wider selection of cosolvents are required to prove or refute the generality of this hypothesis.

4.4.2 Apocytochrome $c$

Judging by an increase in the negative CD band in the vicinity of 222 nm with the rise in urea concentration (see Figure 4.2), apocytochrome $c$ becomes more unfolded as more urea is added to the solution. Consequently, when fitting the urea dependences of the partial molar volume (Figure 4.4b) and compressibility (Figure 4.5b) of apocytochrome $c$ with Equation (4.1) and Equation (4.2), respectively, one may plausibly assume that the hydration number, $n$, increases as the protein, becoming more unfolded, exposes additional atomic groups to the solvent. When fitting the data in Figure 4.4b and Figure 4.5b, we use the linear approximation $n = n_0 + B a_3$ (where $n_0$ is the hydration number in the absence of urea and $B$ is a constant) in Equation (4.1) and Equation (4.2). In the fit, we used the values of $k$, $\Delta V_0$, and $\Delta K_{T0}$ that had been determined for lysozyme. It should be noted, however, that such an approximation may be, potentially, fraught with error and may even skew the results of analysis, since lysozyme and apocytochrome $c$ differ in fractions of charged, polar, and nonpolar residues. The difference in the chemical nature of solvent-accessible atomic groups may, in turn, render the interactions of the two proteins with urea different. While being aware of such a possibility, we proceed below with the fitting of the volumetric data on apocytochrome $c$ with an understanding that the ensuing results may require a further validation.

The fitting of the urea dependences of the partial molar volume (Figure 4.4b) and compressibility (Figure 4.5b) of apocytochrome $c$ yields hydration numbers, $n_0$, of $660 \pm 40$ and $590 \pm 10$, respectively, and the coefficients, $B$, of $0 \pm 6$ and $15 \pm 2 \text{ M}^{-1}$, respectively. It is gratifying that the hydration numbers, $n_0$, determined from the volume and compressibility data are similar although not identical.

An intriguing observation is that the hydration number, $n$, for the unfolded apoprotein with little or no secondary structure ($\sim 630$) is small and comparable to 680, the hydration number estimated for fully folded holocytochrome $c$. The latter can be computed by dividing 6115 Å$^2$, the solvent accessible surface area of cytochrome $c$ [205], by 9 Å$^2$, the effective cross-sectional area of a water molecule. Our determined value of $n$ is consistent with the picture in which apocytochrome $c$, while being unfolded, retains a compact structure with a solvent-accessible surface area comparable to that of the folded holoprotein. This notion is in qualitative agreement with the results of our previous study in which the experimental partial molar adiabatic compressibility, $K_S^A$, of apocytochrome $c$ at 25 °C was found to be twice as high (less negative) as that additively calculated for the fully unfolded polypeptide chain [234]. At present, it is difficult to unequivocally rationalize the lower-than-expected solvent exposure of
the apocytochrome c polypeptide chain. It could be a unique property of apocytochrome c
due to its sequential position of polar and nonpolar residues or it could be a general feature for
a wider class of unfolded proteins. It may also reflect the approximation of the urea binding
properties of apocytochrome c by those of lysozyme which may not be warranted. As mentioned
above, additional studies may be required to verify the compactness of apocytochrome c and
understand the nature of the molecular forces driving the polypeptide contraction.

The values of B determined based on either volume or compressibility are relatively small
suggesting a modest change in hydration and, hence, in solvent-accessible surface area of apoc-
ytochrome c upon an increase in urea concentration. A change in n accompanying an increase in
urea concentration from 0 to 8 M is \(8 \times 15 = 120\), which is only \(\sim 20\%\) of the protein’s hydration
number of \(\sim 630\). Thus, the CD-detected decrease in the \(\alpha\)-helical content of apocytochrome c
is paralleled by an increase in its solvent-accessible surface area.

In general, measuring urea-dependent changes in the volumetric properties of an unfolded
polypeptide in conjunction with Equation (4.1) and Equation (4.2) offers a practical way of
evaluating its effective solvent accessible surface area. This notion may prove useful in estimat-
ing the degree of unfolding of intrinsically unfolded proteins or protein domains as well as other
biologically significant fully or partially unfolded poly- and oligopeptides.

### 4.4.3 Ribonuclease A

The volume- and compressibility-detected denaturation profiles of ribonuclease A presented in
Figure 4.4c and Figure 4.5c, respectively, were fitted using the two-state approximation

\[
X = X_N + \alpha (X_D - X_N)
\]  

(4.3)

where \(\alpha\) is the fraction of denatured protein determined from our CD measurements using LEM
(see Figure 4.3a) and \(X_N\) and \(X_D\) are the pre- and postdenaturational baselines, respectively.

We used Equation (4.1) and Equation (4.2) to fit the baselines in Figure 4.4c and Figure 4.5c,
respectively. The values of \(k\), \(\Delta V_0\), and \(\Delta K_{T0}\) used in the fit were those that had been de-
termined for lysozyme. From the fit, we determined the hydration numbers for the native,
\(n_N\), and denatured, \(n_D\), conformations. The native values, \(n_N\), determined from the volume
and compressibility data are \(740 \pm 80\) and \(670 \pm 25\), respectively. These numbers are close to
754, the hydration number of native ribonuclease A that can be computed as the ratio of its
solvent accessible surface area of \(6790 \text{ Å}^2\) [205] to \(9 \text{ Å}^2\), the cross-sectional surface area of a
water molecule. The observed similarity is gratifying and lends credence to the experimental
protocols and the conceptual foundation of data analysis employed in this work. The hydration
numbers, \( n_D \), for the denatured state determined from the volume and compressibility data are \( 1140 \pm 50 \) and \( 1460 \pm 50 \), respectively. Thus, ribonuclease A exhibits 1.5–2.2 times (on average, 1.9 \( \pm \) 0.4) larger solvent accessible surface area in its unfolded relative to the folded conformation.

It has been reported that the \( m \) values are proportional to changes in solvent accessible surface area, \( \Delta S_A \), accompanying urea-induced protein denaturation [235]. In these studies, \( \Delta S_A \) have been computed as the difference in solvent accessible surface area between the native and fully extended protein conformations [235]. However, proteins do not assume the fully extended conformation even at the harshest denaturing conditions. Thus, while, in principle, the \( m \) values can be used for evaluating \( \Delta S_A \), more realistic correlations between the \( m \) values and \( \Delta S_A \) need to be established for such evaluations to have any predictive power. The volumetric approach described in this work provides realistic \( \Delta S_A \) values for specific protein denaturation events that, in conjunction with their respective \( m \) values, may result in a better parametrization with subsequent applications to \( m \) value based \( \Delta S_A \) calculations.

When extrapolated to 0 M urea, changes in volume and compressibility accompanying urea-induced denaturation of ribonuclease A are \( -21 \pm 3 \) cm\(^3\) mol\(^{-1}\) \((-0.0015 \text{ cm}^3 \text{ g}^{-1}\)\) and \( -(330 \pm 30) \times 10^{-4} \) cm\(^3\) mol\(^{-1}\) bar\(^{-1}\) \((-2.4 \times 10^{-6} \text{ cm}^3 \text{ bar}^{-1} \text{ g}^{-1}\)\), respectively. Volumetric changes accompanying conformational transitions of ribonuclease A have been extensively studied [135, 236–243]. Review of the literature reveals that the reported changes in volume accompanying ribonuclease A denaturation are all negative and range from \(-10\) to \(-70 \) cm\(^3\) mol\(^{-1}\) depending on the research team, solution conditions, and method of measurement. Thus, our measured change in volume of \( -21 \pm 3 \) cm\(^3\) mol\(^{-1}\) is in good agreement with the reported values. On the other hand, our determined change in compressibility of \( -(330 \pm 30) \times 10^{-4} \) cm\(^3\) mol\(^{-1}\) bar\(^{-1}\) differs not only in magnitude but also in sign from the positive changes in compressibility of \( 120 \times 10^{-4} \) to \( 270 \times 10^{-4} \) cm\(^3\) mol\(^{-1}\) bar\(^{-1}\) reported for pressure-induced denaturation of ribonuclease A [236,240]. The disparity is not unexpected given the difference in the nature of the denatured state ensemble (pressure-induced versus urea-induced), the difference in pressure (elevated versus ambient), and a possible error produced by the two-state approximation used in the analysis of pressure-induced denaturation profiles.

If expressed per gram rather than per mole of protein, our measured change in compressibility of \( -2.4 \times 10^{-6} \) cm\(^3\) mol\(^{-1}\) g\(^{-1}\) corresponds to the range of compressibility changes typical for native to partially unfolded protein transitions [112,121,123]. This notion is in agreement with the roughly 2-fold increase in solvent accessible surface area of the protein accompanying its denaturation.
4.4.4 α-Chymotrypsinogen A

The volume- and compressibility-detected denaturation profiles of α-chymotrypsinogen A presented in Figure 4.4d and Figure 4.5d, respectively, were fitted with Equation (4.3). The hydration numbers of the native state, $n_N$, determined from the volume and compressibility data are $1290 \pm 90$ and $980 \pm 25$, respectively. These numbers are similar to 1202, the hydration number of native α-chymotrypsinogen A computed as the ratio of its solvent accessible surface area of $10\,815 \, \text{Å}^2$ [205] to $9 \, \text{Å}^2$, the cross-sectional surface area of a water molecule. The hydration numbers, $n_D$, for the denatured state determined from the volume and compressibility data are $1810 \pm 90$ and $2530 \pm 90$, respectively. Thus, α-chymotrypsinogen A exhibits 1.4–2.6 times (on average, $2.0 \pm 0.6$) larger solvent accessible surface area in its unfolded relative to the folded state.

When extrapolated to 0 M urea, changes in volume and compressibility accompanying urea-induced denaturation of α-chymotrypsinogen A are $180 \pm 5 \, \text{cm}^3 \, \text{mol}^{-1}$ $(0.007 \, \text{cm}^3 \, \text{g}^{-1})$ and $-(190 \pm 50) \times 10^{-4} \, \text{cm}^3 \, \text{mol}^{-1} \, \text{bar}^{-1}$ ($-0.7 \times 10^{-6} \, \text{cm}^3 \, \text{mol}^{-1} \, \text{g}^{-1}$), respectively. In general, changes in volume accompanying protein transitions may be positive or negative depending on the balance between various contributions which are large in magnitude and differ in sign [109, 112, 128, 193]. The positive sign in the change in volume, $\Delta V$, accompanying the urea-induced denaturation of α-chymotrypsinogen A agrees with the sign of $\Delta V$ measured for the pressure- and heat-induced denaturation of the protein [242, 244, 245]. On the other hand, a change in volume accompanying the urea-induced unfolding of α-chymotrypsinogen A that can be determined by extrapolating the pre- and postdenaturational baselines of the denaturation profile detected by a combination of densimetric and dilatometric techniques is extremely negative being on the order of $-350 \, \text{cm}^3 \, \text{mol}^{-1}$ [246]. The origin of such large and negative value is the unrealistically steep baselines of the denaturation profile which may reflect instrumental limitations of the volumetric equipment employed in the study.

A change in compressibility accompanying the pressure-induced denaturation of α-chymotrypsinogen A is positive and equal to $300 \times 10^{-4} \, \text{cm}^3 \, \text{mol}^{-1} \, \text{bar}^{-1}$ [244], while the heat-induced denaturation is associated with an even higher change in compressibility of $(1100 \pm 130) \times 10^{-4} \, \text{cm}^3 \, \text{mol}^{-1} \, \text{bar}^{-1}$ (when extrapolated to 25°C) [245]. These values contrast our negative value of $-(190 \pm 50) \times 10^{-4} \, \text{cm}^3 \, \text{mol}^{-1} \, \text{bar}^{-1}$. The disparity may reflect the differences in the nature of the denatured state ensemble and experimental pressure as well as the two-state approximation used in the analysis of pressure-dependent data. If expressed per gram rather than per mole of protein, our measured change in compressibility of $-0.7 \times 10^{-6} \, \text{cm}^3 \, \text{mol}^{-1} \, \text{g}^{-1}$ corresponds to the range of compressibility changes characteristic of native to partially unfolded protein transitions [112, 121, 193], consistent with a 2-fold increase in solvent accessible surface area of the protein.
4.4.5 The problem of baselines

Inspection of Figure 4.4c,d reveals that the pre- and postdenaturational baselines of the partial molar volume unfolding profiles of ribonuclease A and α-chymotrypsinogen A, although described by Equation (4.1), are, practically, linear. Therefore, their linear extrapolation to 0 M urea does not produce appreciable error in determining the volume change accompanying protein unfolding in pure water.

In contrast, the pre- and postdenaturational baselines in the compressibility profiles of ribonuclease A and α-chymotrypsinogen A, presented in panels c and d of Figure 4.5, respectively, are highly nonlinear. An attempt to linearly extrapolate the baselines to 0 M urea may result in significant error in magnitude and even the sign of a change in compressibility associated with protein unfolding in water (in the absence of urea). In this respect, note that, for both ribonuclease A (Figure 4.5c) and α-chymotrypsinogen A (Figure 4.5d), the compressibility difference between the unfolded and native states is negative in the absence of urea however decreases in magnitude, eventually, changing the sign as the concentration of urea increases. This observation is in line with our predictions based on additive calculations using urea-dependent volumetric data on low-molecular-weight compounds mimicking protein groups [132].

Thus, caution needs to be exercised when analyzing the volumetric profiles of cosolvent-induced protein folding/unfolding transitions. This notion may concern, in addition to volume- and compressibility-monitored protein unfolding profiles, also expansibility data.

4.5 Conclusion

We applied densimetric and ultrasonic velocimetric measurements to determining the partial molar volumes, $V^\circ$, and adiabatic compressibilities, $K_S^\circ$, of lysozyme, apocytochrome c, ribonuclease A, and α-chymotrypsinogen A at urea concentrations ranging from 0 to 8 M. The primary volumetric data were rationalized based on previously described analytical equations that had been derived based on the statistical thermodynamic formalism.

At 25°C, lysozyme does not denature under the influence of urea within the entire range of experimental urea concentrations. We interpreted the urea-dependent volumetric properties of lysozyme in terms of the equilibrium constant, $k$, and changes in volume, $\Delta V_0$, and compressibility, $\Delta K_{T0}$, for an elementary reaction in which urea binds to a protein with a concomitant release of two waters of hydration. Comparison of the resulting values of $k$, $\Delta V_0$, and $\Delta K_{T0}$ with the similar data obtained on small molecules mimicking proteins reveals a lack of cooperative effects involved in urea–protein interactions. Analysis of the volumetric properties of the unfolded protein apocytochrome c suggests that, in the absence of urea, the protein
exhibits a solvent-accessible surface area, $S_A$, comparable to that of folded holocytochrome c. An increase in urea concentration from 0 to 8M causes an additional $\sim 20\%$ increase in $S_A$. In contrast to lysozyme and apocytochrome c, ribonuclease A and $\alpha$-chymotrypsinogen A undergo urea-induced unfolding transitions. The denatured states of ribonuclease A and $\alpha$-chymotrypsinogen A exhibit solvent-accessible surface areas which are roughly twice as large as those of their respective native states.

The statistical thermodynamic formalism adequately describes the urea-dependent volumetric behavior of native and unfolded proteins. In general, urea-dependent changes in volumetric properties offer a practical way for evaluating the effective solvent accessible surface areas of biologically significant unfolded polypeptides.
Chapter 5

Volumetrically Derived Thermodynamic Profile of Interactions of Urea with a Native Protein

“Reprinted with permission from Biochemistry, 55, Ikbae Son, and Tigran V. Chalikian,. Volumetrically Derived Thermodynamic Profile of Interactions of Urea with a Native Protein, 6475–6483. Copyright (2016) American Chemical Society”.

Author’s Contribution: Ikbae Son performed all experiments, analyzed the data, and participated in the preparation of the manuscript.

5.1 Abstract

We report the first experimental characterization of the full thermodynamic profile for binding of urea to a native protein. We measured the volumetric parameters of lysozyme at pH 7.0 as a function of urea within a temperature range of 18–45 °C. At neutral pH, lysozyme retains its native conformation between 0 and 8 M urea over the entire range of temperatures studied. Consequently, our measured volumetric properties reflect solely the interactions of urea with the native protein and do not involve contributions from urea-induced conformational transitions. We analyzed our data within the framework of a statistical thermodynamic analytical model in which urea–protein interactions are viewed as solvent exchange in the vicinity of the protein. The analysis produced the equilibrium constant, $k$, for an elementary reaction of urea–protein
binding with a change in standard state free energy ($\Delta G^\circ = -RT \ln k$) at each experimental temperature. We used the van’t Hoff equation to compute from the temperature dependence of the equilibrium constant, $k$, changes in enthalpy, $\Delta H^\circ$, and entropy, $\Delta S^\circ$, accompanying binding. The thermodynamic profile of urea–protein interactions, in conjunction with published molecular dynamics simulation results, is consistent with the picture in which urea molecules, being underhydrated in the bulk, form strong, enthalpically favorable interactions with the surface protein groups while paying a high entropic price. We discuss ramifications of our results for providing insights into the combined effects of urea, temperature, and pressure on the conformational preferences of proteins.

5.2 Introduction

The marginal stability of globular proteins renders them susceptible to drastic shifts in folded-unfolded equilibria following changes in temperature, pressure, pH, and the chemical composition of solvent [22,87,247–251]. Although cosolvents that increase or decrease protein stability have been employed in protein studies for more than a century, the network of molecular interactions underlying the mode of cosolvent action is still being debated and not very well understood [22,23,60,93,104]. On the other hand, the experimental thermodynamic signatures of stabilizing, destabilizing, and neutral cosolvents, as expressed in preferential binding, $\Gamma_{23}$, and/or hydration, $\Gamma_{21}$, parameters, have been established and linked to the mode of action of specific cosolvents [63,93,99,252,253]. Molecular interpretation of experimental data on $\Gamma_{23}$ and $\Gamma_{21}$ has been conducted within the framework of the Kirkwood-Buff theory [62,254,255], the local bulk partition model [95,106,256], and the solvent exchange model [257–261]. The latter has been originally formulated and derived for the binding of a cosolvent molecule to a binding site on the surface of a solute with the release of a single water molecule to the bulk [260]. In a later development, the model has been reformulated to include a stoichiometric release of any number of water molecules to the binding site [83,131,132]. The modified solvent exchange model has been used in conjunction with the results of cosolvent-dependent volumetric measurements to evaluate the equilibrium constants, $k$, for the binding of the destabilizing cosolvent urea and the stabilizing cosolvent glycine betaine to proteins and their low-molecular weight analogues [67,68,83,131,137]. These studies have sketched a picture in which the specific mode of action of an individual cosolvent is governed by a fine balance between the energetics of direct cosolvent–protein interactions and the cosolvent-dependent energetics of cavity formation (the excluded volume effect) [67,68].

In this work, we continue this line of research and characterize the thermodynamic profile of direct interactions between urea and lysozyme when the protein retains its native conformation over the entire range of temperatures and urea concentrations employed in this study.
We perform urea-dependent measurements of the partial molar volume and adiabatic compressibility of lysozyme within a temperature range of 18–45 °C. From these measurements, we calculate the equilibrium constants for urea–protein association, $k$, as well as the related changes in volume, $\Delta V^\circ$, and adiabatic compressibility, $\Delta K^\circ_S$, for each experimental temperature. We use the van’t Hoff analysis of the temperature dependence of $k$ to evaluate the enthalpic, $\Delta H^\circ$, and entropic, $\Delta S^\circ$, components of the binding-induced changes in standard state free energy ($\Delta G^\circ = -RT \ln k$). We also calculate the binding-induced change in expansibility $[\Delta E^\circ = (\partial \Delta V^\circ / \partial T)_P]$.

The combined analysis of the values of $\Delta G^\circ$, $\Delta H^\circ$, and $\Delta S^\circ$ offers novel molecular insights into the nature of urea–protein interactions allowing one to discriminate between the energetics of direct intermolecular interactions and the order (disorder) introduced into the system. Our results represent the first evaluation of the full thermodynamic profile of the binding of urea to the native conformation of a globular protein. The single-conformation nature of our result is important, because theoretical studies have suggested that interactions of cosolvents, such as urea and TMAO, with proteins may depend not only on the chemical nature and the number of solvent-accessible atomic groups but also on the conformational state of the protein [262, 263]. We compare the resulting data of this investigation with the literature data for binding of urea to unfolded proteins and cyclic dipeptides. The comparison reveals a qualitative but not quantitative agreement between our data and literature data. The observed similarity is intriguing given the variety of experimental techniques and approaches used for determining $\Delta G^\circ$, $\Delta H^\circ$, and $\Delta S^\circ$ and the difference in the nature of the solutes and points to a common feature dominating urea–solute interactions.

5.3 Materials and Methods

5.3.1 Materials

Hen egg white lysozyme was purchased from Sigma-Aldrich Canada, Ltd. (Oakville, ON). The protein was exhaustively dialyzed against distilled water and subsequently lyophilized. The volumetric and spectroscopic studies of lysozyme reported here were performed in a 10 mM sodium phosphate buffer adjusted to pH 7.0. In all densimetric and ultrasonic velocimetric measurements, the protein concentrations were on the order of $\sim 6$ mg/mL. The protein concentration was measured spectrophotometrically using the reported urea-dependent extinction coefficients, $\varepsilon_{280}$ [134].

Aqueous solutions of urea at concentrations of 1, 2, 3, 4, 5, 6, 7, and 8 M were prepared by weighing 10–50 g of urea and adding pre-estimated amounts of the phosphate buffer to achieve the desired molalities, $m$. The molar concentration, $C$, of urea was computed from its molal
value, \( m \), using the equation \( C = \left[ 1/(m \rho_b) + \phi V/1000 \right]^{-1} \), where \( \rho_b \) is the density of pure buffer and \( \phi V \) (\( = 44.1 \text{ cm}^3 \text{ mol}^{-1} \)) is the apparent molar volume of urea [83]. Urea solutions were subsequently used as solvents for preparing protein solutions.

### 5.3.2 Optical spectroscopy

UV light absorbance spectra of the proteins were recorded at 25 °C with a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, ON). CD spectra were recorded using an Aviv model 62A DS spectropolarimeter (Aviv Biomedical, Inc., Lakewood, NJ). All CD spectroscopic measurements were performed in a 1 mm path-length cuvette.

### 5.3.3 High-precision densimetry

All densities were measured at 18, 25, 35, and 45 °C with a precision of \( \pm 1.5 \times 10^{-4} \% \) using a vibrating tube densimeter (DMA-5000, Anton Paar, Gratz, Austria). The apparent molar volumes, \( \phi V \), of water or the protein were calculated from the relationship \( \phi V = M/\rho_0 - (\rho - \rho_0)/(\rho C) \), where \( M \) is the molecular weight of a solute (for lysozyme, \( M = 14.3 \text{ kDa} \)), \( C \) is the molar concentration of a solute, and \( \rho \) and \( \rho_0 \) are the densities of the solution and the solvent (water or urea solution), respectively.

### 5.3.4 Ultrasonic velocimetry

Solution sound velocities, \( U \), were measured at 18, 25, 35, and 45 °C at a frequency of 7.2 MHz using the resonator method and a previously described differential technique [175, 177, 231]. The analysis of the frequency characteristics of the ultrasonic resonator cells required for sound velocity measurements was performed with a Hewlett-Packard (Mississauga, ON) model E5100A network/spectrum analyzer. For the type of ultrasonic resonators used in this work, the accuracies of the sound velocity measurements are approximately \( \pm 1 \times 10^{-4} \% \) [111, 177]. The key acoustic characteristic of a solute derived from ultrasonic measurements is the relative molar sound velocity increment, \( [U] = (U - U_0)/(U_0 C) \), where \( U \) and \( U_0 \) are the sound velocities in the solution and solvent, respectively.

The values of \( [U] \) were combined with the \( \phi V \) values derived from densimetric measurements to calculate the apparent molar adiabatic compressibilities, \( \phi K_S \), from the equation \( \phi K_S = \beta S_0 (2\phi V - 2[U] - M/\rho_0) \), where \( \beta S_0 = (\rho_0 U_0^2)^{-1} \) is the coefficient of adiabatic compressibility of the solvent [111, 178, 179]. The values of \( \rho_0 \), \( U_0 \), and \( \beta S_0 \) were measured for each concentration of urea used in this study. Three to five independent measurements were taken for each evaluation.
of $\phi V$ or $\phi K_S$. The reported values of $\phi V$ or $\phi K_S$ represent the averages of these measurements with the errors calculated as standard deviations.

Given the relatively low protein concentrations used in our study ($\sim 6$ mg/mL), we do not discriminate between the apparent and partial molar volumetric properties. Hence, we treat below the experimentally determined apparent molar volumes, $\phi V$, and compressibilities, $\phi K_S$, of the proteins as being equivalent to their partial molar volumes, $V^0$, and compressibilities, $K_S^0$, respectively.

### 5.4 Results

Figure 5.1 shows the far-UV (panel a) and near-UV (panel b) CD spectra of lysozyme at 18 and 45 °C and 0 and 8 M urea. Inspection of panels a and b of Figure 5.1 reveals that the CD spectra of lysozyme do not undergo any significant temperature- or urea-induced alterations between 18 and 45 °C and 0 to 8 M urea. By extension, we conclude that the protein remains native throughout the entire range of temperatures and urea concentrations employed in this study.

![Figure 5.1: Far-UV (a) and near-UV (b) CD spectra of lysozyme at 18 celsius (circles) and 45 celsius (squares) in the absence (filled) and presence (empty) of 8 M urea.](image)

Figure 5.1: Far-UV (a) and near-UV (b) CD spectra of lysozyme at 18 celsius (circles) and 45 celsius (squares) in the absence (filled) and presence (empty) of 8 M urea.
Figure 5.2 presents the urea dependences of the partial molar volume, \( V^\circ \) (panel a), the relative molar sound velocity increment, \([U]\) (panel b), and the partial molar adiabatic compressibility, \( K^\circ_S \) (panel c), of water at 18, 25, 35, and 45 °C. The excess partial molar volume, \( \Delta V^\circ_1 \), of water (in cubic centimeters per mole) derived from the data plotted in Figure 5.2a can be approximated by:

\[
\Delta V^\circ_1 = -0.0011[\text{urea}] - 0.0011[\text{urea}]^2,
\]

\[
\Delta V^\circ_1 = -0.0046[\text{urea}] - 0.0015[\text{urea}]^2,
\]

\[
\Delta V^\circ_1 = -0.0076[\text{urea}] - 0.0012[\text{urea}]^2,
\]

\[
\Delta V^\circ_1 = -0.0063[\text{urea}] - 0.0017[\text{urea}]^2
\]

at 18, 25, 35, and 45 °C, respectively (where [urea] is the molar concentration of urea). The excess partial molar adiabatic compressibility, \( \Delta K^\circ_{S1} \), of water (in units of \( 10^{-4} \) cubic centimeters per mole per bar) in Figure 5.2b can be approximated by:

\[
\Delta K^\circ_{S1} = -0.10[\text{urea}] + 0.0012[\text{urea}]^2,
\]

\[
\Delta K^\circ_{S1} = -0.09[\text{urea}] + 0.0026[\text{urea}]^2,
\]

\[
\Delta K^\circ_{S1} = -0.09[\text{urea}] + 0.0031[\text{urea}]^2,
\]

\[
\Delta K^\circ_{S1} = -0.08[\text{urea}] + 0.0037[\text{urea}]^2
\]

at 18, 25, 35, and 45 °C, respectively. These parabolic approximations were used below when analyzing the urea dependences of the partial molar volume, \( V^\circ \), and adiabatic compressibility, \( K^\circ_S \), of lysozyme.

Figure 5.3 shows the urea dependencies of the partial molar volume, \( V^\circ \) (panel a), the relative molar sound velocity increment, \([U]\) (panel b), and the partial molar adiabatic compressibility, \( K^\circ_S \) (panel c), of lysozyme at 18, 25, 35, and 45 °C. At all temperatures, the volumetric parameters of lysozyme change monotonically with urea, consistent with the absence of urea-induced conformational alterations.

As shown in panels a and c of Figure 5.3, at all temperatures studied, the partial molar volume, \( V^\circ \), and adiabatic compressibility, \( K^\circ_S \), of lysozyme increase with an increase in urea concentration. This observation, which is in line with previous urea-dependent volumetric parameters of solutes [83, 134], reflects solvent exchange in the vicinity of the protein; urea gradually replaces water from the solvation layer of lysozyme. On the other hand, the increase in \( V^\circ \) and \( K^\circ_S \) with an increase in temperature is a well-documented property of proteins, reflecting the temperature-induced decrease in the extent of solute–solvent interactions [205,264].

The urea-dependent changes in the partial molar volume, \( \Delta V^\circ \), and adiabatic compressibility, \( \Delta K^\circ_S \), of lysozyme at each experimental temperature in panels a and c of Figure 5.3 will be analyzed below within the framework of a previously described statistical thermodynamic model [83,131,132]. In the model, urea–solute interactions are viewed as cosolvent binding to independent and identical binding sites with each binding event being accompanied by a release of \( r \) water molecules to the bulk [83,131,132].
Figure 5.2: Urea dependencies of the partial molar volume (a), the relative molar sound velocity increment (b), and the partial molar adiabatic compressibility (c) of water at 18, 25, 35, and 45 °C. Experimental data in panels a and b were approximated by second-order polynomials (—).

5.5 Discussion

5.5.1 Data analysis

A change in volume, $\Delta V^o$, accompanying the transfer of a solute from water to a concentrated water/cosolvent mixture is given by the following analytical expression [83, 131, 132]:

\[
\Delta V^o = \sum \Delta V_i^o
\]
Figure 5.3: Urea dependencies of the partial molar volume (a), the relative molar sound velocity increment (b), and the partial molar adiabatic compressibility (c) of lysozyme at 18, 25, 35, and 45 °C. Experimental data in panels a and b were approximated by Equations (5.1) and (5.2), respectively (—).

\[
\Delta V^\circ = \Delta V_C - \gamma_1 n \Delta V_1^\circ + \frac{\Delta V \left( \frac{a_3}{a_1} \right) k}{1 + \left( \frac{a_3}{a_1} \right) k} 
\]

(5.1)

where \(a_1\) and \(a_3\) are the activities of water and the cosolvent (urea), respectively; \(n\) is the number of water molecules solvating a solute in pure water (hydration number); \(r\) is the number of water
molecules released to the bulk from the hydration shell of a binding site; $\Delta V_C$ is the differential volume of a cavity enclosing a solute in a concentrated urea solution and water; $k$ is the equilibrium constant for a reaction in which urea binds to a solute replacing $r$ water molecules; $\Delta V = \Delta V^0 + \gamma_1 r \Delta V^o_1 - \gamma_3 \Delta V^o_3$ is the change in volume accompanying an elementary reaction of urea binding to a binding site in a concentrated cosolvent solution; $\Delta V^o$ is the binding volume in an ideal solution; $\Delta V^o_1$ and $\Delta V^o_3$ are the excess partial molar volumes of water and the cosolvent in a concentrated solution, respectively; and $\gamma_1$ and $\gamma_3$ are the correction factors reflecting the influence of the bulk solvent on the properties of solvating water and the cosolvent, respectively.

In Equation (5.1), the activity of water was taken equal to its mole ratio: $a_1 = [W]/([W] + [\text{urea}])$, where the actual concentration of water at each experimental temperature and urea concentration was determined from the relationship $[W] = (\rho - [\text{urea}]M_{\text{urea}})/M_W$, where $\rho$ is the density of the urea solution (data not shown) and $M_{\text{urea}}$ and $M_W$ are the molecular masses of urea and water, respectively. The activity of urea, $a_3$, was assumed to be equal to its molar concentration.

For urea, the number of released water molecules, $r$, is 2 [131]. Correction factors $\gamma_1$ and $\gamma_3$ may adopt values ranging from 0, when the solvent in the vicinity of a solute and the solvent in the bulk are physicochemically indistinguishable, to 1, when the solvent in the solvation shell of a solute is independent of the solvent in the bulk [83, 131]. As a first approximation, correction factor $\gamma_1$ can be assumed to be unity for charged and nonpolar atomic groups, within the hydration shells of which, water is highly constrained and, therefore, relatively insensitive to structural and thermodynamic changes of water in the bulk [83, 132]. On the other hand, water solvating polar atomic groups (that easily exchange their hydrogen bonds between solute and bulk water) should be more responsive to cosolvent-induced structural and thermodynamic changes in water in the bulk, thereby producing a $\gamma_1$ close to 0 [83, 131, 132]. Thus, for a protein, correction factor $\gamma_1$ is roughly equal to the fraction of nonpolar and charged groups in the solvent-accessible surface area of a solute: $\gamma_1 = [(S_n + S_C)/S_A]$ [83, 131, 137]. Cosolvent molecules being, in general, much larger than water molecules retain a considerable portion of their interactions with bulk solvent even when solvating a solute. The physicochemical properties of the cosolvent in the solvation shell of a solute or in the bulk should not be very different with correction factor $\gamma_3$ being close to 0 [83, 131, 137]. Consequently, we use in Equation (5.1) a $\gamma_3$ of 0.

For lysozyme, $\gamma_1$ is 0.62, which can be calculated from the solvent-accessible surface areas of nonpolar, $S_n$, and charged, $S_C$, groups and the total solvent-accessible surface area, $S_A$ [205]. The hydration number of a protein, $n$, can be estimated as the ratio of its solvent-accessible surface area, $S_A$, of 6685 Å$^2$ to 9 Å$^2$ ($S_W$), the effective cross-sectional area of a water molecule. For lysozyme, $n = S_A/S_W = 743$. 
As discussed previously, $\Delta V_C$ in Equation (5.1) may or may not be considered explicitly. If it is not considered explicitly, the cosolvent-induced change in the cavity volume, $\Delta V_C$, if any, appears as an added contribution to the values of $\Delta V^o$ without affecting the values of $k$ determined from fitting the urea-dependent volume data presented in Figure 5.3a [132,134].

To fit the urea-dependent compressibility data shown in Figure 5.3c, we used an expression derived for a change in isothermal compressibility, $\Delta K_T^o$, accompanying the transfer of a solute from water to a concentrated water/cosolvent mixture [131–133]:

$$\Delta K_T^o = \Delta K_{TC} - \gamma_1 n \Delta K_{T1}^o + \frac{\Delta K_T \left( \frac{n}{r} \right) \left( \frac{a_3}{a_1^2} \right) k}{1 + \left( \frac{a_3}{a_1^2} \right) k} + \frac{\Delta V^2 \left( \frac{n}{r} \right) \left( \frac{a_3}{a_1} \right) k}{RT \left[ 1 + \left( \frac{a_3}{a_1^2} \right) k \right]^2}$$  (5.2)

where $\Delta K_{TC} = -(\partial \Delta V_C/\partial P)_T$ is the differential compressibility of the cavity enclosing a solute in water and a concentrated cosolvent solution; $\Delta K_T = \Delta K_{T0} + \gamma_1 r \Delta K_{T1}^o - \gamma_3 \Delta K_{T3}^o$ is the change in compressibility associated with urea binding in a concentrated cosolvent (urea) solution; $\Delta K_{T1}^o$ and $\Delta K_{T3}^o$ are the excess partial molar isothermal compressibilities of water and the cosolvent in a concentrated cosolvent solution, respectively; and $\Delta K_{T0} = -(\partial \Delta V_0/\partial P)_T$ is the change in compressibility caused by urea binding in an ideal solution.

Approximating the experimental points presented in panels a and c of Figure 5.3 with Equations (5.1) and (5.2), respectively, produces the three adjustable parameters of the fit, namely, equilibrium constants, $k$, as well as changes in volume, $\Delta V^o$, and compressibility, $\Delta K_T^o$, accompanying the binding of urea to lysozyme at 18, 25, 35, and 45 °C. The results of the fitting are listed in Table 5.1. The uncertainties in Table 5.1 represent the standard deviations of the mean of measurement repeats. The error arising from the multiple assumptions of the model used to derive Equations (5.1) and (5.2) and solution nonidealities related to high urea concentrations (e.g., deviations of the activity coefficient from unity or aggregation of urea at high concentrations) are not considered. Inspection of data presented in Table 5.1 reveals a number of interesting observations. First, at all temperatures studied, the equilibrium constants, $k$, derived from the volume and compressibility measurements are in good agreement. The agreement lends credence to our experimental protocols and the analytical procedures used to treat urea-dependent volumetric data. Second, the values of $k$, $\Delta V^o$, and $\Delta K_T^o$ all decrease as the temperature increases from 18 to 45 °C. Note that the temperature slope of the change in volume, $\Delta V^o$, is the change in expansibility, $\Delta E^o = (\partial \Delta V^o/\partial T)_P$, accompanying the binding of urea to the protein.

In our analysis, the number of urea-binding sites on the surface of the protein is given by $n/r \approx 370$. This value corresponds to the maximum number of possible binding sites. A 3-fold lower number of urea-binding sites of native lysozyme of 119 was estimated by Makhatadze
and Privalov in their seminal work [162]. Their number has been derived from treating data on the specific heat of protein titration by urea within the framework of the simplest binding model in conjunction with a number of simplifying assumptions [162]. It is virtually impossible to come up with a model- and assumption-free estimate of the number of binding sites. More importantly, inspection of Equations (5.1) and (5.2) reveals that the assumed number of urea-binding sites, while influencing the determined values of $\Delta V^\circ$ and $\Delta K_S^\circ$, has no effect on the value of $k$.

A note of caution is warranted here. As Equations (5.1) and (5.2) have been derived under the simplifying assumption of identical and independent urea-binding sites, the determined equilibrium constants, $k$, produce only an “effective” description of the binding event. The protein surface is highly heterogeneous with respect to the local geometry and chemical nature of solvent-exposed atomic groups. The actual equilibrium constants for individual solvent-exposed atomic groups may be significantly distinct from the effective values determined in this work.

Table 5.1: Urea–Protein Binding Constants Determined from the Volume and Compressibility Data and Changes in Volume, $\Delta V^\circ$, and Compressibility, $\Delta K_S^\circ$, Accompanying the Binding in an Ideal Solution at Various Temperatures

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>18 °C</th>
<th>25 °C</th>
<th>35 °C</th>
<th>45 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$ (M) (volume)</td>
<td>0.17 ± 0.01</td>
<td>0.10 ± 0.03</td>
<td>0.07 ± 0.01</td>
<td>0.04 ± 0.01</td>
</tr>
<tr>
<td>$k$ (M) (comp.)</td>
<td>0.15 ± 0.01</td>
<td>0.12 ± 0.01</td>
<td>0.06 ± 0.01</td>
<td>0.05 ± 0.01</td>
</tr>
<tr>
<td>$\Delta V^\circ$ (cm$^3$ mol$^{-1}$)</td>
<td>0.10 ± 0.01</td>
<td>0.07 ± 0.01</td>
<td>0.04 ± 0.01</td>
<td>0.01 ± 0.03</td>
</tr>
<tr>
<td>$\Delta K_S^\circ$ (cm$^3$ mol$^{-1}$ bar$^{-1}$)</td>
<td>2.23 ± 0.04</td>
<td>1.85 ± 0.04</td>
<td>1.81 ± 0.14</td>
<td>1.58 ± 0.11</td>
</tr>
</tbody>
</table>

Figure 5.4 plots $\ln k$ against $1/T$ (the van’t Hoff plot). In this plot, we used the average values of $k$ derived from the analysis of the volume and compressibility data. The slope of the $\ln k$ versus $1/T$ dependence is related to the enthalpy of direct urea–protein interactions, $\Delta H^\circ$, via the van’t Hoff equation, $\Delta H^\circ = -R[\partial \ln k/\partial (1/T)]_P$. With $\Delta H^\circ$ and a standard state free energy change, $\Delta G^\circ = -RT \ln k$, one calculates a change in entropy, $\Delta S^\circ = (\Delta H^\circ - \Delta G^\circ)/T$, associated with the binding. Table 5.2 presents the values of $\Delta G^\circ$, $\Delta H^\circ$, $\Delta S^\circ$, $\Delta V^\circ$, $\Delta E^\circ$, and $\Delta K_S^\circ$ for the urea–protein interactions at 25 °C. The uncertainties of $\Delta H^\circ$, $\Delta S^\circ$, and $\Delta E^\circ$ data in Table 5.2 represent computer-generated errors of the linear approximation of $\ln k$ versus $T^{-1}$ (van’t Hoff) and $\Delta V^\circ$ versus $T$ plots.

It is pertinent here to briefly discuss the relationship between the equilibrium constant for an elementary reaction of the binding of a cosolvent to a binding site on the surface of a solute (protein), $k$, and the free energy of transfer of the solute from water to a concentrated cosolvent solution. A change in free energy accompanying transfer of a solute from the principal solvent (water) to a solvent/cosolvent (water/urea) mixture is given by the expression [83]
Chapter 5. Thermodynamic Profile of Protein–Urea Interactions

Figure 5.4: Temperature dependence of the equilibrium constant, $k$ (the van’t Hoff plot).

\[
\begin{align*}
\Delta G^\circ \text{ (kcal mol}^{-1}\text{)} &\quad 1.3 \pm 0.1^a \\
\Delta H^\circ \text{ (kcal mol}^{-1}\text{)} &\quad -8.7 \pm 0.4 \\
\Delta S^\circ \text{ (cal mol}^{-1}\text{K}^{-1}\text{)} &\quad -33.4 \pm 1.2 \\
\Delta V^\circ \text{ (cm}^3\text{mol}^{-1}\text{)} &\quad 0.07 \pm 0.02^b \\
\Delta E^\circ \times 10^{-3} \text{ cm}^3\text{mol}^{-1}\text{K}^{-1} &\quad -3.3 \pm 0.2 \\
\Delta K_S^\circ \times 10^{-4} \text{ cm}^3\text{mol}^{-1}\text{bar}^{-1} &\quad -1.85 \pm 0.04^b
\end{align*}
\]

\(^a^\text{Calculated as } \Delta G^\circ = -RT \ln k, \text{ where } k \text{ is the average equilibrium constant at } 25 ^\circ\text{C from Table 5.1.} \)

\(^b^\text{From Table 5.1.}\)

Table 5.2: Thermodynamic Parameters of Urea–Protein Interactions at 25 °C

\[
\Delta G_{tr} = \Delta \Delta G_c - \left(\frac{n}{r}\right) RT \ln \left[ \left( \frac{a_1}{a_{10}} \right)^r + k \left( \frac{a_3}{a_{10}^2} \right) \right] \quad (5.3)
\]

where $\Delta \Delta G_c$ is the differential free energy of cavity formation in a concentrated cosolvent solution and water and $a_{10}$ is the activity of water in a protein solution in the absence of the cosolvent.

At a low protein concentration, $a_{10}$ is close to unity ($a_{10} \approx 1$). Thus, Equation (5.3) simplifies to the form

\[
\Delta G_{tr} = \Delta \Delta G_c - \left(\frac{n}{r}\right) RT \ln(a_1^r + ka_3) \quad (5.4)
\]

Note that the $\Delta \Delta G_f = -(n/r)RT \ln(a_1^r + ka_3)$ term in Equation (5.4) represents the differential free energy of solute–solvent interactions in a concentrated urea solution and water. For a 1:1 binding stoichiometry ($r = 1$), the $\Delta \Delta G_f$ term in Equation (5.4) simplifies to the relationship
derived by Schellman: \( \Delta \Delta G_I = -nRT \ln(a_1 + ka^3) \) [258, 259, 265].

Inspection of Equation (5.4) reveals that the differential free energy of solute–solvent interactions, \( \Delta \Delta G_I \), can be negative even if the equilibrium constant \( k \) is smaller than unity and, consequently, if the standard state binding free energy, \( \Delta G^o = -RT \ln k \), is positive. Furthermore, the sign of the transfer free energy (and, hence, the stabilizing-destabilizing action of a cosolvent) is determined by an interplay between the cosolvent-induced changes in the cavity, \( \Delta \Delta G_C \), and interaction, \( \Delta \Delta G_I \), terms [68]. For urea, the negative \( \Delta \Delta G_I \) term prevails over the positive \( \Delta \Delta G_C \) term, resulting in a negative \( \Delta G_{tr} \) that explains the strongly denaturing action of urea.

### 5.5.2 Thermodynamic parameters of urea–protein interactions

Inspection of the energetic profile of association of urea with lysozyme reveals that the unfavorable change in standard state free energy, \( \Delta G^o \), of 1.3 ± 0.1 kcal mol\(^{-1}\) results from an unfavorable (negative) change in the entropic contribution, \( T \Delta S^o = -10.0 \pm 0.4 \) kcal mol\(^{-1}\), that prevails over a favorable (negative) change in enthalpy, \( \Delta H^o \), of \(-8.7 \pm 0.4 \) kcal mol\(^{-1}\). These values have been determined for the first time for a protein that retains its native conformation throughout the entire experimental range of applied temperatures and urea concentrations. Also for the first time, the experimental data have been analyzed within the framework of a solvent exchange model in which cosolvent may replace any number of water molecules (for urea, the number of replaced water molecules is 2) [67, 83, 131, 132]. Recall that, in the original solvent exchange model, only a single water molecule was posited to be replaced by a cosolvent molecule [259, 260, 265]. While our results cannot be directly related in a strictly quantitative manner to any of the previously reported data (which have been obtained on unfolded proteins or small analogues of proteins), a qualitative comparison is instructive and provides useful insights into the nature of interactions of urea with protein groups.

The energetic parameters reported for urea association with unfolded proteins and small analogues of proteins are in qualitative agreement with our data. Namely, an unfavorable \( \Delta G^o \) of urea binding results from an interplay of a favorable \( \Delta H^o \) and an unfavorable \( \Delta S^o \) [162, 163, 266]. The average values of \( \Delta G^o \), \( \Delta H^o \), and \( \Delta S^o \) reported for binding of urea to unfolded cytochrome \( c \), ribonuclease A, and lysozyme are 1.6 ± 0.2 kcal mol\(^{-1}\), 2.1 ± 0.5 kcal mol\(^{-1}\), and -13 ± 2 cal mol\(^{-1}\) K\(^{-1}\), respectively [162]. These values have been derived from treating calorimetric titration data by a simple binding model [162]. Negative changes in the enthalpy of urea binding have been also reported for the unfolded states of ribonuclease A, trypsin, \( \beta \)-lactoglobulin, ovalbumin, and bovine serum albumin with an average \( \Delta H^o \) of \(-5.5 \) kcal mol\(^{-1}\) [266]. Along the same lines, calorimetric heat of dissolution measurements have revealed that the binding of urea to the peptide moiety of cyclic dipeptides is characterized by \( \Delta G^o \), \( \Delta H^o \), and \( \Delta S^o \) values
of $2.2 \pm 0.3 \text{ kcal mol}^{-1}$, $-8.1 \pm 0.7 \text{ kcal mol}^{-1}$, and $-35 \pm 3 \text{ cal mol}^{-1} \text{K}^{-1}$, respectively, in close quantitative agreement with our data [163].

Thus, results of our study qualitatively agree with reported data on the energetics of urea–protein interactions obtained on the basis of less sophisticated binding models. The qualitative agreement between our results for a native protein and the previously reported data on unfolded proteins and cyclic dipeptides is significant and consistent with the notion that urea interacts with protein groups via localized interactions with comparable affinity, independent of the conformational state or even the size of the solute. This notion appears not to be in agreement with the results of MD simulations, suggesting that cosolvent–protein interactions are modulated, in addition to the chemical nature and the number of solvent-accessible atomic groups, by the conformational state of the protein [262, 263]. In particular, it has been suggested that urea preferentially associates with the unfolded rather than the folded state given a larger nonpolar solvent-accessible surface area of the former [262].

### 5.5.3 Rationalization of the thermodynamic profile of urea–protein interactions

The thermodynamic picture of urea–protein interactions emerging from our results and previously published results is intriguing. The positive sign of $\Delta G^\circ$ is not surprising and characterizes urea–protein interactions as weak binding; consequently, very high concentrations of urea are, generally, required for causing noticeable changes in the conformational preferences of proteins. The negative signs of $\Delta H^\circ$ and $\Delta S^\circ$, on the other hand, are more difficult to understand. The negative sign of $\Delta H^\circ$ has been interpreted as a thermodynamic signature of hydrogen bond(s) formed by urea and polar protein groups [162, 163]. However, within the framework of the solvent exchange model, formation of urea–protein hydrogen bonds should be preceded by disruption of urea–water and protein–water hydrogen bonds with a concomitant formation of water–water hydrogen bonds. Note that urea interacts with more or less comparable affinities with all protein groups, including the peptide backbone and amino acid side chains, via hydrogen bonding and van der Waals interactions [22, 83, 227, 267].

MD simulations from Berne’s group and Grubmüller’s group suggest that protein–urea hydrogen bonds are significantly weaker than protein–water hydrogen bonds (by $\sim 1.6 \text{ kcal mol}^{-1}$) and that the denaturing action of urea is due to stronger dispersion interactions between protein and urea compared to those between protein and water [82, 86, 268]. The computed interaction energies for urea–protein, water–protein, water–urea, and water–water hydrogen bonds are 4.5, 6.2, 4.6, and 6.6 kcal mol$^{-1}$, respectively [86]. Stumpe and Grubmüller have hypothesized that the favorable enthalpy of urea–protein interactions originates from slightly stronger (by $\sim 0.4 \text{ kcal mol}^{-1}$) hydrogen bonds between released water molecules relative to water–protein
hydrogen bonds [86]. The following simplified considerations, however, are not in agreement with this notion. As a probable scenario, a urea molecule may interact with a protein by forming two hydrogen bonds with surface groups and releasing to the bulk two water molecules that, in turn, will form two hydrogen bonds with other water molecules. The resulting change in enthalpy would be positive and equal $2(1.6 - 0.4) = 2.4 \text{ kcal mol}^{-1}$, in contrast to our measured $\Delta H^\circ$ of $-8.7 \pm 0.4 \text{ kcal mol}^{-1}$. Any other scenario of hydrogen bond reshuffling that can be devised would not bring the calculated enthalpy, $\Delta H^\circ$, to any satisfactory agreement with the experimental value.

A large negative change in enthalpy, $\Delta H^\circ$, accompanying the binding of urea to folded and unfolded proteins and small analogues of proteins (cyclic dipeptides) is an experimental reality reported not only by us but also by other research groups [162,163,266]. Given an average hydrogen bond energy of $-4$ to $-5 \text{ kcal mol}^{-1}$, our measured $\Delta H^\circ$ of $-8.7 \pm 0.4 \text{ kcal mol}^{-1}$ implies that roughly two new hydrogen bonds are formed when urea replaces two water molecules from the solvation layer of lysozyme. It is difficult to rationalize this observation solely in terms of reshuffling water–protein, water–urea, water–water, and urea–protein hydrogen bonds. Other factors need to be considered, e.g., the role of dispersion interactions. An intriguing possibility is the “underhydration” of urea in the bulk with some of its polar groups not forming hydrogen bonds as a statistical average. This possibility, accounting for the tendency of urea to self-aggregate in an aqueous solution [23,269], is consistent with the results of MD simulations that suggest that urea–water hydrogen bonds are significantly weaker (by $\sim 2 \text{ kcal mol}^{-1}$) than water–water hydrogen bonds [86]. This notion is also in agreement with the MD simulation results from the Horinek group that suggest that, in water, a urea molecule, of six possible hydrogen bonds, forms only 4.3–4.5 hydrogen bonds depending on the urea concentration [84]. In the vicinity of the protein, however, urea may strongly interact with protein groups by accepting and/or donating hydrogen bonds and by forming stronger dispersion interactions than water, because of its molecular structure, as suggested by a number of theoretical and experimental studies [22, 82, 86, 217, 226, 227, 262, 267, 268]. In fact, strong urea–protein dispersion interactions may be instrumental in facilitating formation of stronger hydrogen bonds between urea and polar protein groups. Additionally, water molecules previously hydrating urea molecules, when released to the bulk, will form stronger water–water hydrogen bonds contributing to a further decrease in enthalpy.

The measured unfavorable change in entropy, $\Delta S^\circ$, cannot be related to the binding-induced release of water molecules, an event that should contribute favorably to a change in entropy accompanying urea–protein binding. However, it is consistent with the hypothesis of underhydrated urea in the bulk that strongly interacts with the surface groups of a protein upon binding. Newly formed urea–protein hydrogen bonds should bring about a decrease in entropy in agreement with experimental results. In addition, partial immobilization of urea molecules, forming more than one hydrogen bond with protein groups, and the related partial “fixing” of
amino acid side chains hydrogen bonded to urea will additionally contribute to a decrease in entropy.

5.5.4 Volumetric considerations

Changes in volume, $\Delta V^\circ$, compressibility, $\Delta K_S^\circ$, and expansibility, $\Delta E^\circ$, associated with the binding of urea to the protein are $0.07 \pm 0.01 \text{ cm}^3 \text{ mol}^{-1}$, $(-1.85 \pm 0.04) \times 10^{-4} \text{ cm}^3 \text{ mol}^{-1} \text{ bar}^{-1}$, and $(-3.3 \pm 0.2) \times 10^{-3} \text{ cm}^3 \text{ mol}^{-1} \text{ K}^{-1}$, respectively (see Section 5.5.1). These parameters in conjunction with our determined values of $\Delta G^\circ$, $\Delta H^\circ$, and $\Delta S^\circ$ are capable of collectively characterizing the pressure and temperature responses of single-site urea–protein interactions:

$$\Delta G = \Delta H^\circ - T \Delta S^\circ + \Delta V^\circ (P - P_0) + \Delta E^\circ \times (T - T_0)(P - P_0) - 0.5\Delta K_S^\circ (P - P_0)^2$$  \hspace{1cm} (5.5)

The latter capability is important given the recent applications of combined effects of pressure and cosolvent to conformational characterizations of proteins [270–272]. Our estimates based on Equation (5.5) reveal that $\Delta G$ increases (becomes more unfavorable) as the temperature increases, while practically not changing with pressure. The latter is a consequence of a very small change in volume, $\Delta V^\circ$, accompanying the urea–protein interaction. These results may look qualitatively inconsistent with the results of all-atom replica exchange MD simulations of Canchi et al. [226] Specifically, the authors found that the $m$-value for urea-induced denaturation of Trp-cage miniprotein increases with temperature and, even more significantly, with pressure [226]. An increase in the $m$-value may imply stronger urea–protein interactions in contrast to our results. This is an interesting point that, clearly, warrants further study. However, it should be noted that the $m$-values are contributed not only by the energetics of direct urea–protein interactions but also by the urea-dependent energetics of cavity formation [67,68]. Also, the $m$-values, generally, correlate with a change in solvent-accessible surface area accompanying protein unfolding with the latter being a nontrivial function of temperature and pressure. In this respect, given its small size (only 20 amino acids), the Trp cage may exhibit uncharacteristic conformational energetics within the temperature-pressure-cosolvent space. As noted above, experimental studies of the temperature and pressure dependencies of $m$-values for different proteins would be useful and clarify the situation. Such studies are underway in our lab.
5.6 Conclusion

We report results of high-precision densimetric and ultrasonic velocimetric measurements in lysozyme solutions in the presence and absence of urea between 18 and 45 °C. The measurements were taken at pH 7, at which the protein remains native over the entire range of temperatures and denaturant concentrations employed in this work. Analysis of the urea dependencies of the partial molar volume, $V^\circ$, and adiabatic compressibility, $K_S^\circ$, of lysozyme within the framework of a statistical thermodynamic model yielded the equilibrium constants, $k$, for an elementary reaction in which urea binds to protein replacing two waters of hydration. On the basis of the van’t Hoff analysis of the temperature dependence of $k$, we evaluated changes in enthalpy, $\Delta H^\circ$, and entropy, $\Delta S^\circ$, in addition to a change in standard state free energy, $\Delta G^\circ$, accompanying urea–protein association. The thermodynamic profile of the binding of urea to a native protein is in qualitative agreement with similar data reported for unfolded proteins and small protein analogues. The observed similarity suggests that urea interacts with protein groups via localized interactions essentially independent of the conformational state or even the size of the solute. We rationalize the standard state free energy, $\Delta G^\circ$, of urea–protein interactions and its enthalpic, $\Delta H^\circ$, and entropic, $T\Delta S^\circ$, contributions, in conjunction with the published MD simulation results, as being consistent with the picture in which urea molecules, being underhydrated in the bulk, form enthalpically stronger interactions with the surface protein groups while paying a high entropic price. We discuss the ramifications of our results for providing insights into the combined effects of urea, temperature, and pressure on the conformational preferences of proteins. More generally, our results underscore the strength and versatility of strategically designed volumetric measurements for providing novel thermodynamic insights into complex molecular phenomena.
Chapter 6

Effect of Urea on Protein–Ligand Association


Author’s Contribution: Ikbae Son analyzed the data and participated in the preparation of the manuscript.

6.1 Abstract

We combine experimental and theoretical approaches to investigate the influence of a cosolvent on a ligand-protein association event. We apply fluorescence measurements to determining the affinity of the inhibitor tri-N-acetylglucosamine for lysozyme at urea concentrations ranging from 0 to 8 M. Notwithstanding that, at room temperature and neutral pH, lysozyme retains its native conformation up to the solubility limit of urea, the affinity of (GlcNAc)$_3$ for the protein steadily decreases as the concentration of urea increases. We analyze the urea dependence of the binding free energy within the framework of a simplified statistical thermodynamics-based model that accounts for the excluded volume effect and direct solute–solvent interactions. The analysis reveals that the detrimental action of urea on the inhibitor–lysozyme binding originates from competition between the free energy contributions of the excluded volume effect and direct urea-solute interactions. The free energy contribution of direct urea-solute interactions narrowly overcomes the excluded volume contribution thereby resulting in urea weakening the protein–ligand association. More broadly, the successful application of the simple model em-
ployed in this work points to the possibility of its use in quantifying the stabilizing/destabilizing action of individual cosolvents on biochemical folding and binding reactions.

6.2 Introduction

Water-miscible cosolvents have been extensively employed in protein studies as a means to control the folding/unfolding equilibria of protein species [87, 247, 253, 261]. A large set of water-soluble cosolvents with a potential to affect protein stability has been identified, while their influence on conformational equilibria of proteins has been characterized by various experimental means. Subsequent theoretical and computational studies have begun to provide us with insights into the molecular processes involved in regulating the stabilizing/destabilizing action of specific cosolvents [22, 60, 93, 258, 260, 273]. More recently, water-miscible cosolvents have been used in “osmotic stress” measurements to modulate the equilibria of folding and binding reactions of proteins and nucleic acids with the resulting data being interpreted in terms of a release or uptake of water molecules to/from the bulk [61, 66, 274, 275]. In osmotic stress measurements, the equilibrium constant, $K$, of a reaction is measured as a function of the activity of cosolvent, $a_3$, or water, $a_1$, with the differential interaction, $\Delta \Gamma_{23}$, and hydration, $\Delta \Gamma_{21}$, parameters being related to changes in $K$ according to $\Delta \Gamma_{23} = (\partial \ln K / \partial \ln a_3)_{T,P} = \Delta n_3 - (N_3/N_1)\Delta n_1$ and $\Delta \Gamma_{21} = (\partial \ln K / \partial \ln a_1)_{T,P} = \Delta n_1 - (N_1/N_3)\Delta n_3$ [61, 63, 93, 96, 276–278]. In these expressions, $N_1$ and $N_3$ are, respectively, the numbers of moles of water and cosolvent in solution, while $\Delta n_1$ and $\Delta n_3$ refer, respectively, to the excess numbers of water and cosolvent around the solute.

Recently, based on a statistical thermodynamics-based formalism, we have derived expressions that link $\Delta \Gamma_{23}$ and $\Delta \Gamma_{21}$ to the relative affinity of cosolvent for the solute(s) [67, 68]:

$$\Delta \Gamma_{23} = \left( \frac{\partial K}{\partial \ln a_3} \right)_{T,P} = -\left( \frac{1}{RT} \right) \left( \frac{\partial \Delta \Delta G_C}{\partial \ln a_3} \right)_{T,P} \left[ \frac{(n_C + n_F)}{r} \right] \left[ \frac{ka_3}{(a_r^r + ka_3)} \right]$$

$$- (n_C + n_F) \left[ \frac{a_r^r}{(a_r^r + ka_3)} \right] \frac{N_3}{N_1}$$

(6.1)
\[ \Delta \Gamma_{21} = \left( \frac{\partial K}{\partial \ln a_1} \right)_{T,P} - \left( \frac{1}{RT} \right) \left( \frac{\partial \Delta \Delta G_C}{\partial \ln a_1} \right)_{T,P} \left\{ \frac{(n_C + n_F)}{r} \right\} \left[ \frac{ka_3}{(a_1^r + ka_3)} \right] \]
\[ + (n_C + n_F) \left[ \frac{a_1^r}{(a_1^r + ka_3)} \right] \frac{N_1}{N_3} \]

where \( \Delta \Delta G_C \) is the differential free energy of cavity formation between the ligand–protein complex and the free ligand and free protein; \( n_C \) and \( n_F \) are the numbers of water-binding sites of the ligand–protein complex and free ligand and free protein, respectively; \( r \) is the number of waters of hydration replaced by cosolvent upon its binding to a solute (for urea, \( r \) equals 2 [83, 131]); and \( k \) is the equilibrium constant of an elementary reaction in which a cosolvent binds to a solute (ligand or protein) replacing \( r \) waters of hydration.

Equations (6.1) and (6.2) allow one to quantify the action of a cosolvent in terms of its influence on the free energy of cavity formation [the first term in Equations (6.1) and (6.2)] and that of direct solute-solvent interactions [the second and third terms in Equations (6.1) and (6.2)]. In this work, we apply this formalism to analyzing the urea-dependence of the binding of the inhibitor tri-N-acetylglucosamine \([\text{GlcNAc}]_3\) to lysozyme. The binding of \([\text{GlcNAc}]_3\) to lysozyme has been extensively characterized structurally and thermodynamically [141–146, 173, 183]. It is an enthalpy-driven reaction that proceeds with an unfavorable change in entropy [173]. In a recent publication, we have applied volumetric measurements to characterization of changes in hydration accompanying the binding of \([\text{GlcNAc}]_3\) to lysozyme in a pH 5.5 acetate buffer [198]. Being a well-characterized system, the lysozyme–\([\text{GlcNAc}]_3\) complex makes an attractive model for deriving a detailed understanding of the molecular determinants of the effect of urea on ligand–protein binding events. Here, we study the urea-dependence of lysozyme–\([\text{GlcNAc}]_3\) association at pH 7 and 25 °C. At neutral pH and room temperature, the protein remains native at concentrations of urea as high as its solubility limit [99]. This property of lysozyme enables us to study the effect of urea on protein binding within a wide concentration range of 0 to 8 M.
6.3 Materials and Methods

6.3.1 Materials

Lysozyme from chicken egg white was purchased from Sigma-Aldrich Canada (Oakville, ON, Canada). It was exhaustively dialyzed against water and, subsequently, lyophilized. Tri-N-acetylglucosamine [tri-(N-acetyl)-chitotriose] was acquired from V-labs (Covington, LA, USA) and used without further purification. All measurements were performed in a pH 7.0 buffer containing 10 mM monosodium phosphate/disodium phosphate and urea at concentrations ranging between 0 and 8 M. Solutions of urea with concentrations of 2, 4, 6, and 8 M were prepared by weighing 10 to 50 g of urea and adding pre-estimated amounts of the phosphate buffer to achieve the desired molalities, m. The molar concentration, C, of urea was computed from its molal value, m, using 
\[
C = \frac{1}{m \rho_b} + \frac{\phi V}{1000} - 1
\]
where \(\rho_b\) is the density of pure buffer and \(\phi V = 44.1 \text{ cm}^3 \text{ mol}^{-1}\) is the apparent molar volume of urea [131]. Urea solutions were subsequently used as a solvent for preparing protein and ligand solutions.

The concentration of lysozyme was determined from light absorbance measured at 25 °C with a Cary 300 Bio spectrophotometer (Varian Canada, Inc., Mississauga, ON, Canada) using the reported urea-dependent extinction coefficients, \(\varepsilon_{280}\) [134].

6.3.2 Fluorescence

Fluorescence intensity measurements were performed at 25 °C using an Aviv model ATF 105 spectrofluorometer (Aviv Associates, Lakewood, NJ, USA). Fluorescence titration profiles were measured by the incremental addition of aliquots of (GlcNAc)_3 to a 10 mm-pathlength cuvette initially containing 1 mL of lysozyme solution. The protein concentrations were between 17 and 34 µM. The initial volume of lysozyme was delivered to the cuvette using a 1 mL Hamilton syringe, while additions of the ligand [(GlcNAc)_3] were made using a 10 µM Hamilton syringe (Hamilton Co., Reno, NV). Each syringe was equipped with a Chaney adapter which allowed a relative delivery accuracy of ±0.1 %. The protein samples were excited at 296 nm and the intensity of emission light was recorded through a monochromator at 345 nm. When calculating the relative fluorescence intensity of lysozyme, we have taken into account its dilution due to addition of ligand solution.
6.4 Results and Discussion

Figure 6.1 plots the relative fluorescence intensity of lysozyme against the ligand-to-protein binding ratio, \( r = [LT]/[PT] \) (where \([LT]\) is the total concentration of \((\text{GlcNAc})_3\) and \([PT]\) is the total concentration of lysozyme) at 0 (panel A), 2 (panel B), 4 (panel C), 6 (panel D), and 8 M (panel E) urea. To determine the association constants for the ligand–protein binding reaction, the experimental points shown in Figure 6.1a to e were approximated by an analytical function representing a one-to-one stoichiometric binding [198]:

\[
F = F_0 + \alpha \Delta F
\]

where \( F \) is the relative fluorescence intensity, \( F_0 \) is the initial value of \( F \) in the absence of the ligand, \( \Delta F \) is the maximum change in \( F \) when the protein becomes saturated by the ligand; \( \alpha \) is the fraction of protein molecules associated with the ligand. The fraction of ligated protein, \( \alpha = [PL]/[PT] \), is given by the relationship:

\[
\alpha = 0.5(r + 1)Y^{-1} - [0.25(r - 1)^2 + \frac{r + 1}{Y} + Y^{-2}]^{0.5}
\]

where \( Y = 2K_b[PT] \); \( K_b = [PL]/([P][L]) \) is the binding constant; \([PL]\) is the concentration of the \((\text{GlcNAc})_3\)–lysozyme complex; \([P]\) is the concentration of free protein; and \([L]\) is the concentration of free ligand.

Fitting the binding profiles illustrated in Figure 6.1a to e to Equation (6.3) yields the values of \( K_b \) for \((\text{GlcNAc})_3\) association with lysozyme at each experimental urea concentration. Table 6.1 presents the binding constants, \( K_b \), and standard state binding free energies, \( \Delta G_b = -RT \ln K_b \). Figure 6.2 plots \( \Delta G_b \) against the concentration of urea. Inspection of data presented in Table 6.1 and Figure 6.2 reveals a steady decrease in the binding affinity with an increase in urea concentration. This observation should be viewed against the backdrop of the protein not undergoing structural changes throughout the entire range of urea concentrations of 0 to 8 M [99]. Thus, the observed decrease in the affinity of \((\text{GlcNAc})_3\) for lysozyme should be caused by the increasingly unfavorable energetics of ligand and protein desolvation as the concentration of urea rises. On the other hand, since the protein does not undergo structural alterations, the energetics of interatomic interactions of lysozyme and \((\text{GlcNAc})_3\) in the binding pocket should not be strongly influenced by the presence of cosolvent in solution. In general, the influence of a cosolvent on folding and binding reactions involving proteins is governed by a tight balance between the cosolvent-induced changes in the energetics of direct solute–solvent interactions and the energetics of cavity formation (the excluded volume effect) [67, 68]. Theoretically, the effect of urea on protein–ligand association can be presented as follows:
Figure 6.1: The relative fluorescence intensity of lysozyme plotted against the (GlcNAc)$_3$-to-lysozyme molar ratio, $r$, at 0 (panel A), 2 (panel B), 4 (panel C), 6 (panel D), and 8 M (panel E) urea. The excitation and emission wavelengths are 296 and 345 nm, respectively. The initial concentrations of lysozyme were 34, 17, 34, 18, and 18 µM at 0, 2, 4, 6, and 8 M urea, respectively. The experimental points were fitted to Equation (6.3) (solid lines).

\[
\Delta G_b(a_3) = \Delta G_{b0} + \Delta G_{trPL}(a_3) - \Delta G_{trP}(a_3) - \Delta G_{trL}(a_3)
\]

(6.5)

in which $\Delta G_{b0}$ is the binding free energy in the absence of urea; and $\Delta G_{trPL}(a_3)$, $\Delta G_{trP}(a_3)$, and $\Delta G_{trL}(a_3)$ are the urea-dependent changes in free energy accompanying the water-to-urea
transfer of the protein–ligand complex, free protein, and free ligand, respectively.

A change in free energy accompanying transfer of a solute from the principal solvent (water) to a solvent–cosolvent (water–urea) mixture is given by $\Delta G_{tr} = \Delta \Delta G_C(a_3) - (n/r)RT\ln(a_1^n + ka_3)$ [83]. Hence, the urea-dependent protein-ligand binding free energy can be presented as follows:

$$\Delta G_b(a_3) = \Delta G_{b0} + \Delta \Delta G_C(a_3) - \frac{n_C - n_F}{r}RT\ln(a_1^n + ka_3) \quad (6.6)$$

Note that differentiation of Equation (6.6) with respect to $a_3$ and $a_1$ yields Equations (6.1) and (6.2) derived for the differential interaction, $\Delta \Gamma_{23}$, and hydration, $\Delta \Gamma_{21}$, parameters, respectively. In other words, Equation (6.6) can also be derived by integration of Equations (6.1) and (6.2). The analysis we perform below with Equation (6.6) is based on the following posits. Firstly, the differential number of water-binding sites, $(n_C - n_F)$, can be estimated as the ratio of the binding-induced change in solvent-accessible surface area, $\Delta S_A$, to 9 Å$^2$, the effective cross-sectional surface area of a water molecule. A change in solvent-accessible surface area, $\Delta S_A$, accompanying the binding of (GlcNAc)$_3$ to lysozyme of $-686 \pm 113$ Å$^2$ has been computed from structural data [198]. With $\Delta S_A$ of $-686 \pm 113$ Å$^2$, we calculate $(n_C - n_F)$ of $-76 \pm 13$ ($-686/9$). Secondly, the initial estimate of the equilibrium constant, $k$, used in the fit is assumed to be close to that determined for urea–lysozyme interactions 0.11 ± 0.01 M [134]. One should keep in mind, however, that the ligand (GlcNAc)$_3$, being a saccharide and not a peptide, may differ from proteins with respect to the equilibrium constant for its interactions with urea. Thirdly, the activity of urea, $a_3$, is taken equal to its molar concentration, $C_3$. Finally, recent scaled particle theory based simulations have revealed that the slope $(\partial \Delta \Delta G_C / \partial C_3)_{T,P}$ for folding and binding reactions is roughly proportional to a change in solvent-accessible surface area, $\Delta S_A$, associated with the reaction with the proportionality coefficient, $(\partial \Delta \Delta G_C / \partial C_3)_{T,P}/\Delta S_A$, being dependent on the size of a cosolvent [68]. This notion is consistent with the earlier results of SPT-based computations of the free energy of cavity formation from the Graziano lab [279–281]. For urea, $(\partial \Delta \Delta G_C / \partial C_3)_{T,P}/\Delta S_A$ is 1.06 ± 0.26 cal mol$^{-1}$ M$^{-1}$ Å$^{-2}$ [68]. In the analysis below, we fix the value of $(\partial \Delta \Delta G_C / \partial C_3)_{T,P}/\Delta S_A$ to 1.06 cal mol$^{-1}$ M$^{-1}$ Å$^{-2}$ and fit the $\Delta G_b$-versus-urea data presented in Figure 6.2 to Equation (6.6). The best-fit value of $k$ is 0.09 ± 0.01 M in close proximity with the initial estimate of 0.11 ± 0.01 M. Inspection of Figure 6.2 reveals a reasonable qualitative and quantitative agree-

<table>
<thead>
<tr>
<th>[urea], M</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_b$, 10$^4$ M$^{-1}$</td>
<td>7.6 ± 0.6</td>
<td>4.1 ± 0.5</td>
<td>2.0 ± 0.3</td>
<td>1.2 ± 0.1</td>
<td>0.6 ± 0.1</td>
</tr>
<tr>
<td>$\Delta G_b$, kcal mol$^{-1}$</td>
<td>6.7 ± 0.5</td>
<td>6.3 ± 0.5</td>
<td>5.9 ± 0.5</td>
<td>5.6 ± 0.4</td>
<td>5.1 ± 0.4</td>
</tr>
</tbody>
</table>

Table 6.1: The (GlcNAc)$_3$–Lysozyme Binding Constants, $K_b$, and Binding Free Energy, $\Delta G_b = -RT \ln K_b$, at Different Urea Concentrations

Chapter 6. Effect of Urea on Protein–Ligand Association
ment between the experimental points and the approximating function which lends credence to the analytical approach employed in our treatment. The approximating function correctly captures the tendency of urea to diminish the affinity of (GlcNAc)₃ for lysozyme. The numerical agreement between the fit and experimental points is also satisfactory in view of the relatively large error of experimental data and the simplified nature of the approximating function.

Figure 6.2: The urea-dependence of the (GlcNAc)₃–lysozyme binding free energy. The experimental points were fitted to Equation (6.6) (solid line).

To unveil the balance of thermodynamic forces governing the effect of urea on the ligand–lysozyme binding event, Figure 6.3 plots the urea-induced changes in the free energies of cavity formation, \( \Delta \Delta G_C = [(\partial \Delta \Delta G_C / \partial C_3)_{T,P} / \Delta S_A] \Delta S_A a_3 \) and that of direct solute-solvent interactions, \( \Delta \Delta G_I(a_3) = -(n_C - n_F)/r \cdot RT \ln(a_1 + ka_3) \), against the concentration of urea. Inspection of Figure 6.3 reveals that the excluded volume effect, as reflected in \( \Delta \Delta G_C \), decreases (becomes more favorable) as the concentration of urea increases. This result is consistent with the picture in which cavity formation becomes more “expensive” as more urea is added to solution, a notion in agreement with previous observations [279, 282]. Thus, the excluded volume effect increasingly favors the associated state (ligand–protein complex) as the urea concentration increases. In contrast, the solute-solvent interaction term, \( \Delta \Delta G_I \), increases (becomes more unfavorable) with an increase in urea favoring the dissociated state (free ligand plus free protein). This observation reflects the thermodynamically more favorable urea-solute interactions relative to water-solute interactions. Importantly, the differential binding free energy between a concentrated urea solution and water, \( \Delta \Delta G_b \), represents a small difference between the two large numbers, \( \Delta \Delta G_C \) and \( \Delta \Delta G_I \), suggesting a strong mutually compensating effect. The interaction term, \( \Delta \Delta G_I \), slightly prevails over the cavity term, \( \Delta \Delta G_C \), resulting in urea weakening the affinity of (GlcNAc)₃ for the protein. For example, at 8 M urea, the values of \( \Delta \Delta G_C \), \( \Delta \Delta G_I \), and \( \Delta \Delta G_b \) equal -5.8, 7.3, and 1.5 kcal mol⁻¹, respectively.
Figure 6.3: The urea-dependence of excluded volume ($\Delta \Delta G_C$) and direct solute–solvent interaction ($\Delta \Delta G_I$) components of the differential (GlcNAc)$_3$–lysozyme binding free energy ($\Delta \Delta G_b$) between a concentrated urea solution and water.

6.5 Conclusion

In the aggregate, our cumulative results lead to a number of important inferences. The effect of urea (and, by extension, other cosolvents) on a protein binding reaction is governed by a close interplay between the free energy contributions of the excluded volume effect and direct solute–solvent interactions. The excluded volume contribution is accounted for by a term that has been determined from SPT-based computations and is proportional to the binding-induced change in solvent-accessible surface area [68]. The free energy contribution of direct solute–solvent interactions is realistically described by a simplified statistical thermodynamics-based analytical approach in which urea-binding sites on the surface of protein and ligand are treated as identical and independent. The success of this simplified algorithm in treating a phenomenon as complex as the effect of cosolvent on a biochemical reaction is encouraging and points to the predictive capabilities of the analytical approach employed in this study.
Chapter 7

General Conclusions and Future Perspectives

7.1 Conclusions

My doctoral thesis is aimed at characterizing the role of protein–solvent interactions in protein–ligand binding events and folding/unfolding conformational transitions of proteins. To quantify non-specific protein–solvent interactions, we measured and analyzed volumetric data within the framework of statistical thermodynamic models. Along these lines, we designed a series of experiments to characterize and quantify changes in protein–solvent interactions accompanying protein–ligand association and protein folding/unfolding transitions and to discriminate between protein–water and protein–cosolvent interactions.

In Chapter 2, we applied high precision densimetric and ultrasonic velocimetric measurements to characterizing the binding of (GlcNAc)_3 to lysozyme at 25°C. This enzyme–inhibitor association event causes changes in volume, \( \Delta V \), and adiabatic compressibility, \( \Delta K_S \), of \(-44 \pm 2 \) cm\(^3\)mol\(^{-1}\) and \((22 \pm 4) \times 10^{-4} \) cm\(^3\)mol\(^{-1}\)bar\(^{-1}\), respectively. We interpreted these results in conjunction with X-ray crystallographic data in terms of changes in hydration of the ligand and the protein. On the basis of our \( \Delta V \) data, we estimated that \( 79 \pm 44 \) water molecules were released to the bulk from the hydration shells of the protein and the ligand. Our \( \Delta K_S \) data suggested a \( 4 \pm 2 \) % decrease in the mean-square fluctuations of the intrinsic volume of the protein, \( \langle \delta V_M^2 \rangle \) (or 2% decrease in \( \delta V_M \)). Thus, the ligand binding stiffens the enzyme and renders it less dynamic compared to the unbound state.

In Chapter 3, we present a volumetric characterization of the association of the cAMP-binding domain of EPAC1 with cAMP. We discuss the conceptual basis for resolving macroscopic prop-
erties (volume and compressibility) into microscopic events (protein hydration and dynamics). Our volumetric analysis performed, in conjunction with the structural data on the cAMP/EPAC complex, revealed that \( \sim 10^3 \) water molecules were released to the bulk upon the complex formation. We found that the holoprotein is more rigid and less dynamic compared to its apo-form as reflected in a 4\% decrease in its intrinsic coefficient of adiabatic compressibility. Our results permitted us to estimate the hydration, \( \Delta S_{\text{hyd}} \), and configurational, \( \Delta S_{\text{conf}} \), contributions to the binding entropy, \( \Delta S_b \). The sign and the magnitude of \( \Delta S_b \) are determined by a subtle balance between the \( \Delta S_{\text{hyd}} \) and \( \Delta S_{\text{conf}} \) terms.

In Chapter 4, we applied volumetric measurements to determining the partial molar volumes, \( V^\circ \), and adiabatic compressibilities, \( K_S^\circ \), of folded and unfolded proteins at urea concentrations ranging from 0 to 8 M. We interpreted the urea-dependent volumetric properties of proteins in terms of the solvent exchange model in which the binding of urea to a protein proceeds with a release of two waters of hydration. Comparison of the urea-binding thermodynamic data for the folded lysozyme with the similar data obtained on small molecules mimicking proteins revealed a lack of cooperative effects involved in protein–urea interactions. Analysis of the volumetric properties of the unfolded protein apocytochrome \( c \) suggested that, in the absence of urea, the protein exhibits a solvent-accessible surface area, \( S_A \), comparable to that of folded holocytochrome \( c \). An increase in urea concentration from 0 to 8 M causes an additional \( \sim 20\% \) increase in \( S_A \).

In contrast to lysozyme and apocytochrome \( c \), ribonuclease A and \( \alpha \)-chymotrypsinogen A undergo urea-induced unfolding transitions. Our data suggest that the unfolded states of ribonuclease A and \( \alpha \)-chymotrypsinogen A exhibit solvent-accessible surface areas which are roughly twice as large as those of their respective native states. In general, urea-dependent changes in volumetric properties offered a practical way for evaluating the effective solvent accessible surface areas of biologically significant unfolded polypeptides.

In Chapter 5, we report results of high-precision densimetric and ultrasonic velocimetric measurements in lysozyme solutions in the presence and absence of urea between 18 and 45 \(^\circ\)C. The measurements were taken at pH 7, at which lysozyme remains native over the entire range of temperatures and at urea concentrations between 0 and 8 M. Analysis of the urea dependencies of the partial molar volume, \( V^\circ \), and adiabatic compressibility, \( K_S^\circ \), of lysozyme yielded the equilibrium constants, \( k \), for an elementary protein–urea binding event. On the basis of the van’t Hoff analysis of the temperature dependence of \( k \), we evaluated the full thermodynamic profile of protein–urea interactions. Changes in standard state free energy, \( \Delta G^\circ \), enthalpy, \( \Delta H^\circ \), and entropy, \( \Delta S^\circ \), for the binding of urea are 1.3 ± 0.1 kcal mol\(^{-1}\), −8.7 ± 0.4 kcal mol\(^{-1}\), and −33.4 ± 1.2 cal mol\(^{-1}\) K\(^{-1}\), respectively. The thermodynamic profile of the binding of urea to a native protein is in qualitative agreement with similar data reported for unfolded proteins and small protein analogues. The similarity suggests that urea interacts with protein groups
via localized interactions essentially independent of the conformational state or even the size of the solute. We analyzed the standard state free energy, $\Delta G^\circ$, of urea-protein interactions and its enthalpic, $\Delta H^\circ$, and entropic, $\Delta S^\circ$, contributions, in conjunction with the published MD simulation results. The analysis revealed that urea molecules are underhydrated in the bulk and form enthalpically strong interactions with the surface protein groups while paying a high entropic price.

In Chapter 6, we combine experimental and theoretical approaches to evaluating the influence of a cosolvent on a protein–ligand association event. We applied fluorescence measurements to determining the affinity of $(\text{GlcNAc})_3$ for lysozyme at urea concentrations ranging from 0 to 8 M. We analyzed the urea dependence of the binding free energy within the framework of a simplified statistical thermodynamics-based model that accounts for the excluded volume effect and direct solute–solvent interactions. The effect of urea on the protein–ligand binding reaction was governed by a close interplay between the free energy contributions of the excluded volume effect and direct solute–solvent interactions. The free energy contribution of direct urea–solute interactions narrowly overcome the excluded volume contribution, thereby, resulting in urea weakening the protein–ligand association. More broadly, the successful application of the simple model employed in this work points to the possibility of its use in quantifying the stabilizing/destabilizing action of individual cosolvents on biochemical folding and binding reactions.

### 7.2 Future Directions

In my thesis, interactions of a destabilizing cosolvent, urea, with globular proteins has been examined by combining the solvent exchange model and volumetric measurements. Here, I propose studies that will expand our general understanding of the role of protein–cosolvent interactions in protein folding/unfolding transitions further characterizing the influence of stabilizing and destabilizing cosolvents.

#### 7.2.1 pH-induced folding and unfolding of lysozyme in urea solution

While globular proteins are generally susceptible to drastic urea-induced shifts in folded–unfolded equilibria, lysozyme maintains its native conformation even at urea concentration as high as 8 M [134]. Lysozyme is also unaffected by a reduction in pH at room temperature [247]. However, the conformational equilibrium of lysozyme shifts to the unfolded state at extremes of pH in the presence of urea [283]. The stability of a protein is affected by the combined effect of protons and urea.
We propose to characterize the acid- and urea-induced conformational transitions by adding aliquots of HCl to lysozyme in binary solutions consisting of urea and water. We will perform pH-dependent measurements of the partial molar volume and adiabatic compressibility of lysozyme at urea concentrations 0 and 8 M. The structural transitions of lysozyme will be monitored by CD spectroscopic measurements.

We will analyze the folding/unfolding transition profiles of lysozyme within the framework of a two-state model to evaluate the equilibrium constant, $K = [F] / [U]$ (where $[F]$ and $[U]$ are concentrations of the folded and unfolded protein states, respectively). Since the binding of protons to ionizable residues causes pH-induced unfolding, the effect of pH on the equilibrium constant, $K(pH)$, can be expressed as [119]:

$$K(pH) = K_0(1 + 10^{pK_a - pH})^\Delta v$$

where $K_0$ is the equilibrium constant in the absence of protons; $\Delta v$ is the difference in the numbers of protons bound to the protein in its final and initial structural states; and $pK_a$ is the effective value of $pK_a$ of abnormally titrable amino acid residues in the unfolded state. The values of $pK_a$ can be used to calculate the differential number of protons bound to the folded state relative to the unfolded state [284].

We will use volumetric measurements data to calculate changes in partial molar volume, $\Delta V$, and compressibility, $\Delta K_T$, associated with the protein unfolding. Assuming a two-state transition of lysozyme unfolding, the pH-dependence of a volumetric observable $X$ can be presented as follows [119]:

$$\Delta X(pH) = \Delta X \frac{K(pH)}{1 + K(pH)}$$

where $\Delta X$ is the change in a volumetric observable associated with the acid-induced unfolding in the presence and absence of urea, respectively. We will apply a statistical thermodynamic model described in Chapter 4 to analyzing pH- and urea-dependent volumetric properties of lysozyme. We will interpret volumetric properties in terms of changes in volume, $\Delta V^\circ$, and compressibility, $\Delta K_S^\circ$, associated with the pH-induced unfolding. The analysis will reveal the balance of thermodynamic forces between the free energy contributions of the excluded volume effect and direct solute–solvent interactions. The comparison of pH-induced unfolding of lysozyme at different urea concentrations will provide insights into the effect of urea on the pH-induced unfolding of lysozyme.
7.2.2 Thermodynamic characterization of the interaction of glycine betaine with a native protein

We have conducted a systematic characterization of interactions of glycine betaine with low-molecular weight model compounds and with native globular proteins [133, 137]. It has been found that cooperative effects are absent in the interactions of GB with folded proteins [137]. Chapter 5 demonstrates that the solvent exchange model can be used in conjunction with the results of cosolvent-dependent volumetric measurements to characterize the thermodynamic profile of interactions between cosolvents and a native proteins. We propose to continue this line research and characterize the full thermodynamic profile of interactions between GB and lysozyme over a wide range of GB concentrations.

We propose to perform GB-dependent measurements of the partial molar volume and adiabatic compressibility of lysozyme within a temperature range of 18–45 °C. The fact that, at pH 7, lysozyme remains native within this temperature range provides an opportunity to study the interactions of GB with the native protein within the entire range of experimentally accessible GB concentrations [137]. From these measurements, we will calculate the equilibrium constants for protein–GB association, $k$, as well as the related changes in volume, $\Delta V^\circ$, and adiabatic compressibility, $\Delta K_S^\circ$, for each experimental temperature. We will use the van’t Hoff analysis of temperature dependence of $k$ to evaluate the enthalpic and entropic components of the binding-induced changes in standard state free energy. This thermodynamic characterization will offer novel molecular insights into the nature of protein–GB interactions with the enthalpic and entropic contributions to the free energy of interaction. Since GB is an effective protein stabilizer while urea is a strong denaturant, the proposed study will enable us to compare the thermodynamic signatures of protein interactions with cosolvents of differing chemical action.
Bibliography


[105] E. S. Courtenay, M. W. Capp, R. M. Saecker, and M. T. Record, “Thermodynamic analysis of interactions between denaturants and protein surface exposed on unfolding: interpretation of urea and guanidinium chloride m-values and their correlation with changes in


