INVESTIGATING THE EFFECTS OF IONIC LIQUIDS ON DNA GQUADRUPLEX AND PROTEIN STRUCTURE USING MOLECULAR DYNAMICS SIMULATIONS

Nicholas J. Paradis
Rowan University

Recommended Citation
Paradis, Nicholas J., "INVESTIGATING THE EFFECTS OF IONIC LIQUIDS ON DNA GQUADRUPLEX AND PROTEIN STRUCTURE USING MOLECULAR DYNAMICS SIMULATIONS" (2022). Theses and Dissertations. 3074.
https://rdw.rowan.edu/etd/3074
INVESTIGATING THE EFFECTS OF IONIC LIQUIDS ON DNA G-QUADRUPLEX AND PROTEIN STRUCTURE USING MOLECULAR DYNAMICS SIMULATIONS

by
Nicholas J Paradis

A Thesis

Submitted to the
Department of Chemistry and Biochemistry
College of Science and Mathematics
In partial fulfillment of the requirement
For the degree of
Master of Science in Pharmaceutical Sciences
at
Rowan University
August 11, 2022

Thesis Chair: Chun Wu, Ph.D., Professor, Department of Chemistry and Biochemistry

Committee Members:
Subash Jonnalagadda, Ph.D., Professor, Chair of the Department of Department of Chemistry and Biochemistry
Gregory A. Caputo, Ph.D., Associate Dean, Department of Chemistry and Biochemistry
Thomas M. Keck, PhD., Associate Professor, Department of Chemistry and Biochemistry
Dedications

I dedicate this work to my father, brother, grandmother and grandfather for their unwavering support in my academic studies and pursuits. You have taught me to chase my dreams when nothing slowed me down but my fears, which you have helped me to quash. I also thank my fellow peers Emily, Griffin, Dylan, Holli, Mike, Lucas, Brian and Phillip for inspiring each other and keeping good cheer in our endeavors together. This motley has a unique aura that I’ve not seen since we’ve gone our separate ways in academia or industry. Our memories are surreal.
Acknowledgements

I thank Dr. Chun Wu for recognizing the talent I harbored that I was too modest to realize. Since the beginning, you’ve helped me embark on an unforgettable journey that opened my eyes to the molecular world, rekindle my inspiration when it sputtered and reinforced my focus when it wavered. From being an unknown undergraduate to an ambitious graduate, teacher’s assistant, and PhD candidate, I thank you for these opportunities. I also like to thank Dr. Thomas Keck and Dr. Gregory Caputo for their excellent teachings, guidance and support that have brought me to new heights. Most importantly, I thank you three for helping me find myself in this world, for conquering my fears, and enabling me to make an impact on this Earth.
Abstract

Nicholas J Paradis
INVESTIGATING THE EFFECTS OF IONIC LIQUIDS ON DNA G-QUADRUPLEX AND PROTEIN STRUCTURE USING MOLECULAR DYNAMICS SIMULATIONS
2021-2022
Chun Wu, Ph.D.
Master of Science in Pharmaceutical Sciences

Nucleic acids and proteins have huge implications in biomedicine and bioengineering, however their instability limits their applicability. Current storage protocols are expensive; using a biocompatible media to store nucleic acids and proteins would reduce costs and increase their applicability. Ionic liquids (ILs) are molten salts that modulate the stability and activity of nucleic acids and proteins. In this thesis, molecular modeling studies of DNA/RNA and protein structure in ILs will be discussed (Chapter 1) and this method will be used to study the IL effects on the structure on the Pu22 c-MYC DNA G-quadruplex (Chapter 2) and the azurin protein (Chapter 3). ILs have been observed to stabilize/destabilize DNA G-quadruplexes linked to cancer oncogene expression, however the structural effects of imidazolium-based ILs on G-quadruplexes remain unknown. Bioengineering of azurin is attractive for soil bioremediation, thus understanding the structural changes induced by TMG-amino acid-based ILs will mediate future IL design for enhancing azurin’s activity. In Chapter 2, molecular dynamics (MD) simulations will elucidate the stabilizing mechanism of four imidazolium-based ILs of increasing hydrophobicity to Pu22, using the G-quadruplex stabilizer TMPyP4 as a molecular probe. In Chapter 3, conventional and replica-exchange MD simulations will provide insight into the enthalpic and entropic change induced by two TMG-AA based ILs on the folded and unfolded azurin conformations.
Table of Contents

Abstract ........................................................................................................................................ v

List of Figures ............................................................................................................................ x

List of Tables ............................................................................................................................. xi

Chapter 1: Investigating Ionic Liquid Interactions with Nucleic Acids and Proteins using Molecular Dynamics Simulations ......................................................... 1

1.1 Introduction ....................................................................................................................... 1

1.2 Methodology .................................................................................................................. 11

1.2.1 MD Simulations ..................................................................................................... 11

1.3 Thesis Outline ............................................................................................................. 13

Chapter 2: To Probe the Binding of TMPyP4 to c-MYC G-Quadruplex in Water and in Imidazolium-Based IL Solvents using Spectroscopy Coupled with Molecular Dynamics Simulations ................................................................. 16

2.1 Introduction .................................................................................................................... 16

2.1.1 G-Quadruplexes (G4s) ....................................................................................... 16

2.1.2 C-MYC Oncogene Pu22 G4 ............................................................................ 16

2.1.3 G4 Stabilizer TMPyP4 and Previous Studies ................................................ 18

2.1.4 Previous Studies of IL Effects on G4 Structure .............................................. 20

2.1.5 Experimental Overview ................................................................................ 22

2.2 Methodology ............................................................................................................... 22
Table of Contents (Continued)

2.2.1 Simulation System Protocols ................................................................. 22
2.2.2 MD Simulation Protocols ..................................................................... 23
2.2.3 Simulation Convergence ........................................................................ 24
2.2.4 Characterization of Structural Fluctuations .......................................... 24
2.2.5 K+ and O6 Positions .............................................................................. 25
2.2.6 Geometric Parameter Calculations ........................................................ 25
2.2.7 Identification of Ligand Binding Modes ................................................... 26
2.2.8 Calculating Atom Contacts between TMPyP4/IL and Pu22 ....................... 26
2.2.9 Representative Ligand Binding Trajectory ............................................. 27
2.2.10 Structural Alignment ............................................................................ 27
2.2.11 Ligand Binding Analysis ...................................................................... 27
2.2.12 Structural Visualization ...................................................................... 27
2.3 Results ....................................................................................................... 27
2.3.1 Pu22 Structure is Stable in Water and IL Solvents ................................. 27
2.3.2 Hydrophobic Side Chains of HMIM-Cl and OMIM-Cl Reduces TMPyP4
    Top and Bottom Binding to Pu22 ................................................................. 31
2.3.3 TMPyP4 Binding Decreases as IL Hydrophobicity Increases ................. 36
2.3.4 Top and Bottom Binding are the Major Binding Modes of TMPyP4 ...... 37
### Table of Contents (Continued)

2.3.5 Binding Mode of TMPyP4 to Pu22 Shows Lack of Intercalation at Top and Bottom G4 Layers ................................................................................................. 38

2.3.6 Atom Contact Plots Reveal Increased IL Binding to Pu22 G4 in ILs as Hydrophobic Chain Length Increases ................................................................... 39

2.4 Discussion ............................................................................................................. 40

2.5 Conclusion ............................................................................................................ 45

Chapter 3: Probing the Thermodynamic Destabilization of Azurin by Two Different Tetramethylguanidinium Amino Acid Ionic Liquids Using Molecular Dynamics Simulations ....................................................................................................................... 47

3.1 Introduction ........................................................................................................... 47

3.1.1 Previous Studies of Amino Acid-Based ILs on Protein Structure ............... 47

3.1.2 Azurin Protein .............................................................................................. 48

3.1.3 Experimental Overview ............................................................................... 49

3.2 Methodology ......................................................................................................... 50

3.2.1 Simulation System Protocols ....................................................................... 50

3.2.2 MD Simulation Protocols ............................................................................ 52

3.2.3 Simulation Convergence .............................................................................. 53

3.2.4 Protein Secondary Structure Information .................................................... 54

3.2.5 Characterization of Structural Conformations ............................................. 54
## Table of Contents (Continued)

3.2.6 Calculating Atom Contacts Between IL Cation/Anion and Protein Polar/Hydrophobic Side Chains.............................................................. 54

3.3 Results.............................................................................................................. 54

3.3.1 Thermal Unfolding Experiments ............................................................... 54

3.3.2 MD Simulations.......................................................................................... 55

3.3.3 Refolding Events in Unfolded Azurin ....................................................... 57

3.3.4 Atom Contact Plot Analysis..................................................................... 59

3.3.5 RMSF Analysis......................................................................................... 63

3.4 Discussion....................................................................................................... 65

3.5 Conclusions..................................................................................................... 67

Chapter 4: Concluding Statements...................................................................... 69

References .......................................................................................................... 71

Appendix A  To Probe the Binding of TMPyP4 to c-MYC G-Quadruplex in Water and in Imidazolium-based IL Solvents Using Spectroscopy Coupled with Molecular Dynamics Simulations ......................................................... 84

Appendix B  Probing the Thermodynamic Destabilization of Azurin by Two Different Tetramethylguanidinium Amino Acid Ionic Liquids Using Molecular Dynamics Simulations ................................................................. 178
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1. Examples of Molecular-Based Therapies</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2. Stabilization of DNA by HMIM-Br</td>
<td>8</td>
</tr>
<tr>
<td>Figure 3. Inhibition Mechanism of Xylanase by EMIM-Acetate</td>
<td>10</td>
</tr>
<tr>
<td>Figure 4. Force Field Parameters and Associated Equations</td>
<td>12</td>
</tr>
<tr>
<td>Figure 5. Pu22 G4, Ligand TMPyP4 and Four Imidazolium-Based ILs</td>
<td>18</td>
</tr>
<tr>
<td>Figure 6. Pu22 G4 RMSF</td>
<td>29</td>
</tr>
<tr>
<td>Figure 7. Pu22 Geometric Parameters in Apo-Form</td>
<td>31</td>
</tr>
<tr>
<td>Figure 8. TMPyP4 and IL Distribution Around Pu22</td>
<td>32</td>
</tr>
<tr>
<td>Figure 9. TMPyP4 Inhibition Mechanism by IL Cations</td>
<td>34</td>
</tr>
<tr>
<td>Figure 10. Effects of Ligand Geometry and G4 Topology on Ligand Binding</td>
<td>41</td>
</tr>
<tr>
<td>Figure 11. TMG-AA ILs in Our Study</td>
<td>50</td>
</tr>
<tr>
<td>Figure 12. Reference Structures in Our MD Simulations</td>
<td>51</td>
</tr>
<tr>
<td>Figure 13. Azurin RMSD in Water and ILs</td>
<td>56</td>
</tr>
<tr>
<td>Figure 14. Last Snapshots of Azurin in Water and ILs</td>
<td>58</td>
</tr>
<tr>
<td>Figure 15. IL Distribution in Last Snapshot Structures</td>
<td>60</td>
</tr>
<tr>
<td>Figure 16. Protein RMSF in Water and ILs</td>
<td>63</td>
</tr>
<tr>
<td>Figure 17. Experimental Protein Unfolding Thermodynamic Parameters</td>
<td>66</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1. Molecular-Based Therapies</td>
<td>2</td>
</tr>
<tr>
<td>Table 2. Literature Review of ILs on Nucleic Acid Stability</td>
<td>7</td>
</tr>
<tr>
<td>Table 3. MD Simulations of ILs on Protein Stability</td>
<td>9</td>
</tr>
<tr>
<td>Table 4. Percent Binding Modes of TMPyP4 to Pu22</td>
<td>37</td>
</tr>
<tr>
<td>Table 5. Atom Contacts Between IL and Pu22</td>
<td>39</td>
</tr>
<tr>
<td>Table 6. Atom Contacts Between Azurin and ILs</td>
<td>61</td>
</tr>
<tr>
<td>Table 7. Tabulated Averaged Protein RMSF Values</td>
<td>64</td>
</tr>
</tbody>
</table>
Chapter 1

Investigating Ionic Liquid Interactions with Nucleic Acids and Proteins using Molecular Dynamics Simulations

1.1 Introduction

Nucleic acids are polymeric strands of deoxyribose (DNA) or ribose (RNA) sugar monomers connected to four possible nitrogenous bases (DNA: A, T, C, G; RNA: A, U, C, G) at the 1’ ribose site and a phosphate group at the 5’ ribose site. Individual monomers are connected via sugar-phosphate covalent bonds and individual polymer strands can form pairs via complementary base-pairing (A-T, A-U, C-G). Nucleic acids hold critical responsibilities in molecular biology, including genetic storage, preservation and replication, protein production, gene regulation and reaction catalysis. Therefore, nucleic acids are considered potentially valuable biopharmaceuticals for biomedical applications. Over the last few decades, a number of DNA-based and RNA-based therapeutics under FDA-approval and in clinical trials have been developed to treat, diagnose and prevent many human diseases, including cancer \(^1\text{-}^3\), viral infections \(^4\) and genetic diseases \(^5\text{-}^7\) (Table 1, Figure 1).
Table 1

Molecular-Based Therapies

<table>
<thead>
<tr>
<th>Therapy</th>
<th>Functions</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>FDA-Approved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ligand-based</td>
<td>Modulate bio-processes</td>
<td>Simple structure</td>
<td>Most proteins “undruggable”</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Treat microbial infection</td>
<td>(atomic level)</td>
<td>↑Toxicity, ↓selectivity, ↓PK</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No immunogenicity</td>
<td>Challenging novel drug synthesis</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chemically-stable</td>
<td>↑Use, ↑drug resistance</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Polar/charged→poor cell uptake</td>
<td></td>
<td>2725</td>
</tr>
<tr>
<td>Protein-based</td>
<td>Modulate bio-process</td>
<td></td>
<td>Complex structure (AA level, 3D conformation)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Treat protein deficiency</td>
<td>↑Affinity, ↑specificity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑Immunity</td>
<td>Reduced off-target</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Drug delivery vehicle</td>
<td>bio-signaling</td>
<td>Chemical instability</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑Dosing, ↑Immunogenicity</td>
<td>230</td>
</tr>
<tr>
<td>DNA-based</td>
<td>Modulate gene expression</td>
<td>Simple structure</td>
<td>↑Genotoxicity, ↑Immunogenicity, ↑Infection</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gene-editing</td>
<td>(NT level)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑Immunity</td>
<td>Simple modification</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chemical instability</td>
<td></td>
</tr>
<tr>
<td>RNA-based</td>
<td>Modulate protein and RNA</td>
<td>Simple structure</td>
<td>↑Immunogenicity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and RNA expression</td>
<td>(NT level)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biosensor</td>
<td>No genotoxicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>↑Immunity</td>
<td>Simple modification</td>
<td>Chemical instability</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Charged→poor cell uptake</td>
<td>14</td>
</tr>
</tbody>
</table>

*Note.* PK: Pharmacokinetics. NT: Nucleotide. AA: Amino acid
Figure 1

Examples of Molecular-Based Therapies

Note. (A) Ligand-based therapy. (B) Protein-based therapy. (C) DNA-based therapy. (D) RNA-based therapy.

DNA/RNA-based therapeutics offer several advantages over small molecule-based and protein-based therapeutics, such as molecular simplicity, ease-of-modification and the ability to treat genetic diseases at the DNA and RNA level. However, nucleic acids are inherently unstable and thus severely limit their applicability. Specifically, nuclease-mediated and water-mediated self-hydrolysis of its sugar-phosphate backbone is the primary cause of nucleic acid degradation in-vivo and in storage settings. More importantly, reliably storing nucleic acids for therapeutic applications is a major logistical
challenge; current storage methods involve keeping nucleic acids under ultra-cold temperatures (-20°C and below) that is unfeasible at the global scale. Developing a storage medium for nucleic acids at room temperature could significantly increase their applicability in both research and patient treatment settings, as well as significantly reducing storage costs.

Proteins are long-chain polymers composed of twenty possible amino acids linked together via peptide bonding. Following ribosomal translation of mRNA in the cytoplasm of eukaryotic cells, nascent protein undergoes extensive modification and begins to fold from the primary amino acid sequence into higher-order secondary, tertiary and sometimes quaternary structures via hydrogen bonding between amino acid residues. Proteins in tertiary structure gain biological function, and the degree and type of function can vary depending on the orientation of their amino acid groups in the conformational fold. Proteins are critical in maintaining and fulfilling biological processes, serving as enzymes, storage and transport vehicles, ligand receptors, immunogenic agents, components in molecular scaffolds and hormones. Proteins thus carry immense potential in biotechnology, including the fields of biomedicine (see Table 1) and bioengineering. In biomedicine, the global protein therapeutics industry amassed ~$283 billion in revenue in 2020 and is forecasted to grow to ~$560 billion by 2030, thus increased usage of protein therapeutics is expected. However, liquid protein formulations are plagued by an increased degradation rate via hydrolysis of the protein backbone, leading to a shortened shelf life, protein aggregation, reduced protein efficacy and increased immunogenicity risks. In bioengineering, enhancing enzyme activity is desired for increasing final product yield in a reaction process. However, enzymes are usually
unstable and poorly soluble in organic solvents. More importantly, protein application is dependent on their efficient extraction and purification, which often utilize organic solvents. Finding a biocompatible media that can increase protein solubility and maintain protein structure and function is crucial for protein applications in biomedicine and bioengineering.

Ionic liquids (ILs) are molten salt compounds that hold several advantages over traditional molecular solvents, including high customizability, high polarity, high-conductivity, high thermostability, non-volatility and low toxicity, making them relatively useful for studying various chemical and biological phenomena. In aqueous media, different cation-anion combinations of varying chemical nature (i.e., hydrophobicity, polarity) produce ILs of asymmetrical shape, decreasing the cation-anion electrostatic interactions and decreases IL molecular packing, behaving less like a solid. This in-turn drastically lowers IL melting temperature ($T_m < 100^\circ C$) and causes cations and anions to dissociate and interact with other molecular species, be it biomolecule or the bulk solvent. Ethylammonium nitrate compounds are considered to be the first ILs developed in 1914 by Paul Walden. Since then, three generations of ILs have been made; First-generation ILs have high thermal stability and low volatility, yet are sensitive to air and water (i.e. aluminum chloride, 1-ethylpyridinium bromide); Second-generation ILs have increased resistance to air and water, yet are highly toxic and not very biodegradable (i.e. cations: dialkylimidazolium, phosphonium, ammonium; anions: tetrafluoroborates); Third-generation ILs are more biocompatible and have better biodegradability (i.e. cations: choline, dihydrogen phosphate; anions: amino acids, serine, aspartate). In fact, choline-based ILs are increasingly being used in
biocompatibility studies with biomolecules. The effects of IL interactions with biomolecules (i.e., nucleic acids and proteins) are critical to enhance our understanding of biomolecular structural dynamics under room temperature conditions.

Molecular dynamics (MD) simulation is a powerful computational tool that provides insight into biomolecular structure dynamics. This tool simulates a biological system over time to investigate detailed molecular interactions using the laws of physics. Moreover, MD simulations can provide insight into biomolecule structural stability in IL solvents. Several studies used MD simulations coupled with spectroscopic techniques (i.e. CD, UV-Vis spectroscopy) to probe the stabilization or destabilization mechanism of ILs on nucleic acids and proteins.
## Table 2

**Literature Review of ILs on Nucleic Acid Stability**

<table>
<thead>
<tr>
<th>DNA/RNA structure model (PDB if available)</th>
<th>IL</th>
<th>Method(s)</th>
<th>IL effect</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calf-thymus B-DNA (PDB:1BNA)</td>
<td>[Ch][A]</td>
<td>Spectroscopy</td>
<td>Stabilized DNA and prevented intercalation of EB and DAPI in DNA grooves</td>
<td>Sahoo et al. [28]</td>
</tr>
<tr>
<td></td>
<td>A = Gly, Ala, Pro</td>
<td>MD simulations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calf thymus DNA (PDB:425D)</td>
<td>[C][Br]</td>
<td>Spectroscopy</td>
<td>Stabilized DNA via dehydration</td>
<td>Jumbri et al. [29]</td>
</tr>
<tr>
<td></td>
<td>C = EMIM, BMIM, HMIM</td>
<td>MD simulations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drew Dickerson B-DNA (PDB:1BNA)</td>
<td>[C][A]</td>
<td>MD simulations</td>
<td>Stabilized DNA</td>
<td>Carodoso et al. [30]</td>
</tr>
<tr>
<td></td>
<td>C = BOXA, BMIM, BPYR, BPYRR, Ch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A = BF4, PF4</td>
<td>Spectroscopy</td>
<td></td>
<td>Chandran et al. [31]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MD simulations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drew-Dickerson B-DNA (N/A)</td>
<td>[BMIM][Cl]</td>
<td>Spectroscopy</td>
<td>Stabilized DNA</td>
<td></td>
</tr>
<tr>
<td>22mer human telomeric G-quadruplex DNA</td>
<td>[TMG][A]</td>
<td>Spectroscopy</td>
<td>Destabilized G4 DNA in more hydrophobic anions</td>
<td>Sarkar et al. [32]</td>
</tr>
<tr>
<td>(PDB:143D)</td>
<td>A = Ac, OC, DHP, TFAC, PFOC</td>
<td>MD simulations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24mer human telomeric G-quadruplex DNA</td>
<td>[Gua][(C2F5)3F3P]</td>
<td>Spectroscopy</td>
<td>Stabilized G4 DNA</td>
<td>Satpathi et al. [33]</td>
</tr>
<tr>
<td>(PDB:2GKU)</td>
<td></td>
<td>MD simulations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yeast tRNA (PDB: 1EHZ)</td>
<td>[Ch][A]</td>
<td>Spectroscopy</td>
<td>Stabilized tRNA and prevented intercalation of EB in DNA grooves</td>
<td>Tulsiyian et al. [34]</td>
</tr>
<tr>
<td></td>
<td>A = Asp, Glu</td>
<td>MD simulations</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. [C]: Variable cation. [A]: Variable anion.*
Figure 2

*Stabilization of DNA by HMIM-Br*

Note. From Jumbri et al. 29.
Table 3

*MD Simulations of ILs on Protein Stability*

<table>
<thead>
<tr>
<th>Protein structure (PDB if available)</th>
<th>IL</th>
<th>Method(s)</th>
<th>IL effect</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>mCherry RFP (PDB:2H5Q)</td>
<td>[TMG][A] A = Ala, Asp</td>
<td>Spectroscopy MD simulations</td>
<td>Destabilized protein (Ala more destabilizing than Asp)</td>
<td>Borrell et al. [35]</td>
</tr>
<tr>
<td>Trp mini-cage (PDB:1L2Y)</td>
<td>[EMIM][A] A = Ala, Met, Trp</td>
<td>MD Simulations</td>
<td>Stabilized protein</td>
<td>Chevrot et al. [36]</td>
</tr>
<tr>
<td>Human serum albumin protein (PDBs:4IW2, 1E7I)</td>
<td>[BMIM][BF4] [Ch][DHP]</td>
<td>MD Simulations</td>
<td>[BMIM][BF4] destabilized protein, [Ch][DHP] caused stable protein swelling</td>
<td>Jager and Pfaendtner [37]</td>
</tr>
<tr>
<td>T. maritima xylanase (PDB:1VBR)</td>
<td>[EMIM][Ac]</td>
<td>MD Simulations</td>
<td>Destabilized protein and decreased catalytic activity in high IL concentration</td>
<td>Manna and Ghosh [38]</td>
</tr>
<tr>
<td>B. amyloquefaciens alpha amylase (PDB:3BH4)</td>
<td>[HMIM][Cl]</td>
<td>MD Simulations</td>
<td>Induced protein structural changes and decreased catalytic activity</td>
<td>Mohammadyazdani et al. [128]</td>
</tr>
<tr>
<td>Chicken egg white lysozyme (PDB:1AKI)</td>
<td>[EMIM][A] A = EtSO4, Et2PO4</td>
<td>Spectroscopy MD simulations</td>
<td>Decreased protein Tm; [EMIM][Et2PO4] induced greater protein instability</td>
<td>Singh et al. [39]</td>
</tr>
<tr>
<td>Human insulin aspart (PDB:1ZEFH)</td>
<td>[Ch][A] A = Pro, Gly, Alan</td>
<td>MD Simulations</td>
<td>[Ch][Pro] stabilized protein and [Ch][Ala] and [Ch][Gly] destabilized protein</td>
<td>Sundaram et al. [40]</td>
</tr>
</tbody>
</table>

*Note.* [C]: Variable cation; [A]: Variable anion.
**Figure 3**

*Inhibition Mechanism of Xylanase by EMIM-Acetate*

*Note.* EMIM+ cations obstruct the catalytic site region (red and blue licorice) and reduce the rate of hydrolysis of xylan hemicellulose into xylose.

Additionally, using MD simulations can help bridge the gap between how different IL properties affect nucleic acid and protein structure. Once the IL property and its mechanism of structural effect is understood, that property can then be modified to enhance the desired structural effect.
1.2 Methodology

1.2.1 MD Simulations

MD simulation is a computational tool that is based Newton’s Law and the equations of motion\textsuperscript{27}. MD simulations integrate Newton’s laws of motion and the energy function to solve the force exerted on each atom and thus predict atomic motion over time ($F = ma$), where $F$ is the force exerted on each atom, $m$ is the atom’s mass and $a$ is the atom’s acceleration. Integrating acceleration ($a$), the atom’s velocity over time can be obtained. Subsequently, by integrating velocity, the atom’s position over time can be determined. MD repeatedly iterates these calculations to update an atom’s position every timestep (i.e., 2 femtoseconds). To represent the atomic forces, force fields are used. Force fields are a series of equations and constants that describe molecular parameters (i.e., molecular geometry, point charges) of a given molecule (i.e., drug ligand, ion, solvent, nucleic acid, protein) and are used to calculate the potential energy of different interactions occurring in a simulation system. Force field parameters include bond length, bond rotation, bond angle and dihedral angles, as well as Coulombic interactions including van der Waals and electrostatic interactions\textsuperscript{27} (Figure 4).
To represent the bulk solvent (i.e., water), an explicit or implicit solvent model is used depending on if two or more systems are significantly different (explicit model) or more similar to each other (implicit model). An explicit solvent model includes solvent molecules in the simulation system, providing greater accuracy at increased computational cost. An implicit solvent model treats solvent molecules as a continuous medium of a given dielectric constant, providing slightly less accuracy at decreased computational cost. The built simulation system is then run until equilibrium is reached. Throughout the simulation, snapshots are taken at intervals of time (i.e., 1 ns) and composed into a trajectory, containing the three-dimensional coordinates of all atoms involved in each snapshot. Each snapshot in the trajectory involves integrating Newton’s equations of motion as described above to predict the overall atomic motion. Additionally, different biological phenomena can be observed based on the length of simulation time that passes, ranging from simple bond length vibration (~1 fs) to complex membrane protein folding (~1000 s). Most atomic-level MD simulations are ~1
microsecond in length, which we can observe lipid diffusion and rapid protein folding. Thus, the current technology limits our ability to observe more complex biological phenomena using MD simulations.

Post-simulation analysis can be conducted to help settle our inquiries (i.e., is a nucleic acid stable in the simulated system?). Trajectory clustering abundance is used to identify major structural conformations. Each snapshot within the trajectory is aligned to the reference structure, where like-conformations are grouped together into populated cluster families based on their root mean-squared deviation (RMSD) of the backbone to the reference conformation. This technique can be used to identify common final structural conformations (at the end of MD) as well as preferred ligand binding modes.

1.3 Thesis Outline

In Chapter 2, the effect of four imidazolium-based ILs of increasing hydrophobicity (EMIM-Cl, BMIM-Cl, HMIM-Cl and OMIM-Cl) on the structure of a 22mer human c-MYC parallel-stranded DNA G-quadruplex (Pu22 G4) was determined using several 1 microsecond MD simulations. Additionally, the effect of each IL on the binding of an anticancer compound TMPyP4 to Pu22 was investigated. TMPyP4 is known to bind and stabilize Pu22 and decrease c-MYC oncogene expression. ILs have also been observed to modulate G4 structure stability, serving as a potential solution for stabilizing and storing nucleic acid molecules for biomedical applications. Yet, no high-resolution structure of TMPyP4-Pu22 complex has been solved. Furthermore, the effect of imidazolium-based ILs on G4 structure has not been studied. Additionally, the effect of IL on ligand binding to G4s is unknown. To attempt to answer these questions, simulation systems were built containing Pu22 +/- TMPyP4 in water (control) and Pu22
+/− TMPyP4 in IL (experiment). During post-simulation analysis, the control system showed that Pu22 was stable in water. Additionally, the major binding modes of TMPyP4 to Pu22 are top and bottom binding, based on trajectory clustering and confirmed with CD spectroscopy results. Two major findings were observed in the experiment systems. First, the Pu22 structure was stable in each IL, which was confirmed by negligible changes in Pu22 root mean-squared fluctuations (RMSF) and geometric parameters. Second, TMPyP4 binding to Pu22 was reduced mostly by OMIM-Cl, followed by HMIM-Cl, BMIM-Cl and EMIM-Cl ILs. By calculating the number of atom contacts between TMPyP4/IL and Pu22, the number of IL contacts increased as TMPyP4 contacts decreased in the order EMIM<BMIM<HMIM<OMIM. Further decomposition of IL-Pu22 contacts revealed hydrophobic contacts increasing in the same order, while polar contacts stayed consistent between each IL, suggesting TMPyP4 binding is likely reduced via increasing IL hydrophobicity.

In Chapter 3, the destabilization mechanism of two amino acid-based ILs (TMG-Asp and TMG-Ser) on the structure of the azurin protein was elucidated using several 1 microsecond MD simulations. Azurin works in-concert with other proteins in soil denitrification, making it attractive for biotechnological applications. Enhancing azurin’s activity could significantly reduce soil nitrate concentrations and reduce potential groundwater pollution. By quantifying IL-azurin protein interactions, the driving force for azurin destabilization can be identified to enhance our understanding in developing future TMG-AA ILs for augmenting protein activity. To determine if the ILs are capable of protein unfolding, conventional MD simulation systems containing folded azurin+water+IL at room temperature were carried out. From these, it was clear that
azurin was stable in water and in both ILs. Folded azurin was then denatured into unfolded azurin using replica exchange MD (REMD) simulations. Unfolded azurin systems were created under the same conditions as the folded systems and then ran. Unfolded azurin conformation did not significantly change in IL, with some refolding of native and non-native protein secondary structures being observed. To elucidate the main driver for stabilizing unfolded azurin conformation, the interactions between cation/anion and polar/hydrophobic amino acid side chains were calculated for both folded and unfolded systems. It was clear that IL anion-polar interactions dominate over IL anion-hydrophobic interactions, with Asp anions showing a significant increase in interactions with unfolded azurin than with folded azurin. Protein RMSF data suggests that unfolded azurin structure becomes more disordered in IL.
Chapter 2

To Probe the Binding of TMPyP4 to c-MYC G-Quadruplex in Water and in Imidazolium-Based IL Solvents using Spectroscopy Coupled with Molecular Dynamics Simulations

2.1 Introduction

2.1.1 G-Quadruplexes (G4s)

G4s are non-canonical nucleic acid secondary structures comprised of DNA or RNA. G4s are characterized by two or more guanine tetrads arranged in planar formation and are stabilized by Π-Π-stacking and Hoogsteen hydrogen bonding interactions and electrostatic coordination between monovalent cations (i.e. K⁺, Na⁺) and O6 guanine atoms within the G4 core. G4s can have more than one conformation based on the directionality of their 5’ and 3’ strand ends (parallel, antiparallel or hybrid) and their loop conformation (diagonal, propeller or lateral). G4s are known to form near telomeric sequences and within c-MYC, KRAS and BCL2 oncogene promoter regions, which are involved in cancer pathogenesis. Stabilizing cancer-associated G4s with small drug molecules demonstrated decreased cancer cell growth, thus these G4s are promising targets in cancer chemotherapy.

2.1.2 C-MYC Oncogene Pu22 G4

The human c-MYC oncogene regulates the transcription of genes which promote cell growth, proliferation and differentiation (i.e. hTERT, ODC, etc.). In fact, overexpression of c-MYC oncogene has been observed in various cancers, including
colon\textsuperscript{53}, B-cell lymphomas\textsuperscript{54,55}, prostate\textsuperscript{56}, breast\textsuperscript{57}, pancreatic\textsuperscript{47,58} and small lung cancers\textsuperscript{54,59}. Up to 90\% of the transcription processes of the c-MYC oncogene is regulated by the nuclear hypersensitivity element III\textsubscript{1} (NHE III\textsubscript{1}), located within the oncogene’s promoter region. G4s of varying sequence length, of which the 22mer c-MYC G4 (Pu22) is known to form within NHE III\textsubscript{1} and modulate transcription of the c-MYC oncogene\textsuperscript{44,60}. Numerous small molecules (ligands) were observed to stabilize Pu22 G4 and downregulate cancer proliferation, including quindolines, flavones, piperazines and perylenes\textsuperscript{61-67}. Thus, the Pu22 G4 is an attractive target for cancer chemotherapy. To elucidate the detailed interactions between ligands and G4, the NMR solution structures of Pu22 parallel G4 without ligand (\textit{Figure 5}, PDB id: 1XAV)\textsuperscript{68} and complexed with the anti-cancer ligand Phen DC-34 (PDB id: 5W77)\textsuperscript{69} have been solved in the last twenty years.
Figure 5

Pu22 G4, Ligand TMPyP4 and Four Imidazolium-Based ILs


2.1.3 G4 Stabilizer TMPyP4 and Previous Studies

TMPyP4 is a small, aromatic, planar macrocyclic ligand with a high intrinsic positive charge (+4). The TMPyP4 structure is characterized by a central porphyrin ring of four 1H-pyrrole rings and four positively-charged 1-methylpyridinium side chain groups. In terms of G4 drug design, 1) the TMPyP4 central porphyrin ring forms Π-Π stacking interactions with G4 layers and/or nucleotides in the 5’ or 3’ terminal loop regions; 2) the TMPyP4 side chains form electrostatic interactions with the negatively charged phosphate groups in the major and minor groove regions of the G4 backbone.
TMPyP4 is known to both stabilize and destabilize G4 structures in this manner $^{49,59,70}$. Han et al. observed that the melting temperature ($T_m$) of a 32mer oligonucleotide G4 increased upon binding to the ligands TMPyP4 and TMPyP2 ($T_m \approx 57^\circ C$) compared to G4 in 100 mM K$^+$ buffer solution alone ($T_m \approx 50^\circ C$), indicating that TMPyP4 and TMPyP2 similarly stabilized the G4 structure $^{49}$. However, another study reported that TMPyP4, but not TMPyP2, was able to downregulate c-MYC oncogene and hTERT1 gene and protein expression, thus downregulating telomerase catalysis, decreasing cancer cell proliferation $^{59,71}$ and decreasing tumor growth, suggesting that only TMPyP4 could bind and stabilize the c-MYC G4. TMPyP4 is also able to destabilize G4s $^{72}$. One study reported binding of TMPyP4 to the parallel RNA G4 containing r(GGGGCC) repeats in C9orf72 gene associated with neurodegenerative diseases (i.e., Lou Gehrig’s disease) $^{73}$; the melting temperature of the TMPyP4-G4 complex was lower ($T_m \approx 67^\circ C$) compared to the un-complexed G4 ($T_m > 95^\circ C$), indicating strand separation. TMPyP4 also acts as an antagonist for nucleic acid proteins. TMPyP4 blocked binding of pathogenic proteins ASF/SF2 and hnRNPA1 to the G4 in a dose-dependent manner $^{72}$. TMPyP4 is thus a promising ligand for studying G4 structure and probing interactions of other small molecules to G4 structures.

TMPyP4 has been observed to stabilize both c-MYC G4 and telomeric G4 and decrease expression of c-MYC oncogene and regulator genes, such as cytochrome p450, in HeLa cancer cells in-vivo $^{71,74-76}$. The high-resolution structure of TMPyP4 complexed with a human telomeric parallel G4 ($\textbf{Figure A1}$, PDB id: 2HRI) $^{77}$ revealed that TMPyP4 engages in non-covalent Π-Π stacking interactions with the top G4 layer and the 5’ loop region of telomeric G4 and stabilizes its structure. Interestingly, TMPyP4 is a highly
conjugated chromophore and its binding to nucleic acids such as G4s can be monitored using UV-Vis absorbance spectroscopy, which is indicated by a red shift in the G4 spectra caused by pi conjugation between the TMPyP4 porphyrin ring and nitrogenous bases of DNA or RNA. In the literature, TMPyP4 can act as a chemical probe to investigate the detailed binding interactions between G4s and other ligands. Yet, although significant advances have been made to elucidate the detailed molecular interactions between TMPyP4 and telomeric G4 and between Phen DC-34 and c-MYC Pu22 G4, a high-resolution complex structure of TMPyP4 with Pu22 G4 does not yet exist, thus its detailed interaction mechanism remains unknown.

### 2.1.4 Previous Studies of IL Effects on G4 Structure

The high degree of tunability of ILs have demonstrated that different combinations of cations and anions can produce different solvent environments that can modulate nucleic acid stability. Recently, two studies utilized MD simulations to elucidate the stabilizing and destabilizing effects of ILs on G4 structure, based on G4 sequence length, G4 structural conformation and IL anion type. In the first study, Satpathi et al. used MD simulations to elucidate the stabilizing mechanism of IL guanidinium tris(pentafluoroethyl)-trifluorophosphate (Gua-FEP) on the NMR solution structure of a 24mer human telomeric G4 (PDB id: 2GKU). Based on the combined action of both Gua$^+$ cation and FEP$^-$ anion, Gua$^+$ is small enough to enter the G4 core while FEP$^-$ interacts with the surface of the G4 DNA backbone. Specifically, Gua$^+$ likely coordinates its positive charge with O6 guanine atoms and thus lowers the net charge of G4, stabilizing its structural conformation. In the second study, Sarkar et al. used MD simulations to investigate the anion-dependent destabilizing mechanism of two ILs
containing the cation 1,2,3,3,-tetramethyl guanidinium (TMG⁺) and either the anions acetate (AC⁻) or octanoate (OC⁻) on the NMR solution structure of a 22mer human telomeric antiparallel DNA G4 (PDB id: 143D). Essentially, while TMG⁺ cations interact with the DNA phosphate backbone of 143D, AC⁻ and OC⁻ anions congregate at the G4 layers and decrease Hoogsteen hydrogen bonding and Π-Π stacking interactions between G4 layers, disrupting a critical stabilizing feature of G4s. During our study, the effects of ILs on ligand binding to G4s had not been investigated. Additionally, the effect of ILs on c-MYC DNA G4s has not yet been probed (Figure 5). Lastly, only the anion-dependent effects on G4 structural stability have been investigated, thus the cation-dependent effects on G4 stability have yet to be uncovered.

To date, no studies have probed the effects of imidazolium-based ILs on G4 structure. However, EMIM-Cl and BMIM-Cl demonstrated preservation of duplex B-DNA structure. Chandran et al. used MD simulations to probe the dehydration of a B-DNA dodecamer duplex by BMIM-Cl. Interestingly, BMIM⁺ cations invaded the minor groove region of DNA and formed hydrophobic and electrostatic interactions with the DNA structure, while simultaneously disrupting the first hydration shell surrounding the DNA structure. BMIM⁺ formed a hydrophobic barrier around the DNA structure and reduced hydrolysis of the DNA sugar-phosphate backbone via the hydrophobic effect, preserving the DNA structure. Following this logic, we speculated that increasing the hydrophobicity of the imidazolium cation (i.e. HMIM-Cl and OMIM-Cl) could self-aggregate at the G4 binding interface to preserve its structure and likely reduce binding of small molecule ligands to G4s.
2.1.5 Experimental Overview

Coupling spectroscopy and MD simulations has provided insight into the detailed molecular interactions of ILs to G4s $^{32,33}$. In this study, we combined spectroscopy and MD simulation methods to investigate 1). The detailed binding interactions between TMPyP4 and the NMR solution structure of c-MYC Pu22 parallel DNA G4 (PDB id: 1XAV) in water solvent only; 2). the cation-dependent hydrophobic effects of four imidazolium-based ILs of increasing hydrophobic chain length (EMIM-Cl, BMIM-Cl, HMIM-Cl and OMIM-Cl) to the Pu22 G4 structure and their effects on binding of TMPyP4 to Pu22. Indeed, CD and UV-Vis absorbance spectroscopy showed a 2:1 binding ratio of TMPyP4 to Pu22 in water solvent, with MD simulations suggesting TMPyP4 binds mainly to the top and bottom G4 layers of Pu22. Both results are consistent with literature findings $^{72,86}$. CD spectroscopy suggests Pu22 G4 is stable in the four ILs. Interestingly, each IL exhibited reduced binding of TMPyP4 to Pu22 G4 with increasing hydrophobic chain length. MD simulations elucidated the mechanism of reduced TMPyP4 binding to Pu22 G4 is driven by the hydrophobic effect of ILs enacted on G4 structure.

2.2 Methodology

2.2.1 Simulation System Protocols

The NMR solution structure of Pu22 c-MYC parallel G4 without ligand (PDB id: 1XAV) $^{68}$ was downloaded from the Protein Data Bank (PDB). We set up twelve simulation systems (Table A2) using a three-step approach: 1). The solute is represented by Pu22 in the unbound state or one TMPyP4 molecule complexed with Pu22. TMPyP4
ligand was set in one of two initial starting positions, being placed above the 5’ terminus
and beneath the 3’ terminus of G4 using a displacement distance of ~15 Å (See Figure
A2); 2). Neutralization of the solute net charge by counter ions; 3). Solvation of the solute
in a water box of truncated octahedron with a 10 Å separation distance for water
molecules and a concentration of 0.1 M potassium chloride (KCl) or 0.5 M IL salt. The
Pu22 DNA structure was represented by the OL15 version of AMBER nucleic acid force
field containing several corrections for torsion angle parameters (parm99bsc0 + χ_{OL4} +
\epsilon/\zeta_{OL1} + \beta_{OL1})^{87-90}. Water molecules were represented by the TIP3P water model force
field. Ions were represented by the Cheatham group K+ force field model^{91,92}. The force
field for TMPyP4 has already been developed in our previous study^{93}. All other
parameters for EMIM, BMIM, HMIM and OMIM cations were determined by the
GAFF2 force field^{94,95}. Between 2 and 15 independent runs of 1000 ns were run for each
simulation system.

2.2.2 MD Simulation Protocols

AMBER16 software simulation package was used to run each simulation system.
The protocol has been covered in-detail in our previous studies^{96-98} and is explained
briefly here. First, all unbound DNA-ligand systems underwent a pre-run of 1000 ps each
at 500 K to randomize the ligand position and orientation before undergoing a production
run at 300 K, while fixing the G4 receptor position. A total of 74 independent runs were
performed (thirty-four independent runs for the G4 system in water and 10 independent
runs each for G4 in EMIM-Cl, BMIM-Cl, HMIM-Cl and OMIM-Cl) each at 300K with
random initial velocities. Each independent run at 300 K was ran with a 1.0 ns MD
simulation included under the NPT ensemble (constant temperature and pressure) for
equilibrating the system density and production dynamics for the succeeding NVT ensemble mode (constant volume and temperature). SHAKE algorithm was used for constraining all hydrogen-containing bonds with an enabled 2.0 fs time step in these simulations. The Particle-mesh Ewald method was used for treating long-range electrostatic interactions under periodic boundary conditions (with a charge grid spacing of distance ~1.0 Å, the fourth-order of the B-spline charge interpolation and direct sum tolerance of $10^{-5}$). Short-range electrostatic interactions were described using a cut-off distance of 10 Å and long-range van der Waals interactions were based off the uniform density approximation. Non-bonded forces were calculated using a two-stage RESPA approach to save computational resources, where short-range and long-range forces were updated every one and two steps, respectively. The temperature was maintained with the Langevin thermostat using a coupling constant of 2.0 ps. All trajectories were saved at 50.0 ps intervals for post-simulation analysis.

2.2.3 Simulation Convergence

To validate if each simulation system had attained steady-state equilibrium, the root mean-squared deviation (RMSD) was calculated for all heavy atoms in the DNA backbone of Pu22 G4 (receptor RMSD, R-RMSD) and TMPyP4 ligand (ligand RMSD, L-RMSD) against their initial starting structures in each trajectory.

2.2.4 Characterization of Structural Fluctuations

To determine the effects of IL interactions with G4 structure, the root mean-squared fluctuation (RMSF) was calculated for each nucleotide in Pu22 G4 and averaged over the total length of each trajectory in water and each IL solvent.
2.2.5 K+ and O6 Positions

The relative distance between K+ cations and O6-guanine atoms in G4 layers is an important parameter for assessing G4 structure stability. The coordination between partially positive K+ atoms and partially negative O6-guanine atoms in G4 layers stabilizes G4 structure in two ways: 1). Reduces the negative charges in G4s; 2). Reduces electrostatic repulsion between adjacent O6 guanine atoms within the G4 core. The K+ - O6-guanine distances can be perturbed by binding interactions of small molecule ligands and solvent, which could displace K+ positions and alter G4 structure and stability. Thus, for Pu22 G4 alone (unbound) and Pu22 G4 complexed with TMPyP4 (top, bottom and side binding positions of TMPyP4), we computed 1). K+ - O6 guanine atom distances in the top, middle and bottom G4 layers and 2). Inter- and intra-layer O6-O6 guanine atom distances in water and each IL solvent.

2.2.6 Geometric Parameter Calculations

The rise, H-rise and H-twist geometric parameters were calculated between each adjacent G4 layer of Pu22 (i.e., top layer-middle layer, middle layer-bottom layer) to determine the inter-G4 layer geometry. The Rise parameter characterizes the distance (in Å) between guanine base centers between a lower G4 layer and the G4 layer directly above it. The H-rise parameter characterizes the projection of the rise of the G4 layer relative to the vertical axis (i.e., Z-axis) generated by the two K+ cations in the Pu22 G4 core. The H-twist parameter characterizes the rotational angle for each nitrogenous base with respect to the vertical axis. H-twist calculates the amount of clockwise rotation needed for the bottom of two G4 layers to align with the G4 layer above it, with the top G4 layer serving as a reference.
2.2.7 Identification of Ligand Binding Modes

After confirming the stability of Pu22 G4 in water and IL solvents, clustering analysis was done to determine the major binding modes of TMPyP4 to Pu22 G4. Visual inspection and confirmation was done for each cluster and were assigned to top, bottom or side binding. The DNA backbone of stable Pu22-TMPyP4 complexes was aligned from each trajectory using the least square fitting method. The Daura algorithm was used to categorize each aligned Pu22-TMPyP4 complex into different cluster families, with a 2 Å pairwise RMSD cutoff for TMPyP4 only to calculate the number of neighboring structures for each structure. Structure families were constructed by combining the structure with the largest number of neighboring structures (i.e., the centroid structure, representing each family) together with its neighbors. This was done for each structure in water and each IL solvent. Families were assigned a binding mode by visual inspection and were combined into superfamilies of the same binding mode (top, bottom and side binding).

2.2.8 Calculating Atom Contacts between TMPyP4/IL and Pu22

The total atomic contacts between TMPyP4/IL and Pu22 G4 were calculated using a 2.5 Å distance cutoff. Total IL-Pu22 G4 contacts were decomposed into apolar IL-Pu22 G4 and polar IL-Pu22 G4 contacts to determine the IL-induced mechanism of reduced TMPyP4 binding to Pu22 G4. The nonpolar alkyl side chain of each IL represents apolar IL and the polar imidazolium ring of each IL represents polar IL.
2.2.9 **Representative Ligand Binding Trajectory**

The ligand binding trajectory was calculated and generated for representative top, bottom and side binding mode trajectories of TMPyP4 to Pu22 in water and each IL solvent.

2.2.10 **Structural Alignment**

To elucidate the impact of G4 structural conformation on TMPyP4 binding, superimposition of the DNA backbones of Pu22 NMR solution structures in the unbound (PDB id: 1XAV) and bound states (PDB id: 5W77) was performed using the RMSD calculator tool of VMD 1.9.3 program.

2.2.11 **Ligand Binding Analysis**

To gain insight into the detailed binding interactions between TMPyP4 and Pu22 G4, we generated ligand interaction diagrams of TMPyP4 to the representative top and bottom binding structures of Pu22 in water solvent using the 2-D ligand interaction diagram program of Maestro 10.7 software from the Schrodinger 2016 package software suite.

2.2.12 **Structural Visualization**

All structure figures were generated using VMD 1.9.3 program.

2.3 **Results**

2.3.1 **Pu22 Structure is Stable in Water and IL Solvents**

To check if each system reached convergence, receptor RMSD of Pu22 G4 (R-RMSD) and ligand RMSD for TMPyP4 (L-RMSD) was calculated and averaged over all trajectories in water and each IL solvent. Plots were generated for average R-RMSD and
L-RMSD values (Figures A3) and for each trajectory for R-RMSD (Figures A10a-A10e) and L-RMSD (Figures A11a-A11e). The average values and their standard deviations were tabulated (Table A3). Indeed, R-RMSD convergence was achieved within 50 ns in each simulation system for water and each IL, with the average R-RMSD being consistent between water and IL systems (2.8±0.1 to 2.9±0.2 Å). This suggests that Pu22 G4 was stable in the water and IL solvents. Likewise, L-RMSD was achieved within 50 ns for each trajectory, despite the average values not being consistent between water and each IL solvent (20.6±1.1 to 35.7±3.4 Å). This suggests TMPyP4 may assume different binding poses to Pu22 G4.

To investigate the structural fluctuations of Pu22 in water and each IL, the RMSF of each nucleotide in Pu22 (R-RMSF) was calculated as described in the methods section. An average RMSF plot was generated (Figure 6) and the average values and standard deviations were tabulated (Table A3).
Figure 6

*Pu22 G4 RMSF*

![Graph showing RMSF values for different bases and solvents.](image)

**Note.** Averaged Pu22 RMSF in water (30 trajectories) and in each IL (10 trajectories per IL, 40 trajectories total). Non-G-tetrad residues are highlighted in yellow. See Table A3 for average and standard deviation values.

Indeed, the average R-RMSF values are consistent between the water and IL systems (1.5±1.0 to 2.1±1.7 Å), despite relatively larger fluctuations being observed in the first three nucleotides starting from the 5’ end. This suggests the Pu22 G4 structure exhibits minimal fluctuations in water and IL solvents.

To more rigorously characterize Pu22 structural stability, the geometric parameters Rise, H-Rise, and H-Twist were calculated for the unbound and bound structures (TMPyP4 to top, bottom and side binding modes) of Pu22 in water and each IL.
as described in the methods section. Average and standard deviation values have been tabulated \( (Table\ A4) \). Not surprisingly, the geometric parameter values are highly consistent for Pu22 G4 in water and each IL. The \( K^+ - O6 \) guanine distance and \( O6 - O6 \) guanine atom distance values were calculated and tabulated for each system and the average and standard deviation values have been tabulated \( (Tables\ A5a-A5e\ and\ A6a-A6e) \). Additionally, the average \( K^+ - O6 \) guanine distances for Pu22 in water are shown graphically \( (Figure\ 7) \). Indeed, the values are consistent for Pu22 in water and each IL. From the RMSD, RMSF and geometric parameter calculations, we conclude that the Pu22 G4 is stable in water and each of the four ILs.
Note. (A): Oxygen is represented by a colored ball and K$^+$ cations are represented by a yellow ball; Side view of the three G4 layers (B); Distance (Å) of the oxygen from each residue to the nearest K$^+$ cations (C); Distance (Å) of each oxygen relative to the nearest side (D).

2.3.2 Hydrophobic Side Chains of HMIM-Cl and OMIM-Cl Reduces TMPyP4 Top and Bottom Binding to Pu22

From spectroscopy analysis (data not shown), ILs with longer alkyl chain lengths (HMIM-Cl and OMIM-Cl) greatly reduce TMPyP4 binding to Pu22 G4. Distribution
diagrams of TMPyP4 ligand and IL to Pu22 G4 over all trajectories were generated for each system (Figure 8).

**Figure 8**

*TMPyP4 and IL Distribution Around Pu22*

<table>
<thead>
<tr>
<th></th>
<th>SIDE</th>
<th>TOP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WAT</strong></td>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
</tr>
<tr>
<td><strong>EMIM</strong></td>
<td><img src="image3" alt="Diagram" /></td>
<td><img src="image4" alt="Diagram" /></td>
</tr>
<tr>
<td><strong>BMIM</strong></td>
<td><img src="image5" alt="Diagram" /></td>
<td><img src="image6" alt="Diagram" /></td>
</tr>
<tr>
<td><strong>HMIM</strong></td>
<td><img src="image7" alt="Diagram" /></td>
<td><img src="image8" alt="Diagram" /></td>
</tr>
</tbody>
</table>
Note. An atom distance cutoff of 2.5Å was used to define atom contacts with Pu22. TMPyP4, IL and Cl\textsuperscript{-} are represented as blue, red and green balls, respectively. Atom contacts are shown every 20 ns.

The distribution of TMPyP4 and IL was also mapped to an electrostatic surface potential representation of Pu22 to elucidate the detailed binding mechanism of IL to Pu22 G4 (Figure A7). To investigate the detailed inhibition mechanism of TMPyP4 binding the top G4 layer (its primary major binding mode), we exhibit the distribution of each IL to the top G4 layer of Pu22 (Figure 9).
Figure 9

**TMPyP4 Inhibition Mechanism by IL Cations**

<table>
<thead>
<tr>
<th></th>
<th>SIDE</th>
<th>TOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMIM</td>
<td><img src="EMIM.png" alt="Image" /></td>
<td><img src="EMIM_TOP.png" alt="Image" /></td>
</tr>
<tr>
<td>BMIM</td>
<td><img src="BMIM.png" alt="Image" /></td>
<td><img src="BMIM_TOP.png" alt="Image" /></td>
</tr>
<tr>
<td>HMIM</td>
<td><img src="HMIM.png" alt="Image" /></td>
<td><img src="HMIM_TOP.png" alt="Image" /></td>
</tr>
<tr>
<td>OMIM</td>
<td><img src="OMIM.png" alt="Image" /></td>
<td><img src="OMIM_TOP.png" alt="Image" /></td>
</tr>
</tbody>
</table>

*Note.* IL cation distribution at G13 in the top G4 layer of the representative top-binding mode. Carbon and nitrogen atoms in IL are colored green and blue. Atom distance cutoff is 2.5 Å. IL contacts are shown every 20 ns.
The distribution of TMPyP4 ligand seemed to vary significantly between each system (Figure 8). In water, TMPyP4 clearly binds at the top, bottom and side binding sites on Pu22 G4. When Pu22 is exposed to each IL, however, TMPyP4 clustering density decreases at each binding site in order of increasing IL hydrophobic chain length. Simultaneously, an increase in IL hydrophobic clustering was observed at the top and bottom G4 layers, the most obvious case being in OMIM-Cl. Specifically, TMPyP4 clustering is dense above the top G4 layer and between the 5’ loop. In IL, TMPyP4 clustering at the top G4 layer is most dense in EMIM-Cl and decreases further in order of BMIM-Cl, HMIM-Cl and OMIM-Cl, with OMIM-Cl showing almost no TMPyP4 binding. In OMIM-Cl, TMPyP4 exhibits almost no binding to the top and bottom G4 layers and is mostly clustered at the side binding sites of Pu22, highlighting the antagonistic effect of OMIM-Cl. This antagonistic mechanism is likely due to the OMIM$^+$ cations, as the Cl$^-$ anions don’t show a clustering pattern between the four ILs. For instance, Cl$^-$ interacts with the top G4 layer in EMIM-Cl, BMIM-Cl and HMIM-Cl, but not OMIM-Cl, which instead clusters at the 5’ loop region.

From these findings, each IL likely reduces TMPyP4 binding to Pu22 by engaging in both electrostatic and hydrophobic interactions with Pu22 (Figure 9). First, electrostatic interactions occur between the polar IL imidazole ring (the positively-charged nitrogen atom) and the DNA backbone of Pu22 (negatively-charged phosphate group); each IL was observed to do this. Second, hydrophobic interactions occur between the IL alkyl chain and the conjugated nitrogenous bases of nucleotides in the top G4 layer of Pu22; hydrophobic interactions become more significant in HMIM-Cl and OMIM-Cl. Interestingly, a cluster of OMIM-Cl molecules were observed to form over the top G4
layer, strengthening hydrophobic interactions between individual OMIM-Cl molecules and with the top G4 layer. This may suggest the formation of a hydrophobic barrier, to which TMPyP4 would be unable to access the top G4 layer. The hydrophobic effect is not as significant in the other three ILs, as the length of their alkyl chains are likely not long enough to cover the G4 layer, instead engaging at the DNA backbone.

2.3.3 TMPyP4 Binding Decreases as IL Hydrophobicity Increases

The total number of atom contacts between TMPyP4 ligand and Pu22 G4 in water and each IL were calculated as outlined in the methods section (Figure A4). This was to confirm the decrease in the fraction of TMPyP4 binding to Pu22 G4 in each IL, especially in HMIM-Cl and OMIM-Cl. Indeed, a decrease in TMPyP4 contacts averaged over all the trajectories was observed for Pu22 in ILs with increasing hydrophobicity, thus it’s likely that the IL hydrophobic effect may play some role in reducing TMPyP4 binding (Table A7). As cation IL hydrophobicity increases, the percentage of unbound Pu22 G4 cluster abundance increases (WAT: 21.8%; EMIM: 45.9%; BMIM: 56.1%; HMIM: 70.1% and OMIM: 82.0%) and the percent bound cluster abundance decreases (WAT: 78.2%; EMIM: 54.1%; BMIM: 43.9%; HMIM: 29.9% and OMIM: 18.0%) (Table 4), suggesting that ILs probably play a role in antagonizing TMPyP4 binding to Pu22.
Table 4

Percent Binding Modes of TMPyP4 to Pu22

<table>
<thead>
<tr>
<th>Solvent</th>
<th>%Unbound</th>
<th>%Bound</th>
<th>%Top</th>
<th>%Bottom</th>
<th>%Side</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAT</td>
<td>21.8</td>
<td>78.2</td>
<td>48.0</td>
<td>23.3</td>
<td>6.5</td>
</tr>
<tr>
<td>EMIM</td>
<td>45.9</td>
<td>54.1</td>
<td>31.4</td>
<td>10.7</td>
<td>9.7</td>
</tr>
<tr>
<td>BMIM</td>
<td>56.1</td>
<td>43.9</td>
<td>16.8</td>
<td>14.6</td>
<td>7.7</td>
</tr>
<tr>
<td>HMIM</td>
<td>70.1</td>
<td>29.9</td>
<td>23.3</td>
<td>0.0</td>
<td>5.3</td>
</tr>
<tr>
<td>OMIM</td>
<td>82.0</td>
<td>18.0</td>
<td>8.7</td>
<td>3.8</td>
<td>2.9</td>
</tr>
</tbody>
</table>

2.3.4 Top and Bottom Binding are the Major Binding Modes of TMPyP4

To determine the major binding modes of TMPyP4 to Pu22 G4, clustering analysis was performed for each system in water and the four ILs. We expect TMPyP4 to have two major binding modes to Pu22 G4 (top and bottom G4 layers), based on CD spectroscopy analysis (data not shown). Indeed, we observed top and bottom binding of TMPyP4 to Pu22 in water and most of the ILs, as well as a side binding transition state. Cluster abundance revealed TMPyP4 binding to Pu22 in the order top>bottom>side binding, with top binding being the most abundant binding mode (Table 4). Not surprisingly, top and bottom binding cluster abundances decreased in ILs of increasing hydrophobicity.

The percent abundance of each TMPyP4 binding mode is tabulated (Table 4) and the representative structures of the most abundant clusters are shown for each system (Figure A8). With Pu22 in water serving as a reference, Pu22 in IL exhibited decreased top binding (WAT: 48.0%; EMIM: 31.4%; BMIM: 16.8%; HMIM: 23.3% and OMIM:
8.7%), bottom binding (WAT: 23.3%; EMIM: 10.7%; BMIM: 14.6%; HMIM: 0% and OMIM: 3.8%). and mostly side binding (WAT: 6.5%; EMIM: 9.7%; BMIM: 7.7%; HMIM: 5.3% and OMIM: 2.9%). Therefore, TMPyP4 primarily engages in binding to the top and bottom G4 layers of Pu22. Based on its low percent cluster abundance, the side-binding mode is likely a transition state prior to top or bottom binding.

2.3.5 Binding Mode of TMPyP4 to Pu22 Shows Lack of Intercalation at Top and Bottom G4 Layers

We probed the detailed binding interactions between TMPyP4 and the top and bottom binding sites of Pu22 G4 (Figure A14). Based on the extensive conjugation in the porphyrin ring (four 1H-pyrrole rings) and the intrinsic positive charge in the four side chains (four N-methyl-4-pyridyl groups; 1+ per side chain, +4 total), we expect TMPyP4 to engage in both binding modes to Pu22 via a mixture of hydrophobic and electrostatic interactions. Hydrophobic interactions likely comprise Π-Π stacking interactions between the porphyrin ring of TMPyP4 and the DNA nitrogenous bases in Pu22. Simultaneously, electrostatic interactions likely comprise positively-charged nitrogen atoms in the TMPyP4 side chains and negatively-charged oxygen atoms in the phosphate groups of the DNA backbone.

For top-binding, 1H-pyrrole groups of TMPyP4 were found to engage in Π-Π stacking interactions with the DNA nucleotides in the top G4 layer of Pu22. However, the porphyrin ring seems to be strained and nonplanar, suggesting these interactions are not fully optimized. Additionally, two of the 1H-pyrrole rings seem to be engaged in Π-Π stacking interactions with nucleotide A3 in the 5’ stem loop. The TMPyP4 N-methyl-4-pyridyl side groups have various interactions within the top G4 layer, which includes
both hydrophobic interactions (\(\Pi-\Pi\) stacking with nucleotide G13) and electrostatic interactions (phosphate groups in the groove region of DNA). Surprisingly in bottom binding, no \(\Pi-\Pi\) stacking interactions were observed between TMPyP4 and the bottom G4 layer, as TMPyP4 was unable to intercalate here.

### 2.3.6 Atom Contact Plots Reveal Increased IL Binding to Pu22 G4 in ILs as Hydrophobic Chain Length Increases

We generated atom contact plots of each IL to Pu22 averaged over all trajectories (Figure A9) and for each trajectory (Figures A12a-A12d). The average values were also tabulated (Table 5).

#### Table 5

<table>
<thead>
<tr>
<th>Atom Contacts Between IL and Pu22</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>EMIM</td>
</tr>
<tr>
<td>BMIM</td>
</tr>
<tr>
<td>HMIM</td>
</tr>
<tr>
<td>OMIM</td>
</tr>
</tbody>
</table>

*Note. Total = Polar IL + Apolar IL contacts to Pu22 G-quadruplex. Atom contacts were averaged across each trajectory. See Figure A9 for graphs.*

The total IL atom contacts were further decomposed into IL apolar-Pu22 contacts and IL polar-Pu22 contacts to determine the driving force for IL binding to Pu22 G4. For the total atom contacts, ILs of increasing hydrophobicity have greater interactions with Pu22, with OMIM-Cl having the highest average atom contacts (33±1), followed by HMIM-Cl (23±1), BMIM-Cl (18±1) and EMIM-Cl (14±1). For apolar and polar atom
contacts, IL polar-Pu22 interactions dominate in ILs with shorter hydrophobic chain length (EMIM-Cl and BMIM-Cl) and IL-apolar interactions dominate in ILs with longer hydrophobic chains (OMIM-Cl). For EMIM-Cl, IL-polar (9.04±0.5) dominates over IL-apolar (5.44±0.4). For BMIM-Cl, IL-polar (10.14±0.6) dominates over IL-apolar (7.82±0.7). For HMIM-Cl, no significant difference was observed between IL-polar (11.34±0.6) and IL-apolar (11.46±0.9). For OMIM-Cl, IL-apolar (19.11±0.9) dominates over IL-polar (13.44±0.6). It’s thus probable that IL-apolar interactions play a significant role in reducing TMPyP4 binding to Pu22.

2.4 Discussion

Pu22 is a highly thermostable DNA G4 that is the dominant of four conformers (containing 22, 23, 24 or 25 nucleotides) within the c-MYC silencing region. Stabilization of Pu22 G4 with TMPyP4 elicited a decrease in the expression of the c-MYC oncogene, demonstrating that Pu22 is a good drug target for cancer therapy and TMPyP4 a good chemotherapeutic candidate. MD simulations in our study (1XAV G4 and TMPyP4 ligand) confirmed that TMPyP4 engages primarily in the top and bottom binding modes to Pu22. These results are consistent with those in another study by Calabrese et al. (5W77 G4 with Phen DC-34 ligand) and in our previous study (5W77 G4 with DC-34 ligand). However, II-II stacking interactions between TMPyP4 and the top and bottom G4 layers were likely not optimal. We did observe subtle differences between this study and our previous study. Upon superimposing the representative 5W77-DC-34 top and bottom binding structures with our representative 1XAV-TMPyP4 top and bottom binding structures, the resulting RMSD of the DNA backbones was ~3Å, indicating they show good alignment.
Figure 10

Effects of Ligand Geometry and G4 Topology on Ligand Binding

<table>
<thead>
<tr>
<th></th>
<th>MD</th>
<th>5W77</th>
<th>SUPERIMPOSED</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SIDE</strong></td>
<td><img src="image1" alt="MD Side" /></td>
<td><img src="image2" alt="5W77 Side" /></td>
<td><img src="image3" alt="Superimposed Side" /></td>
</tr>
<tr>
<td><strong>TOP</strong></td>
<td><img src="image4" alt="MD Top" /></td>
<td><img src="image5" alt="5W77 Top" /></td>
<td><img src="image6" alt="Superimposed Top" /></td>
</tr>
<tr>
<td><strong>SIDE</strong></td>
<td><img src="image7" alt="MD Side" /></td>
<td><img src="image8" alt="5W77 Side" /></td>
<td><img src="image9" alt="Superimposed Side" /></td>
</tr>
<tr>
<td><strong>BOTTOM</strong></td>
<td><img src="image10" alt="MD Bottom" /></td>
<td><img src="image11" alt="5W77 Bottom" /></td>
<td><img src="image12" alt="Superimposed Bottom" /></td>
</tr>
</tbody>
</table>

*Note.* Using the last snapshot structures of Pu22 with TMPyP4 (PDB id: 1XAV) and DBD1 (PDB id: 5W77). TMPyP4 (green and orange for top and bottom binding) and DBD1 (cyan and pink for top and bottom binding) are shown. Superimposed 1XAV and...
5W77 are shown in red and blue, respectively.

Yet, the ligands did not show good alignment. The structure of DC-34 is clearly smaller than TMPyP4, allowing it to lay almost planar with the top and bottom G4 layers. In comparison, TMPyP4 lays only partly planar to the top G4 layer and cannot even access the bottom G4 layer. This is likely due to the bulky chemical structure of TMPyP4, thus affecting its binding mode and geometry to Pu22.

G4 structure conformation can also be responsible for TMPyP4 binding mode and its geometry. In the above example, 1XAV has an overhanging 5’ and 3’ loop which restricts access to the top and bottom G4 layers, respectively; the overhanging loops combined with the bulkiness of TMPyP4 likely explain its observed binding modes. If the 5’ and 3’ loops were projected away from the top and bottom G4 layers, TMPyP4 may be able to lay planar with these layers. In another study, the X-ray crystal structure of a 22mer human telomeric G4 was solved with two TMPyP4 molecules bound to the top and side regions (PDB id: 2HRI) (Figure A1). 2HRI shows some similarities and differences to our MD simulation G4 complex structure. 1XAV and 2HRI are similar in that they both have 22 NT G4 sequences. However, 1XAV is a monomeric c-MYC-associated single-stranded G4 containing 2 K+ cations, whereas 2HRI is a monomer of a dimeric, double-stranded human telomeric-associated G4 containing 3 K+ cations. Of greater importance is the difference in G4 conformation; 1XAV has an overhanging 5’ loop and an underhanging 3’loop as mentioned before, which partially restricts access to the top and bottom G4 layers; in 2HRI, the 5’ loop restricts access to the top G4 layer and
the 3’ loop and a bottom-binding G4 monomer restricts access to the bottom G4 layer. 2HRI also exhibits a 2:1 binding ratio of TMPyP4 to the top and side binding sites.

TMPyP4 was used as a molecular probe to determine the detailed binding mode for each IL to Pu22 G4. The IL binding mode was driven primarily by hydrophobic interactions as well as electrostatic interactions and Pu22 structure was stable in each IL. From previous findings in the literature, it’s possible that the hydrophobic effect induced by each IL stabilized the Pu22 structure. We used MD simulations to investigate the detailed interactions of the four ILs with Pu22. The two studies mentioned earlier also used MD simulations to determine the detailed interactions between other ILs and G4s, which modulated their structural stability. In the first study, Sarkar et al. revealed that octanoate anions from the TMG-octanoate IL engage with the G4 layer of G-22 (PDB id: 143D) and reduce both Hoogsteen hydrogen bonding and inter-G4 layer Π-Π stacking interactions, destabilizing G-22. In our study, while HMIM-Cl and OMIM-Cl engaged in hydrophobic interactions with the top G4 layer of Pu22, we did not observe destabilization of the Pu22 structure. In the second study, Satpathi et al. revealed that from the Gua-FEP IL, the Gua+ cation is small enough to enter the core of G-24 (PDB id: 2GKU) whereas the FEP- anion binds to the DNA surface. More importantly, the Gua+ cations likely coordinate its positive charge with the partially-negative charges of the O6-guanine atoms and decrease the net negative charge inside the G4 core, increasing its stability. However, the imidazolium cations in this study are too large to enter the G4 core of Pu22 G4, instead binding to the top G4 layer. The Cl- anion doesn’t seem to play a major role in modulating Pu22 structure stability (Figure 8); Cl- experiences negative charge-charge repulsion from the phosphates in the DNA backbone, thus it can only
interact with G4 layer guanines. These interactions are unlikely to occur in the presence of HMIM$^+$ and OMIM$^+$ cations. The results here reinforce the idea that IL tunability can produce different effects on biomolecular structure.

Interestingly, the experimental results that formed the basis of our study suggest that the four ILs formed micelles around TMPyP4 and reduced binding to Pu22 G4 in this manner. However, we did not observe micelle formation by any of the ILs in our MD simulations. This is likely because micelle formation is a complex and long process, with which our MD simulations were unlikely to capture during the 1000 ns runs. Thus, extending the MD simulation time may allow for micelle formation.

The hydrophobic effect is likely the main driver for the binding interactions between each IL and Pu22 G4 and is a reported theme in IL binding to nucleic acids. Ding et al. used experimental methods alone to deduce the driving force of BMIM-Cl binding to calf thymus DNA (ctDNA) while antagonizing the binding of the chromophore ethidium bromide (EtBr), an intercalator of DNA. Using fluorescence spectroscopy, when EtBr binds to nucleic acids such as ctDNA, it exhibits a signature fluorescence signal as it interacts with the nitrogenous bases of DNA via Π-Π stacking interactions; a reduced fluorescence signal was observed in ctDNA with EtBr in BMIM-Cl solvent compared to ctDNA with EtBr in buffer alone, suggesting BMIM-Cl antagonized EtBr binding. $^3$P NMR and IR spectroscopy showed strong hydrophobic interactions between the BMIM$^+$ alkyl chain and nitrogenous bases and electrostatic interactions between the BMIM$^+$ polar region and the DNA backbone, with isothermal titration microcalorimetry confirming the hydrophobic effect is the primary driver of
BMIM-Cl binding to ctDNA. These results are our reasoning for using imidazolium cations of increasing hydrophobicity to antagonize TMPyP4 binding to Pu22 G4.

Intriguingly, imidazolium-based ILs may provide a good alternative for stabilizing nucleic acids such as the SARS-CoV-2 mRNA vaccine. The vaccine is a single-stranded mRNA and is not stable under room temperature conditions, caused by self-hydrolysis of its RNA sugar-phosphate backbone\textsuperscript{104}. Storing the vaccine under extremely cold conditions (-20 to -80°C) would slow this reaction down, however cold chain storage facilities are costly to efficiently maintain and are thus not accessible to most countries\textsuperscript{105}. While lyophilization of the vaccine has been proposed\textsuperscript{106,107}, the freeze-drying process in-particular is also costly and time-consuming. BMIM-cation-mediated dehydration of B-DNA, which was observed by Chandran et al.\textsuperscript{31} to stabilize its structure, may provide a clue for reducing phosphodiester bond hydrolysis of mRNA. Testing this potential solution would make for an interesting future study.

2.5 Conclusion

From this study, MD simulations were used to 1). Elucidate the detailed binding interactions between TMPyP4 and Pu22 G4 at the molecular level; 2). Investigate the cation-dependent effects of EMIM-Cl, BMIM-Cl, HMIM-Cl and OMIM-Cl on the structure of Pu22 G4; 3). Determine the binding mode of these four ILs on Pu22 G4 using TMPyP4 as a molecular probe. Clustering results confirmed two binding modes (top and bottom binding) of TMPyP4 to Pu22 G4 observed from CD spectroscopy, where TMPyP4 primarily engages in binding to the top G4 layer. Negligible difference in RMSF and geometry parameter values of Pu22 in water and each IL indicate that Pu22 G4 is stable in water and each of the ILs. Atom contact plots reveal that increasing the IL
cation hydrophobic chain length further reduces TMPyP4 binding to Pu22 G4, which is consistent with UV-Vis spectroscopy results. MD simulations indicate that ILs bind to Pu22 G4 primarily via hydrophobic interactions and electrostatic interactions; the hydrophobic interactions occur between the hydrophobic alkyl chain with the top G4 layer and electrostatic interactions occur between the polar imidazolium ring and the phosphates in the DNA backbone. Nucleic acid stability studies on mRNA using these four ILs should be conducted on other nucleic acids, such as the SARS-CoV-2 mRNA vaccine.
Chapter 3

Probing the Thermodynamic Destabilization of Azurin by Two Different Tetramethylguanidinium Amino Acid Ionic Liquids Using Molecular Dynamics Simulations

3.1 Introduction

3.1.1 Previous Studies of Amino Acid-Based ILs on Protein Structure

The three-dimensional structure of a protein is directly correlated to its function, with small perturbations in structure leading to a loss of protein function. The tunability of ILs make them apt for modulating protein structure and conformation as cosolvents in aqueous solution. ILs are thus capable of stabilizing or destabilizing protein structure depending on their chemical makeup. For instance, imidazolium-based ILs are reported to denature proteins and cause their unfolding, however short chain imidazolium ILs (i.e. BMIM-Cl) were reported to induce protein refolding and choline-based ILs are touted for promoting protein stability. However, the search for more biocompatible ILs for studying protein structure has led to the development of amino acid-based ILs.

Amino acid-based ILs (AA ILs) have also gained recent notice as biocompatible solvents for modulating protein structure. The amino acid residue (i.e. aspartate, leucine, alanine, serine, etc.) is usually incorporated as a variable anion to investigate the effect of AA-based ILs on protein structure. The choice in cation can vary, although tetramethylguanidinium (TMG) has been proposed to be used for applications in
biotechnology, such as studying protein activity, and has been shown to be both nontoxic and biodegradable \(^{122}\). A study by Singh et al. demonstrated that TMG-based ILs coupled with several different anions (i.e. salicylate, nitrate, acetate, valerate, etc.) were non-toxic towards HEK293 and the cancer cell culture DLD-1 and were biodegradable in the presence of the bacteria *Pseudomonas putida* and *Bacillus subtilis* \(^{122}\). TMG-based ILs have also been shown to destabilize protein structure. For instance, TMG-lactate IL has been shown to destabilize silk protein by disrupting the hydrogen-bonding network in its beta-sheet motif \(^{123}\). More importantly, Borrel et al. \(^{35}\) coupled spectroscopy with MD simulations to elucidate the protein unfolding mechanism of AA-based ILs TMG-Asp and TMG-Ala on the red fluorescent protein mCherry. UV-Vis and fluorescence spectroscopy revealed a more dramatic blue shifting and increased Stokes shift of the spectrum of mCherry in the presence of TMG-Ala compared to in TMG-Asp, indicating TMG-Ala exhibited a stronger destabilization effect. MD simulations revealed that the Ala anion engages in more contacts with mCherry than the Asp anion, suggesting increasing the cation hydrophobicity may lead to greater destabilization of mCherry. However, Ala alone is insufficient for unfolding mCherry, thus it’s suggested Ala and TMG work in-concert to destabilize and unfold mCherry. From these findings, the observed effects on protein structure dynamics can help optimize TMG-AA ILs for protein-based biotechnological applications.

### 3.1.2 Azurin Protein

Azurin is a tetrameric, mixed structure, Cu\(^{2+}\)-containing redox protein from *Pseudomonas aeruginosa*. Azurin mediates electron transfer reactions between nitrite reductase and other enzymes involved in denitrification \(^{124}\). Specifically, the Cu\(^{2+}\) ion is
held in-place by a single alpha helix, six beta-sheets and random coil motif and aids in stabilizing the protein tertiary structure. Azurin-mediated denitrification has been exploited in reducing excess nitrate into nitrogen gas within soil, preventing nitrate from entering and contaminating groundwater. Understanding the protein structure dynamics of azurin would assist in modulating its activity and support its application in biotechnology. Several spectroscopy studies have been performed to understand the unfolding/folding effects of different ILs on azurin structure. Recently, Acharyya et al. observed the unfolding mechanism of imidazolium-based ILs on the tertiary structure of azurin, which suggested that azurin destabilization increased in ILs of increasing hydrophobicity (BMIM-Cl<HMIM-Cl<OMIM-Cl). In contrast, two studies observed stabilization of azurin in choline dihydrogen phosphate (Chol-Dhp). Understanding these results can be further improved using MD simulations, which is a powerful tool in providing insight into the stabilizing/destabilizing effects of ILs on protein structure.

3.1.3 Experimental Overview

In this study, we utilize MD simulations to investigate the thermodynamic destabilization of azurin by TMG-AA ILs containing a carboxylic acid side chain (TMG-Asp) and an alcohol side chain (TMG-Ser) (Figure 11).
Figure 11

*TMG-AA ILs in Our Study*

Based on fluorescence spectroscopy, the unfolded fraction of azurin protein observed was quite similar between TMG-Asp and TMG-Glu and between TMG-Ser and TMG-Thr, with TMG-Asp and TMG-Ser exhibiting lower protein unfolding temperatures than TMG-Glu and TMG-Thr, respectively. Thus, MD simulations were performed for folded/unfolded azurin in the presence of TMG-Asp and TMG-Ser. MD simulation results are verified with experimental results to confirm the different interactions between AAs and protein structure.

3.2 Methodology

3.2.1 Simulation System Protocols

MD simulations were performed to analyze the stability of folded and unfolded azurin with and without the presence of TMG-Asp and TMG-Ser ILs. To save computational resources, the other two ILs based on this study (TMG-Glu and TMG-Thr) were not included in our MD simulations as they are chemically similar to TMG-Asp and TMG-Ser, respectively. A total of six systems were built: Folded azurin with water only,
with TMG-Asp, and with TMG-Ser, and unfolded azurin with water only, with TMG-Asp and with TMG-Ser (Table B1). Each system from here on is named by the protein conformation (folded/unfolded) and solvent (water/TMG-Asp/TMG-Ser). Each folded system was built using the solved high-resolution X-Ray crystal structure of azurin monomer (PDB id: 4MFH) obtained from the Protein Data Bank (Figure 12A).

**Figure 12**

*Reference Structures in Our MD Simulations*

Note. (A) Folded azurin. (B) Unfolded azurin.

Complete secondary structure information has been placed in the supporting document (Figure B1). Each unfolded system was built from an unfolded structure (Figure 12B) that was generated using Replica Exchange Molecular Dynamics (REMD) simulation, which started from the initial folded structure at 300K and increased the temperature up to 500K, yielding the unfolded protein conformation. To investigate the role of alcohol and acid side chains on azurin structure and to reduce computation time,
the ILs TMG-Asp and TMG-Ser were chosen. Azurin protein was prepared using
Maestro’s Protein Preparation Wizard program\textsuperscript{129}. Pre-processing (adding missing
hydrogens, optimization of pH 7 protonated state of side chains), geometry optimization
of global azurin structure utilized the default parameters for restricted minimization.
ff14SB force field was used to represent the prepared azurin protein and its disulfide
bonds. AMBER16 software was used to characterize the IL ions TMG, Ser and Asp, and
the GAFF2 force field was used to represent them. The MOL2 files of TMG, Asp and Ser
ions have been placed in the supporting document (\textbf{Figure B2-B4}). Each IL system was
populated with enough cations and anions to comprise 0.1 M TMG-Asp or TMG-Ser. IL
molecule positions were randomly generated surrounding the protein while using a
relaxation protocol to increase randomization of IL positions. Each system was generated
in a TIP3P water box of truncated octahedron with a 10 Å cutoff. The system net charge
was neutralized by adding sufficient counter ions. The detailed information of each
system was tabulated (\textbf{Table B1}).

\textbf{3.2.2 MD Simulation Protocols}

AMBER16 simulation package was used to perform simulations of each system,
using default simulation protocols. The four systems containing ILs (folded-TMG-Ser,
folded-TMG-Asp, unfolded-TMH-Ser and unfolded-TMG-Asp) underwent a pre-run of
1000 ps at 500 K to randomize the position and orientation of TMG, Ser and Asp ions
prior to initiating the production run at 300 K. The protein position was fixed in each
simulation system during the pre-run. The production run contained a 1.0 ns MD
simulation under the NPT ensemble (constant pressure and temperature) to bring the
system density to equilibrium. This was followed by 999 ns of MD under the equivalent
NVT ensemble (constant volume and temperature). A validation plot demonstrating the density of the bulk water of each system reaches ~1 g/mL following the 1.0 ns NSP equilibration simulation has been generated and placed in the supporting document (Figure B5). The SHAKE algorithm with a 2.0 fs time step was used to treat all bonds involving hydrogen atoms. The Particle-mesh Ewald method\textsuperscript{100} was used to treat long-range electrostatic interactions under periodic boundary conditions (using a charge grid spacing of ~ 1.0 Å, the fourth order of B-spline charge interpolation, with the direct sum tolerance of $10^{-5}$). In contrast, short-range non-bonded interactions were characterized using a 10 Å cutoff and long-range van der Waals interactions were characterized based on a uniform density approximation. Non-bonded forces were calculated using a two-stage RESPA approach to decrease computational resources\textsuperscript{101}. Short- and long-range forces were calculated every one and two steps, respectively. The system temperature was maintained using the Langevin thermostat with a coupling constant of 2.0 ps. System trajectories were saved every 50.0 ps for post-simulation analysis.

3.2.3 Simulation Convergence

RMSD values for the protein were calculated for each system. Protein RMSD plots were generated for folded/unfolded azurin in water, TMG-Asp and TMG-Ser against the respective reference structure based on azurin’s initial backbone conformation. Another protein RMSD plot was generated for unfolded azurin in water and each IL against the folded azurin crystal structure.

53
3.2.4 Protein Secondary Structure Information

To investigate if protein refolding occurred during unfolded azurin trajectories, secondary structure type was calculated for each amino acid over the entire trajectory in folded and unfolded azurin systems using the Timeline plug-in tool from VMD 1.9.3. Secondary structure plots were generated for the folded systems (Figure B7) and unfolded systems (Figure B8). The folded system plots were used to compare against the unfolded system plots to determine if refolding of native secondary structure occurred in unfolded azurin trajectories.

3.2.5 Characterization of Structural Conformations

RMSF values were calculated for each amino acid in folded azurin systems to characterize the fluctuations in azurin protein conformation in water and each IL.

3.2.6 Calculating Atom Contacts Between IL Cation/Anion and Protein Polar/Hydrophobic Side Chains

Atom contact plots were generated for the four IL-containing systems using a distance cutoff of 2.5 Å to characterize the number of 1). TMG-protein contacts, 2). Asp/Ser-protein contacts, 3). Protein polar side chain-IL contacts and 4). Protein hydrophobic side chain-IL contacts.

3.3 Results

3.3.1 Thermal Unfolding Experiments

From fluorescence spectroscopy, TMG-Ser and TMG-Thr significantly destabilized azurin’s conformation. No significant difference was observed between TMG-Asp and TMG-Glu. Interestingly, a small difference was observed between TMG-
Ser and TMG-Thr, in which TMG-Ser caused a smaller $T_m$ than TMG-Thr for azurin. This may be due to the added hydrophobicity on the Thr side chain (added methyl group). Herein lies our motivation for simulating TMG-Asp and TMG-Ser over the other two ILs.

### 3.3.2 MD Simulations

RMSD plots of the protein backbone over 1000 ns were generated for the folded and unfolded azurin states. *(Figure 13).* The unfolded azurin RMSD plots against the folded azurin crystal structure were used to compare both conformational states and are in the supporting document *(Figure B6).*
At room temperature, no significant change in azurin’s conformation was observed in water, TMG-Asp or TMG-Ser systems, when compared to the folded reference conformation (*Figure 13A*), which is reinforced by the average RMSD values of the last 300 ns for each system (water: 1.3 Å; TMG-Ser: 2.8 Å; TMG-Asp: 1.9 Å). Generally, RMSD values greater than 3 Å indicate protein instability, but this was not observed in *Figure 13B*; notice how the RMSD increase within TMG-Ser and TMG-Asp
are relatively minor compared with the rest of the trajectory. From this, azurin’s structure seemed most stable in water and no protein unfolding was observed in the presence of TMG-Ser or TMG-Asp, which is consistent with experimental results. For unfolded azurin, the respective RMSD plot exhibited a relaxation time around 200 ns and each system reached convergence (water: 11.8 Å; TMG-Ser: 14.0 Å; TMG-Asp: 13.1 Å). (Figure 13B). The trends observed between the folded and unfolded state RMSD plots are consistent; water induced the least conformational change in azurin, with no significant changes occurring with TMG-Ser and TMG-Asp.

3.3.3 Refolding Events in Unfolded Azurin

Intriguingly, the MD simulations for the unfolded azurin systems showed partial refolding of azurin’s structure (Figure 14). The extent of refolding varied between water and each IL when compared to the unfolded reference structure (Figure 12B).
In water, β-sheets and α-turns formed; in TMG-Ser, β-sheets and α-helices formed; in TMG-Asp, α-helices and α-turns formed. Plot diagrams showing the partial refolding events for both folded azurin (Figure B10) and unfolded azurin (Figure B11) in water and each IL were generated and placed in the supporting document. Partial refolding was observed in native protein regions in the unfolded azurin MD simulations (Figure B11). For instance, β-sheets formed (water: K85-D93; TMG-Ser: T21) as well as
α-helices (TMG-Asp: M56-T61). Additionally, partial refolding in non-native protein regions were also observed. In water alone, refolding events formed transient α-helices (F97-K101) and partial isolated β-bridges (G121-L125). In TMG-Ser, formation of α-helices is more significant (L86-E91 and T96-K101). In TMG-Asp, formation of α-helices is not as significant (I81-L96). Many 3-10 helices formed under each solvent condition, which is not surprising, considering they are purported to be intermediates in protein folding dynamics. Albeit partial refolding was observed, full refolding of azurin (~7 ms) is not observable using our current computational resources (1 μs).

### 3.3.4 Atom Contact Plot Analysis

To determine the molecular distribution of each IL over azurin, structure models depicting TMG, Ser and Asp ions interacting with folded and unfolded azurin conformations are shown (Figure 15). Intermolecular atom contact plots over each system trajectory were generated (Figure B12) and the average atom contacts over the last 300 ns of these contact plots were calculated and tabulated (Table 6).
Figure 15

*IL Distribution in Last Snapshot Structures*

<table>
<thead>
<tr>
<th>Folded</th>
<th>TMG-Ser</th>
<th>TMG-Asp</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Folded TMG-Ser" /></td>
<td><img src="image2.png" alt="Folded TMG-Asp" /></td>
<td><img src="image3.png" alt="Folded TMG-Asp" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unfolded</th>
<th>TMG-Ser</th>
<th>TMG-Asp</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image4.png" alt="Unfolded TMG-Ser" /></td>
<td><img src="image5.png" alt="Unfolded TMG-Asp" /></td>
<td><img src="image6.png" alt="Unfolded TMG-Asp" /></td>
</tr>
</tbody>
</table>

*Note.* TMG cations (blue) and Ser/Asp anions (red). Protein is colored by residue type.
Table 6

*Atom Contacts Between Azurin and ILs*

<table>
<thead>
<tr>
<th>System</th>
<th>Prot-TMG</th>
<th>Prot-Ser/Asp</th>
<th>Polar-IL</th>
<th>Apolar-IL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Folded-TMG-Ser</td>
<td>23±4</td>
<td>11±2</td>
<td>30±4</td>
<td>7±1</td>
</tr>
<tr>
<td>Unfolded-TMG-Ser</td>
<td>19±4</td>
<td>11±3</td>
<td>24±2</td>
<td>7±1</td>
</tr>
<tr>
<td>Folded-TMG-Asp</td>
<td>30±4</td>
<td>33±7</td>
<td>57±7</td>
<td>8±1</td>
</tr>
<tr>
<td>Unfolded-TMG-Asp</td>
<td>23±5</td>
<td>62±11</td>
<td>76±11</td>
<td>11±3</td>
</tr>
</tbody>
</table>

*Note.* Contacts were averaged over the last 300 ns of each trajectory.

Indeed, we observed ionic interactions (TMG-anionic protein and Asp-cationic protein) and polar interactions (Ser-polar protein) between IL and protein side chain residues (*Figure 15*). Generally, IL-polar protein side chain interactions dominate over IL-apolar protein side chain interactions. A comparison of the molecular distribution of each IL over the folded and unfolded azurin protein conformation was then made. For folded azurin, ILs were only able to sample the outer protein surface. In the Folded-TMG-Ser system, TMG and Ser ions were more evenly distributed over the outer protein surface (*Figure 15*). However, TMG-protein contacts dominate over Ser-protein contacts (Protein-TMG: 23; Protein-Ser: 11, *Table 6*). As Ser is a polar AA, Ser sees more polar-protein than hydrophobic-protein side chain interactions (Polar-IL: 30; Apolar-IL: 7, *Table 6*). In the Folded-TMG-Asp system, both TMG and Asp ions were observed to cluster at the bottom protein region (*Figure 15*) and their contacts with the protein are almost equal (Protein-TMG: 30; Protein-Asp: 33, *Table 6*). As with Ser, Asp engages in more interactions with polar protein side chains (Polar IL: 57; Apolar-IL: 7, *Table 6*).
Unlike the folded conformation, the unfolded conformation of azurin exposes the hydrophobic protein interior to the surrounding solvent and thus increases the available surface area for molecular sampling of protein by IL; thus, we expect an increase in the number of TMG, Ser and Asp atom contacts with unfolded azurin compared with folded azurin. For Unfolded-TMG-Ser system, TMG ions appear to sample only one side of azurin, whereas Ser ions sample both sides. Again, TMG ions dominate over Ser ions in protein interactions (Protein-TMG: 19; Protein-Ser: 11, Table 6). As before, Ser samples more polar protein side chains than hydrophobic protein side chains (Polar-IL: 24; Apolar-IL: 7, Table 6). In the Unfolded-TMG-Asp system, the molecular distribution of ILs sparks greater interest. The unfolded azurin conformation includes a section of α-helices situated above the rest of the protein, creating a region for unique molecular sampling. Three TMG molecules were observed to cluster around this α-helical strand as well as the rest of the structure, whereas Asp molecules clustered at the bottom protein region (Figure 15). Interestingly, Asp dominates TMG for protein side chain interactions (Protein-TMG: 23; Protein-Asp: 62, Table 6), where Unfolded Protein-Asp contacts exhibit the highest protein contact number. As before, Asp engages in more polar than hydrophobic protein side chain interactions (Polar-IL: 76; Apolar-IL: 11, Table 6). For both folded and unfolded azurin conformations, Asp engages in more contacts than Ser for azurin. For TMG-Ser, TMG contacts decrease and Ser interactions did not change between unfolded and folded azurin states. For TMG-Asp, TMG contacts again decrease and Asp interactions increased significantly with unfolded azurin than folded azurin. This is strange, as we expected greater molecular sampling and an increase in atom contacts for each IL to the protein.

62
From these results, polar-IL interactions dominate over hydrophobic-IL interactions, suggesting that polar-IL interactions may play some role in modulating both azurin protein conformations.

3.3.5 RMSF Analysis

To investigate the degree of entropic change in azurin’s structure when exposed to different solvents, a protein RMSF diagram for each residue in the folded and unfolded azurin conformations in each solvent was generated (Figure 16). The average RMSF values for each system have also been tabulated (Table 7).

Figure 16

Protein RMSF in Water and ILs

Note. RMSF is over 128 amino acids. A) Folded azurin. B) Unfolded azurin.
Table 7

*Tabulated Averaged Protein RMSF Values*

<table>
<thead>
<tr>
<th>System</th>
<th>Folded (Å)</th>
<th>Unfolded (Å)</th>
<th>Unfolded-Folded (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>1.1±0.5</td>
<td>4.9±1.7</td>
<td>3.8</td>
</tr>
<tr>
<td>TMG-Ser</td>
<td>1.4±0.7</td>
<td>7.0±2.1</td>
<td>4.7</td>
</tr>
<tr>
<td>TMG-Asp</td>
<td>1.3±1.0</td>
<td>6.1±2.6</td>
<td>5.7</td>
</tr>
</tbody>
</table>

In unfolded azurin, TMG-Ser induced the greatest conformational change (7.0 Å, change = 4.7 Å), followed by TMG-Asp (6.1 Å, change = 5.7 Å) and water (4.9 Å, change = 3.8 Å), with the change reflecting the difference between the RMSF of folded and unfolded azurin conformations in each solvent. However, in folded azurin, the RMSF values are nearly consistent for water (1.1 Å), TMG-Ser (1.4 Å) and TMG-Asp (1.3 Å), suggesting TMG-Ser and TMG-Asp do not cause large fluctuations in the folded azurin structure. Despite this, some regions show relatively high fluctuation. In folded azurin (*Figure 16*), TMG-Ser induced high fluctuations in AA residues 8-18 (RMSF = 6 Å) and TMG-Asp induced a slightly smaller fluctuations in AA residues 35-43 (RMSF = 4 Å). In all three solvents, protein structure fluctuations are relatively stable after residue 43 (~1 to 2 Å), though slightly higher fluctuations were observed in the N-terminus (~3.5 Å) and C-terminus (~2 Å). In unfolded azurin, almost no RMSF overlap is observed in the three solvents, since the protein secondary structure composition varies following exposure to different solvents. Structure stability is observed up to the first 24 residues, with more significant fluctuations observed in the N-terminus (~3 Å) and C-terminus (~9 Å). Water exhibits the greatest stability for unfolded azurin compared to TMG-Ser and TMG-Asp.
Generally, unfolded azurin exhibits greater entropy when subjected to TMG-Ser and TMG-Asp, which is confirmed by experimental measurements\textsuperscript{131}.

3.4 Discussion

Thermal unfolding experiments reveal that TMG-AA ILs destabilize azurin in aqueous environment, due to simultaneous action of both cation (TMG) and anion (Ser/Asp). The destabilizing effect is seemingly stronger in TMG-AA ILs where the anion has an alcohol side chain (TMG-Ser) and is weaker when the anion contains a carboxylic acid side chain (TMG-Asp). To understand how TMG-AA ILs destabilize azurin’s structure conformation, MD simulations were performed for both folded and unfolded azurin in the presence of TMG-Ser and TMG-Asp. At room temperature, MD simulations show that folded azurin is stable when subjected to both TMG-Ser and TMG-Asp and does not show unfolding, which is consistent with experimental findings\textsuperscript{131}.

The MD simulations also suggest how IL cations and anions interact with the folded and unfolded azurin conformations. (Figures 15 and B13) For the folded state, we observed that Asp shows more protein interactions than Ser, which also involves stronger ionic interactions with Asp than weaker polar interactions with Ser. Therefore, this suggests that the folded azurin state is more stable in TMG-Asp than in TMG-Ser. For the unfolded state, both Ser and Asp interact with the protein structure. Yet, Asp shows significantly greater interactions with the unfolded state compared to the folded state, whereas Ser shows similar contact counts between the two conformational states. Experimental results suggests\textsuperscript{131} that TMG-Asp both increase the change in the energy of unfolding ($\Delta H_{\text{unf}}^\circ$) and the change in entropy ($\Delta S_{\text{unf}}^\circ$), indicating moderate protein destabilization, whereas TMG-Ser saw no change in $\Delta H_{\text{unf}}^\circ$ but increased $\Delta S_{\text{unf}}^\circ$, indicating stronger protein destabilization (Figure 17). From both MD simulation and experimental results, we can speculate that Ser stabilizes both the folded and unfolded states to a similar extent as does
water, thus TMG-Ser will not change $\Delta H_{\text{unf}}^\circ$. However, Asp interacts with the folded and unfolded states differently due to its stronger ionic interactions with azurin, thus the stabilizing effects between these states are different with respect to water, therefore TMG-Asp will change $\Delta H_{\text{unf}}^\circ$.

**Figure 17**

*Experimental Protein Unfolding Thermodynamic Parameters*

Protein RMSF reveal that the folded azurin conformation is likely the same in water, TMG-Ser and TMG-Asp *(Figure 16)* and that the average RMSF values for folded azurin are the same in these three solvents *(Table 7)*. Thus, MD simulations suggest that folded azurin is not destabilized in IL. In contrast, unfolded azurin conformations are obviously different in each solvent *(Figure 16)*. The unfolded structure appears to be more disorganized in IL, showing greater disorganization in TMG-Asp than in TMG-Ser. The average RMSF values *(Table 7)* are greater in each IL than in water, with TMG-Ser having the highest RMSF. The MD simulation results seem to support the experimental results, suggesting that both ILs increase the entropy of the unfolded conformation.
MD simulations show that the different interactions observed between IL and azurin are likely anion-dependent (i.e. IL-protein interactions). The alcohol side chain on TMG-Ser engages in polar protein interactions and increases the entropy of the unfolded conformation (increases $\Delta S_{\text{unf}}$) while simultaneously stabilizing both unfolded and folded conformations (no appreciable change in $\Delta H_{\text{unf}}$), leading to more favorable, reduced unfolding free energies and reduced $T_m$. In contrast, the carboxylic acid side chain on TMG-Asp engages in strong, ionic protein interactions and increases the entropy of the unfolded conformation (increases $\Delta S_{\text{unf}}$), yet TMG-Asp induces greater stability in folded azurin versus unfolded azurin, leading to increased change in unfolding enthalpy (increased $\Delta H_{\text{unf}}$) and higher values of unfolding free energies and $T_m$ compared with azurin in TMG-Ser.

3.5 Conclusions

MD simulations were performed on the folded and unfolded conformations of azurin in TMG-Ser and TMG-Asp ILs to investigate their effects on protein structure. MD simulations clearly show that TMG interacts with both the folded and unfolded states similarly, whereas Ser and Asp engage in polar and ionic interactions, respectively. Moreover, from the folded to the unfolded state, the number of Asp-protein contacts doubled, whereas the number of Ser-protein contacts were consistent, indicating that Asp leads to greater stabilization of azurin than Ser. Interestingly, we also observed a higher helical content in the unfolded conformation, which suggests that TMG-Ser and TMG-Asp perhaps affect the global protein dielectric environment. From both MD simulations and experimental results, Ser likely stabilizes both folded and unfolded azurin conformations to the same extent, whereas Asp clearly stabilizes the folded state more
than the unfolded state. RMSF results show noticeable conformational fluctuations in the unfolded state in the presence of ILs, suggesting they increase the unfolded conformational entropy.
Chapter 4

Concluding Statements

From the two studies, it’s clear that ILs can effectively modulate the structural conformation of DNA and proteins. In our findings from Chapter 2, we observed that imidazolium-based ILs of increasing hydrophobicity are able to reduce ligand binding to a DNA G4, while retaining its structural stability. Following this logic, the effects of biocompatible ILs with hydrophobic cations on RNA structure (i.e., RNA G4s, RNA-based therapeutics, ribozymes, etc.) should be studied, as RNA is even more chemically unstable than DNA and holds great therapeutic potential. As an example, hydrophobic biocompatible ILs could possibly protect the secondary and tertiary structures of the SARS-CoV-2 mRNA vaccine from hydrolytic cleavage at room temperature storage conditions via the hydrophobic effect. Additionally, such ILs could be included as a stabilizer in the delivery of RNA therapeutics to human patients (i.e., ASOs, siRNA) to increase their half-life and potency in-vivo. In Chapter 3, we observed that TMG-AA-based ILs engage in very specific interactions with the folded and unfolded states of azurin protein. Indeed, this effect was also observed in a previous study involving the laccase redox enzyme and TMG-Ser IL, where TMG cation-sulfonate and Ser anion-T1 and T3 Cu$^{2+}$ electrostatic interactions were observed and are likely responsible for disrupting the electron flow mechanism and abolishing laccase activity. This idea could be extended to inhibit the activity of viral proteins, such as the SARS-CoV-2 RNA-dependent RNA polymerase (RdRp) responsible for viral replication. RdRp contains two Mg$^{2+}$ cations critical to its activity; perturbing Mg$^{2+}$ positioning with TMG-AA-based ILs
could inactivate RdRp and stop viral replication, thus serving as a potentially promising viral therapeutic.
References


20 Walden, P. *Molecular weights and electrical conductivity of several fused salts.* 405-422 (Bull. Imp. Acad. Sci., 1914).


Amber 2016 (University of California, San Francisco, 2016).


Appendix A

To Probe the Binding of TMPyP4 to c-MYC G-Quadruplex in Water and in Imidazolium-based IL Solvents Using Spectroscopy Coupled with Molecular Dynamics Simulations

Table A1

Sequence of Pu22 G4

| Pu22 | 5’-TGAG¹G¹TG²G²TAG³G³TG⁴G⁴TAA-3’ |

Note. Top, middle, and bottom G-tetrads are colored red, green, and orange respectively.
Table A2

Simulation System Parameters

<table>
<thead>
<tr>
<th>ID</th>
<th>System</th>
<th># of runs</th>
<th># of Waters</th>
<th>Box Size (Å)*</th>
<th># of ILs</th>
<th># of Cl− ions</th>
<th># of K+ ions</th>
<th>NPT eq. (ns)</th>
<th>NVT (ns)</th>
<th>Total Time (µs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rec-Only</td>
<td>2</td>
<td>5570</td>
<td>64.9</td>
<td>0</td>
<td>13</td>
<td>32</td>
<td>1</td>
<td>1000</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Lig-Only</td>
<td>2</td>
<td>4228</td>
<td>57.7</td>
<td>0</td>
<td>10</td>
<td>6</td>
<td>1</td>
<td>1000</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Lig-Top</td>
<td>15</td>
<td>5570</td>
<td>64.9</td>
<td>0</td>
<td>13</td>
<td>28</td>
<td>1</td>
<td>1000</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>Lig-Bot</td>
<td>15</td>
<td>7233</td>
<td>70.1</td>
<td>0</td>
<td>13</td>
<td>28</td>
<td>1</td>
<td>1000</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>Lig-Top-HMIM</td>
<td>5</td>
<td>7369</td>
<td>75.5</td>
<td>64</td>
<td>49</td>
<td>0</td>
<td>1</td>
<td>1000</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>Lig-Bot-HMIM</td>
<td>5</td>
<td>10114</td>
<td>81.7</td>
<td>64</td>
<td>49</td>
<td>0</td>
<td>1</td>
<td>1000</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>Lig-Top-BMIM</td>
<td>5</td>
<td>7486</td>
<td>75.0</td>
<td>64</td>
<td>49</td>
<td>0</td>
<td>1</td>
<td>1000</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>Lig-Bot-BMIM</td>
<td>5</td>
<td>8960</td>
<td>78.8</td>
<td>64</td>
<td>49</td>
<td>0</td>
<td>1</td>
<td>1000</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>Lig-Top-EMIM</td>
<td>5</td>
<td>7660</td>
<td>74.9</td>
<td>64</td>
<td>49</td>
<td>0</td>
<td>1</td>
<td>1000</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>Lig-Bot-EMIM</td>
<td>5</td>
<td>9241</td>
<td>78.7</td>
<td>64</td>
<td>49</td>
<td>0</td>
<td>1</td>
<td>1000</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>Lig-Top-OMIM</td>
<td>5</td>
<td>7104</td>
<td>75.3</td>
<td>64</td>
<td>49</td>
<td>0</td>
<td>1</td>
<td>1000</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>Lig-Bot-OMIM</td>
<td>5</td>
<td>8811</td>
<td>79.4</td>
<td>64</td>
<td>49</td>
<td>0</td>
<td>1</td>
<td>1000</td>
<td>5</td>
</tr>
</tbody>
</table>

Note. Simulations were run with PDB id: 1XAV and ligand TMPyP4. NPT ensemble: Constant pressure and temperature. NVT ensemble: Constant volume and temperature. Box size refers to volume of water box.
Table A3

Average RMSD and RMSF Values for Pu22 and TMPyP4

<table>
<thead>
<tr>
<th></th>
<th>WAT</th>
<th>EMIM</th>
<th>BMIM</th>
<th>HMIM</th>
<th>OMIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-RMSD (Å)</td>
<td>2.8±0.1</td>
<td>2.8±0.1</td>
<td>2.9±0.2</td>
<td>2.8±0.2</td>
<td>2.8±0.2</td>
</tr>
<tr>
<td>L-RMSD (Å)</td>
<td>20.6±1.1</td>
<td>27.8±2.1</td>
<td>19.5±2.1</td>
<td>36.7±2.8</td>
<td>35.7±3.4</td>
</tr>
<tr>
<td>R-RMSF (Å)</td>
<td>1.5±1.0</td>
<td>1.6±1.3</td>
<td>1.7±1.2</td>
<td>1.7±1.3</td>
<td>2.1±1.7</td>
</tr>
</tbody>
</table>

*Note.* R-RMSD: Receptor (Pu22) RMSD. L-RMSD: Ligand (TMPyP4) RMSD. R-RMSF: Receptor (Pu22) RMSF. See Figures A3 and A5.
Table A4

G4 Layer Geometry Parameters

<table>
<thead>
<tr>
<th>Layers</th>
<th>Parameter</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>3:2</td>
<td>Rise¹</td>
<td>3.48±0.08</td>
<td>3.47±0.07</td>
<td>3.48±0.08</td>
<td>3.47±0.07</td>
</tr>
<tr>
<td>3:2</td>
<td>H-Rise¹</td>
<td>3.46±0.08</td>
<td>3.45±0.07</td>
<td>3.46±0.08</td>
<td>3.46±0.07</td>
</tr>
<tr>
<td>3:2</td>
<td>H-Twist²</td>
<td>24.05±2.47</td>
<td>24.32±2.60</td>
<td>23.72±2.70</td>
<td>24.18±2.60</td>
</tr>
<tr>
<td>2:1</td>
<td>Rise¹</td>
<td>3.42±0.08</td>
<td>3.43±0.08</td>
<td>3.44±0.08</td>
<td>3.43±0.08</td>
</tr>
<tr>
<td>2:1</td>
<td>H-Rise¹</td>
<td>3.40±0.08</td>
<td>3.41±0.08</td>
<td>3.42±0.08</td>
<td>3.41±0.08</td>
</tr>
<tr>
<td>2:1</td>
<td>H-Twist²</td>
<td>26.29±3.04</td>
<td>25.94±2.98</td>
<td>26.12±3.13</td>
<td>25.91±3.10</td>
</tr>
</tbody>
</table>

Note. ¹Distance measured in Å. ²Angle measured in degree. 1-3: top, middle and bottom G4 layers, respectively in water.
Table A5A

Pu22 Guanine Tetrad Oxygen-to-Potassium Distance in Water

<table>
<thead>
<tr>
<th>Residue:O6:K⁺</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6:K1</td>
<td>2.61±0.01</td>
<td>2.60±0.01</td>
<td>2.60±0.01</td>
<td>2.58±0.00</td>
</tr>
<tr>
<td>G8:O6:K1</td>
<td>2.65±0.01</td>
<td>2.73±0.01</td>
<td>2.63±0.01</td>
<td>2.66±0.01</td>
</tr>
<tr>
<td>G13:O6:K1</td>
<td>2.63±0.01</td>
<td>2.65±0.00</td>
<td>2.60±0.01</td>
<td>2.61±0.01</td>
</tr>
<tr>
<td>G17:O6:K1</td>
<td>2.63±0.01</td>
<td>2.62±0.00</td>
<td>2.63±0.01</td>
<td>2.63±0.00</td>
</tr>
<tr>
<td>G5:O6:K1</td>
<td>2.80±0.02</td>
<td>2.75±0.02</td>
<td>2.82±0.02</td>
<td>2.75±0.02</td>
</tr>
<tr>
<td>G9:O6:K1</td>
<td>2.68±0.02</td>
<td>2.65±0.02</td>
<td>2.69±0.01</td>
<td>2.72±0.01</td>
</tr>
<tr>
<td>G14:O6:K1</td>
<td>2.79±0.01</td>
<td>2.72±0.01</td>
<td>2.83±0.02</td>
<td>2.85±0.06</td>
</tr>
<tr>
<td>G18:O6:K1</td>
<td>2.68±0.01</td>
<td>2.69±0.02</td>
<td>2.68±0.02</td>
<td>2.68±0.01</td>
</tr>
<tr>
<td>G5:O6:K2</td>
<td>2.81±0.03</td>
<td>2.72±0.03</td>
<td>2.83±0.02</td>
<td>2.78±0.01</td>
</tr>
<tr>
<td>G9:O6:K2</td>
<td>2.87±0.02</td>
<td>3.04±0.08</td>
<td>2.87±0.05</td>
<td>2.87±0.03</td>
</tr>
<tr>
<td>G14:O6:K2</td>
<td>2.79±0.01</td>
<td>2.75±0.03</td>
<td>2.79±0.01</td>
<td>2.84±0.03</td>
</tr>
<tr>
<td>G18:O6:K2</td>
<td>2.82±0.01</td>
<td>3.02±0.02</td>
<td>2.79±0.03</td>
<td>2.78±0.02</td>
</tr>
<tr>
<td>G6:O6:K2</td>
<td>2.65±0.01</td>
<td>2.65±0.02</td>
<td>2.64±0.01</td>
<td>2.65±0.01</td>
</tr>
<tr>
<td>G10:O6:K2</td>
<td>2.66±0.01</td>
<td>2.71±0.01</td>
<td>2.65±0.02</td>
<td>2.61±0.02</td>
</tr>
<tr>
<td>G15:O6:K2</td>
<td>2.67±0.01</td>
<td>2.67±0.01</td>
<td>2.66±0.02</td>
<td>2.64±0.01</td>
</tr>
<tr>
<td>G19:O6:K2</td>
<td>2.66±0.01</td>
<td>2.66±0.01</td>
<td>2.68±0.01</td>
<td>2.64±0.01</td>
</tr>
</tbody>
</table>

*Note.* G4 layer colors (top (red); middle (green); bottom (blue)).
Table A5B

Pu22 Guanine Tetrall Oxygen-to-Potassium Distance in EMIM-Cl

<table>
<thead>
<tr>
<th>Residue:O6 :K⁺</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6:K1</td>
<td>2.61±0.01</td>
<td>2.59±0.00</td>
<td>2.60±0.02</td>
<td>2.59±0.01</td>
</tr>
<tr>
<td>G8:O6:K1</td>
<td>2.65±0.01</td>
<td>2.66±0.00</td>
<td>2.66±0.01</td>
<td>2.66±0.00</td>
</tr>
<tr>
<td>G13:O6:K1</td>
<td>2.63±0.01</td>
<td>2.61±0.01</td>
<td>2.62±0.02</td>
<td>2.62±0.01</td>
</tr>
<tr>
<td>G17:O6:K1</td>
<td>2.63±0.01</td>
<td>2.62±0.01</td>
<td>2.63±0.01</td>
<td>2.63±0.01</td>
</tr>
<tr>
<td>G5:O6:K1</td>
<td>2.80±0.02</td>
<td>2.76±0.05</td>
<td>2.82±0.08</td>
<td>2.79±0.03</td>
</tr>
<tr>
<td>G9:O6:K1</td>
<td>2.68±0.02</td>
<td>2.67±0.01</td>
<td>2.66±0.01</td>
<td>2.68±0.01</td>
</tr>
<tr>
<td>G14:O6:K1</td>
<td>2.79±0.01</td>
<td>2.82±0.01</td>
<td>2.78±0.00</td>
<td>2.76±0.01</td>
</tr>
<tr>
<td>G18:O6:K1</td>
<td>2.68±0.01</td>
<td>2.70±0.01</td>
<td>2.69±0.03</td>
<td>2.68±0.01</td>
</tr>
<tr>
<td>G5:O6:K2</td>
<td>2.81±0.03</td>
<td>2.76±0.01</td>
<td>2.81±0.04</td>
<td>2.80±0.04</td>
</tr>
<tr>
<td>G9:O6:K2</td>
<td>2.87±0.02</td>
<td>2.83±0.03</td>
<td>2.89±0.03</td>
<td>2.86±0.03</td>
</tr>
<tr>
<td>G14:O6:K2</td>
<td>2.79±0.01</td>
<td>2.83±0.02</td>
<td>2.80±0.02</td>
<td>2.81±0.03</td>
</tr>
<tr>
<td>G18:O6:K2</td>
<td>2.82±0.01</td>
<td>2.79±0.02</td>
<td>2.85±0.06</td>
<td>2.82±0.02</td>
</tr>
<tr>
<td>G6:O6:K2</td>
<td>2.65±0.01</td>
<td>2.64±0.00</td>
<td>2.64±0.01</td>
<td>2.65±0.01</td>
</tr>
<tr>
<td>G10:O6:K2</td>
<td>2.66±0.01</td>
<td>2.67±0.03</td>
<td>2.64±0.01</td>
<td>2.64±0.01</td>
</tr>
<tr>
<td>G15:O6:K2</td>
<td>2.67±0.01</td>
<td>2.63±0.00</td>
<td>2.68±0.03</td>
<td>2.64±0.01</td>
</tr>
<tr>
<td>G19:O6:K2</td>
<td>2.66±0.01</td>
<td>2.65±0.01</td>
<td>2.65±0.03</td>
<td>2.63±0.02</td>
</tr>
</tbody>
</table>

Note. G4 layer colors (top (red); middle (green); bottom (blue)).
Table A5C

Pu22 Guanine Tetrad Oxygen-to-Potassium Distance in BMIM-Cl

<table>
<thead>
<tr>
<th>Residue:O6:K⁺</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6:K1</td>
<td>2.61±0.01</td>
<td>2.61±0.01</td>
<td>2.59±0.01</td>
<td>2.60±0.01</td>
</tr>
<tr>
<td>G8:O6:K1</td>
<td>2.65±0.01</td>
<td>2.65±0.01</td>
<td>2.64±0.01</td>
<td>2.65±0.00</td>
</tr>
<tr>
<td>G13:O6:K1</td>
<td>2.63±0.01</td>
<td>2.63±0.02</td>
<td>2.61±0.01</td>
<td>2.63±0.01</td>
</tr>
<tr>
<td>G17:O6:K1</td>
<td>2.63±0.01</td>
<td>2.63±0.01</td>
<td>2.65±0.01</td>
<td>2.62±0.02</td>
</tr>
<tr>
<td>G5:O6:K1</td>
<td>2.80±0.02</td>
<td>2.77±0.04</td>
<td>2.83±0.03</td>
<td>2.80±0.03</td>
</tr>
<tr>
<td>G9:O6:K1</td>
<td>2.68±0.02</td>
<td>2.68±0.02</td>
<td>2.69±0.02</td>
<td>2.66±0.00</td>
</tr>
<tr>
<td>G14:O6:K1</td>
<td>2.79±0.01</td>
<td>2.78±0.03</td>
<td>2.78±0.03</td>
<td>2.76±0.03</td>
</tr>
<tr>
<td>G18:O6:K1</td>
<td>2.68±0.01</td>
<td>2.67±0.02</td>
<td>2.67±0.03</td>
<td>2.69±0.02</td>
</tr>
<tr>
<td>G5:O6:K2</td>
<td>2.81±0.03</td>
<td>2.83±0.03</td>
<td>2.82±0.02</td>
<td>2.76±0.02</td>
</tr>
<tr>
<td>G9:O6:K2</td>
<td>2.87±0.02</td>
<td>2.86±0.04</td>
<td>2.89±0.02</td>
<td>2.85±0.01</td>
</tr>
<tr>
<td>G14:O6:K2</td>
<td>2.79±0.01</td>
<td>2.83±0.02</td>
<td>2.82±0.01</td>
<td>2.81±0.02</td>
</tr>
<tr>
<td>G18:O6:K2</td>
<td>2.82±0.01</td>
<td>2.82±0.04</td>
<td>2.77±0.02</td>
<td>2.89±0.09</td>
</tr>
<tr>
<td>G6:O6:K2</td>
<td>2.65±0.01</td>
<td>2.65±0.02</td>
<td>2.64±0.01</td>
<td>2.64±0.01</td>
</tr>
<tr>
<td>G10:O6:K2</td>
<td>2.66±0.01</td>
<td>2.64±0.02</td>
<td>2.63±0.01</td>
<td>2.66±0.01</td>
</tr>
<tr>
<td>G15:O6:K2</td>
<td>2.67±0.01</td>
<td>2.65±0.01</td>
<td>2.64±0.01</td>
<td>2.65±0.02</td>
</tr>
<tr>
<td>G19:O6:K2</td>
<td>2.66±0.01</td>
<td>2.65±0.01</td>
<td>2.66±0.01</td>
<td>2.66±0.02</td>
</tr>
</tbody>
</table>

*Note.* G4 layer colors (top (red); middle (green); bottom (blue)).
Table A5D

Pu22 Guanine Tetrad Oxygen-to-Potassium Distance in HMIM-Cl

<table>
<thead>
<tr>
<th>Residue:O6:K+</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding*</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6:K1</td>
<td>2.61±0.01</td>
<td>2.58±0.01</td>
<td>2.60±0.01</td>
<td></td>
</tr>
<tr>
<td>G8:O6:K1</td>
<td>2.65±0.01</td>
<td>2.66±0.01</td>
<td>2.64±0.02</td>
<td></td>
</tr>
<tr>
<td>G13:O6:K1</td>
<td>2.63±0.01</td>
<td>2.61±0.02</td>
<td>2.61±0.01</td>
<td></td>
</tr>
<tr>
<td>G17:O6:K1</td>
<td>2.63±0.01</td>
<td>2.64±0.03</td>
<td>2.62±0.00</td>
<td></td>
</tr>
<tr>
<td>G5:O6:K1</td>
<td>2.80±0.02</td>
<td>2.80±0.03</td>
<td>2.75±0.04</td>
<td></td>
</tr>
<tr>
<td>G9:O6:K1</td>
<td>2.68±0.02</td>
<td>2.64±0.01</td>
<td>2.66±0.01</td>
<td></td>
</tr>
<tr>
<td>G14:O6:K1</td>
<td>2.79±0.01</td>
<td>2.78±0.05</td>
<td>2.83±0.04</td>
<td></td>
</tr>
<tr>
<td>G18:O6:K1</td>
<td>2.68±0.01</td>
<td>2.67±0.04</td>
<td>2.73±0.02</td>
<td></td>
</tr>
<tr>
<td>G5:O6:K2</td>
<td>2.81±0.03</td>
<td>2.78±0.03</td>
<td>2.75±0.03</td>
<td></td>
</tr>
<tr>
<td>G9:O6:K2</td>
<td>2.87±0.02</td>
<td>2.84±0.03</td>
<td>2.84±0.04</td>
<td></td>
</tr>
<tr>
<td>G14:O6:K2</td>
<td>2.79±0.01</td>
<td>2.76±0.03</td>
<td>2.76±0.03</td>
<td></td>
</tr>
<tr>
<td>G18:O6:K2</td>
<td>2.82±0.01</td>
<td>2.87±0.07</td>
<td>2.78±0.02</td>
<td></td>
</tr>
<tr>
<td>G6:O6:K2</td>
<td>2.65±0.01</td>
<td>2.62±0.02</td>
<td>2.61±0.01</td>
<td></td>
</tr>
<tr>
<td>G10:O6:K2</td>
<td>2.66±0.01</td>
<td>2.66±0.02</td>
<td>2.63±0.01</td>
<td></td>
</tr>
<tr>
<td>G15:O6:K2</td>
<td>2.67±0.01</td>
<td>2.68±0.02</td>
<td>2.70±0.02</td>
<td></td>
</tr>
<tr>
<td>G19:O6:K2</td>
<td>2.66±0.01</td>
<td>2.67±0.02</td>
<td>2.69±0.01</td>
<td></td>
</tr>
</tbody>
</table>

*Note. G4 layer colors (top (red); middle (green); bottom (blue)). Bottom binding was not observed for this system, thus no data is presented.
Table A5E

*Pu22 Guanine Tetrad Oxygen-to-Potassium Distance in OMIM-Cl*

<table>
<thead>
<tr>
<th>Residue:Oxy :K⁺</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6:K1</td>
<td>2.61±0.01</td>
<td>2.59±0.01</td>
<td>2.59±0.00</td>
<td>2.60±0.02</td>
</tr>
<tr>
<td>G8:O6:K1</td>
<td>2.65±0.01</td>
<td>2.63±0.01</td>
<td>2.66±0.00</td>
<td>2.65±0.01</td>
</tr>
<tr>
<td>G13:O6:K1</td>
<td>2.63±0.01</td>
<td>2.60±0.01</td>
<td>2.63±0.00</td>
<td>2.62±0.01</td>
</tr>
<tr>
<td>G17:O6:K1</td>
<td>2.63±0.01</td>
<td>2.63±0.01</td>
<td>2.63±0.00</td>
<td>2.63±0.01</td>
</tr>
<tr>
<td>G5:O6:K1</td>
<td>2.80±0.02</td>
<td>2.82±0.04</td>
<td>2.78±0.05</td>
<td>2.79±0.04</td>
</tr>
<tr>
<td>G9:O6:K1</td>
<td>2.68±0.02</td>
<td>2.68±0.04</td>
<td>2.65±0.02</td>
<td>2.67±0.01</td>
</tr>
<tr>
<td>G14:O6:K1</td>
<td>2.79±0.01</td>
<td>2.79±0.04</td>
<td>2.76±0.02</td>
<td>2.78±0.02</td>
</tr>
<tr>
<td>G18:O6:K1</td>
<td>2.68±0.01</td>
<td>2.69±0.02</td>
<td>2.68±0.04</td>
<td>2.67±0.01</td>
</tr>
<tr>
<td>G5:O6:K2</td>
<td>2.81±0.03</td>
<td>2.82±0.02</td>
<td>2.77±0.04</td>
<td>2.79±0.02</td>
</tr>
<tr>
<td>G9:O6:K2</td>
<td>2.87±0.02</td>
<td>2.81±0.02</td>
<td>2.87±0.02</td>
<td>2.83±0.01</td>
</tr>
<tr>
<td>G14:O6:K2</td>
<td>2.79±0.01</td>
<td>2.78±0.01</td>
<td>2.79±0.00</td>
<td>2.81±0.02</td>
</tr>
<tr>
<td>G18:O6:K2</td>
<td>2.82±0.01</td>
<td>2.76±0.02</td>
<td>2.82±0.02</td>
<td>2.84±0.01</td>
</tr>
<tr>
<td>G6:O6:K2</td>
<td>2.65±0.01</td>
<td>2.64±0.01</td>
<td>2.64±0.02</td>
<td>2.64±0.01</td>
</tr>
<tr>
<td>G10:O6:K2</td>
<td>2.66±0.01</td>
<td>2.64±0.02</td>
<td>2.65±0.01</td>
<td>2.65±0.01</td>
</tr>
<tr>
<td>G15:O6:K2</td>
<td>2.67±0.01</td>
<td>2.65±0.03</td>
<td>2.67±0.02</td>
<td>2.65±0.03</td>
</tr>
<tr>
<td>G19:O6:K2</td>
<td>2.66±0.01</td>
<td>2.66±0.02</td>
<td>2.67±0.02</td>
<td>2.65±0.02</td>
</tr>
</tbody>
</table>

*Note.* G4 layer colors (top (red); middle (green); bottom (blue)).
Table A6A

Pu22 Guanine Tetrads Oxygen-to-Oxygen Distance in Water

<table>
<thead>
<tr>
<th>Residue:O6-Residue:O6</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6-G8:O6</td>
<td>3.28±0.01</td>
<td>3.27±0.01</td>
<td>3.27±0.03</td>
<td>3.24±0.02</td>
</tr>
<tr>
<td>G8:O6-G13:O6</td>
<td>3.24±0.02</td>
<td>3.24±0.01</td>
<td>3.25±0.03</td>
<td>3.12±0.02</td>
</tr>
<tr>
<td>G13:O6-G17:O6</td>
<td>3.25±0.02</td>
<td>3.17±0.02</td>
<td>3.29±0.01</td>
<td>3.22±0.00</td>
</tr>
<tr>
<td>G17:O6-G4:O6</td>
<td>3.25±0.02</td>
<td>3.31±0.01</td>
<td>3.27±0.02</td>
<td>3.25±0.02</td>
</tr>
<tr>
<td>G5:O6-G9:O6</td>
<td>3.31±0.01</td>
<td>3.23±0.01</td>
<td>3.31±0.04</td>
<td>3.39±0.02</td>
</tr>
<tr>
<td>G9:O6-G14:O6</td>
<td>3.32±0.01</td>
<td>3.41±0.03</td>
<td>3.31±0.02</td>
<td>3.50±0.04</td>
</tr>
<tr>
<td>G14:O6-G18:O6</td>
<td>3.33±0.02</td>
<td>3.36±0.02</td>
<td>3.34±0.03</td>
<td>3.55±0.07</td>
</tr>
<tr>
<td>G18:O6-G5:O6</td>
<td>3.30±0.01</td>
<td>3.27±0.01</td>
<td>3.26±0.01</td>
<td>3.31±0.02</td>
</tr>
<tr>
<td>G6:O6-G10:O6</td>
<td>2.92±0.01</td>
<td>2.91±0.01</td>
<td>2.94±0.01</td>
<td>2.95±0.01</td>
</tr>
<tr>
<td>G10:O6-G15:O6</td>
<td>2.93±0.00</td>
<td>2.95±0.01</td>
<td>2.94±0.01</td>
<td>2.95±0.01</td>
</tr>
<tr>
<td>G15:O6-G19:O6</td>
<td>2.92±0.01</td>
<td>2.91±0.02</td>
<td>2.92±0.02</td>
<td>2.96±0.01</td>
</tr>
<tr>
<td>G19:O6-G6:O6</td>
<td>3.25±0.02</td>
<td>3.31±0.01</td>
<td>3.27±0.02</td>
<td>3.25±0.02</td>
</tr>
<tr>
<td>G4:O6-G5:O6</td>
<td>3.22±0.02</td>
<td>3.19±0.01</td>
<td>3.19±0.01</td>
<td>3.19±0.01</td>
</tr>
<tr>
<td>G8:O6-G9:O6</td>
<td>3.33±0.01</td>
<td>3.45±0.02</td>
<td>3.31±0.03</td>
<td>3.25±0.01</td>
</tr>
<tr>
<td>G13:O6-G14:O6</td>
<td>3.25±0.01</td>
<td>3.27±0.01</td>
<td>3.25±0.01</td>
<td>3.26±0.02</td>
</tr>
<tr>
<td>G17:O6-G18:O6</td>
<td>3.34±0.01</td>
<td>3.46±0.02</td>
<td>3.37±0.03</td>
<td>3.28±0.01</td>
</tr>
<tr>
<td>G5:O6-G6:O6</td>
<td>3.29±0.01</td>
<td>3.33±0.02</td>
<td>3.26±0.02</td>
<td>3.23±0.01</td>
</tr>
<tr>
<td>G9:O6-G10:O6</td>
<td>3.31±0.01</td>
<td>3.34±0.01</td>
<td>3.28±0.02</td>
<td>3.25±0.02</td>
</tr>
<tr>
<td>G14:O6-G15:O6</td>
<td>3.32±0.01</td>
<td>3.40±0.01</td>
<td>3.33±0.01</td>
<td>3.29±0.01</td>
</tr>
<tr>
<td>G18:O6-G19:O6</td>
<td>3.31±0.03</td>
<td>3.28±0.02</td>
<td>3.31±0.01</td>
<td>3.30±0.01</td>
</tr>
</tbody>
</table>

*Note.* G4 layer colors (top (red); middle (green); bottom (blue)).
Table A6B

**Pu22 Guanine Tetrad Oxygen-to-Oxygen Distance in EMIM-Cl**

<table>
<thead>
<tr>
<th>Residue:O6-Residue:O6</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6-G8:O6</td>
<td>3.28±0.01</td>
<td>3.27±0.01</td>
<td>3.26±0.02</td>
<td>3.24±0.02</td>
</tr>
<tr>
<td>G8:O6-G13:O6</td>
<td>3.24±0.02</td>
<td>3.17±0.02</td>
<td>3.22±0.01</td>
<td>3.17±0.02</td>
</tr>
<tr>
<td>G13:O6-G17:O6</td>
<td>3.25±0.02</td>
<td>3.23±0.02</td>
<td>3.24±0.01</td>
<td>3.22±0.02</td>
</tr>
<tr>
<td>G17:O6-G4:O6</td>
<td>3.25±0.02</td>
<td>3.25±0.03</td>
<td>3.23±0.01</td>
<td>3.19±0.03</td>
</tr>
<tr>
<td>G5:O6-G9:O6</td>
<td>3.31±0.01</td>
<td>3.29±0.02</td>
<td>3.30±0.03</td>
<td>3.38±0.03</td>
</tr>
<tr>
<td>G9:O6-G14:O6</td>
<td>3.32±0.01</td>
<td>3.33±0.02</td>
<td>3.30±0.03</td>
<td>3.39±0.03</td>
</tr>
<tr>
<td>G14:O6-G18:O6</td>
<td>3.33±0.02</td>
<td>3.41±0.06</td>
<td>3.33±0.02</td>
<td>3.43±0.03</td>
</tr>
<tr>
<td>G18:O6-G5:O6</td>
<td>3.30±0.01</td>
<td>3.29±0.03</td>
<td>3.29±0.02</td>
<td>3.34±0.03</td>
</tr>
<tr>
<td>G6:O6-G10:O6</td>
<td>2.92±0.01</td>
<td>2.91±0.01</td>
<td>2.95±0.01</td>
<td>2.93±0.02</td>
</tr>
<tr>
<td>G10:O6-G15:O6</td>
<td>2.93±0.00</td>
<td>2.93±0.01</td>
<td>2.93±0.02</td>
<td>2.92±0.01</td>
</tr>
<tr>
<td>G15:O6-G19:O6</td>
<td>2.92±0.01</td>
<td>2.96±0.01</td>
<td>2.92±0.01</td>
<td>2.93±0.02</td>
</tr>
<tr>
<td>G19:O6-G6:O6</td>
<td>3.25±0.02</td>
<td>3.25±0.03</td>
<td>3.23±0.01</td>
<td>3.19±0.03</td>
</tr>
<tr>
<td>G4:O6-G5:O6</td>
<td>3.22±0.02</td>
<td>3.23±0.03</td>
<td>3.20±0.01</td>
<td>3.21±0.01</td>
</tr>
<tr>
<td>G8:O6-G9:O6</td>
<td>3.33±0.01</td>
<td>3.33±0.02</td>
<td>3.32±0.01</td>
<td>3.30±0.02</td>
</tr>
<tr>
<td>G13:O6-G14:O6</td>
<td>3.25±0.01</td>
<td>3.31±0.03</td>
<td>3.24±0.02</td>
<td>3.25±0.00</td>
</tr>
<tr>
<td>G17:O6-G18:O6</td>
<td>3.34±0.01</td>
<td>3.31±0.01</td>
<td>3.33±0.04</td>
<td>3.31±0.01</td>
</tr>
<tr>
<td>G5:O6-G6:O6</td>
<td>3.29±0.01</td>
<td>3.31±0.03</td>
<td>3.28±0.03</td>
<td>3.25±0.02</td>
</tr>
<tr>
<td>G9:O6-G10:O6</td>
<td>3.31±0.01</td>
<td>3.22±0.02</td>
<td>3.30±0.04</td>
<td>3.28±0.02</td>
</tr>
<tr>
<td>G14:O6-G15:O6</td>
<td>3.32±0.01</td>
<td>3.31±0.01</td>
<td>3.33±0.01</td>
<td>3.28±0.01</td>
</tr>
<tr>
<td>G18:O6-G19:O6</td>
<td>3.31±0.03</td>
<td>3.24±0.01</td>
<td>3.32±0.03</td>
<td>3.28±0.01</td>
</tr>
</tbody>
</table>

*Note.* G4 layer colors (top (red); middle (green); bottom (blue)).
### Table A6C

**Pu22 Guanine Tetrads Oxygen-to-Oxygen Distance in BMIM-Cl**

<table>
<thead>
<tr>
<th>Residue:O6-Residue:O6</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6-G8:O6</td>
<td>3.28±0.01</td>
<td>3.28±0.02</td>
<td>3.26±0.05</td>
<td>3.26±0.02</td>
</tr>
<tr>
<td>G8:O6-G13:O6</td>
<td>3.24±0.02</td>
<td>3.21±0.02</td>
<td>3.20±0.03</td>
<td>3.22±0.02</td>
</tr>
<tr>
<td>G13:O6-G17:O6</td>
<td>3.25±0.02</td>
<td>3.23±0.05</td>
<td>3.25±0.03</td>
<td>3.22±0.01</td>
</tr>
<tr>
<td>G17:O6-G4:O6</td>
<td>3.25±0.02</td>
<td>3.24±0.01</td>
<td>3.22±0.03</td>
<td>3.27±0.02</td>
</tr>
<tr>
<td>G5:O6-G9:O6</td>
<td>3.31±0.01</td>
<td>3.29±0.04</td>
<td>3.33±0.03</td>
<td>3.30±0.05</td>
</tr>
<tr>
<td>G9:O6-G14:O6</td>
<td>3.32±0.01</td>
<td>3.34±0.02</td>
<td>3.33±0.03</td>
<td>3.33±0.05</td>
</tr>
<tr>
<td>G14:O6-G18:O6</td>
<td>3.33±0.02</td>
<td>3.35±0.02</td>
<td>3.38±0.02</td>
<td>3.34±0.03</td>
</tr>
<tr>
<td>G18:O6-G5:O6</td>
<td>3.30±0.01</td>
<td>3.30±0.03</td>
<td>3.26±0.02</td>
<td>3.30±0.05</td>
</tr>
<tr>
<td>G6:O6-G10:O6</td>
<td>2.92±0.01</td>
<td>2.93±0.02</td>
<td>2.95±0.00</td>
<td>2.92±0.01</td>
</tr>
<tr>
<td>G10:O6-G15:O6</td>
<td>2.93±0.00</td>
<td>2.94±0.00</td>
<td>2.93±0.02</td>
<td>2.94±0.01</td>
</tr>
<tr>
<td>G15:O6-G19:O6</td>
<td>2.92±0.01</td>
<td>2.92±0.02</td>
<td>2.93±0.01</td>
<td>2.92±0.01</td>
</tr>
<tr>
<td>G19:O6-G6:O6</td>
<td>3.25±0.02</td>
<td>3.24±0.01</td>
<td>3.22±0.03</td>
<td>3.27±0.02</td>
</tr>
<tr>
<td>G4:O6-G5:O6</td>
<td>3.22±0.02</td>
<td>3.26±0.03</td>
<td>3.21±0.01</td>
<td>3.20±0.02</td>
</tr>
<tr>
<td>G8:O6-G9:O6</td>
<td>3.33±0.01</td>
<td>3.33±0.03</td>
<td>3.31±0.01</td>
<td>3.31±0.01</td>
</tr>
<tr>
<td>G13:O6-G14:O6</td>
<td>3.25±0.01</td>
<td>3.26±0.02</td>
<td>3.25±0.02</td>
<td>3.25±0.03</td>
</tr>
<tr>
<td>G17:O6-G18:O6</td>
<td>3.34±0.01</td>
<td>3.33±0.02</td>
<td>3.33±0.02</td>
<td>3.36±0.03</td>
</tr>
<tr>
<td>G5:O6-G6:O6</td>
<td>3.29±0.01</td>
<td>3.29±0.02</td>
<td>3.25±0.03</td>
<td>3.30±0.02</td>
</tr>
<tr>
<td>G9:O6-G10:O6</td>
<td>3.31±0.01</td>
<td>3.30±0.01</td>
<td>3.27±0.02</td>
<td>3.28±0.02</td>
</tr>
<tr>
<td>G14:O6-G15:O6</td>
<td>3.32±0.01</td>
<td>3.31±0.01</td>
<td>3.31±0.03</td>
<td>3.32±0.03</td>
</tr>
<tr>
<td>G18:O6-G19:O6</td>
<td>3.31±0.03</td>
<td>3.31±0.01</td>
<td>3.30±0.03</td>
<td>3.29±0.03</td>
</tr>
</tbody>
</table>

*Note.* G4 layer colors (top (red); middle (green); bottom (blue)).
### Table A6D

Pu22 Guanine Tetrad Oxygen-to-Oxygen Distance in HMIM-Cl

<table>
<thead>
<tr>
<th>Residue:O6-Residue:O6</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding*</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6-G8:O6</td>
<td>3.28±0.01</td>
<td>3.21±0.07</td>
<td>3.26±0.02</td>
<td></td>
</tr>
<tr>
<td>G8:O6-G13:O6</td>
<td>3.24±0.02</td>
<td>3.23±0.01</td>
<td>3.24±0.02</td>
<td></td>
</tr>
<tr>
<td>G13:O6-G17:O6</td>
<td>3.25±0.02</td>
<td>3.21±0.02</td>
<td>3.24±0.00</td>
<td></td>
</tr>
<tr>
<td>G17:O6-G4:O6</td>
<td>3.25±0.02</td>
<td>3.19±0.05</td>
<td>3.24±0.02</td>
<td></td>
</tr>
<tr>
<td>G5:O6-G9:O6</td>
<td>3.31±0.01</td>
<td>3.28±0.02</td>
<td>3.30±0.01</td>
<td></td>
</tr>
<tr>
<td>G9:O6-G14:O6</td>
<td>3.32±0.01</td>
<td>3.31±0.01</td>
<td>3.30±0.01</td>
<td></td>
</tr>
<tr>
<td>G14:O6-G18:O6</td>
<td>3.33±0.02</td>
<td>3.29±0.06</td>
<td>3.38±0.05</td>
<td></td>
</tr>
<tr>
<td>G18:O6-G5:O6</td>
<td>3.30±0.01</td>
<td>3.39±0.07</td>
<td>3.34±0.03</td>
<td></td>
</tr>
<tr>
<td>G6:O6-G10:O6</td>
<td>2.92±0.01</td>
<td>2.92±0.02</td>
<td>2.92±0.01</td>
<td></td>
</tr>
<tr>
<td>G10:O6-G15:O6</td>
<td>2.93±0.00</td>
<td>2.94±0.02</td>
<td>2.93±0.01</td>
<td></td>
</tr>
<tr>
<td>G15:O6-G19:O6</td>
<td>2.92±0.01</td>
<td>2.91±0.01</td>
<td>2.93±0.01</td>
<td></td>
</tr>
<tr>
<td>G19:O6-G6:O6</td>
<td>3.25±0.02</td>
<td>3.19±0.05</td>
<td>3.24±0.02</td>
<td></td>
</tr>
<tr>
<td>G4:O6-G5:O6</td>
<td>3.22±0.02</td>
<td>3.21±0.02</td>
<td>3.21±0.02</td>
<td></td>
</tr>
<tr>
<td>G8:O6-G9:O6</td>
<td>3.33±0.01</td>
<td>3.32±0.02</td>
<td>3.28±0.02</td>
<td></td>
</tr>
<tr>
<td>G13:O6-G14:O6</td>
<td>3.25±0.01</td>
<td>3.24±0.04</td>
<td>3.32±0.02</td>
<td></td>
</tr>
<tr>
<td>G17:O6-G18:O6</td>
<td>3.34±0.01</td>
<td>3.41±0.03</td>
<td>3.37±0.01</td>
<td></td>
</tr>
<tr>
<td>G5:O6-G6:O6</td>
<td>3.29±0.01</td>
<td>3.29±0.02</td>
<td>3.32±0.04</td>
<td></td>
</tr>
<tr>
<td>G9:O6-G10:O6</td>
<td>3.31±0.01</td>
<td>3.29±0.04</td>
<td>3.26±0.03</td>
<td></td>
</tr>
<tr>
<td>G14:O6-G15:O6</td>
<td>3.32±0.01</td>
<td>3.29±0.05</td>
<td>3.24±0.04</td>
<td></td>
</tr>
<tr>
<td>G18:O6-G19:O6</td>
<td>3.31±0.03</td>
<td>3.29±0.02</td>
<td>3.25±0.02</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* G4 layer colors (top (red); middle (green); bottom (blue)). Bottom binding was not observed for this system, thus no data is presented.
Table A6E

*Pu22 Guanine Tetrado Oxygen-to-Oxygen Distance in OMIM-Cl*

<table>
<thead>
<tr>
<th>Residue:O6-Residue:O6</th>
<th>Apo</th>
<th>Top Binding</th>
<th>Bottom Binding</th>
<th>Side Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4:O6-G8:O6</td>
<td>3.28±0.01</td>
<td>3.25±0.01</td>
<td>3.27±0.02</td>
<td>3.27±0.02</td>
</tr>
<tr>
<td>G8:O6-G13:O6</td>
<td>3.24±0.02</td>
<td>3.24±0.01</td>
<td>3.26±0.01</td>
<td>3.24±0.01</td>
</tr>
<tr>
<td>G13:O6-G17:O6</td>
<td>3.25±0.02</td>
<td>3.24±0.01</td>
<td>3.25±0.01</td>
<td>3.24±0.02</td>
</tr>
<tr>
<td>G17:O6-G4:O6</td>
<td>3.25±0.02</td>
<td>3.22±0.02</td>
<td>3.26±0.03</td>
<td>3.28±0.01</td>
</tr>
<tr>
<td>G5:O6-G9:O6</td>
<td>3.31±0.01</td>
<td>3.34±0.06</td>
<td>3.28±0.03</td>
<td>3.29±0.02</td>
</tr>
<tr>
<td>G9:O6-G14:O6</td>
<td>3.32±0.01</td>
<td>3.36±0.03</td>
<td>3.32±0.02</td>
<td>3.32±0.03</td>
</tr>
<tr>
<td>G14:O6-G18:O6</td>
<td>3.33±0.02</td>
<td>3.35±0.06</td>
<td>3.32±0.03</td>
<td>3.32±0.02</td>
</tr>
<tr>
<td>G18:O6-G5:O6</td>
<td>3.30±0.01</td>
<td>3.35±0.06</td>
<td>3.34±0.05</td>
<td>3.32±0.01</td>
</tr>
<tr>
<td>G6:O6-G10:O6</td>
<td>2.92±0.01</td>
<td>2.92±0.01</td>
<td>2.91±0.01</td>
<td>2.92±0.01</td>
</tr>
<tr>
<td>G10:O6-G15:O6</td>
<td>2.93±0.00</td>
<td>2.94±0.01</td>
<td>2.92±0.02</td>
<td>2.93±0.02</td>
</tr>
<tr>
<td>G15:O6-G19:O6</td>
<td>2.92±0.01</td>
<td>2.91±0.01</td>
<td>2.90±0.03</td>
<td>2.92±0.01</td>
</tr>
<tr>
<td>G19:O6-G6:O6</td>
<td>3.25±0.02</td>
<td>3.22±0.02</td>
<td>3.26±0.03</td>
<td>3.28±0.01</td>
</tr>
<tr>
<td>G4:O6-G5:O6</td>
<td>3.22±0.02</td>
<td>3.21±0.04</td>
<td>3.23±0.02</td>
<td>3.22±0.02</td>
</tr>
<tr>
<td>G8:O6-G9:O6</td>
<td>3.33±0.01</td>
<td>3.29±0.03</td>
<td>3.31±0.02</td>
<td>3.31±0.02</td>
</tr>
<tr>
<td>G13:O6-G14:O6</td>
<td>3.25±0.01</td>
<td>3.25±0.03</td>
<td>3.26±0.03</td>
<td>3.24±0.04</td>
</tr>
<tr>
<td>G17:O6-G18:O6</td>
<td>3.34±0.01</td>
<td>3.33±0.06</td>
<td>3.37±0.03</td>
<td>3.34±0.01</td>
</tr>
<tr>
<td>G5:O6-G6:O6</td>
<td>3.29±0.01</td>
<td>3.26±0.01</td>
<td>3.30±0.02</td>
<td>3.28±0.02</td>
</tr>
<tr>
<td>G9:O6-G10:O6</td>
<td>3.31±0.01</td>
<td>3.27±0.04</td>
<td>3.29±0.02</td>
<td>3.29±0.02</td>
</tr>
<tr>
<td>G14:O6-G15:O6</td>
<td>3.32±0.01</td>
<td>3.25±0.02</td>
<td>3.29±0.03</td>
<td>3.32±0.02</td>
</tr>
<tr>
<td>G18:O6-G19:O6</td>
<td>3.31±0.03</td>
<td>3.25±0.02</td>
<td>3.30±0.02</td>
<td>3.29±0.02</td>
</tr>
</tbody>
</table>

*Note.* G4 layer colors (top (red); middle (green); bottom (blue)).
Table A7

*Average Atom Contacts Between TMPyP4 and Pu22 in Water and IL*

<table>
<thead>
<tr>
<th></th>
<th>WAT</th>
<th>EMIM</th>
<th>BMIM</th>
<th>HMIM</th>
<th>OMIM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>34±1</td>
<td>34±1</td>
<td>25±1</td>
<td>19±1</td>
<td>15±1</td>
</tr>
</tbody>
</table>

*Note.* See Figure A4 for plot diagrams.
Figure A1

*Telomeric G4 -TMpyP4 (1:2) Complex*

*Note.* Structure PDB id: 2HRI.
Figure A2

Starting Positions of Apo-Form and TMPyP4-Bound Structures

| Apo-form | Unbound top-binding | Unbound bottom-binding |

Note. 5` and 3` ends of DNA and K+ ions are represented as red, blue and yellow spheres, respectively.
Figure A3

Averaged Receptor and Ligand RMSD Over Each Trajectory

**Note.** Receptor RMSD (orange, nucleic acid backbone) and ligand RMSD (blue, ligand heavy atoms). See Table A3.
Figure A4

Average Atom Contacts Between TMPyP4 and Pu22

Note. See Table A7.
Figure A5

Average Pu22 RMSF in Water and Each IL

Note. Water RMSF averaged over 30 trajectories and IL RMSF averaged over 40 trajectories (10 trajectories per IL). Non-G-tetrad residues are highlighted. See Table A3.
Guanine Tetrad Oxygen-K$^+$ Distance Parameters in Pu22

Note. Of Pu22 apo-form. (A&B) Oxygen (red, green and red balls) and K$^+$ cations (yellow ball). (C) Oxygen-K$^+$ Distances (Å). (D) Oxygen-Oxygen Distances (Å).
Figure A7

Electrostatic contact mapping of TMPyP4 to Pu22.

<table>
<thead>
<tr>
<th></th>
<th>SIDE</th>
<th>TOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAT</td>
<td><img src="image1" alt="WAT SIDE" /></td>
<td><img src="image2" alt="WAT TOP" /></td>
</tr>
<tr>
<td>EMIM</td>
<td><img src="image3" alt="EMIM SIDE" /></td>
<td><img src="image4" alt="EMIM TOP" /></td>
</tr>
<tr>
<td>BMIM</td>
<td><img src="image5" alt="BMIM SIDE" /></td>
<td><img src="image6" alt="BMIM TOP" /></td>
</tr>
<tr>
<td>HMIM</td>
<td><img src="image7" alt="HMIM SIDE" /></td>
<td><img src="image8" alt="HMIM TOP" /></td>
</tr>
<tr>
<td>OMIM</td>
<td><img src="image9" alt="OMIM SIDE" /></td>
<td><img src="image10" alt="OMIM TOP" /></td>
</tr>
</tbody>
</table>
Note. Red, white and blue colors represent negative, neutral and positively-charged surface regions of Pu22 and TMPyP4, respectively. TMPyP4 is represented in licorice. For IL, the charged nitrogen atom and terminal carbon atom in the hydrophobic chain are represented as blue and red balls, respectively. 5’ and 3’ ends of Pu22 are represented as larger red and blue balls, respectively.
**Figure A8**

*Bound Cluster Abundance Structures*

<table>
<thead>
<tr>
<th></th>
<th>WAT</th>
<th>EMIM</th>
<th>BMIM</th>
<th>HMIM</th>
<th>OMIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOP</td>
<td>48.0%</td>
<td>31.4%</td>
<td>16.8%</td>
<td>23.3%</td>
<td>8.7%</td>
</tr>
<tr>
<td>BOTTOM</td>
<td>23.3%</td>
<td>10.7%</td>
<td>14.6%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>SIDE</td>
<td>6.5%</td>
<td>9.7%</td>
<td>7.7%</td>
<td>5.3%</td>
<td>2.9%</td>
</tr>
</tbody>
</table>

*Note.* Unbound clusters not shown (16.4%, 48.2%, 60.9%, 71.4% and 84.6% for water, EMIM-Cl, BMIM-Cl, HMIM-Cl and OMIM-Cl, respectively). TMPyP4 (blue sticks) and K+ cations (yellow balls). 5’ end (red ball) and 3’ end (blue ball).
Figure A9

Decomposition of IL Atom Contacts with Pu22

Note. Distance cutoff for atom contacts is 2.5 Å. See Table 7.
Figure 10A

Pu22-TMPyP4 Receptor RMSD in Water
Figure A10B

Pu22-TMPyP4 Receptor RMSD in EMIM-Cl
Figure A10C

Pu22-TMPyP4 Receptor RMSD in BMIM-Cl
Figure A10d

Pu22-TMPyP4 Receptor RMSD in HMIM-Cl
Figure A10e

Pu22-TMPyP4 Receptor RMSD in OMIM-Cl
Figure A11a

Pu22-TMPyP4 Ligand RMSD in Water
Figure A11b

Pu22-TMPyP4 Ligand RMSD in EMIM-Cl
Figure A11c

Pu22-TMPyP4 Ligand RMSD in BMIM-Cl
Figure A11d

Pu22-TMPyP4 Ligand RMSD in HMIM-Cl
Figure A11e

Pu22-TMPyP4 Ligand RMSD in OMIM-Cl
Figure A12a

Total Atom Contacts Between EMIM and Pu22
Figure A12b

*Total Atom Contacts Between BMIM and Pu22*
Figure A12c

Total Atom Contacts Between HMIM and Pu22
Figure A12d

Total Atom Contacts Between OMIM and Pu22
Figure A13a

Last Snapshot Structures of TMPyP4 Binding to Pu22 in Water
Note. T2 and T3 represent trajectories with TMPyP4 initially placed above or below 1XAV, respectively.
Figure A13b

Last Snapshot Structures of TMPyP4 Binding to Pu22 in EMIM-Cl
Note. T2 and T3 represent trajectories with TMPyP4 initially placed above or below 1XAV, respectively.
Figure A13c

*Last Snapshot Structures of TMPyP4 Binding to Pu22 in BMIM-Cl*

<table>
<thead>
<tr>
<th>T2</th>
<th>01 – UNBOUND</th>
<th>02 – TOP</th>
<th>03 – SIDE</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Diagram" /></td>
<td><img src="image2.png" alt="Diagram" /></td>
<td><img src="image3.png" alt="Diagram" /></td>
<td><img src="image4.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>
Note. T2 and T3 represent trajectories with TMPyP4 initially placed above or below 1XAV, respectively.
Figure A13d

_Last Snapshot Structures of TMPyP4 Binding to Pu22 in HMIM-Cl_

<table>
<thead>
<tr>
<th>T2</th>
<th>01 – SIDE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>![Image 1]</td>
</tr>
<tr>
<td></td>
<td>![Image 2]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>02 – SIDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image 3]</td>
</tr>
<tr>
<td>![Image 4]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>04 – SIDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image 5]</td>
</tr>
<tr>
<td>![Image 6]</td>
</tr>
</tbody>
</table>
Note. T2 and T3 represent trajectories with TMPyP4 initially placed above or below 1XAV, respectively.
Figure A13e

*Last Snapshot Structures of TMPyP4 Binding to Pu22 in OMIM-Cl*
Note. T2 and T3 represent trajectories with TMPyP4 initially placed above or below 1XAV, respectively.
Figure A14a

Representative Trajectories of TMPyP4 Binding Modes in Water
Figure A14b

Representative Trajectories of TMPyP4 Binding Modes in EMIM-Cl
Figure A14c

Representative Trajectories of TMPyP4 Binding Modes in BMIM-Cl
Figure A14d

Representative Trajectories of TMPyP4 Binding Modes in HMIM-Cl
Figure A14e

Representative Trajectories of TMPyP4 Binding Modes in OMIM-Cl
Figure A15a

Cluster Abundances of Major TMPyP4 Binding Poses in Water
Figure A15b

Cluster Abundances of Major TMPyP4 Binding Poses in EMIM-Cl

TOP – 56%
Figure A15c

Cluster Abundances of Major TMPyP4 Binding Poses in BMIM-Cl

TOP – 38%

BOT – 25%
Figure A15d

Cluster Abundances of Major TMPyP4 Binding Poses in HMIM-Cl
Figure A15e

Cluster Abundances of Major TMPyP4 Binding Poses in OMIM-Cl

TOP – 33%

BOT – 11%
Figure A16

Ligand-Interaction Diagrams of TMPyP4 Binding Poses in Water
Figure A17a

Decomposition of EMIM Atom Contacts to Pu22

Note. Total (blue), polar (red) and apolar (green) contacts.
Figure A17b

*Decomposition of BMIM Atom Contacts to Pu22*

*Note. Total (blue), polar (red) and apolar (green) contacts.*
Figure A17c

Decomposition of HMIM Atom Contacts to Pu22

*Note.* Total (blue), polar (red) and apolar (green) contacts.
Figure A17d

Decomposition of OMIM Atom Contacts to Pu22

*Note.* Total (blue), polar (red) and apolar (green) contacts.
Figure A18a

*MOL2 File of AMBER GAFF2 Force Field of EMIM Cation*

```mol2
@<TRIPOS>MOLECULE
EMI
  19 19 1 0 0
SMALL
rc
@<TRIPOS>ATOM
  1 N1     -0.6120  0.0180  -0.3140 na 1 EMI  0.008630
  2 C1     0.4090  -0.8010  -0.2000 cc 1 EMI  0.021561
  3 H1     0.3570  -1.8650  -0.2920 h5 1 EMI  0.230735
  4 N2     1.5070  -0.1170  0.0400 na 1 EMI  0.075521
  5 C2     1.1740  1.2200  0.0830 cc 1 EMI  -0.146036
  6 H2     1.9030  1.9800  0.2600 h4 1 EMI  0.236232
  7 C3    -0.1470  1.3030  0.1360 cd 1 EMI  -0.195966
  8 H3    -0.7930  2.1520  0.1870 h4 1 EMI  0.260468
  9 C4     2.8530  0.6680  0.2220 c3 1 EMI  -0.153548
 10 H41    3.5060  -0.2740  1.5420 h1 1 EMI  0.125860
 11 H42    2.8050  -1.7420  0.1370 h1 1 EMI  0.125860
 12 H43    3.2190  -0.3990  1.5020 h1 1 EMI  0.125860
 13 C5    -2.0160  0.3690  0.5600 c3 1 EMI  -0.011914
 14 H51   -2.3820  0.2650  1.3550 h1 1 EMI  0.097919
 15 H52   -2.0000  1.3840  0.9300 h1 1 EMI  0.097919
 16 C6    -2.8710  0.2490  0.6940 c3 1 EMI  0.049271
 17 H61   -3.8870  0.4400  0.4590 hc 1 EMI  0.049271
 18 H62   -2.8960  0.7700  1.0630 hc 1 EMI  0.049271
 19 H63   -2.5030  0.8940  1.4830 hc 1 EMI  0.049271
@<TRIPOS>BOND
  1  1  2  2
  2  1  7  1
  3  1 13  1
  4  2  3  1
  5  2  4  1
  6  4  5  1
  7  4  9  1
  8  5  6  1
  9  5  7  2
 10  7  8  1
 11  9 10  1
 12  9 11  1
 13  9 12  1
 14 13  14  1
 15 13  15  1
 16 13  16  1
```

170
17 16 17 1
18 16 18 1
19 16 19 1
@<TRIPOS>SUBSTRUCTURE
1 EMI 1 TEMP 0 **** **** 0 ROOT
Figure A18b

*MOL2 File of AMBER GAFF2 Force Field of BMIM Cation*

@<TRIPOS>MOLECULE

BMI
    25  25  1  0  0
SMALL
rc

@<TRIPOS>ATOM

1  N1  2.5180 -0.2110 -0.1500 na  1  BMI   0.068344
2  C1  2.2980  1.1380 -0.3250 cc  1  BMI  -0.144228
3  H1  3.0550  2.2930  0.0820 h4  1  BMI   0.236564
4  C2  1.0350  1.3820  0.0570 cd  1  BMI  -0.204915
5  H2  0.4780  2.2930  0.0820 h4  1  BMI   0.050024
6  N2  0.4940  0.1830  0.4660 na  1  BMI   0.050024
7  C3  1.4140 -0.7450  0.3250 cc  1  BMI   0.003407
8  H3  1.2850 -1.7790  0.5630 h5  1  BMI   0.236564
9  C4  3.7710 -0.9160 -0.4370 c3  1  BMI  -0.135806
10 H41 3.6450 -1.9620 -0.2040 h1  1  BMI   0.120974
11 H42 4.5610 -0.5040  0.1730 h1  1  BMI   0.120974
12 H43 4.0110 -0.8050 -1.4830 h1  1  BMI   0.120974
13 C5  0.8890 -0.0260  0.9430 c3  1  BMI  -0.035423
14 H51 0.9000 -0.9750  1.4630 h1  1  BMI   0.085340
15 H52 1.0890  0.7480  1.6710 h1  1  BMI   0.085340
16 C6  1.9040 -0.0060 -0.1970 c3  1  BMI   0.006080
17 H61 -1.8410  0.9440 -0.7210 hc  1  BMI   0.025831
18 H62 -1.6480 -0.7800 -0.9150 hc  1  BMI   0.025831
19 C7  -3.3300 -0.2190  0.3180 c3  1  BMI   0.018698
20 H71 -3.3860 -1.1670  0.8470 hc  1  BMI   0.022863
21 H72 -3.5720  0.5540  1.0430 hc  1  BMI   0.022863
22 C8  -4.3620 -0.2030 -0.8090 c3  1  BMI  -0.095267
23 H81 -5.3600 -0.3560 -0.4160 hc  1  BMI   0.036119
24 H82 -4.1660 -0.9880 -1.5330 hc  1  BMI   0.036119
25 H83 -4.3550  0.7470 -1.3350 hc  1  BMI   0.036119

@<TRIPOS>BOND

1  1  2  1
2  1  7  2
3  1  9  1
4  2  3  1
5  2  4  2
6  4  5  1
7  4  6  1
@<TRIPOS>SUBSTRUCTURE
  1 BMI       1 TEMP       0 **** ****       0 ROOT
Figure A18c

MOL2 File of AMBER GAFF2 Force Field of HMIM Cation

@<TRIPOS>MOLECULE

HMI
  31 31 1 0 0
SMALL
rc

@<TRIPOS>ATOM

1  N1   1.6580  0.2780  0.4560 na   1 HMI   0.041636
2  C1   2.5320 -0.7020  0.4050 cc   1 HMI   0.008087
3  H1   2.3870 -1.6770  0.8190 h5   1 HMI   0.237705
4  N2   3.6140 -0.3030 -0.2280 na   1 HMI   0.062298
5  C2   3.4270  1.0080 -0.6070 cc   1 HMI  -0.124091
6  H2   4.1750  1.5570 -1.1370 h4   1 HMI   0.229224
7  C3   2.2070  1.3690 -0.1820 cd   1 HMI  -0.224479
8  H3   1.6840  2.2960 -0.2690 h4   1 HMI   0.268494
9  C4   4.8160 -1.1020 -0.4830 c3   1 HMI  -0.136071
10 H41  5.6650 -0.6340 -0.0070 h1   1 HMI   0.120441
11 H42  4.9790 -1.1740 -1.5480 h1   1 HMI   0.120441
12 H43  4.6720 -2.0900 -0.0730 h1   1 HMI   0.120441
13 C5   0.3100  0.2130  1.0570 c3   1 HMI  -0.030796
14 H51  0.3060 -0.6370  1.7250 h1   1 HMI   0.086961
15 H52  0.1920  1.1030  1.6590 h1   1 HMI   0.086961
16 C6   -0.7880  0.0950  0.0030 c3   1 HMI   0.019647
17 H61  -0.7330  0.9440 -0.6730 hc   1 HMI   0.020476
18 H62  -0.6140 -0.7960 -0.5940 hc   1 HMI   0.020476
19 C7   -2.1770  0.0340  0.6460 c3   1 HMI  -0.006702
20 H71  -2.2240 -0.8130  1.3270 hc   1 HMI   0.013924
21 H72  -2.3360  0.9250  1.2490 hc   1 HMI   0.013924
22 C8   -3.2980 -0.0840 -0.3890 c3   1 HMI  -0.012792
23 H81  -3.2500  0.7630 -1.0710 hc   1 HMI   0.011511
24 H82  -3.1370 -0.9750 -0.9940 hc   1 HMI   0.011511
25 C9   -4.6900 -0.1470  0.2440 c3   1 HMI   0.029319
26 H91  -4.7380 -0.9930  0.9250 hc   1 HMI   0.006576
27 H92  -4.8500  0.7430  0.8480 hc   1 HMI   0.006576
28 C10  -5.8060 -0.2650 -0.7930 c3   1 HMI  -0.068757
29 H101 -6.7780 -0.3060 -0.3140 hc   1 HMI   0.022353
30 H102 -5.6930 -1.1640 -1.3920 hc   1 HMI   0.022353
31 H103 -5.8060  0.5860 -1.4690 hc   1 HMI   0.022353

@<TRIPOS>BOND

1  1  2  2
2  1  7  1
3  1  13  1
4  2  3  1
5  2  4  1
6  4  5  1
7  4  9  1
8  5  6  1
9  5  7  2
10  7  8  1
11  9  10  1
12  9  11  1
13  9  12  1
14  13  14  1
15  13  15  1
16  13  16  1
17  16  17  1
18  16  18  1
19  16  19  1
20  19  20  1
21  19  21  1
22  19  22  1
23  22  23  1
24  22  24  1
25  22  25  1
26  25  26  1
27  25  27  1
28  25  28  1
29  28  29  1
30  28  30  1
31  28  31  1

@<TRIPOS>SUBSTRUCTURE
  1 HMI  1 TEMP  0 **** ****  0 ROOT
### Figure A18d

**MOL2 File of AMBER GAFF2 Force Field of OMIM Cation**

```mol2
@<TRIPOS>MOLECULE

OMI
37 37 1 0 0
SMALL
rc

@<TRIPOS>ATOM

<table>
<thead>
<tr>
<th>Atom</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Charge</th>
<th>Mass</th>
<th>Type</th>
<th>Solvent</th>
<th>Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>3.918</td>
<td>-0.970</td>
<td>-0.232</td>
<td>na</td>
<td></td>
<td>OMI</td>
<td>0.041800</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>3.396</td>
<td>0.158</td>
<td>-0.661</td>
<td>cc</td>
<td></td>
<td>OMI</td>
<td>0.044210</td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>3.560</td>
<td>0.581</td>
<td>-1.630</td>
<td>h5</td>
<td></td>
<td>OMI</td>
<td>0.223946</td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td>2.651</td>
<td>0.693</td>
<td>0.280</td>
<td>na</td>
<td></td>
<td>OMI</td>
<td>0.028435</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>2.695</td>
<td>-0.137</td>
<td>1.377</td>
<td>cc</td>
<td></td>
<td>OMI</td>
<td>-0.193522</td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td>2.165</td>
<td>0.086</td>
<td>2.277</td>
<td>h4</td>
<td></td>
<td>OMI</td>
<td>0.236738</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>3.485</td>
<td>-1.175</td>
<td>1.060</td>
<td>cd</td>
<td></td>
<td>OMI</td>
<td>-0.111976</td>
<td></td>
</tr>
<tr>
<td>H3</td>
<td>3.779</td>
<td>-2.030</td>
<td>1.627</td>
<td>h4</td>
<td></td>
<td>OMI</td>
<td>0.222528</td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>4.813</td>
<td>-1.855</td>
<td>-0.982</td>
<td>c3</td>
<td></td>
<td>OMI</td>
<td>-0.153200</td>
<td></td>
</tr>
<tr>
<td>H41</td>
<td>4.358</td>
<td>-2.829</td>
<td>-1.074</td>
<td>h1</td>
<td></td>
<td>OMI</td>
<td>0.126342</td>
<td></td>
</tr>
<tr>
<td>H42</td>
<td>4.974</td>
<td>-1.438</td>
<td>-1.965</td>
<td>h1</td>
<td></td>
<td>OMI</td>
<td>0.126342</td>
<td></td>
</tr>
<tr>
<td>H43</td>
<td>5.757</td>
<td>-1.931</td>
<td>-0.464</td>
<td>h1</td>
<td></td>
<td>OMI</td>
<td>0.126342</td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>1.888</td>
<td>1.957</td>
<td>0.163</td>
<td>c3</td>
<td></td>
<td>OMI</td>
<td>-0.149880</td>
<td></td>
</tr>
<tr>
<td>H51</td>
<td>2.513</td>
<td>2.650</td>
<td>-0.384</td>
<td>h1</td>
<td></td>
<td>OMI</td>
<td>0.135849</td>
<td></td>
</tr>
<tr>
<td>H52</td>
<td>1.780</td>
<td>2.338</td>
<td>1.168</td>
<td>h1</td>
<td></td>
<td>OMI</td>
<td>0.135849</td>
<td></td>
</tr>
<tr>
<td>C6</td>
<td>0.532</td>
<td>1.783</td>
<td>-0.517</td>
<td>c3</td>
<td></td>
<td>OMI</td>
<td>0.005999</td>
<td></td>
</tr>
<tr>
<td>H61</td>
<td>0.680</td>
<td>1.391</td>
<td>-1.521</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>0.027959</td>
<td></td>
</tr>
<tr>
<td>H62</td>
<td>0.122</td>
<td>2.782</td>
<td>-0.638</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>0.027959</td>
<td></td>
</tr>
<tr>
<td>C7</td>
<td>-0.464</td>
<td>0.909</td>
<td>0.250</td>
<td>c3</td>
<td></td>
<td>OMI</td>
<td>0.043733</td>
<td></td>
</tr>
<tr>
<td>H71</td>
<td>-0.580</td>
<td>1.300</td>
<td>1.259</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>0.010847</td>
<td></td>
</tr>
<tr>
<td>H72</td>
<td>-0.073</td>
<td>-0.101</td>
<td>0.349</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>0.010847</td>
<td></td>
</tr>
<tr>
<td>C8</td>
<td>-1.834</td>
<td>0.850</td>
<td>-0.430</td>
<td>c3</td>
<td></td>
<td>OMI</td>
<td>-0.007381</td>
<td></td>
</tr>
<tr>
<td>H81</td>
<td>-1.716</td>
<td>0.471</td>
<td>-1.444</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>0.009885</td>
<td></td>
</tr>
<tr>
<td>H82</td>
<td>-2.230</td>
<td>1.858</td>
<td>-0.524</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>0.009885</td>
<td></td>
</tr>
<tr>
<td>C9</td>
<td>-2.838</td>
<td>-0.023</td>
<td>0.325</td>
<td>c3</td>
<td></td>
<td>OMI</td>
<td>-0.018253</td>
<td></td>
</tr>
<tr>
<td>H91</td>
<td>-2.952</td>
<td>0.356</td>
<td>1.339</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>-0.006873</td>
<td></td>
</tr>
<tr>
<td>H92</td>
<td>-2.440</td>
<td>-1.032</td>
<td>0.418</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>-0.006873</td>
<td></td>
</tr>
<tr>
<td>C10</td>
<td>-4.210</td>
<td>-0.080</td>
<td>-0.351</td>
<td>c3</td>
<td></td>
<td>OMI</td>
<td>0.025258</td>
<td></td>
</tr>
<tr>
<td>H101</td>
<td>-4.096</td>
<td>-0.457</td>
<td>-1.366</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>0.003940</td>
<td></td>
</tr>
<tr>
<td>H102</td>
<td>-4.607</td>
<td>0.929</td>
<td>-0.444</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>0.003940</td>
<td></td>
</tr>
<tr>
<td>C11</td>
<td>-5.216</td>
<td>-0.952</td>
<td>0.402</td>
<td>c3</td>
<td></td>
<td>OMI</td>
<td>0.024815</td>
<td></td>
</tr>
<tr>
<td>H111</td>
<td>-5.331</td>
<td>-0.574</td>
<td>1.416</td>
<td>hc</td>
<td></td>
<td>OMI</td>
<td>0.005960</td>
<td></td>
</tr>
</tbody>
</table>
```

176
<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>H12</td>
<td>-4.8200</td>
<td>-1.9610</td>
<td>0.4940</td>
<td>hc</td>
<td>1</td>
<td>OMI</td>
<td>0.005960</td>
</tr>
<tr>
<td>34</td>
<td>C12</td>
<td>-6.5840</td>
<td>-1.0030</td>
<td>-0.2780</td>
<td>c3</td>
<td>1</td>
<td>OMI</td>
<td>-0.077680</td>
</tr>
<tr>
<td>35</td>
<td>H121</td>
<td>-7.2740</td>
<td>-1.6280</td>
<td>0.2790</td>
<td>hc</td>
<td>1</td>
<td>OMI</td>
<td>0.020091</td>
</tr>
<tr>
<td>36</td>
<td>H122</td>
<td>-7.0220</td>
<td>-0.0130</td>
<td>-0.3540</td>
<td>hc</td>
<td>1</td>
<td>OMI</td>
<td>0.020091</td>
</tr>
<tr>
<td>37</td>
<td>H123</td>
<td>-6.5070</td>
<td>-1.4090</td>
<td>-1.2830</td>
<td>hc</td>
<td>1</td>
<td>OMI</td>
<td>0.020091</td>
</tr>
</tbody>
</table>

@<TRIPOS>BOND
1 1 2 2
2 1 7 1
3 1 9 1
4 2 3 1
5 2 4 1
6 4 5 1
7 4 13 1
8 5 6 1
9 5 7 2
10 7 8 1
11 9 10 1
12 9 11 1
13 9 12 1
14 13 14 1
15 13 15 1
16 13 16 1
17 16 17 1
18 16 18 1
19 16 19 1
20 19 20 1
21 19 21 1
22 19 22 1
23 22 23 1
24 22 24 1
25 22 25 1
26 25 26 1
27 25 27 1
28 25 28 1
29 28 29 1
30 28 30 1
31 28 31 1
32 31 32 1
33 31 33 1
34 31 34 1
35 34 35 1
36 34 36 1
37 34 37 1

@<TRIPOS>SUBSTRUCTURE
1 OMI 1 TEMP 0 **** **** 0 ROOT
Appendix B

Probing the Thermodynamic Destabilization of Azurin by Two Different Tetramethylguanidinium Amino Acid Ionic Liquids Using Molecular Dynamics Simulations

Table B1

*Simulation System Parameters*

<table>
<thead>
<tr>
<th>ID</th>
<th>Protein/Solvent</th>
<th># of run</th>
<th># of Water molecules</th>
<th># of TMG Ions</th>
<th># of ASP Ions</th>
<th># of SER Ions</th>
<th>Box Size (Å)*</th>
<th>NPT eq. (ns)</th>
<th>NVT (ns)</th>
<th>Total time (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Water</td>
<td>1</td>
<td>5394</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>64.9</td>
<td>1</td>
<td>999</td>
<td>1000</td>
</tr>
<tr>
<td>2</td>
<td>TMG-Asp</td>
<td>1</td>
<td>5980</td>
<td>48</td>
<td>47</td>
<td>0</td>
<td>72.0</td>
<td>1</td>
<td>999</td>
<td>1000</td>
</tr>
<tr>
<td>3</td>
<td>TMG-Ser</td>
<td>1</td>
<td>5866</td>
<td>48</td>
<td>0</td>
<td>47</td>
<td>71.3</td>
<td>1</td>
<td>999</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>Folded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Water</td>
<td>1</td>
<td>8859</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>75.4</td>
<td>1</td>
<td>999</td>
<td>1000</td>
</tr>
<tr>
<td>5</td>
<td>TMG-Asp</td>
<td>1</td>
<td>10927</td>
<td>48</td>
<td>47</td>
<td>0</td>
<td>72.0</td>
<td>1</td>
<td>999</td>
<td>1000</td>
</tr>
<tr>
<td>6</td>
<td>TMG-Ser</td>
<td>1</td>
<td>11261</td>
<td>48</td>
<td>0</td>
<td>47</td>
<td>71.3</td>
<td>1</td>
<td>999</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>Unfolded</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Triclinic box equivalent to the true truncated octahedron box shape.
Figure B1

**Secondary Structure of Azurin Protein**
Figure B2

*MOL2 File of AMBER GAFF2 Force Field of TMG Cation*

```tripos
@<TRIPOS>MOLECULE
TMG
22 21 1 0 0
SMALL
rc

@<TRIPOS>ATOM
1 C1 -1.3380 -1.4490  0.6900 c3  1 TMG  -0.097721
2 H11 -0.4530 -1.7080  1.2480 h1  1 TMG  0.089306
3 H12 -2.1630 -1.3580  1.3840 h1  1 TMG  0.089306
4 H13 -1.5640 -2.2310 -0.0220 h1  1 TMG  0.089306
5 N1  -1.1630 -0.1590  0.0150 nh  1 TMG  -0.013680
6 C2  -2.4050  0.4990 -0.3930 c3  1 TMG  -0.212516
 7 H21 -2.2190  1.2000 -1.1930 h1  1 TMG  0.121231
 8 H22 -3.0790 -0.2600 -0.7610 h1  1 TMG  0.121230
 9 H23 -2.8810  1.0050  0.4410 h1  1 TMG  0.121230
10 C3  -0.0000  0.4870 -0.0000 cz  1 TMG  0.274795
11 N2  -0.0000  1.8180 -0.0000 nh  1 TMG  -0.652146
12 H2  0.8100  2.3370 -0.2540 hn  1 TMG  0.383347
13 H3  -0.8100  2.3370  0.2540 hn  1 TMG  0.383347
14 N3  1.1630 -0.1590 -0.0150 nh  1 TMG  -0.022725
15 C4  1.3380 -1.4490 -0.6900 c3  1 TMG  -0.109503
16 H41  2.1630 -1.3570 -1.3840 h1  1 TMG  0.092823
17 H42  1.5640 -2.2310  0.0220 h1  1 TMG  0.092823
18 H43  0.4530 -1.7080 -1.2480 h1  1 TMG  0.092823
19 C5  2.4050  0.4990  0.3930 c3  1 TMG  -0.201842
20 H51  3.0790 -0.2600  0.7610 h1  1 TMG  0.119522
21 H52  2.8810  1.0050 -0.4410 h1  1 TMG  0.119522
22 H53  2.2190  1.2000  1.1930 h1  1 TMG  0.119522

@<TRIPOS>BOND
  1  1  2 1
  2  1  3 1
  3  1  4 1
  4  1  5 1
  5  5  6 1
  6  5 10 2
  7  6  7 1
  8  6  8 1
  9  6  9 1
 10 10 11 1
 11 10 14 1
```

180
@<TRIPOS>SUBSTRUCTURE
  1 TMG   1 TEMP   0 **** **** 0 ROOT
**Figure B3**

*MOL2 File of AMBER GAFF2 Force Field of Serine Anion*

```<TRIPOS>ATOM
1  N   0.7400  1.8470  0.0580 n3     1 ASE -1.064631
2  H1  0.7910  1.9190 -0.9450 hn     1 ASE  0.360084
3  H   1.6940  1.6720  0.3180 hn     1 ASE  0.360084
4  CA  0.0240  0.6100  0.3690 c3     1 ASE  0.259781
5  HA -0.1440  0.5880  1.4450 h1     1 ASE  0.035054
6  C   0.8250 -0.6790  0.0230 c      1 ASE  0.726077
7  O   2.0350 -0.5410 -0.1400 o      1 ASE -0.759762
8  CB -1.3550  0.6390 -0.2920 c3     1 ASE  0.063054
9  HB2 -1.8600  1.5630 -0.0240 h1     1 ASE  0.043644
10 HB3 -1.2260  0.6490 -1.3790 h1     1 ASE  0.043644
11 OG -2.1660 -0.4270  0.0990 oh     1 ASE -0.694529
12 HG -1.5720 -1.1850  0.0890 ho     1 ASE  0.413741
13 O1  0.1530 -1.1850 -0.0230 o      1 ASE -0.786241
</TRIPOS>ATOM```

```<TRIPOS>BOND
1 1 2 1
2 1 3 1
3 1 4 1
4 4 5 1
5 4 6 1
6 4 8 1
7 6 7 1
8 6 13 1
9 8 9 1
10 8 10 1
11 8 11 1
12 11 12 1
</TRIPOS>BOND```

```<TRIPOS>SUBSTRUCTURE
1 ASE 1 TEMP 0 **** **** 0 ROOT
</TRIPOS>SUBSTRUCTURE```
Figure B4

*MOL2 File of AMBER GAFF2 Force Field of Aspartate Anion*

```molden
@<TRIPOS>MOLECULE
AAS
  15 14 1 0 0
SMALL
rc

@<TRIPOS>ATOM
  1 CA  -0.5070  0.0520  -0.4690 c3  1 AAS  0.095413
  2 HA  -0.3920  -0.1410  -1.5270 hx  1 AAS  0.018505
  3 C   -1.9540  -0.2940  -0.0090 c  1 AAS  0.687662
  4 O   -2.3740  -1.3740  -0.3880 o  1 AAS -0.730116
  5 CB   0.5690  -0.6970   0.3240 c3  1 AAS -0.041940
  6 HB2  0.3580  -0.5800   1.3870 hc  1 AAS  0.018010
  7 HB3  0.4870  -1.7500   0.0990 hc  1 AAS  0.018010
  8 CG   2.0250  -0.2190   0.0840 c  1 AAS  0.738847
  9 OD1  2.8950  -1.0660   0.0880 o  1 AAS -0.727337
 10 OD2  2.1670  1.0280  -0.0540 o  1 AAS -0.774131
 11 O1  -2.4740  0.5810   0.7010 o  1 AAS -0.756042
 12 N1  -0.3070  1.5200  -0.2470 n4  1 AAS -0.234037
 13 H1  -0.6370  2.0610  -1.0250 hn  1 AAS  0.232801
 14 H3   0.7330  1.6400  -0.1010 hn  1 AAS  0.227089
 15 H2  -0.9070  1.7340  0.5430 hn  1 AAS  0.227266

@<TRIPOS>BOND
  1  1  2  1
  2  1  3  1
  3  1  5  1
  4  1  12 1
  5  3  4  1
  6  3 11 1
  7  5  6  1
  8  5  7  1
  9  5  8  1
 10  8  9  1
 11  8 10 1
 12 12 13 1
 13 12 14 1
 14 12 15 1

@<TRIPOS>SUBSTRUCTURE
  1 AAS  1 TEMP  0 **** **** 0 ROOT
```

183
Figure B5

Validation of Bulk Water Density

Note. Folded-TMG-Asp (purple-dashed), Unfolded-TMG-Asp (green-solid), Folded-TMG-Ser (black-dashed) and Unfolded-TMG-Ser (red-dotted) over the 1.0 ns equilibration MD under NPT.
Figure B6

RMSD of Unfolded Azurin

Note. With reference to the folded X-ray crystal conformation in water (water), TMG-Ser (green) and TMG-Asp (red).
Figure B7

Secondary Structure of Folded Azurin Over 1000 ns

Note. (A) In water; (B) in TMG-Ser; (C) in TMG-Asp. Green represents β-turns, yellow represents an extended β-sheet configuration, olive represents isolated bridges, pink represents α-helices and dark blue represents 3-10 α-helices.
Figure B8

Secondary Structure of Unfolded Azurin Over 1000 ns

Note. (A) In water. (B) In TMG-Ser. (C) In TMG-Asp. Green represents β-turns, yellow represents an extended β-sheet configuration, olive represents isolated bridges, pink represents α-helices and dark blue represents 3-10 α-helices.
Figure B9

Secondary Structure Refolding in Unfolded Azurin

<table>
<thead>
<tr>
<th>Superimposed Conformations</th>
<th>Water</th>
<th>TMG-Ser</th>
<th>TMG-Asp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superimposed Secondary Structure</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Folded and unfolded proteins (top row) were superimposed on the protein residues that recovered their α-helical or β-strand characteristics (bottom row). Folded and unfolded azurin is shown in red and blue ribbons, respectively.
Electrostatic Potential Surface of Azurin

Note. (A) Folded-Ser. (B): Unfolded-Ser. (C): Folded-Asp. (D): Unfolded-Asp. TMG and Ser/Asp anions are colored in blue and red spheres, respectively.